

Optimality and Efficiency of Neighbouring Design

K.N. Ponnuswamy and C. Santharam¹
University of Madras, Madras
(Received : May, 1993)

SUMMARY

This paper deals with the optimality and efficiency of Nearest Neighbour Balanced Block Designs (NNBD) using Generalised Least Square method of estimation for different correlated models (AR(1), AR(2), MA(1), MA(2) and ARMA (1, 1)). NNBD turns out to be optimal for AR(1) model and the performance of NNBD is quite satisfactory for the remaining models. The efficiency of the proposed design in comparison to regular block design is substantial for the above correlated models.

Key words: Auto-regressive model, Moving average model, ARMA model, Correlated errors, Generalised least square estimation, Optimality and efficiency.

1. Introduction

In the analysis of the comparative experiments, it is usually assumed that the observations are independently distributed. The usual methods of local control adopted in comparative experiments such as randomized block design, latin square, BIBD and PBIBD serve the purpose well, when fertility variation is smooth and well known. These methods perform poorly in more realistic situations where fertility variation changes in an unknown fashion. Under these circumstances, it may be useful to adjust for fertility in a more continuous manner such as making use of the inter-dependence among neighbouring observations (Gill and Shukla, [3]).

Usually in comparative experiments the treatments are allocated at random to plots within a block ignoring the order of allocation of the treatments. However in neighbour design the ordering of treatments within blocks and number of times that any two treatments appear in adjacent plots within block are important. These designs are known as one dimensional designs. These designs may be block designs in which the neighbours of a given plot are those of the immediate left and right.

¹ Loyola College, Madras

Street and Street [9] has considered the use of BIBD (one dimensional NNBIBD) for use in seed orchards with $t = (a \text{ prime number})$ of form $6u + 1$, (u is a non negative integer), b, k, r and λ are the parameters. Universal optimality of NNBD using ARMA models has been introduced by Santharam and Ponnuswamy [8]. Efficiencies of NNBD using ARMA models have also been studied by Santharam and Ponnuswamy [8].

In the present paper, we have taken NNBD with the parameter structure $t = 5, b = 5, k = 5, r = 5, \lambda = 1$ and investigated the optimality and efficiency of NNBD for

$$\rho = 0.1 \text{ (0.1) } 0.9 \quad \text{and} \quad (\rho_1, \rho_2) = (.1, .1), (.2, .2) \dots (.9, .9)$$

(where ρ_1, ρ_2 and ρ are the correlation coefficients) when errors behave according to AR(1), MA(1), ARMA (1, 1), AR(2) and MA(2) models.

2. Universal Optimality

Keifer [5] introduced the universal optimality criterion which includes the well known D, A and E optimality criteria as special cases. The sufficient conditions for a design d^* in Δ (where Δ is a class of unary block designs) to be universally optimal are:

(i) the matrix C_d^* is completely symmetric in the sense that all the diagonal elements are the same and all the off diagonal elements are the same.

(ii) $\text{tr}(C_d^*) \geq \text{tr}(C_d)$, for all d in Δ

In some cases, it is not possible to find a design with specified parameters having completely symmetric information matrix. In such cases, it is expected that a design with information matrix nearly completely symmetric and satisfying condition (ii) is nearly optimal.

AR(1), MA(1), and ARMA(1, 1) in
Nearest Neighbour Balanced Block Design (NNBD)

Let Δ be a class of unary block designs for t treatments in which each treatment is applied to r plots, the tr plots being arranged in b blocks, each $k \leq t$, let y be a $tr \times 1$ random vector corresponding to the observations. Consider a fixed effect additive linear model

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

where X is the observation-treatment incidence matrix of order $tr \times t$; Z is the observation-block incidence matrix of order $tr \times b$; τ and β are vectors of

treatment and block effects, respectively; ε is a random error vector representing local variation in soil fertility with $E(\varepsilon) = 0$ and $\text{Var}(\varepsilon) = \sigma_\varepsilon^2 \Omega$; η is an additional error vector with $E(\eta) = 0$ and $\text{Var}(\eta) = \sigma_\eta^2 I$, representing other sources of variability in plots which are independent of local fertility. Therefore, $\text{Var}(y) = \sigma_\varepsilon^2 \Omega + \sigma_\eta^2 I = V$, say.

The model (1) is called an error-in-variables model (Besag [1]) and is closely related to the smooth 'trend plus error' model of Wilkinson *et al.* [11]). This is a general model which gives a better fit in situations where the error structure is non-stationary (Besag [1]; Wilkinson *et al.* [11]; Patterson [6]).

The five correlation models considered for the error vector ε in (1) are the first-order autoregressive model, the moving average model, autoregressive-moving average model, second order auto-regressive and moving average models.

3. Information Matrix

Consider the model $Y = X\tau + Z\beta + \varepsilon + \eta$

The generalised least squares normal equations for estimating τ and β are

$$\begin{aligned} X'V^{-1}X\hat{\tau} + X'V^{-1}Z\hat{\beta} &= X'V^{-1}y \\ Z'V^{-1}X\hat{\tau} + Z'V^{-1}Z\hat{\beta} &= Z'V^{-1}y \end{aligned}$$

Reduced normal equations for $\hat{\tau}$ can be written as

$$C\hat{\tau} = Qy$$

where

$$C = X'V^{-1}X - X'V^{-1}Z(Z'V^{-1}Z)^{-1}Z'V^{-1}X$$

and

$$Q = X'V^{-1} - X'V^{-1}Z(Z'V^{-1}Z)^{-1}Z'V^{-1}X$$

The matrix C (Gill and Shukla [3]) is called the information matrix of a design for treatment parameters. To emphasize the dependence of information matrix on design, we write it as C_d for $d \in \Delta$, row sums and column sums of C_d are zero for each $d \in \Delta$. We assume that $\text{rank}(C_d) = t-1$ for each $d \in \Delta$, so that all the treatment contrasts are estimable.

4. Deviation from Exact Optimality

Since $\text{tr}(C_d)$ is constant over the class Δ , the departure from exact optimality of a design can be measured by the departure of C_d from complete symmetry. This departure is measured by two indices E_D and E_A called D-efficiency and A-efficiency, respectively by Cheng and Wu [2].

Let $\gamma_{d1}, \dots, \gamma_{d,t-1}$ be the non-zero eigenvalues of C_d . The design is universally optimal if all γ_{di} 's are equal. However, such a design for a given set of parameters may not exist. The information matrix of the hypothetical optimal design would have eigenvalue $\gamma = (t-1)^{-1}(\gamma_{d1} + \dots + \gamma_{d,t-1})$ with multiplicity $t-1$. The product $\Pi\gamma_{di} = \gamma_{d1} \dots \gamma_{d,t-1}$ is maximum for this design and is equal to γ^{t-1} . The D-efficiency of a design d relative to the hypothetical universally optimal design can be defined as $E_D = \gamma^{-(t-1)}\Pi\gamma_{di}$. Similarly, the harmonic mean of the γ_{di} 's is a maximum for the hypothetical design and is equal to γ . The A-efficiency of d is defined similarly by $E_A = (t-1)\gamma^{-1}(\Sigma\gamma_{di}^{-1})^{-1}$.

Conversely, an inefficient design is one with highly dissimilar γ_{di} 's. This happens when parameters e_i , μ_{ij} and g_{ij} are different to the maximum possible extent. This is the case with a regular design which has the same spatial configuration of treatments in every block. Therefore, D-efficiency and A-efficiency of a regular design are the lower bounds of the efficiencies.

Table 1 and 2 show the A-efficiency, E_A and the D-efficiency E_D of complete nearest neighbour block design with $t = 5$, $r = 5$ when errors behave under autoregressive, moving average and ARMA processes. These lower bounds are the respective values for regular designs with the same number of treatments and replications. The ratio $\sigma_n^2/\sigma_\epsilon^2 = \alpha$ takes values 0 and $\frac{1}{2}$. An empirical value of α obtained from a series of trials was 0.42 (Patterson [6]).

Conclusion

In this paper we have investigated the optimality of block designs when observations within a block are assumed to be correlated and generalized least square method of estimation has been used, when the errors follow AR(1), MA(1), ARMA (1, 1), AR(2) and MA(2) models. The gain in efficiencies (A and D) of NNBD over regular block design is substantial for all the above models.

Table 1 : A-efficiency and D-efficiency of some nearest neighbour balanced designs ;
lower bounds of efficiencies in parentheses

		(t=5 r=5)															
Error Model	$\rho=0.1$	$\rho=0.2$	$\rho=0.3$	$\rho=0.4$	$\rho=0.5$	$\rho=0.6$	$\rho=0.7$	$\rho=0.8$	$\rho=0.9$								
	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$						
AR(1)																	
E _A	1.000 (.990)	1.000 (.983)	1.000 (.919)	1.000 (.965)	1.000 (.877)	1.000 (.915)	1.000 (.807)	1.000 (.745)	1.000 (.961)	1.000 (.684)	1.000 (.853)	1.000 (.618)	1.000 (.820)	1.000 (.558)	1.000 (.785)		
E _D	1.000 (.980)	1.000 (.925)	1.000 (.848)	1.000 (.933)	1.000 (.758)	1.000 (.890)	1.000 (.666)	1.000 (.844)	1.000 (.578)	1.000 (.924)	1.000 (.497)	1.000 (.796)	1.000 (.699)	1.000 (.364)	1.000 (.654)		
MA(1)																	
E _A	1.000 (.987)	1.000 (.994)	0.998 (.978)	0.977 (.882)	0.994 (.950)	0.943 (.798)	0.986 (.913)	0.885 (.706)	0.974 (.871)	0.807 (.621)	0.957 (.828)	0.721 (.550)	0.938 (.788)	0.647 (.499)	0.917 (.752)	0.598 (.471)	0.899 (.725)
E _D	0.998 (.975)	1.000 (.989)	0.987 (.898)	0.996 (.956)	0.955 (.777)	0.988 (.902)	0.889 (.636)	0.973 (.834)	0.784 (.497)	0.949 (.758)	0.650 (.385)	0.916 (.301)	0.879 (.620)	0.417 (.247)	0.841 (.566)	0.357 (.221)	0.809 (.525)
ARMA (1,1)	$\rho_1=.1$ $\rho_2=.1$	$\rho_1=.2$ $\rho_2=.2$	$\rho_1=.3$ $\rho_2=.3$	$\rho_1=.4$ $\rho_2=.4$	$\rho_1=.5$ $\rho_2=.5$	$\rho_1=.6$ $\rho_2=.6$	$\rho_1=.7$ $\rho_2=.7$	$\rho_1=.8$ $\rho_2=.8$	$\rho_1=.9$ $\rho_2=.9$								
E _A	1.000 (.957)	1.000 (.841)	1.000 (.931)	0.985 (.685)	1.000 (.858)	0.945 (.526)	0.998 (.776)	0.888 (.392)	0.995 (.689)	0.819 (.287)	0.993 (.609)	0.751 (.214)	0.990 (.537)	0.696 (.165)	0.988 (.472)	0.661 (.133)	0.987 (.414)
E _D	1.000 (.917)	1.000 (.707)	1.000 (.868)	0.998 (.467)	0.998 (.741)	0.893 (.272)	0.995 (.611)	0.789 (.147)	0.991 (.491)	0.671 (.076)	0.986 (.392)	0.564 (.041)	0.981 (.317)	0.485 (.024)	0.977 (.257)	0.437 (.016)	0.974 (.211)

**Table 2 : A-efficiency and D-efficiency of some nearest neighbour balanced designs ;
lower bounds of efficiencies in parentheses**

Error Model	(t=5 r=5)													
	$\rho_1 = .1$ $\rho_2 = .1$	$\rho_1 = .2$ $\rho_2 = .2$	$\rho_1 = .3$ $\rho_2 = .3$	$\rho_1 = .4$ $\rho_2 = .4$	$\rho_1 = .5$ $\rho_2 = .5$	$\rho_1 = .6$ $\rho_2 = .6$	$\rho_1 = .7$ $\rho_2 = .7$	$\rho_1 = .8$ $\rho_2 = .8$	$\rho_1 = .9$ $\rho_2 = .9$	$\alpha = 0$	$\alpha = .5$	$\alpha = 1$		
AR(2)														
E_A	1.000 (.959)	1.000 (.852)	1.000 (.707)	1.000 (.552)	1.000 (.355)	0.978 (.683)	.999 (.281)	0.963 (.571)	.998 (.180)	.945 (.448)	.998 (.101)	.924 (.312)	.996 (.043)	0.902 (.164)
E_D	1.000 (.921)	1.000 (.878)	1.000 (.761)	.978 (.326)	.999 (.633)	.956 (.155)	.998 (.505)	.927 (.104)	.997 (.384)	.892 (.052)	.995 (.272)	.854 (.230)	.993 (.171)	.813 (.081)
MA(2)														
E_A	1.000 (.959)	1.000 (.828)	0.997 (.954)	0.986 (.758)	0.997 (.881)	0.935 (.662)	0.987 (.716)	0.976 (.531)	0.976 (.494)	0.873 (.609)	0.976 (.472)	0.855 (.573)	0.947 (.460)	0.933 (.394)
E_D	1.000 (.920)	1.000 (.688)	0.994 (.911)	0.973 (.574)	0.994 (.774)	0.874 (.617)	0.974 (.513)	0.811 (.282)	0.952 (.431)	0.766 (.244)	0.950 (.371)	0.898 (.328)	0.721 (.212)	0.872 (.138)

5. Comparison of Efficiency

In this section a quantitative measure of efficiency of the proposed designs in comparison to randomized block designs is derived. For this we assume that the covariance matrix is known except for some constant multiplier. Since the covariance matrix has to be estimated from the data, the efficiencies given here are the upper limits attainable. We compare the average variance of an elementary treatment contrast $\tau_i - \tau_j$ in both cases.

For a regular block design the variance of an elementary treatment contrast, estimated by ordinary least squares methods, is given by

$$V_1 = 2r^{-1} \sigma^2(1 - \bar{\rho})$$

where σ^2 is the variance of an observation, $\bar{\rho}$ is the average correlation between observations from any two plots within a block, the average being taken over all possible randomizations; r is the number of replications.

If observations within a block follow an errors-in-variable autoregressive model, then (Williams [10])

$$\sigma^2 = \sigma_\varepsilon^2(1 + \alpha(1 - \rho^2)(1 - \rho^2)^{-1})$$

and

$$\bar{\rho} = \frac{2\rho}{\{(1 + \alpha(1 - \rho^2))(t - 1)(1 - \rho)\}} \left\{ 1 - \frac{1 - \rho^t}{t(1 - \rho)} \right\}$$

so that

$$V_1 = \frac{2\sigma_\varepsilon^2}{r(1 - \rho^2)} \left[1 + \alpha(1 - \rho^2) - \frac{2}{(t - 1)(1 - \rho)} \left\{ 1 - \frac{(1 - \rho^t)}{t(1 - \rho)} \right\} \right]$$

When errors follow a first order moving average model, then

$$V_1 = 2r^{-1} \sigma_\varepsilon^2 (1 + \rho^2) \left(1 + \frac{\alpha}{1 + \rho^2} - \frac{2}{(1 + \rho^2)t} \right)$$

When errors follow first order ARMA model, then

$$V_1 = 2r^{-1} \sigma_\varepsilon^2 r_0 \left(1 + \frac{\alpha}{r_0} - \frac{2}{t(t-1)} \sum_{s=1}^{(t-1)} (t-s)\gamma_s \right)$$

where $\alpha = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$

Table 3 : Efficiency of nearest neighbour balanced block designs in comparison to randomized block designs

		(t=5 r=5)																
Error Model	$\rho=0.1$	$\rho=0.2$	$\rho=0.3$	$\rho=0.4$	$\rho=0.5$	$\rho=0.6$	$\rho=0.7$	$\rho=0.8$	$\rho=0.9$									
	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$	$\alpha=0$	$\alpha=.5$						
AR(1)	1.096	1.040	1.016	1.088	1.036	1.154	1.061	1.237	1.092	1.343	1.128	1.468	1.168	1.617	1.214	1.793	1.265	
MA(1)	1.019	1.010	1.062	1.030	1.129	1.059	1.218	1.114	1.325	1.132	1.439	1.186	1.552	1.229	1.651	1.269	1.718	1.302
ARMA (1, 1)	$\rho_1=.1$	$\rho_1=.2$	$\rho_1=.3$	$\rho_1=.4$	$\rho_1=.5$	$\rho_1=.6$	$\rho_1=.7$	$\rho_1=.8$	$\rho_1=.9$									
	$\rho_2=.1$	$\rho_2=.2$	$\rho_2=.3$	$\rho_2=.4$	$\rho_2=.5$	$\rho_2=.6$	$\rho_2=.7$	$\rho_2=.8$	$\rho_2=.9$									
	1.044	1.018	1.184	1.070	1.432	1.162	1.796	1.285	1.927	1.439	2.285	1.625	2.558	1.842	2.722	2.091	2.982	2.360

For generalized least squares estimation, the average variance of an elementary treatment contrast, estimated from a design d , is

$$V_2 = 2(t-1)^{-1} \sum_{i=1}^{t-1} \gamma_{di}^{-1}$$

where γ_{di} 's are non-zero values of C_d (Kempthorne [4]).

We define the efficiency of a design d relative to a randomised block design as V_1/V_2 . Table 3 shows this efficiency of complete nearest neighbour balanced block design with $t = 5$, $r = 5$.

Conclusion

The gain in efficiency of NNBD over regular block design is substantial for all the models considered when the correlation between adjacent plots (ρ) increases. Among the models considered for the elementary treatment contrast, the NNBD under ARMA (1, 1) model shows the highest efficiency.

REFERENCES

- [1] Besag, J., 1977. Errors-in-variables estimation for gaussian lattice schemes. *J. Roy. Statist. Soc.*, B 39, 73-78.
- [2] Cheng, C.S. and Wu, C.F., 1981. Nearly incomplete block designs, *Biometrika*, 68, 493-500.
- [3] Gill, P.S. and Shukla, G.K., 1985. Efficiency of nearest neighbour balanced block designs for correlated observations. *Biometrika*, 72, 3, 639-644.
- [4] Kempthorne, O., 1956. The efficiency factor of an incomplete block design, *Ann. Math. Statist.*, 27, 846-849.
- [5] Keifer, J., 1975. Construction and optimality of generalized youden designs. In: *A Survey of Statistical Designs and Linear Models* (J.N. Srivastava, eds.) 333-353, Amsterdam, North Holland.
- [6] Patterson, H.D., 1983. Discussion of paper by G.N. Wilkinson *et al.*, *J. Roy. Statist. Soc.*, B 45, 178-180.
- [7] Santharam, C. and Ponnuswamy, K.N., 1995. Universal optimality of nearest neighbour balanced block designs using ARMA models, *Statistica* (to appear).
- [8] Santharam, C. and Ponnuswamy, K.N., 1995. On the efficiencies of nearest neighbour balanced block designs with correlated error structure, *Biometrika* (to appear).
- [9] Street, A.P. and Street, D.J., 1987. *Combinatorics of Experimental Design*. Oxford University Press, New York.
- [10] Williams, R.M., 1952. Experimental designs for serially correlated observations. *Biometrika*, 39, 151-167.
- [11] Wilkinson, G.N., Eckert, S.R., Hancock, T.W. and Mayo, O., 1983. Nearest neighbour (NN) analysis of field experiments (with discussion). *J. Roy. Statist. Soc.*, B 45, 151-211.