Contributions to Balanced Arrays of Strength Five

D.V. Chopra

Department of Mathematics and Statistics
Wichita State University
Wichita, KS 67260-0033 (USA)

Abstract. In this paper we derive some inequalities on the existence of two-symbol balanced arrays (B-arrays) of strength five. We then apply these inequalities to obtain an upper bound on the number of constraints for these arrays, and provide an illustrative example.

1. Introduction and Preliminaries

A matrix T of size $(m \times N)$ with two elements (say, 0 and 1) is called a 2-symbol balanced array (B-array) of strength t $(0 < t \leq m)$ with m constraints (rows), N runs (treatment-combinations, columns) if in each $(t \times N)$ submatrix T^* of T, every $(t \times 1)$ vector $\underset{\sim}{\alpha}$ of weight i (the weight of $\underset{\sim}{\alpha}$ is the number of 1's in it, $0 \leq i \leq t$) appears as a column of T^* exactly μ_i times. Sometimes B-array T is denoted by $\{m, N, t, 2; \mu_0, \mu_1, ..., \mu_t\}$ where m, N, t, and $\mu_i (i = 0, 1, 2, ..., t)$ are called the parameters of T. This definition can be easily generalized to a B-array with s symbols. It is quite obvious that $N = \sum_{i=0}^{t} \binom{t}{i} \mu_i$.

The existence and construction of B-arrays for a given $m (m > t)$ and an arbitrary index set $\underline{\mu}' = (\mu_0, u_1, ..., u_t)$ is clearly a nontrivial problem. To find the maximum value of m for a given $\underline{\mu}'$ is an important problem both in combinatorial mathematics and statistical design of experiments. Such problems for orthogonal arrays (O-arrays) and B-arrays have been discussed, among others, by Bose and Bush [1], Seiden and Zemach [13], Rafter and Seiden [9], Saha, Mukerjee and Kageyama [11], Chopra and Diòs [5], etc. etc.

It is well-known that O-arrays and the incidence matrices of incomplete block designs as well as those of BIB designs are special cases of B-arrays. B-arrays for various values of t, under certain conditions, have been extensively used to construct balanced fractional factorial designs of various resolutions, with rows and columns of T representing factors and treatment-combinations respectively while the two symbols 0 and 1 representing the levels of the corresponding factor. In order to learn more about the applications of B-arrays to combinatorics and statistical design of experiments, the interested reader may consult the list of references at the end of this paper, and also further references given therein.

In this paper we restrict ourselves to the case when $t = 5$, and obtain further results on the existence of such B-arrays with arbitrary values of m and $\underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_t)$. We then describe how these results can be used to obtain the maximum number of constraints m for a given $\underline{\mu}'$. For certain situations we will demonstrate, through an illustrative example, that the inequalities presented here give sharper bounds on m as compared to those obtained by applying earlier results.

2. Main Results and Their Discussion

We first state some results, which are easy to derive, for later use for arrays with t.

Lemma 2.1. A B-array T with $\underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_t)$ and $m = t$ always exists.

Lemma 2.2. A B-array T of strength t and with $\underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_t)$ is also of strength t' $(0 < t' \le t)$ with its new index set $\underline{\mu}'^{*} = (A_j, t';\ j = 0, 1, 2, \ldots, t'$ with $A_{j,t'} = \sum_{i=0}^{t-t'} \binom{t-t'}{i} \mu_{i+j}$, where $j = 0, 1, 2, \ldots, t')$.

Definition 2.1. Two columns of a B-array T with m rows and with $\underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_t)$ are said to have j coincidences (where $0 \le j \le m$) if the symbols appearing in the corresponding j positions are the same.

Lemma 2.3. Consider a B-array $T(m \times N)$ with $\underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_t)$. If l is the number of 1's in some column (say, the first one) of T, then the following $(t + 1)$ equalities are true:

$$\left.\begin{aligned} &\sum_{j=0}^{m} x_j = N - 1 \\ &\sum j^k x_j = \sum_{i=1}^{k} b_{i,k}\ B_i \end{aligned}\right] k = 1, 2, \ldots, t \tag{2.1}$$

where $B_k = \sum_{i=0}^{k} \binom{l}{i}\binom{m-l}{k-i}(A_{i,k} - 1)$, $k = 1, 2, \ldots, t$, $b_{i,k}$ are appropriate constants and $A_{i,k} = \mu_i$ for $k = t$, and x_j denotes the number of columns of T, the first column being excluded, having exactly j coincidences with the first one. The constants $A_{i,k}$ in (2.1) are linear functions of the elements of the vector $\underline{\mu}'$, and thus B_k's are polynomial functions involving l, m, and vector $\underline{\mu}'$ elements. Next we state some results, without proofs, from Mitrinović [8], to be used with (2.1) to obtain the desired inequalities.

Result: Let a_k, b_k, c_k, $(k = 1, 2, \ldots, n)$ be non-negative reals. Then the following are true:

(a) $$(\sum a_k b_k)^2 \leq \sum a_k^2 \sum b_k^2 \qquad (2.2a)$$

(b) $$(\sum a_k b_k c_k)^4 \leq \sum a_k^4 \sum b_k^4 (\sum c_k^2)^2 \qquad (2.2b)$$

(c) Hölder inequality: $\sum a_k^{\frac{1}{p}} b_k^{\frac{1}{q}} \leq (\sum a_k)^{\frac{1}{p}} (\sum b_k)^{\frac{1}{q}}$; where $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. (2.2c)

(d) Minkowski's inequality: $[\sum (a_k + b_k + c_k)^p]^{\frac{1}{p}} \leq$
$\leq (\sum a_k^p)^{\frac{1}{p}} + (\sum b_k^p)^{\frac{1}{p}} + (\sum c_k^p)^{\frac{1}{p}}$, p being >1. (2.2d)

Theorem 2.1. Consider a B-array $T\{m, N, t = 5; \underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_5)\}$. Then the following results are true:

(a) $$\left[\sum_{i=1}^{4} b_{i,4} B_i\right]^2 \leq \left[\sum_{i=1}^{5} b_{i,5} B_i\right] \left[\sum_{i=1}^{3} b_{i,3} B_i\right] \ldots \qquad (2.3a)$$

(b) $$\left[\sum_{i=1}^{3} b_{i,3} B_i\right]^2 \leq \left[\sum_{i=1}^{4} b_{i,4} B_i\right] \left[\sum_{i=1}^{2} b_{i,2} B_i\right] \ldots \qquad (2.3b)$$

(c) $$\left[\sum_{i=1}^{3} b_{i,3} B_i\right]^3 \leq \left[\sum_{i=1}^{5} b_{i,5} B_i\right] \left[\sum_{i=1}^{2} b_{i,2} B_i\right]^2 \ldots \qquad (2.3c)$$

(d) $$\left[\sum_{i=1}^{2} b_{i,2} B_i\right]^4 \leq (N-1)^2 \left[\sum_{i=1}^{5} b_{i,5} B_i\right] \left[\sum_{i=1}^{3} b_{i,3} B_i\right] \ldots \qquad (2.3d)$$

(e) $$\left[\sum_{i=1}^{4} b_{i,4}\, B_i\right]^3 \leq \left[\sum_{i=1}^{5} b_{i,5}\, B_i\right]^2 \left[\sum_{i=1}^{2} b_{i,2}\, B_i\right] \ldots \tag{2.3e}$$

(f) $$\left[\sum_{i=1}^{1} b_{i,1}\, B_i + \sum_{i=1}^{2} b_{i,2}\, B_i + \sum_{i=1}^{3} b_{i,3}\, B_i\right] \leq$$
$$\leq \left[(N-1)\sum_{i=1}^{4} b_{i,4}\, B_i\right]^{\frac{1}{2}} + \left[(N-1)\sum_{i=1}^{2} b_{i,2}\, B_i\right]^{\frac{1}{2}} +$$
$$+ \left[\sum_{i=1}^{2} b_{i,2}\, B_i \sum_{i=1}^{4} b_{i,4}\, B_i\right]^{\frac{1}{2}} \ldots \tag{2.3f}$$

Proof outline: We can obtain the above six results by using suitable substitutions for a_k, b_k, and c_k in (2.2a-d), and substituting the values of $\sum j^k x_j$ from (2.1) in terms of l and the parameters of the array T. To obtain (2.3a) and (2.3b), we use the following substitutions in (2.2a): $a_k = k\sqrt[\frac{5}{12}]{x_k}$, $b_k = k\sqrt[\frac{3}{2}]{x_k}$ will lead us to (2.3a), and $a_k = k^2\sqrt{x_k}$, $b_k = k\sqrt{x_k}$ will establish (2.3b). Next we set $a_k = k^{\frac{3}{4}}\, x_k^{\frac{1}{4}}$, $b_k = k^{\frac{5}{4}}\, x_k^{\frac{1}{4}}$, and $c_k = k\, x_k^{\frac{1}{2}}$ in (2.2b) to obtain (2.3c); while setting $a_k = k^{\frac{3}{4}}\, x_k^{\frac{1}{4}}$, $b_k = k^{\frac{5}{4}}\, x_k^{\frac{1}{4}}$, $c_k = x_k^{\frac{1}{2}}$ in (2.2b) will give us (2.3d). To derive (2.3e), we employ Hölder inequality where we take $p = 3$ (and therefore $q = \frac{3}{2}$), $a_k = k^2 a_k$, $b_k = k^5 x_k$. Minkowski's inequality is used to get (2.3f) by taking $p = 2$, $a_k = \sqrt{x_k}$, $b_k = k^2\sqrt{x_k}$, and $c_k = k\sqrt{x_k}$.

Remarks: 1. For computational ease, we provide the values of $b_{i,k}$ for $0 < k \leq 5$ and $1 \leq i \leq k$ occurring in (2.1) which we obtained in the process of deriving (2.1). These are $(k = 1;\ b_{1,1} = 1)$, $(k = 2;\ b_{1,2} = 1,\ b_{2,2} = 2)$, $(k = 3;\ b_{1,3} = 1,\ b_{2,3} = 6,\ b_{3,3} = 6)$, $(k = 4;\ b_{1,4} = 1,\ b_{2,4} = 14,\ b_{3,4} = 36,\ b_{4,4} = 24)$, and $(k = 5;\ b_{1,5} = 1,\ b_{2,5} = 30,\ b_{3,5} = 150,\ b_{4,5} = 240,\ b_{5,5} = 120)$.
2. Also, we give explicit expressions for $A_{j,t'} (0 < t' \leq 5)$ in terms of μ_i's $(i = 0, 1, \ldots, 5)$, $A_{j,1} = \sum_{i=0}^{4}\binom{4}{i}\,\mu_{i+j}$ for $j = 0, 1$; $A_{j,2} = \sum_{i=0}^{3}\binom{3}{i}\,\mu_{i+j}$ for $j = 0, 1, 2$; $A_{j,3} = \sum_{i=0}^{2}\binom{2}{i}\,\mu_{i+j}$ with $j = 0, 1, 2, 3$; $A_{j,4} = \sum_{i=0}^{1}\binom{1}{i}\,\mu_{i+j}$ for $j = 0, 1, 2, 3, 4$; and $A_{j,5} = \mu_j$ when $J = 0, 1, 2, 3, 4, 5$.

It is not difficult to check that (2.3a-2.3f) are functions of m and l for an array T with a given $\underline{\mu}'$. If no information on l is available, we can always attach a vector with $l = 0$ or $l = m$ to the B-array T under consideration which will only affect the values of μ_0 and/or μ_5. A computer program can be easily prepared to check these inequalities. For a given $\underline{\mu}'$ and l, if any of these six inequalities is contradicted for $m = m^* + 1$ (say), then m^* is an upper bound for the number of constraints. It means there exists no balanced array for the given $\underline{\mu}'$ and l and with $m \geq m^* + 1$.

Next we present an example to show the usefulness of the results derived in this paper.

Example 1. Consider an array T with $\underline{\mu}' = (1, 0, 1, 1, 1, 1)$, and, therefore, $N = 27$. We take $l = 0$ (say). Using (2.3a - 2.3f), we find conditions (2.3a, 2.3b, 2.3c, 2.3e) are all contradicted for $m = 7$. For the sake of illustration, we give the LHS and RHS for (2.3a): LHS=6,315,169.00 , RHS=6,206,977.00. Clearly the (2.3a) is contradicted, and thus $m \leq 6$. Next, when we use the corresponding condition from Chopra and Diòs [5], we observe that each $m \leq 11$ satisfy it. Thus, the conditions presented here provide sharper bounds on m in this case.

Remark: We do not claim the uniform superiority of the results given here, and the above example is merely presented to show there can be arrays for which we get sharper bounds on the number of constraints as compared to the one obtained by using the result of Chopra and Diòs [5].

References

[1] R.C. Bose and K.A. Bush. Orthogonal arrays of strength two and three, Am. Math. Statist. 23(1952), 508-524.

[2] C.S. Chang. Optimality of some weighing and 2^m factorial designs, Ann. Statist. 8 (1980), 436-444.

[3] D.V. Chopra. On balanced arrays with two symbols, *Ars* Combinatoria 20A (1985), 59-63.

[4] D.V. Chopra. On arrays with some combinatorial structure, Discrete Mathematics 138(1995), 193-198.

[5] D.V. Chopra and R. Diòs. Some combinatorial investigations on balanced arrays of strength three and five, Congressus Numerantium 123(1997), 173-179.

[6] M. Lakshmanamurti. On the upper bound of $\sum_{i=1}^{n} X_i^m$ subject to the conditions $\sum_{i=1}^{n} X_i = 0$ and $\sum_{i=1}^{n} X_i^2 = n$, Math. Student 18(1950), 111-116.

[7] J.Q. Longyear. Arrays of strength s on two symbols, Jour. Statist. Plann. and Inf. 10(1984), 227-239.

[8] D.S. Mitrinović. Analytic Inequalities, Springer-Verlag, New York, 1970.

[9] J.A. Rafter and E. Seiden. Contributions to the theory and construction of balanced arrays, Ann. Statist. 2(1974), 1256-1273.

[10] C.R. Rao. Hypercubes of strength d leading to confounded designs in factorial experiments, Bull. Calcutta Math. Soc. 38(1946), 66-73.

[11] G.M. Saha, R. Mukerjee and S. Kageyama. Bounds on the number of constraints for balanced arrays of strength t, Jour. Statist. Plann. and Inf. 18(1988), 255-265.

[12] E. Seiden. On the maximum number of constraints of an orthogonal array, Ann. Math. Statist. 26(1955), 132-135.

[13] E. Seiden and R. Zemach. On orthogonal arrays, Ann. Math. Statist. 27(1966), 1355-1370.

[14] W.D. Wallis. Combinatorial Designs, Marcel-Dekker Inc., New York, 1988.

[15] S. Yamamoto, M. Kuwada, and F. Yuan. On the maximum number of constraints for s symbol balanced arrays of strength t, Commun. Statist. Theory Meth. 14(1985), 2447-2456.