Theor Appl Genet (1990) 79:849-852

THEORETICAL AND APPLIED GENETICS

Asymptotically efficient partial diallel crosses

A.S. Arya [1] and P. Narain [2]

[1] Department of Mathematics and Statistics, Haryana Agricultural University, Hisar 125 004, India
[2] Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi 110 012, India

Received March 25, 1989; Accepted December 12, 1989
Communicated by B. R. Murty

Summary. The asymptotic efficiency of partial diallel crosses is defined, and circulant plans having such a property are identified. In the absence of an optimal plan to suit his requirements, the breeder should opt for suggested plans to derive maximum information for given resources.

Key words: Partial diallel cross – Circulant plan – Combining ability – Optimal and asymptotically efficient plans

Introduction

Partial diallel experiments are successfully being used to screen a large number of inbred lines or homozygous parents for their combining abilities at manageable and low levels of resources. These experiments give information on the genetic architecture of the breeding material to aid in selecting appropriate strategies for further breeding programs. For example, one might like to select the best four lines, out of the available stock, to develop a four-way cross. A partial diallel cross (PDC) consists in raising a set of $NS/2$ crosses in some specified manner, where each of the N lines is made into crosses with S other lines, where N and S are integers (≥ 2) such that both are not odd.

Such plans were constructed using different association schemes, such as triangular and rectangular (Fyfe and Gilbert 1963), group divisible (Hinkelmann and Kempthorne 1963; Arya and Narain 1977), simple rectangular lattice (Arya and Narain 1977), and truncated triangular association schemes (Narain and Arya 1981). The drawback with these plans is that they can only be constructed for very limited combinations of N and S depending upon the factorization of N and, as such, no plans are available when N is prime. Plans based on cyclic association schemes, known as circulant designs (Kempthorne and Curnow 1961; Curnow 1963; Arya 1983) can, however, be constructed for all combinations of N and S, except when both are odd. While plans given by Kempthorne and Curnow (1961) are termed central circulant plans, those of Arya (1983) are known as border circulant plans.

A PDC, with $NS/2$ crosses, constitutes a $S/(N-1)$th fraction of the complete diallel cross (CDC) consisting of $N(N-1)/2$ crosses. Two PDCs will, therefore, in general differ in their efficiencies. In order to make the best use of available resources, one would naturally like to select a PDC plan that is optimum in some sense. But the search for such optimal plans can only be empirical by trying out different values of N and S, and working out the average variance of general combining ability (GCA) comparisons. In this paper, therefore, we introduce a concept of asymptotic efficiency (AE), which can help judge the superiority of one plan over the other.

Measure of efficiency

Let the yield of the cross $(i \times j)$ be modelled as

$$Y_{ij} = m + g_i + g_j + e_{ij} \quad (i, j = 1, \ldots, N; \quad i<j) \tag{1}$$

where m is the general mean, g_i and g_j are the GCA effects of the line i (e.g., male) and j (female), and e_{ij} is a residual comprising the specific combining ability (SCA) effects and the experimental error. The e_{ij}'s are assumed to be independently distributed with mean zero and variance σ^2. The normal equations for estimating the g's are, then:

$$A\hat{g} = Q, \tag{2}$$

where $A=(a_{ij})$, $\hat{g}'=(\hat{g}_1,\dots,\hat{g}_N)$, $Q'=(Q_1,\dots,Q_N)$; $a_{ii}=S$ for all i, $a_{ij}=a_{ji}=1$ if cross $(i\times j)$ is carried out or 0 otherwise; $Q_i=\sum_{j(i)} Y_{ij}-(2G/N)$ wherein $\sum_{j(i)}$ refers to summing over lines j crossed with line i, and G is the total of the character over all crosses.

The average variance over all the possible comparisons $(\hat{g}_i-\hat{g}_j)$ of the GCA effects is known to be (Appendix to Kempthorne and Curnow 1961; Arya 1983):

$$\text{Av.var.}=\text{Av.var.}\ (\hat{g}_i-\hat{g}_j)=(2NSa^o-1)\,\sigma_o^2/S(N-1) \quad (3)$$

where a^o is the diagonal element in A^{-1} and σ_o^2 is the proper error variance. This average variance thus depends on N, S and a^o. Since a^o depends upon the matrix of the sampled crosses, the average variance given in Eq. (3) will, in general, be different for two PDC plans, e.g., A and B, even if they have the same N and the same S. The ratio

$$E=\frac{\text{av. var. under plan B}}{\text{av. var. under plan A}} \quad (4)$$

can, therefore, be taken as a measure of efficiency of plan A compared with plan B, provided the two are comparable. To emphasize, we regard the two plans as comparable only when they have the same N and the same S. As such, two plans using the same number of crosses ($NS/2$) are, in general, not comparable. The average variance is found to be more sensitive to variations in S than in N. Kempthorne and Curnow (1961), Levings and Dudley (1963), and Kearsey (1965) have addressed the problem on this basis. In this study we will take plan A as the best or optimum if it leads to the least average variance among the class of all comparable plans with the same values of N and S. Thus, the reciprocal of the average variance can be taken as the efficiency point of a given plan.

Now, in order to search for the optimum plan one has to compute the average variance of all the PDCs for a given N and S. This, however, leads to quite cumbersome and huge computations, especially where N and S are large. For example, when $N=31$ and $S=12$, one has to deal with as many as 5,005 PDCs. All optimum plans listed in Mathur and Narain (1976) and the literature reviewed herein were identified empirically in this manner, using Eqs. (3) and (4) either as such or some functions thereof. However, in view of Arya (1989), the above 5,005 PDCs can be classified into only 335 distinct classes, which is still a large number to handle. Hence, it is necessary to introduce some additional criterion that could help us in selecting a better plan with relative ease.

Asymptotic efficiency

The efficiency factors (E's) of PDC plans over some standard ones can be studied from three different angles: (i) for patterns in values of E for all N and S; (ii) for patterns in average variance of the two plans for variations in S while keeping N fixed; and (iii) for patterns in average variance of the two plans for variations in N, S being kept fixed.

Of these, if we concentrate on the pattern given in (iii) and find that, for fixed S, the average variance exhibits a consistently decreasing trend for increase in N, we are led to otpimality in the asymptotic sense. Compared to some given standard, then, increase in N would lead to increase in efficiency of the given plan, but makes the standard plan more inefficient. The increase in efficiency with increase in N, for all S may, therefore, be regarded as asymptotic efficiency (AE). We therefore have the following definitions.

Definition 1. A plan showing increasing efficiency with increase in N, for all S, is said to be asymptotically efficient (AE).

Definition 2. An AE plan with maximum efficiency for all N and all S is said to be the best asymptotically efficient (BAE) plan.

Asymptotic efficiency of border circulant plans

The efficiency factors of border plans over central ones, along with their average variances, are listed in Table 1 for several values of N and S. Studying this table from the three different angles as given in the 'Asymptotic efficiency' section above, reveals the following.

Under (i) it will be noted that $E\geq 1$ for all combinations of N and S, implying that the border plan is always better than the central one. Furthermore, the efficiency is higher for smaller values of S and for larger values of N as well. This indicates that the choice of the border plan, as opposed to the central one, will safeguard against the low precision obtainable otherwise for lower values of S and large N, for a given amount of resources (number of crosses). In other words, for given input, more information would be obtained by choosing the border plan as opposed to the central one.

Regarding point (ii), irrrespective of N and any plan used, the average variance decreases with increase in S, and attains its minimum of $2\,\sigma^2/(N\text{-}2)$ when $S=N-1$, a case with the CDC. But as we move down from a higher value of S to a lower one, the increase in average variance is more steep under the central plan than under the border one. This fact makes the latter plan far more efficient than the former for lower values of S, as already pointed out under (i).

It is important for the breeder to know the minimum number of crosses necessary to provide the desired information efficiently since, in many cases, crosses are extremely difficult to make. This requires a judicious decision for the value of S. Murty et al. (1967) and Anand

Table 1. Average variances/σ^2 (V_1 for border plan, V_2 for central plan) and efficiency ($E = V_2/V_1$) of border plan over central one for some values of N and S

N odd and S even

N	$S=4$			$S=6$			$S=8$			$S=10$			$S=12$		
	V_1	V_2	E	V_1	V_2	E	V_1	V_2	E	V_1	V_2	E	V_1	V_2	E
7	0.713	0.713	1.00												
9	0.674	0.806	1.20	0.408	0.408	1.00									
11	0.670	0.910	1.36	0.402	0.426	1.06	0.288	0.288	1.00						
13	0.662	1.005	1.52	0.397	0.453	1.14	0.288	0.294	1.02	0.223	0.223	1.00			
15	0.658	1.105	1.68	0.395	0.481	1.22	0.284	0.302	1.06	0.223	0.225	1.01	0.182	0.182	1.00
17	0.655	1.204	1.84	0.393	0.510	1.30	0.283	0.313	1.11	0.222	0.229	1.03	0.182	0.184	1.01
19	0.652	1.303	2.00	0.392	0.538	1.37	0.282	0.325	1.15	0.221	0.234	1.06	0.182	0.185	1.02
21	0.650	1.403	2.16	0.391	0.565	1.45	0.281	0.337	1.20	0.220	0.239	1.08	0.181	0.187	1.03
23	0.649	1.502	2.32	0.390	0.594	1.52	0.281	0.349	1.24	0.220	0.245	1.11	0.181	0.190	1.05
25	0.647	1.602	2.48	0.389	0.622	1.60	0.280	0.360	1.28	0.219	0.251	1.14	0.180	0.193	1.07
35	0.642	2.101	3.27	0.386	0.764	1.98	0.278	0.418	1.50	0.218	0.280	1.29	0.179	0.210	1.17
55	0.638	3.100	4.86	0.384	1.048	2.73	0.277	0.537	1.94	0.217	0.340	1.57	0.178	0.244	1.37
75	0.636	4.099	6.44	0.383	1.334	3.48	0.276	0.655	2.38	0.216	0.400	1.85	0.178	0.279	1.57
95	0.635	5.099	8.03	0.382	1.619	4.24	0.275	0.774	2.81	0.216	0.461	2.14	0.177	0.314	1.77
99	0.635	5.299	8.34	0.382	1.676	4.39	0.275	0.798	2.90	0.216	0.473	2.19	0.177	0.321	1.81

N even and S odd

N	$S=5$			$S=7$			$S=9$			$S=11$		
	V_1	V_2	E	V_1	V_2	E	V_1	V_2	E	V_1	V_2	E
8	0.717	0.717	1.00									
10	0.556	0.556	1.00	0.337	0.337	1.00						
12	0.529	0.606	1.15	0.338	0.347	1.03	0.251	0.251	1.00			
14	0.516	0.657	1.28	0.343	0.362	1.06	0.251	0.255	1.02	0.201	0.201	1.00
16	0.521	0.705	1.35	0.339	0.379	1.12	0.251	0.260	1.04	0.201	0.202	1.01
18	0.519	0.744	1.43	0.335	0.397	1.19	0.253	0.267	1.06	0.200	0.204	1.02
20	0.515	0.805	1.56	0.333	0.415	1.25	0.251	0.375	1.49	0.200	0.208	1.04
22	0.514	0.853	1.66	0.333	0.432	1.30	0.250	0.283	1.14	0.201	0.212	1.05
24	0.513	0.903	1.76	0.333	0.449	1.35	0.249	0.292	1.17	0.200	0.216	1.08
30	0.510	1.502	2.06	0.331	0.502	1.52	0.248	0.316	1.28	0.198	0.229	1.16
36	0.508	1.201	2.64	0.330	0.555	1.68	0.241	0.341	1.38	0.198	0.242	1.23
48	0.506	1.501	2.96	0.329	0.662	2.01	0.246	0.390	1.59	0.197	0.269	1.37
60	0.505	1.800	3.57	0.328	0.768	2.34	0.245	0.440	1.79	0.197	0.296	1.51
70	0.504	2.050	4.07	0.327	0.858	2.62	0.245	0.481	1.96	0.196	0.319	1.62
80	0.504	2.300	4.57	0.327	0.947	2.90	0.245	0.523	2.14	0.196	0.341	1.74
90	0.503	2.550	5.07	0.327	1.036	3.17	0.245	0.564	2.31	0.196	0.364	1.86
100	0.503	2.799	5.57	0.327	1.125	3.44	0.244	0.606	2.48	0.196	0.387	1.97

and Murty (1969) suggested a desirable value of $S = N/2$, which later on was corrected by Bray (1971) to be near to 10 (his Table 5). The same picture emerges from our Table 1 under V_1 where the relative decrease in average variance beyond $S = 10$ becomes quite low. Hence, a value of S from 10 to 12 is quite satisfactory, regardless of how large N is insofar as estimation of GCA is concerned.

When we come to criterion (iii) used for defining the AE and BAE plans above, it is found that, for fixed S, the average variance exhibits a decreasing trend for increase in N under the border plan, but behaves just the opposite under the central one. Hence, increase in N leads to increase in the efficiency of the border plan but makes the other plan more inefficient. Out of all possible circulant plans, the border and the central ones present, respectively, the extreme examples of asymptotically efficient and inefficient plans. However, for other plans, the trend is not so clear.

Arya (1983) noticed that plans based on two or three associate class group divisible schemes were among the two or three variance circulant plans and, in general, were rated to be the best. Hence, if we consider a list of border plans, with less efficient replaced by the best ones, the resulting list will satisfy the condition of the BAE plan.

The asymptotically efficient plans are such that their efficiency increases with inclusion of a larger number of parents. The advantage for choosing such a plan for unexplored values of N and S is thus obvious. By doing so, the breeder is essentially expected to get more information than by choosing an arbitrary plan without working out its efficiency which, in general, may be quite a cumbersome task.

References

Anand IJ, Murty BR (1969) Serial analysis of combinating ability in diallel and fractional diallel crosses in linseed. Theor Appl Genet 39:88–94

Arya AS (1983) Circulant plants for partial diallel crosses. Biometrics 39:43–52

Arya AS (1989) On equivalence and enumeration of circulant partial diallel crosses. Sankhyā Ser B 51:184–195

Arya AS, Narain P (1977) Partial diallel crosses based on some association schemes with three and four associate classes. Sankhyā Ser B 39:394–399

Bray RA (1971) Quantitative evaluation of the circulant partial diallel cross. Heredity 27:189–204

Curnow RN (1963) Sampling the diallel cross. Biometrics 19:287–306

Fyfe JL, Gilbert NEG (1963) Partial diallel crosses. Biometrics 19:278–286

Hinkelmann K, Kempthorne O (1963) Two classes of group divisible partial diallel crosses. Biometrika 50:281–291

Kearsey MJ (1965) Biometrical analysis of a random mating population: A comparison of five experimental designs. Heredity 20:205–235

Kempthorne O, Curnow RN (1961) The partial diallel crosses. Biometrics 17:229–250

Levings CS III, Dudley JW (1963) Evaluation of certain mating designs for estimation of genetic variance in autotetraploid alfalfa. Crop Sci 3:532–535

Mathur SN, Narain P (1976) Some optimal plans for partial diallel crosses. Ind J Genet 36:301–308

Murty BR, Arunachalam V, Anand IJ (1967) Diallel and partial diallel analysis of some yield factors in *Linum usitatissimum*. Heredity 22:35–41

Narain P, Arya AS (1981) Truncated triangular association scheme and related partial diallel crosses. Sankhyā Ser B 43:93–103