United Nations Educational Scientific and Cultural Organization
and
International Atomic Energy Agency

THE ABDUS SALAM INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

A NOTE ON DISCRIMINATING EQUALLY OPTIMAL SEMI-LATIN SQUARES FOR SIXTEEN TREATMENTS IN BLOCKS OF SIZE FOUR

P.E. Chigbu [1]
Department of Statistics, University of Nigeria, Nsukka, Nigeria
and
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

Abstract

A semi-Latin square for sixteen treatments in blocks of size four is like a 4×4 Latin square except that there exist four treatments in each cell and each of the sixteen treatments occurs once in each row and once in each column. In the literature, three of this class of squares has been found to be A-, D- and E-optimal while an analytic approach has been adopted to further distinguish these optimal ones with the view of identifying the best for experimentation. With this analytic approach the 'best' square was identified; however, it neither provided a common basis for the discrimination of the three squares nor the further classification of the other two good squares. In this paper, therefore, a numerical approach, which basically involves the computation of the generalized inverses of the information matrices of these squares, is adopted. Each of the generalized inverses satisfies the Moore-Penrose inverse properties. Thereafter, a square is considered most preferable among others if it has the maximum number of minimum variance of simple treatment contrasts as well as the minimum number of distinct pairwise treatment variances. Above all, a mini-league table for the three squares is ascertained.

MIRAMARE – TRIESTE
August 2004

[1] Regular Associate of the Abdus Salam ICTP. pechigbu@yahoo.com, pchigbu@ictp.trieste.it

1 Introduction

A semi-Latin square for sixteen treatments in blocks of size four is an arrangement of sixteen treatments in a (4×4) array in such a way that each row-column intersection contains four treatments while no treatment occurs more than once in each row and each column; this is simply a $(4 \times 4)/4$ semi-Latin square: see, for example, Bailey (1992) and Bailey and Chigbu (1997). The statistical uses and methods of analyzing semi-Latin squares are well-documented in the literature: see, for example, Preece and Freeman (1983) and Bailey (1992). Indeed, they are analyzed as the well-known incomplete-block designs where each row-column intersection is a block. The typical semi-Latin squares we are discriminating in this work are Γ_1, Γ_2 and Γ_3, given in Figures 1, 2 and 3, respectively. The quotient block design of each square is connected and so all simple contrasts are estimable. However, their inherent information matrices are not of full rank.

t_1 t_2 t_3 t_4	t_5 t_6 t_7 t_8	t_9 t_{10} t_{11} t_{12}	t_{13} t_{14} t_{15} t_{16}
t_5 t_{10} t_{15} t_8	t_1 t_{14} t_{11} t_4	t_{13} t_2 t_7 t_{16}	t_9 t_6 t_3 t_{12}
t_9 t_{14} t_7 t_{16}	t_{13} t_{10} t_3 t_{12}	t_1 t_6 t_{15} t_8	t_5 t_2 t_{11} t_4
t_{13} t_6 t_{11} t_{12}	t_9 t_2 t_{15} t_{16}	t_5 t_{14} t_3 t_4	t_1 t_{10} t_7 t_8

Figure 1: Γ_1

t_1 t_2 t_3 t_4	t_5 t_6 t_7 t_8	t_9 t_{10} t_{11} t_{12}	t_{13} t_{14} t_{15} t_{16}
t_5 t_6 t_{11} t_{16}	t_1 t_2 t_{15} t_{12}	t_{13} t_{14} t_3 t_8	t_9 t_{10} t_7 t_4
t_9 t_{10} t_{15} t_8	t_{13} t_{14} t_{11} t_4	t_1 t_2 t_7 t_{16}	t_5 t_6 t_3 t_{12}
t_{13} t_{14} t_7 t_{12}	t_9 t_{10} t_3 t_{16}	t_5 t_6 t_{15} t_4	t_1 t_2 t_{11} t_8

Figure 2: Γ_2

The squares of Figures 1 and 2 were originally found by Bailey (1992) while Chigbu (1995, 1999) established their A-, D- and E-optimality as well as that of Figure 3. They have the same A-, D- and E-optimal values. Thus, in discriminating these equally optimal squares, we note that each of them is equireplicate with constant block size and surely the number of blocks containing any pair of treatments depends on the associate class to which the pair belongs. Each of the three squares is partially balanced with respect to a given association scheme. Their

t_1	t_2	t_5	t_6	t_9	t_{10}	t_{13}	t_{14}
t_3	t_4	t_7	t_8	t_{11}	t_{12}	t_{15}	t_{16}
t_5	t_{10}	t_1	t_{14}	t_{13}	t_2	t_9	t_6
t_{15}	t_8	t_{11}	t_4	t_7	t_{16}	t_3	t_{12}
t_9	t_{14}	t_{13}	t_{10}	t_1	t_6	t_5	t_2
t_7	t_{16}	t_3	t_{12}	t_{15}	t_4	t_{11}	t_8
t_{13}	t_6	t_9	t_2	t_5	t_{14}	t_1	t_{10}
t_{11}	t_{12}	t_{15}	t_{16}	t_3	t_8	t_7	t_4

Figure 3: Γ_3

associate classes, which seem to give some impression of differences in them, are adaptable from their concurrences given in Bailey (1992) and Chigbu (1995, 1999).

Recently, an analytic approach, which basically involved deriving the inverse, K, of some algebraic expression of the information matrix, L, of each of the squares given by $K = \frac{8}{3}\{L^2 - \frac{9}{4}L + \frac{13}{8}I\}$, where I is a conformable identity matrix; and then comparing the variances of the simple contrasts calculated therefrom was adopted to identify Γ_2 as the 'best'. Other results obtained showed different lowest variances of simple contrasts for Γ_1 and Γ_3 and this did not make it quite convenient to further classify and/or discuss their sameness: see Chigbu (2003). Indeed, the analytic approach did not induce a proper ordering among the squares and of course among their corresponding information matrices. Here, we set out to further discriminate these squares with the view of not only identifying the most preferable one for experimentation but also ascertaining the sameness or otherwise of the squares using a common basis. Here, the 'common basis' refers to the condition that the squares under consideration have the same minimum and maximum values of variances of simple treatment contrasts.

2 Methods

Generally, given a matrix, A, of order $(m \times n)$ in some algebraic space of matrices, the Moore-Penrose generalized matrix inverse of A is a unique matrix, A^+, of order $(n \times m)$ which satisfies the following properties:

1. $AA^+A = A$;
2. $A^+AA^+ = A^+$;
3. $(AA^+)' = AA^+$;
4. $(A^+A)' = A^+A$:

see, for example, Penrose (1955), Rao and Mitra (1971) and Ben-Israel and Greville (1977) as well as for other important theories of Moore-Penrose inverses.

Usually, in analyzing an incomplete-block design, each and every treatment contrast, is of possible interest and needs to be estimated and/or compared. The main interest is usually on the variance of the estimator of the contrasts.

Let $\tau' = (\tau_i, \tau_j)$, $t' = (t_i, t_j)$, $c' = (1, -1)$, $c'\tau$ and $c't = (t_i - t_j)$ $\forall$ $i < j$, be a vector of the i^{th} and j^{th} treatments, a vector of the estimates of the i^{th} and j^{th} treatments, the vector of coefficients of a simple contrast which sums to zero, a simple treatment contrast and the estimate of a simple treatment contrast, respectively, for a given connected incomplete-block design whose information matrix is not of full rank, then the variance of the estimate of the simple contrast, $Var(c't)$, is $\sigma^2 c'L^+c$, where L^+ is the conformable generalized inverse of the information matrix, L, of the design. Ignoring the constant, σ^2, the variance of the estimate of a given contrast involving the i^{th} and j^{th} treatments, say, is given by $({L_{ii}}^+ - {L_{ij}}^+ - {L_{ji}}^+ + {L_{jj}}^+)$, where ${L_{ij}}^+$ is the $(ij)^{th}$ entry of L^+, as also given in Chigbu (2003). However, it can easily be shown that when a non-zero multiple of the all-one matrix is added to the information matrix of a connected design, the result is a non-singular matrix. Its inverse is a generalized inverse of the information matrix.

In this work, each generalized inverse, L^+, obtainable by firstly adding an all-one matrix, J, to the information matrix, L, of each of the squares under consideration and then calculating the inverse of the sum of L and J satisfies the above Moore-Penrose inverse properties with respect to the L's and even the $(L + J)$'s matrices. In some algebraic sense, the all-one matrix, in conjunction with an identity matrix of the same size, span some subspace of the real vector space associated with each design. Though, we shall not dwell on this in this work, the all-one matrix is indeed analogous to the sum of all the zero-one matrices of order sixteen that make up the association scheme on the set of sixteen treatments of each of the semi-Latin squares: see, for example, Cameron et al (2003).

The inverse of the information matrices could simply be found using any statistical computing package.

3 Results and Discussion

On the whole, 120 variances of simple treatment contrasts for each semi-Latin square were calculated, compared and used for this discrimination. The maximum and minimum values of variances, corrected to four places of decimal, for Γ_1, Γ_2 and Γ_3 are 0.7500 and 0.5000, respectively. This result is unlike the analytic approach where it was found that the minimum value of variance for Γ_1 is equal to 0.6042 while that of Γ_2 and Γ_3 is 0.5000. The mean of all the 120 values of variance for each of Γ_1, Γ_2 and Γ_3 is 0.6667, which is equal to $\frac{2}{3}$. In other words, the average of the variances of all simple contrasts, known as the efficiency of a design or the E'-optimality, which is equal to 0.6667, is the same for all the squares. Further results on the computed variance values are given below and in Figure 4 as frequency distribution tables.

These tables of frequency distribution of the 120 values of variances for each of the squares would make it easy for an experimenter to appreciate the criteria for discriminating the squares.

Square Γ_1:

There are two simple contrasts with minimum variance of 0.5000 and they are: $(t_1 - t_4)$ and $(t_5 - t_8)$. On the other hand, there exist thirteen simple contrasts with maximum variance of 0.7500 which are :$(t_1 - t_5)$, $(t_1 - t_8)$, $(t_1 - t_12)$, $(t_1 - t_{16})$, $(t_4 - t_5)$, $(t_4 - t_8)$, $(t_4 - t_{12})$, $(t_4 - t_{16})$, $(t_5 - t_{12})$, $(t_5 - t_{16})$, $(t_8 - t_{12})$, $(t_8 - t_{16})$ and $(t_{12} - t_{16})$.

X	F
0.5000	2
0.5990	8
0.6042	4
0.6250	8
0.6406	16
0.6458	8
0.6614	32
0.6666	5
0.6876	8
0.7240	8
0.7292	8
0.7500	13

$for\ \ \Gamma_1$;

X	F
0.5000	4
0.6250	16
0.6562	64
0.6668	12
0.7500	24

$for\ \ \Gamma_2$;

X	F
0.5000	2
0.5990	8
0.6042	4
0.6250	8
0.6406	16
0.6458	8
0.6614	32
0.6666	5
0.6876	8
0.7240	8
0.7292	8
0.7500	13

$for\ \ \Gamma_3$.

Figure 4: Frequency (F) Distribution of values (X) of the variance of the simple contrasts for Γ_1, Γ_2 and Γ_3

Square Γ_2:

There are four simple contrasts with minimum variance of 0.5000 and they are: $(t_1 - t_2)$, $(t_5 - t_6)$, $(t_9 - t_{10})$ and $(t_{13} - t_{14})$; while the number of simple contrasts with maximum variance of 0.7500 is twenty four and they are: $(t_1 - t_5)$, $(t_1 - t_6)$, $(t_1 - t_9)$, $(t_1 - t_{10})$, $(t_1 - t_{13})$, $(t_1 - t_{14})$, $(t_2 - t_5)$, $(t_2 - t_6)$, $(t_2 - t_9)$, $(t_2 - t_{10})$, $(t_2 - t_{13})$, $(t_2 - t_{14})$, $(t_5 - t_9)$, $(t_5 - t_{10})$, $(t_5 - t_{13})$, $(t_5 - t_{14})$, $(t_6 - t_9)$, $(t_6 - t_{13})$, $(t_6 - t_{14})$, $(t_9 - t_{13})$, $(t_9 - t_{14})$, $(t_{10} - t_{13})$ and $(t_{10} - t_{14})$.

Square Γ_3:

There exist two simple contrasts with minimum variance of 0.5000 just like Γ_1. They are: $(t_1 - t_4)$ and $(t_5 - t_8)$. On the other hand, there exist thirteen simple contrasts with maximum variance of 0.7500 and they are :$(t_1 - t_5)$, $(t_1 - t_8)$, $(t_1 - t_12)$, $(t_1 - t_{16})$, $(t_4 - t_5)$, $(t_4 - t_8)$, $(t_4 - t_{12})$, $(t_4 - t_{16})$, $(t_5 - t_{12})$, $(t_5 - t_{16})$, $(t_8 - t_{12})$, $(t_8 - t_{16})$ and $(t_{12} - t_{16})$.

Thus, it can easily be seen from the foregoing results that square Γ_2 has the greater number of simple contrasts with minimum variance than squares Γ_1 and Γ_3; and therefore would be considered different and most preferable for experimentation. Moreover, Cameron et al (2003) gave a criterion for the optimality of designs analyzed as the ones in this work which simply states that a design is optimal if the number of its distinct pairwise treatment variances is fewest when compared with those of the others in the same class with it. Now, as a further step to ascertaining the statistical sameness of Γ_1 and Γ_3 different from the preference of Γ_2 for experimentation, it can easily be seen from the frequency distribution tables of Figure 4, that the number of distinct pairwise treatment variances for Γ_2 is five while that of Γ_1 and Γ_3 are twelve each. The number of pairwise treatment variances for Γ_2 is the fewest among them all and therefore optimal among the three squares under consideration based on this optimality criterion. It is noteworthy that the number of distinct pairwise treatment variances for the ideal balanced incomplete-block design is just one. On the other hand, Γ_1 and Γ_3 are the same, statistically.

4 Conclusion

In consistence with earlier results, among the three squares under consideration, Γ_2 is the most preferable for experimentation while Γ_1 and Γ_3 are the same in many respects and especially with respect to the two discriminating criteria in this work.

Unlike the analytic procedure, the three squares now have a common basis for comparison since each and every one of them have the same minimum and maximum values of variances of simple treatment contrasts.

Furthermore, the sameness of Γ_1 and Γ_3 can now be easily seen in Figures 1 and 3 due to the style of labeling their treatments with t_i's ($i = 1, 2, \ldots, 16$) in which treatment 8 (t_8) and treatment 4 (t_4) are swapped between columns three and four of rows three and four of one of them to get the other; in Figure 4 as their frequency distributions are exactly the same; and as their generalized inverses in this work are exactly the same.

Acknowledgements. My thanks go to the International Atomic Energy Agency and UNESCO for hospitality at the Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy. This work was done during the period of my visiting ICTP in 2003 as a Regular Associate and the financial support of the Swedish International Development Cooperation Agency (SIDA) is very much acknowledged.

References

[1] Bailey, R.A., *Efficient semi-Latin squares*, Statistica Sinica, **2**, 1992, 413–437.

[2] Bailey, R.A. and Chigbu, P.E., *Enumeration of semi-Latin squares*, Discrete Maths., **167/168**, 1997, 73–84.

[3] Ben-Israel, A. and Greville, T.N.E., *Generalized inverses: theory and applications*, Wiley, New York, 1977.

[4] Cameron, P., Dobcsányi, P., Morgan, J.P. and Soicher, L. (eds.), *The external representation of block designs*, Version **1.0rc1**: http://localhost/ext-rep.html, 2003.

[5] Chigbu, P.E., *Semi-Latin squares: methods for enumeration and comparison*, Ph.D. Thesis, University of London, 1995.

[6] Chigbu, P.E., *Optimal semi-Latin squares for sixteen treatments in blocks of size four*, Journ. Nig. Stat. Assoc., **13**, 1999, 11–25.

[7] Chigbu, P.E., *The 'Best' of the three optimal* $(4 \times 4)/4$ *semi-Latin squares*, Sankhyā: The Indian Journal of Statistics, **65 (3)**, 2003, 641–648.

[8] Penrose, R., *A generalized inverse for matrices*, Proc. Cambridge Phil. Soc., **51**, 1955, 406–413.

[9] Preece, D.A. and Freeman, G.H., *Semi-Latin squares and related designs*, J. Roy. Statist. Soc., **Ser. B 45**, 1983, 267–277.

[10] Rao, C.R. and Mitra, S.K., *Generalized inverse of matrices and its applications*, Wiley, New York, 1971.