On the Constraints of Some Balanced Arrays

D.V. Chopra
Department of Mathematics and Statistics
Wichita State University
Wichita, KS
USA 67208-1595

ABSTRACT. In this paper we obtain some combinatorial inequalities involving the parameters of a balanced array (B-array) T of strength four and with two levels. We discuss the usefulness of these inequalities in obtaining an upper bound for the number of constraints of T, and briefly describe the importance of these arrays in design of experiments as well as in combinatorics.

1 Introduction and Preliminaries

An array T with m rows (constraints), N columns (runs, treatment-combinations) and with s levels (symbols) is merely a matrix T of size $(m \times N)$ whose elements are (say) $0, 1, 2, \ldots, s-1$. T is called a binary array if $s = 2$. Here we restrict ourselves to binary arrays. The weight of a column $\underline{\alpha}$ of T, denoted by $w(\underline{\alpha})$, is the number of 1's in $\underline{\alpha}$, and clearly $0 \leq \omega(\underline{\alpha}) \leq m$. T is said to be of strength $t (0 < t \leq m)$ if in every sub-matrix T^* $(t \times N)$ of T, the vectors of weight $i (0 \leq i \leq t)$ appear with a frequency (say) μ_i, with μ_i depending only on i. Imposing further combinatorial constraint on T, we obtain the following definition of a balanced array (B-array).

Definition 1.1. T is called a B-array of strength t if in every $(t \times N)$ submatrix T^* of T, every $(t \times 1)$ vector of weight $i (0 \leq i \leq t)$ appears a constant number μ_i (say) times. The vector $\underline{\mu}' = (\mu_0, \mu_1, \mu_2, \ldots, \mu_t)$ is called the index set of T, and we denote T sometimes by B-array $(m, N; \underline{\mu}', t, s = 2)$. Obviously

$$N = \sum_{i=0}^{t} \binom{t}{i} \mu_i$$

Thus the number of runs in T is a linear function of the μ_i's. In this paper we restrict ourselves to arrays with $t = 4$, but the results can be extended to general t without much difficulty. For $t = 4$, clearly $N = \mu_0 + 4\mu_1 + 6\mu_2 + 4\mu_3 + \mu_4$. B-arrays have been extensively used in the constructions of symmetrical as well as asymmetrical factorial designs. B-arrays with $\mu_i = \mu$ for each i give rise to orthogonal arrays (O-arrays), and the incidence matrices of balanced incomplete block designs (BIBD) are special kind of B-arrays with $t = 2$. Those who are interested to gain further insight into the importance of B-arrays may consult the list of references given at the end.

The existence and construction of B-arrays with $m > t$ and $\underline{\mu}' = (\mu_0, \mu_1, \ldots, \mu_t)$ is a non-trivial problem. Furthermore, for a given $\underline{\mu}'$, the construction of these arrays with maximum possible value of m is an important problem both in combinatorics and design of experiments. Such problems for O-arrays for a given μ and m have been studied, among others, by Bose and Bush (1952), Seiden and Zemach (1966). The corresponding problems for B-arrays have been investigated by Chopra (1985), Chopra and Dio's (1989), Longyear (1984), Rafter and Seiden (1974), etc. In this paper we obtain further results on the existence of B-arrays of strength four and with arbitrary values of m and $\underline{\mu}'$. We will describe how the results obtained here can be used to obtain an upper bound for m for a given value of $\underline{\mu}'$. The generalization of these results to B-arrays with $t = 2l$ is quite straightforward, but the resulting notation is both messy and cumbersome. For B-arrays with $t = (2l + 1)$, one can obtain similar results by considering the fact that these are also arrays with $t = 2l$.

2 Main results with Discussion

It is easy to establish the results given in the next three lemmas.

Lemma 2.1. *A B-array T with $m = t = 4$ always exists.*

Lemma 2.2. *A B-array $T(m, N; (\mu_0, \mu_1, \mu_2, \mu_3, \mu_4), t = 4)$ is also of strength t' where $0 < t' \leq 4$. Considered as an array of strength 3, 2 and 1 its index sets are $\{a_i; a_i = \mu_i + \mu_{i+1}$ where $i = 0, 1, 2, 3\}$, $\{b_j; b_j = a_j + a_{j+1}$ where $j = 0, 1, 2\}$, and $\{c_k; c_k = b_k + b_{k+1}$ where $k = 0, 1\}$ respectively.*

Definition 2.1. Two columns of an $(m \times N)$ binary B-array T are said to have i coincidences $(0 \leq i \leq m)$ if the symbols appearing in these two columns in exactly i of the rows are identical.

Lemma 2.3. *Consider a B-array $T(m \times N)$ with $s = 2$ and $\underline{\mu}' = (\mu_0, \mu_1, \mu_2, \mu_3, \mu_4)$. If $\omega(\underline{\alpha}) = l$, where $\underline{\alpha}$ denotes some column (say, the first) of T, then the following results hold: (Here x_j is the number of columns of weight j).*

$$\sum_{j=0}^{m} x_j = N - 1 \; A_1 \text{ (say)} \tag{2.1}$$

$$\sum_{j=0}^{m} j x_j = \sum_{i=0}^{1} \binom{l}{i}\binom{m-l}{1-i}(c_i - 1) = A_2 \text{ (say)} \tag{2.2}$$

$$\sum j^2 x_j = 2\sum_{i=0}^{2} \binom{l}{i}\binom{m-l}{2-i}(b_i - 1) + \sum_{i=0}^{1} \binom{l}{i}\binom{m-l}{1-i}$$
$$= A_3 \text{ (say)} \tag{2.3}$$

$$\sum j^3 x_j = 6\sum_{i=0}^{3} \binom{l}{i}\binom{m-l}{3-i}(a_i - 1) + 6\sum_{i=0}^{2} \binom{l}{i}\binom{m-l}{2-i}(b_i - 1)$$
$$+ \sum \binom{l}{i}\binom{m-l}{1-i}(c_i - 1) = A_4 \text{ (say)} \tag{2.4}$$

$$\sum j^4 x_j = 24\sum_{i=0}^{4} \binom{l}{i}\binom{m-l}{4-i}(\mu_i - 1) + 36\sum_{i=0}^{3} \binom{l}{i}\binom{m-l}{3-i}(a_i - 1)$$
$$+14\sum_{i=0}^{2} \binom{l}{i}\binom{m-l}{2-i}(b_i - 1) + \sum_{i=0}^{1} \binom{l}{i}\binom{m-l}{1-i}(c_i - 1) = A_5 \text{ (say)} \tag{2.5}$$

Next, we state some classical inequalities from Mitrinovic' (1970) for later use to derive results on the existence of B-arrays.

Result I. For any finite sequence of positive numbers

$$a = (a_1, a_2, \ldots, a_n), \text{ we have } \frac{n}{\sum \frac{1}{a_i}} \le (a_1 a_2 \ldots a_n)^{\frac{1}{n}}$$

Result II. If $a = (a_1, a_2, \ldots, a_n)$ and $b = (b_1, b_2, \ldots, b_n)$ are two real sequences such that $a_1 \le a_2 \le \cdots \le a_n$ and $b_1 \le b_2 \le \cdots \le b_n$, or $a_1 \ge a_2 \ge \cdots \ge a_n$ and $b_1 \ge b_2 \ge \cdots \ge b_n$, then the following inequality (called the Čebyšev inequality) is true

$$\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} b_i\right) \le \frac{1}{n}\sum_{i=1}^{n} a_i b_i$$

Remark: The above inequality can be generalized to more than two such sequences.

Result III. If $a_1 > a_2 > \cdots > a_n > 0$, then the following are true a)

$$\left(\sum_{k=1}^{n} a_k\right)\left(\sum(3k^2+k)a_k\right) - 4\left(\sum ka_k\right)^2 > 0$$

b)

$$5\left(\sum_{k=1}^{n} ka_k\right)^2 - 2\left(\sum a_k\right)\left(\sum ka_k\right) - 3\left(\sum a_k\right)\left(\sum k^2 a_k\right) > 0.$$

Theorem 3.1. *Consider a B-array T with* $\underline{\mu} = (\mu_0, \mu_1, \mu_2, \mu_3, \mu_4)$ *and* $m \geq 5$. *Then for T to exist, we must have*

$$\left[ln \sum_{p,q,r,s=1}^{5} A_p A_q A_r A_s \text{ where } p < q < r < s\right] \geq ln5 + \frac{4}{5}\sum_{k=1}^{5} lnA_k \quad (2.6)$$

Where A_i*'s are as defined in Lemma 2.3.*

Proof: By using (2.1) - (2.5), and Result I with $n = 5$, $a_i = A_i$ ($i = 1, 2, 3, 4$, and 5), we obtain (2.6) after some simplification.

Theorem 3.2. *For a B-array T* $(m, N, \underline{\mu}, t = 4, s = 2)$ *to exist, we must have*

$$A_1(A_4 + A_5) + A_2(A_1 + A_5) + A_3(A_1 + A_4) \leq 2A_1^2 + 2A_2A_4 + 2A_3A_5 \quad (2.7)$$

Proof: Here we use Čebyšev's inequality for the two sequences $A_1 \leq A_2 \leq A_3$, and $A_1 \leq A_4 \leq A_5$ made out of A_i's given in (2.1) - (2.5) we obtain

$$\frac{A_1 + A_2 + A_3}{3}\,\frac{A_1 + A_4 + A_5}{3} \leq \frac{A_1^2 + A_2A_4 + A_3A_5}{3}$$

Simplifying this leads us to (2.7).

Theorem 3.3. *If T is a B-array of strength four with* $\underline{\mu}' = (\mu_0, \mu_1, \mu_2, \mu_3, \mu_4)$ *and with* $m \geq 5$, *the following results are true*

a)

$$\left(\sum_{k=1}^{5} A_k\right)\left(\sum_{k=1}^{5}(3k^2+k)A_{6-k}\right) > 4\left(\sum_{k=1}^{5} kA_{6-k}\right)^2 \quad (2.8)$$

b)

$$5\left(\sum_{k=1}^{5} kA_{6-k}\right)^2 > 2\left(\sum A_k\right)\left(\sum kA_{6-k}\right) + 3\left(\sum A_k\right)\left(\sum k^2 A_{6-k}\right) \tag{2.9}$$

Proof: It is quite clear that A_i's ($i = 1, 2, 3, 4$, and 5) satisfy $A_5 > A_4 > A_3 > A_2 > A_1 > 0$. Thus replacing a_k's in Result III in (a), and (b) by A_{6-k}'s from (2.1) - (2.5), and letting k takes values 1 through 5, we obtain (2.8) and (2.9) respectively.

Remark 1. The inequalities (2.6) - (2.9) are very useful in checking the existence of a B-array T for a given m and $\underline{\mu}$. A contradiction obtained in any one inequality will mean that T does not exist for the given values of m and $\underline{\mu}$. However, we must stress that if all the inequalities are satisfied does not imply that B-array exists necessarily. Thus conditions (2.6) - (2.9) are necessary conditions for the existence of B-arrays.

Remark 2. In order to obtain an upper bound on m, for a given $\underline{\mu}$, a computer program can be easily prepared to check conditions (2.6) - (2.9) for the given values of $m \geq 5$. If $m = m^*$ is the least value of m for which all the conditions (2.6) - (2.9) are satisfied, then m^* is an upper bound for T.

References

[1] E.F. Beckenbach and R. Bellman, *Inequalities*, Springer-Verlag, New York, 1961.

[2] R.C. Bose and K.A. Bush, Orthogonal arrays of strength two and three, *Ann. Math. Statist.* **23** (1952), 508–524.

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

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

[5] D.V. Chopra and R. Dios, On the existence of some balanced arrays, *The Jour. of Combinatorial Mathematics and Combinatorial Computing* **6** (1989), 177–182.

[6] D.V. Chopra, Some combinatorial inequalities on the existence of balanced arrays, *The Jour. of Combinatorial Mathematics and Combinatorial Computing* **8** (1990), 9–12.

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

[8] D.S. Mitrinovic', *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, Some combinatorial problems of arrays and applications to design of experiments, in *A survey of combinatorial theory*, edited by J.N. Srivastava, *et. al.*, North-Holland Publishing Co., 1973, 349–359.

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

[12] E. Seiden, On the problem of construction of orthogonal arrays, *Ann. Math. Statist.* **25** (1954), 151–156.

[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] Herbert S. Wilf, *Finite sections of some classical inequalities*, Springer-Verlag, New York, 1970.