Current Trends on Biostatistics & Biometrics
DOI: 10.32474/CTBB.2020.02.000131
ISSN: 2644-1381
Review article

Semi-Balanced Array and Designs for Double Cross

Sharma M K[1]* and Mohammad O Ibrahim[2]*

[1]Department of Statistics, Addis Ababa University, Addis Ababa, Ethiopia

[2]Department of Statistics, Addis Ababa University, Addis Ababa, Ethiopia

***Corresponding author:** M.K. Sharma, Department of Statistics, Addis Ababa University, Addis Ababa, Ethiopia

Received: January 10, 2020 **Published:** January 22, 2020

Abstract

For studying quantitative inheritance and developing hybrid programs, it is useful to examine single cross or triallel cross or double cross. In this article we are presenting the method of construction of partial double cross designs using Semi-balanced arrays.

Keywords: Complete diallel cross; partial double cross

Introduction

Mating design involving partial diallel crosses lines play very important role to study the genetic properties of a set of inbred lines in plant breeding experiments. A partial double cross design is obtained by crossing two unrelated F_1 hybrids symbolized as $(i*j)$ and $(k*l)$, where *i, j, k*, and *l* are denoting the grandparents and no two of them are the same and $(i*j)$ and $(k*l)$ are two F_1's {see Rawlings and Cockerham (1962b) and Hinkelmann (1968, 1975)}. For four lines *i, jj, k, and l,* the three distinct cross are $(i*j)(k*l)$ $(i*k)(j*l)$ and $(i*l)(j*k)$. The total number of distinct double cross for p lines is $3\binom{p}{4}$. In a double cross the hybrid effects consists of average effect of lines, two lines effects, three lines effects and four lines specific effects. In the absence of dominance and epistatic interaction effects, ordering of lines is not important. Therefore, the total number of double cross without considering the order of the lines is $\binom{p}{4}$.

Let n_c denote the total number of crosses (experimental units) involved in an partial double cross experiment. Generally partial double cross experiments are conducted using a completely randomized design (CRD) or a randomized complete block (RCB) design involving some or all n_c crosses as treatments. The number of crosses in such a mating design increases rapidly with increase in the number of lines. It leads to an overall inefficient experiment. It is for this reason that the use of incomplete block design as environment design is needed for partial double cross experiments (see for references Singh et al. (2012)).

Parsad *et al.* (2005) constructed optimal block designs for partial double cross experiments by using balanced incomplete block designs and nested balanced incomplete block designs of Morgan et al. (2001). Sharma and Tadesse (2016) constructed double cross designs for even and odd value of *p* by using initial block of unreduced balanced incomplete block designs of Bose et al. (1953) and initial block of row –column designs by and Gupta and Choi (1998), respectively [1-6].

In this paper we are using semi-balanced arrays for the construction of block design for partial double crosses which follows the condition of Winkelman(1963) We have considered the experimental model that includes the gca effects, apart from block effects, but no specific combining ability effects. The paper is organized as follows: in section 2, definition has been provided. In section 3 and 4, experimental model of the design in 1-way setting and relation between semi balanced array $(p(p-1)/2, p, p, 2)$ and partial double crosses and optimality has been discussed.

Definitions

Definition: The double cross has been defined by Rawlings and Cockerham (1962 b) as a cross between two unrelated F_1

hybrids, say denoted by $(i*j)$ and $(k*l)$, where $i \neq j \neq k \neq l \neq i$, are denoting the gr and parentsand no two of them are same. Ignoring reciprocal crosses, with *p* grandparents, there will be 3 $\binom{p}{4}$ different possible double crosses from p lines.k

Definition: According to Hinkemann (1963) a set of matings is said to be a partial double cross if the following conditions are satisfied:

(i) Every line occurs exactly r times in the set,

(ii) Every four-way cross occurs either once or not at all.

Definition: According to Hinkemann (1963) a set of matings is said to be a partial double cross in strict sense, if it satisfies the condition of Definition 2.2 and

(iii) every single cross occurs once or not at all in the set.

Experimental Modelfor 1-Way- Heterogeneity

Let *d* be a block design for an 4- allel cross experiment involving *p* inbred lines, *b* blocks each of size *k.* This means that there are k crosses in each of the blocks of *d*. Further, let s_{di} denotes the number of replications of the i[th] line in different crosses in $d[i=1,2,...,p]$. Evidently, $\Sigma_{si} = mbk = mn$ and $n = bk$, the total number of observations. Now we take the following additive experimental model for the observations obtained from design *d* [7-14].

$$y = \mu 1_n + \Delta_1^{'} g + \Delta_2^{'} \beta + e \quad (3.1)$$

where **y** is an $n*1$ vector of observations, 1 is an $n*1$ vector of ones, $\ddot{A}_1'$ is an n × p dessign matrix for lines and $\ddot{A}_2'$ is an $n*b$ design matrix for blocks, that is, the $(h,l)^{th}$ element of $\ddot{A}_1'$ (also of $\ddot{A}_2'$) is 1 if the h^{th} observation pertains to the l^{th} line (also of block) and is zero otherwise. μ is a general mean, **g** is a $(p*1)$ vector of line parameters, β is a $b*1$ vector of block parameters and **e** is an $n*1$ vector of residuals. It is assumed that vector β is fixed and **e** is normally distributed with mean **0** and Var $(e) = \sigma^2 I$ with Cov $(\beta, e) = 0$, where **I** is the identity matrix of conformable order.

For the analysis of proposed design *d,* the method of least squares leads to the following reduced normal equations for estimating the linear functions of the gca effects of lines under model (3.1).

$$C_d = G_d - N_d K_d^{-1} N_d^{-1} = (c_{ij}) \; (i,j = 1,2,....p) \quad (3.2)$$

where $G_d = \Delta_1 \Delta_2^{'} = g_{dij}, g_{dij} = s_{dij}$ and for $i \neq i^{'}$, g_{dij} is the number of crosses in d in which the lines *i* and *i′* appear together. $N_d = \Delta_1 \Delta_2^{'} = (n_{dij}), n_{dij}$ is the number of times the line *i* occurs in block *j* of *d* and $k_d = \Delta_2 \Delta_2^{'}$ is the diagonal matrix of block sizes.

A design *d* will be called connected if and only if all elementary comparison among gca effects are estimable using *d* or rank (C) must be equal to $(v-1)$. We denote by $D(p,b,k)$, the class of all such connected block design {*d*} with *p* lines, *b* blocks each of size *k*.

Theorem

Following Kiefer (1975) **f**or any design $d \varepsilon D(p,b,k)$

$$tr(c_d) \leq k^{-1} b \{4k(k-1-2x) + px(x+1)\}$$

where $x = \left[{mk}/{p} \right]$ and for a square matrix *A*, tr (*A*) stands for the sum of the diagonal elements of *A*.

Kiefer (1975) showed that a design is universally optimal in a relevant class of competing design if (i) the information matrix (the $\mathbf{C}_d$ – matrix) of the design is completely symmetric in the sense that $\mathbf{C}_d$ has all its diagonal elements equal and all of its off-diagonal elements equal, and (ii) the matrix $\mathbf{C}_d$ has maximum trace in the class of competing designs, that is, such a design minimize the average variance of the best linear unbiased estimators of all elementary contrasts among the parameters of interest i.e. the general combining ability in our context. On the basis of the theorem 3.1 and Kiefer criterion of the universal optimality, we can state following Theorem.

Theorem

For any design $d * \varepsilon D(p,b,k)$ be a block design for 4-allel crosses satisfying

(i) $tr(c_{d*}) = k^{-1} b \{4k(k-1-2x) + px(x+1)\}$

(ii) $c_{d*} = (p-1)^{-1} k^{-1} b \{4k(k-1-2x) + px(x+1)\} (I_p - 1_p 1^{'}{}_p {}^{P}/_{P})$ is completely symmetric.

where **I**p is an identity matrix of order p and **1**p **1**′p is a $P \times P$ matrix of all ones. Furthermore, using d^* all elementary contrasts among *gca* effects are estimated with variance

$[2b^{-1}(p-1)k / \{4k(k-1-2x) + px(x+1)\}]\sigma^2$ Then *d** is universally optimal in $D(p,b,k)$, and in particular minimizes the average variance of the best linear unbiased estimator of all elementary contrasts among the general combining ability effects.

Relation Between Semi Balanced Array ((*p*(*p*-1)/2,*p*,*p*2) and Partial Double Crosses and Optimality

Rao (1973) gave method of construction of semi balanced array $(p(p-1)/2, p, p, 2)$ Consider now a semi-balanced $(p(p-1)/2, p, p, 2)$ where p is an odd prime or power of odd prime. A semi-balanced array can be divided into $\left({p-1}/{2} \right)$ groups. If we identify the elements of groups as lines of a double cross experiment and ignoring the first row of each group and perform cross among the first element of the second row of each group, we get the first double cross of the mating design. Similarly perform cross among second element of the second row with each group; we get second double cross of the mating design. Thus repeating the above operation for third element of first row with the corresponding elements of each group and so on, we get first row of the mating design for double cross experiment. Similarly we can repeat the operation over other elements of other rows. Thus we obtain the base of mating design for partial double cross experiment which contains $p(p-1)$ double cross in $(p-1)$ rows and p columns and where each cross is replicated $r=(p-1)/2$ times in mating design. By bifurcating the mating design into r equal parts i.e. taking alternative blocks together such as 1st and 3rd, 2nd and 4th and so on , we obtain r designs for partial double cross experiments with parameters $v=2p, b=2$, and $k=p$ and where each line is replicated 2r times in each r design. These designs fulfill the condition of definition 2.3.

The design $d_k \in D(p,b,k)$ where $k=1,2,3$, and 4. Now following (3.2) the information matrix of C_{dk} , where $k=1,2,3$, and 4.

$$C_{dk} = 2(p-2(I_p - \frac{p^2-4p+1}{2p(p-2)} 1_p 1'_p) \quad (5.1)$$

where $\boldsymbol{I}_p$ is an identity matrix of order p and $\mathbf{1}_p$ is a unit column vector of ones. Clearly C_{dk} given by (6.1) is completely symmetric and $t_r(c_{dk})$ is $2(p-1)(p-2)$ which is greater than the upper bound given in Theorem 3.2 (i). Hence the designs $d_k \in D(p,b,k)$ is A-optimal and also minimizes the average variance in comparison to a variance given in (3.2).

Example

For the construction of semi-balanced array (36, 9, 9, 2), we take GF (3^2). The minimum function of GF (3^2) is $X^2 + X = 2$ and power cycle GF (3^2) is $X^0 = 1, X = X, X^2 = 2X+1, X^3 = 2X+2, X^4 = 2$, where x is primitive root of GF (3^2). The residue classes $0,1,2,X,X+1,X+2,2X,2X+1,2X+2$ form a field. We write the 9 elements of GF (3^2) as $0,\pm1,\pm2,\pm X and \pm X+1$ and hence using the formula (4.1), the four key sets are as given below.

1. $0,1,2,X,X+1,X+2,2X,2X+1,2X+2$
2. $0,2,1,2X+2,2X+1,X,X+2,X+1$
3. $0,X,2X,2X+1,1,X+1,X+2,2X+2,2$
4. $0,X+1,2X+2,1,X+2,2X,2,X,2X+1$

where second, third and fourth vectors are obtained from the first on multiplying by 2, x and $X+1$, respectively. Writing these columns vertically and generating the other columns by the addition of elements GF (3^2) as indicated in (4.3). We obtain 36 columns as shown below which is divided into four groups.

Group 1

0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
1	2	0	$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$
2	0	1	$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$
x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$	0	1	2
$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$	1	2	0
$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$	2	0	1
$2x$	$2x+1$	$2x+2$	0	1	2	x	$x+1$	$x+2$
$2x+1$	$2x+2$	$2x$	1	2	0	$x+1$	$x+2$	x
$2x+2$	$2x$	$2x+1$	2	0	1	$x+2$	x	$x+1$

Group 2

0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
2	0	1	$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$
1	2	0	$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$
$2x$	$2x+1$	$2x+2$	0	1	2	x	$x+1$	$x+2$
$2x+2$	$2x$	$2x+1$	2	0	1	$x+2$	x	$x+1$
$2x+1$	$2x+2$	$2x$	1	2	0	$x+1$	$x+2$	x
x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$	0	1	2
$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$	2	0	1
$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$	1	2	0

Group3

0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$	0	1	2
$2x$	$2x+1$	$2x+2$	0	1	2	x	$x+1$	$x+2$
$2x+1$	$2x+2$	$2x$	1	2	0	$x+1$	$x+2$	x
1	2	0	$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$
$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$	1	2	0
$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$	2	0	1
$2x+2$	$2x$	$2x+1$	2	0	1	$x+2$	x	$x+1$
2	0	1	$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$

Group 4

0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$	1	2	0
$2x+2$	$2x$	$2x+1$	2	0	1	$x+2$	x	$x+1$
1	2	0	$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$
$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$	2	0	1
$2x$	$2x+1$	$2x+2$	0	1	2	x	$x+1$	$x+2$
2	0	1	$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$
x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$	0	1	2
$2x+1$	$2x+2$	$2x$	1	2	0	$x+1$	$x+2$	x

Now we relabeled the lines 0-8, using the correspondence $0 \to 0, 1 \to 1, 2 \to 2, X \to 3, X+1 \to 4, X+2 \to 5, 2X \to 6, 2X+1 \to 7, 2X+2 \to 8$, we obtain relabeled groups as given below.

Group 1 (relabeled)

0	1	2	3	4	5	6	7	8
1	2	0	4	5	3	7	8	6
2	0	1	5	3	4	8	6	7
3	4	5	6	7	8	0	1	2
4	5	3	7	8	6	1	2	0
5	3	4	8	6	7	2	0	1
6	7	8	0	1	2	3	4	5
7	8	6	1	2	0	4	5	3
8	6	7	2	0	1	5	3	4

Group2 (relabeled)

0	1	2	3	4	5	6	7	8
2	0	1	5	3	4	8	6	7
1	2	0	4	5	3	7	8	6
6	7	8	0	1	2	3	4	5
8	6	7	2	0	1	5	3	4
7	8	6	1	2	0	4	5	3
3	4	5	6	1	8	0	1	2
5	3	4	8	6	7	2	0	1
4	5	3	7	8	6	1	2	0

Group 3 (relabeled)

0	1	2	3	4	5	6	7	8
3	4	5	6	7	8	0	1	2
6	7	8	0	1	2	3	4	5
7	8	6	1	2	0	4	5	3
1	2	0	4	5	3	7	8	6
4	5	3	7	8	6	1	2	0
5	3	4	8	6	7	2	0	1
8	6	7	2	0	1	5	3	4
2	0	1	5	3	4	8	6	7

Group 4(relabeled)

0	1	2	3	4	5	6	7	8
4	5	3	7	8	6	1	2	0
8	6	7	2	0	1	5	3	4
1	2	0	4	5	3	7	8	6
5	3	4	8	6	7	2	0	1
6	7	8	0	1	2	3	4	5
2	0	1	5	3	4	8	6	7
3	4	5	6	7	8	0	1	2
7	8	6	1	2	0	4	5	3

Now ignoring the first row of each of the four group and perform cross between the first elements of the second row of each group, we get the first double cross of the mating design. Similarly perform cross between second element of the second row with corresponding elements from each group; we get second double cross of the mating design. Thus, repeating the above operation for third element of the first row with the corresponding elements of each group and so on, we get first row of the mating design for double cross experiment. Similarly, we can repeat the operation on the elements of other rows. Thus, we obtain the base of mating design consisting 72 double cross in 8 rows and 9 columns, where each cross is replicated 4 times. The above design will serve as a base for constructing designs for partial double crosses. Considering rows as blocks and bifurcating the block design into two equal parts i.e. taking B_1, B_2, B_3, and B_4 blocks in one part and B_5, B_6, B_7, and B_8 blocks in second part. Now we can divide the blocks B_1, B_2, B_3, and B_4 and B_5, B_6, B_7, and B_8 into four parts as given below:

In first part B_1 and B_3, in second part B_3 and B_4, in third part B_5 and B_7 and in fourth part B_6 and B_8, in this way we get four more designs d_1, d_2, d_3 and d_4 with same parameters $v = 18, b = 2, k = 9 and r = 1$, respectively. These designs fulfill the condition of definition 2.3.

The block design constructed using semi –balanced array (36, 9, 9, 2) are as given below. The trace and variance of the proposed A- optimal design $d_k\left(k = 1.....4\right)$ is 56 and variances is 0.2856, respectively while trace and variance of optimal design due to theorem 3.2, is 40 and 0.4, respectively. Hence our proposed designs are universally optimal.

Base Partial Double Cross mating design

$B1$(1,2)(3,4)	(2,0)(4,5)	(0,1)(5,3)	(4,5)(6,7)	(5,3)(7,8)	(3,4)(8,6)	(7,8)(0,1)	(8,6)(1,2)	(6,7)(2,0)
$B2$(2,1)(6,8)	(0,2)(7,6)	(1,0)(8,7)	(5,4)(0,2)	(3,5)(1,0)	(4,3)(2,1)	(8,7)(3,5)	(6,8)(4,3)	(7,6)(5,4)
$B3$(3,6)(7,1)	(4,7)(8,2)	(5,8)(6,0)	(6,0)(1,4)	(7,1)(2,5)	(8,2)(0,3)	(0,3)(4,7)	(1,4)(5,8)	(2,5)(3,6)
$B4$(4,8)(1,5)	(5,6)(2,3)	(3,7)(0,4)	(7,2)(4,8)	(8,0)(5,6)	((6,1)(3,7)	(1,5)(7,2)	(2,3)(8,0)	(0,4)(6,1)
$B5$(5,7)(4,6)	(3,8)(5,7)	(4,6)(3,8)	(8,1)(7,0)	(6,2)(8,1)	(7,0)(6,2)	(2,4)((1,3)	(0,5)(2,4)	(1,3)(0,5)
$B6$(6,3)(5,2)	(7,4)(3,0)	(8,5)(4,1)	(0,6)(8,5)	(1,7)(6,3)	(2,8)(7,4)	(3,0)(2,8)	(4,1)(0,6)	(5,2)(1,7)
$B7$(7,5)(8,3)	(8,3)(6,4)	(6,4)(7,5)	(1,8)(2,6)	(2,6)(0,7)	(0,7)(1,8)	(4,2)(5,0)	(5,0)((3,1)	(3,1)(4,2)
$B8$(8,4)(2,7)	(6,5)(0,8)	(7,3)(1,6)	(2,7)(5,1)	(0,8)(3,2)	(1,6)(4,0)	(5,1)(8,4)	(3,2)(6,5)	(4,0)(7,3)

Derived designs from base design

Design d_1

B_1 (12)(34) (20)(45) (01)((53) (45)(67) (53)(78) (34)(86) (78)(01) (86)(12) (67)(20)
B_2 (36)(71) (47)(82)) (58)(60)) (60)(14) (71)(25) (82)(03) (03)(47) (14)(58) (25)(36)

Design d_2

B_1 (21)(68) (02)((76) (10)(87) (54)(02) (35)(10) (43)(21) (87)(35) (68)(43) (76)(54)
B_2 (48)(15) (56)(23) (37)(04) (72)(48) (80)(56) (61)(37) (15)(72) (23)(80) (04)(61)

Design d_3

B_1 (57)(46) (38)(57) ((46)(38) (81)(70) (62)(81) (70)(62) (24)((13) (05)(24) (13)(05)
B_2 (75)(83) (83)(64) (64)(75) (18)(26) (26)(07) (07)((18) (42)(50) (50)(31) (31)(42)

Design d_4

B_1 (63)(52) (74)(30) (85)(41) ((06)(85) (17)(63) (28)(74) (30)(28) (41)(06) (52)(17)
B_2 (84)(27) (65)(08) (73)(16) (27)(51) (08)(32) (16)(40) (51)(84) (32)((65) (40)(73)

Table 1: Parameters of new A- optimal block designs.

p	n	v	b	k	Array	Fraction
32	18	18	2	9	(36, 9, 9, 2)	Feb-21
11	22	22	2	11	(55, 11, 11, 2)	Jan-36
13	26	26	2	13	(78, 13, 13, 2)	Feb-55
17	34	34	2	17	(136,17,17,2)	2/105
19	38	38	2	19	(171,19,19,2)	Jan-68
23	46	46	2	23	(253,23,23,23,2)	1/105
33	54	54	2	27	(351,27,27,2)	1/150
29	58	58	2	29	(406,29,29,2)	2/351

Conclusion

In this article, we have constructed block designs for partial double cross experiments using semi-balanced array $\left(p\left(p-1\right)/2, p.p, 2\right)$,where p is an odd prime or power of an odd prime. In these designs each line is replicated 2r times. These designs are found to be A- optimal and follow the condition of Hinkelmann (1963).

References

1. Bose RC, Shrikhande SS, Bhattacharya KN (1953) On the construction of group divisible incomplete block designs. Ann Math Stat 24: 167-195.
2. Griffing B (1956) Concept of General and Specific Combining ability in relation to diallel crossing systems. Aust. J Biol Sci 9: 463-493.
3. Gupta S, Choi KC (1998) Optimal row-column designs for complete diallel crosses, Communication in Statistics-Theory and Methods 27(11): 2827-2835.
4. Hinkelmann K (1963) Design and analysis of muti-way genetic cross experiments. Ph.D thesis submitted to IOWA State University of Science and Technology, USA.
5. Hinkelmann K (1968) Partial quadriallel crosses. Theory Appl Genetics 38: 85-89.
6. Hinkelmann K (1975) Design of genetical experiment. In a Survey of Statistical design and linear models, (Ed) JN Srivastava, (PP).243-269. Amesterdam, North Hollond, Netherlands.
7. Kiefer J (1975) Construction and optimality of generalized Youden designs. In a Survey of Statistical Design and Linear Models. (ED) J N Srivastava. (PP)333-353. Amsterdam, North Holland, Netherlands.
8. MorganJP, Preece DA, Rees, DH (2001) Nested balanced incomplete block designs. Discrete Mathematics 231: 351-389.
9. Parsad R, Gupta, VK, Gupta, Sudhir (2005) Optimal Designs for Diallel and Double Cross Experiments. Utilitas Mathematica 68: 11-32.
10. Rao CR (1961) Combinatorial arrangements analogus to orthogonal arrays. Sankhya 23: 283-286.
11. Rao C R (1973) Some Combinatorial Problems of Arrays and Application to Design of Experiments. A survey to Design of experiments, North Holland, Netherlands, 29: 349-359.
12. Rawlings JO, Cockerham (1962) Analysis of double cross hybrid populations. Biometrics 18(2): 229-224.
13. Sharma M K, Tadesse M (2016) Optimal block designs for double cross experiments. Communication in Statistics-Theory and Methods 45(15): 4392-4396.
14. Singh M, Gupta S, Parsad R (2012) Genetic Cross Experiments, John Wiley & Sons Inc.

DOI: 10.32474/CTBB.2020.02.000131