Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 77(1) 2023 35–42

Linear Integer Programming and its Innovative Applications in Design of Experiments and Sample Surveys

Baidya Nath Mandal[1] and V.K. Gupta[2]

[1]*ICAR-Indian Agricultural Research Institute, Gauria Karma, Hazaribag*
[2]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi*

Received 07 April 2023; Revised 17 May 2023; Accepted 18 May 2023

SUMMARY

Linear integer programming is a widely used optimization technique to solve various real life problems. The purpose of this article is to present some innovative applications of linear integer programming in the area of design of experiments and sample surveys. It is demonstrated how construction problems of various block designs and different classes of sampling plans can be solved using linear integer programming formulations.

Keywords: Linear integer programming; Block designs; Balanced; Controlled Sampling; Distance balance; Inclusion probability.

Prologue

Today, we are all united not only in our desire to pay our respect to Late Dr. Daroga Singh during his birth centenary year, but rather in our need to do so because of his extraordinary appeal in the community of statisticians across the globe. He has always been held in high esteem.

This paper is a tribute in honour and loving memory of Daroga Singh with whom I (second author) had an opportunity to work as a student as well as faculty at ICAR-Indian Agricultural Statistics Research Institute (IASRI), New Delhi. Right from our student days to the entire professional career, he had been a source of strength and inspiration to all of us. Indeed, the first author was fortunate to learn sampling theory from the textbook written by him and from the teaching of his students who currently holds eminent positions in various institutes of national and global importance. It gives us immense pleasure to know that Indian Society of Agricultural Statistics has decided to bring out a Special Issue of the Journal of the Indian Society of Agricultural Statistics in memory of Daroga Singh, former Director of IASRI on his birth centenary.

By giving us an opportunity to contribute to the Special Issue, we have been given a chance to say thank you, Daroga Singh, for the way you illuminated statistical sciences in general and theory and applications of survey sampling in particular and our lives by your messages and training. We want you to know that life without you is very, very difficult, though the strength of the messages that you gave us over the years has gradually provided us strength to move forward. The days that we spent under your guidance and what we had learned from you will always remain in our hearts as our most cherished treasure.

We express our proud thankfulness to God for allowing us of our generation to be associated with this towering personality who not only made monumental contributions towards the advancement of statistical sciences, but also remained, at the same time, so down to earth and so compassionate simple person easily accessible. Daroga Singh was the very essence of simplicity, of wisdom, of dedication, of duty, of sincerity, of humbleness, of compassion, of friendship, of care. He was always a helping hand to all. His eagle eye to look at the data and work at ground level was unparalleled. He was a pillar of support.

We miss him deeply. His fond memories survive.

1. INTRODUCTION

Linear Integer Programming (LIP) refers to constrained optimization techniques where objective function is a linear function of integer decision variables subject to linear inequality constraints. A LIP formulation in general can be stated as follows:

$$\text{Maximize } \phi = \mathbf{a}'\mathbf{x}$$
$$\text{subject to } \mathbf{A}\mathbf{x} \le \mathbf{b}$$
$$\mathbf{x} \in \mathbf{Z}^n \qquad (1)$$

Corresponding author: V.K. Gupta
E-mail address: vkgupta1751@yahoo.co.in

where the $\mathbf{x} \in \mathbf{Z}^n$ is a vector of n integer decision variables i.e., $\mathbf{x} = (x_1, x_2, ..., x_n)'$ with x_i, $i = 1, 2, ..., n$ being a non-negative integer variable. In formulation (1), $\mathbf{a}$ is $n \times 1$ vector of coefficients in the objective function, $\mathbf{A}$ is a matrix of order $m \times n$ of coefficients of the constraints and $\mathbf{b}$ is a $m \times 1$ vector and $\mathbf{a}$, $\mathbf{A}$ and $\mathbf{b}$ are all known. A solution of (1) can be obtained by either cutting planes algorithms, branch and bound method, branch and cut method, branch and price method or relaxation and decomposition techniques (Genova and Guliashki, 2011). Solvers such as Cplex and Symphony can be used to solve an LIP formulation on a computer. For more details on LIP, one can see Schrijver (1998); Conforti *et al.* (2014).

LIP is widely used in various kinds of practical problems such as warehouse location problem, machinery selection problem, capital budgeting problem, network and graph problems, maximum flow problems, set covering problems, matching problems, spanning trees problems and many scheduling problems (Chen *et al.*, 2011). The purpose of this article is to give an exposure to the readers about some selected applications of LIP in design of experiments and sample surveys. More specifically, applications of LIP to obtain various classes of block designs and controlled sampling plans will be shown.

2. APPLICATIONS IN CONSTRUCTION OF BLOCK DESIGNS

Construction of various classes of block designs such as Balanced Incomplete Block (BIB) Designs, nearly BIB designs including Regular Graph designs, Semi-Regular Graph designs, Balanced Treatment Incomplete Block (BTIB) designs, Balanced Bipartite Block (BBPB) designs are important problems and various algebraic methods are available to construct these designs. These methods work for specific parametric settings for any class of these designs. Construction problems of these designs can be set as a LIP formulation. The advantage of this method is that it is uniformly applicable across all permissible parametric settings for any class of designs.

Before giving the LIP formulation for a block design construction, some preliminaries are needed. Let $\mathbf{N} = (n_{ij})$ denote the $v \times b$ treatment-block incidence matrix of a binary incomplete block design where v denotes the number of treatments and b, the number of blocks and n_{ij} denote the number of times ith ($i = 1, 2, ..., v$) treatment appears in jth ($j = 1, 2, ..., b$) block. Let $\mathbf{NN'}$ denote the $v \times v$ concurrence matrix of the design where diagonal elements represent the number of replications of the treatments and off-diagonal elements gives the concurrences between pairs of treatments, that is, the number of blocks in which pairs of treatment appear together. To be more specific,

$$\mathbf{NN'} = \begin{pmatrix} r_1 & \lambda_{12} & \cdots & \lambda_{1v} \\ \lambda_{21} & r_2 & \cdots & \lambda_{2v} \\ \vdots & \ddots & \vdots & \vdots \\ \lambda_{v1} & \cdots & \lambda_{v,v-1} & r_v \end{pmatrix} \tag{2}$$

where r_i denotes the number of replications of ith ($i = 1, 2, ..., v$) treatment, $\lambda_{ii'}$ denotes the number of concurrences between treatment i and i' ($i \neq i' = 1, 2, ..., v$). Interestingly, the structure of $\mathbf{NN'}$ matrix is known for various classes of block designs mentioned above. For example, for a BIB design $r_i = r\ \forall\ i$ and $\lambda_{ii'} = \lambda$, $\forall\ i \neq i' = 1, 2, ..., v$, for a nearly BIB design $r_i = r\ \forall\ i$ and $\lambda_{ii'} = \lambda$ or $\lambda + 1$, $\forall\ i \neq i' = 1, 2, ..., v$ and for a BTIB design, $r_2 = ... = r_v = r$ (say) and $\lambda_{1i'} = \lambda_{i'1} = \lambda_1\ \forall\ i' = 2, 3, ..., v$, $\lambda_{ii'} = \lambda\ \forall\ i \neq i' = 2, ..., v$. Further, we assume that the block size for each block is k here and the designs are binary, i.e., elements of $\mathbf{N}$ are 0 and 1. Clearly, $\mathbf{N}'\mathbf{1}_v = k\mathbf{1}_b$ where $\mathbf{1}_t$ denotes a $t \times 1$ vector of ones. One can easily generalize the approach to get designs with unequal block sizes.

Now, we are ready for construction of an incomplete block design belonging to above mentioned classes using LIP. For doing so, $\mathbf{N}$ matrix will be obtained from a given $\mathbf{NN}'$ matrix in v steps so that $\mathbf{N} = (\mathbf{N}_1 \vdots \mathbf{N}_2 \vdots \cdots \vdots \mathbf{N}_v)'$ where $\mathbf{N}'_i$ denotes the ith ($i = 1, 2, ..., v$) row of $\mathbf{N}$ matrix and is clearly of order $1 \times b$. In each step, one row of the $\mathbf{N}$ will be obtained such that the entire $\mathbf{N}$ matrix gives the desired $\mathbf{NN}'$ matrix and thus, the desired design. Further, for a given matrix $\mathbf{A}$ with n rows, let $\mathbf{A}^{(i)}$ be the sub-matrix taking first i rows of the matrix and let $\mathbf{A}_{(-i)}$ be the sub-matrix after removing the ith row from the matrix. In other words, $\mathbf{A}^{(i)} = (\mathbf{A}_1 \vdots \cdots \vdots \mathbf{A}_i)'$ and $\mathbf{A}_{(-i)} = (\mathbf{A}_1 \vdots \cdots \vdots \mathbf{A}_{i-1} \vdots \mathbf{A}_{i+1} \vdots \cdots \vdots \mathbf{A}_n)'$, $i = 2, 3, ..., n - 1$ with $\mathbf{A}_{(-1)} = (\mathbf{A}_2 \vdots \cdots \vdots \mathbf{A}_n)'$. In other words,. Now we describe the steps to get a block design using LIP.

Step 1: Obtain $\mathbf{N}'_1$ by filling 1 in r_1 positions at random out of b positions.

Step i,($i = 2, 3, ..., v$): Compute $\mathbf{k}^{(i-1)} = \mathbf{N}^{(i-1)\prime}\mathbf{1}_{i-1}$ and $w_j = 1/k_j^{(i-1)}$ if $k_j^{(i-1)} > 0$, otherwise $w_j = 1$, $j = 1, 2, ..., b$. Denote $\mathbf{w} = (w_1, w_2, ..., w_b)'$. Solve the LIP problem:

$$\text{Maximize } \phi = \mathbf{w}'\mathbf{x}$$
$$\text{subject to } \mathbf{1}_b'\mathbf{x} = r_i$$
$$\mathbf{N}^{(i-1)}\mathbf{x} = \boldsymbol{\lambda}_i \quad (3)$$
$$\mathbf{x} \le k\mathbf{1}_b - \mathbf{k}^{(i-1)}$$

where $\boldsymbol{\lambda}_i = (\lambda_{1i}, \lambda_{2i}, ..., \lambda_{i-1,i})'$ and $\mathbf{x}$ is $b \times 1$ vector of binary integer variables. There are two possibilities here.

Substep 1: A solution of (3) exists. Let an optimum solution be $\mathbf{x}_o$. Then update incidence matrix as $\mathbf{N}^{(i)} = (\mathbf{N}^{(i-1)'} \vdots \mathbf{x}_o)'$ and if $i < v$, go to next i, else return $\mathbf{N}^{(v)}$.

Substep 2: A solution of (3) does not exist. This means $\mathbf{N}^{(i-1)}$ does not lead to a desired $\mathbf{N}^{(v)}$. This calls for an alternate $\mathbf{N}^{(i-1)}$. To obtain such a matrix, delete any one (say, mth) row from $\mathbf{N}^{(i-1)}$ at random. Store mth row of $\mathbf{N}^{(i-1)}$ in a matrix $\mathbf{T}$ and set mth row of $\mathbf{N}^{(i-1)}$ as $\mathbf{0}'_b$. Compute $\mathbf{k}^{(i-1)} = \mathbf{N}^{(i-1)'}\mathbf{1}_{(i-1)}$ and $w_j = 1/k_j^{(i-1)}$ if $k_j^{(i-1)} > 0$, otherwise $w_j = 1$, $j = 1, 2, ..., b$. Solve the LIP problem:

$$\text{Maximize } \phi = \mathbf{w}'\mathbf{x}$$
$$\text{subject to } \mathbf{1}_b'\mathbf{x} = r_m$$
$$\mathbf{N}_{(-m)}^{(i-1)}\mathbf{x} = \boldsymbol{\lambda}_{(m,i)} \quad (4)$$
$$\mathbf{x} \le \mathbf{k}\mathbf{1}_b - \mathbf{k}^{(i-1)}$$
$$\mathbf{T}\mathbf{x} < r_m\mathbf{1}_s$$

where $\boldsymbol{\lambda}_{(m,i)} = (\lambda_{1m}, \lambda_{2m}, ..., \lambda_{m-1,m}, \lambda_{m+1,m}, ..., \lambda_{im})'$ If there exists an optimum solution, say, x_o, then update incidence matrix $\mathbf{N}^{(i-1)}$ with mth row as $\mathbf{x}_o$ and go to next i. Otherwise, repeat this substep till an alternate solution is found. If an alternate solution is not found even after t_1 (a preassigned value, say 100) repetitions then start afresh with Step 1.

Remark 1. *For allocation of a given treatment to blocks, the objective function gives higher weightage to those blocks which have received least number of other treatments before that step. The first two constraints in the formulation (3) ensure desired treatment replications and concurrences. Last constraint ensures that block sizes do not exceed beyond k. Formulation (4) has an extra constraint which is to ensure that a deleted row does not recur as a solution.*

Remark 2. *If all v steps are completed then the above method produces* $\mathbf{N}^{(v)}$ *which is a desired incidence matrix of an incomplete block design with parameters as v, b, k and replications and concurrences as in* $\mathbf{NN}'$ *matrix. Different classes of block designs mentioned earlier can be constructed using this approach.*

2.1 Examples of block design construction

In this section we shall provide some examples of block designs constructed using the approach presented in Section 2.

2.1.1 Construction of BIB design

Consider construction of a BIB design with parameters $v = 8, b = 14, r = 7, k = 4, \lambda = 3$. First solution was presented by Bose (1939). Here we present a solution in Table 1 using our proposed approach. First few steps to generate this design is illustrated below.

In Step 1, $\mathbf{N}'_1$ is obtained by allocating first treatment to any of the 7 blocks out of the 14 blocks. Let $\mathbf{N}'_1 = (1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 1)$. For allocating treatment 2 in step 2, we compute $\mathbf{k}^{(1)} = (1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 1)'$ which gives $\mathbf{w} = \mathbf{1}_{14}$. So to allocate treatment 2, we solve the following formulation:

$$\text{Maximize } \phi = x_1 + x_2 + ...x_{14}$$
$$\text{subject to } x_1 + x_2 + ... + x_{14} = 7$$
$$x_1 + x_2 + x_3 + x_4 + x_9 + x_{13} + x_{14} = 3$$
$$x_1 \le 3 \quad x_2 \le 3 \quad x_3 \le 3 \quad x_4 \le 3$$
$$x_5 \le 4 \quad x_6 \le 4 \quad x_7 \le 4 \quad x_8 \le 4$$
$$x_9 \le 3 \quad x_{10} \le 4 \quad x_{11} \le 4 \quad x_{12} \le 4$$
$$x_{13} \le 3 \quad x_{14} \le 3$$

An optimal solution to above formulation is $\mathbf{x}'_0 = (0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1)'$ and thus, we get $\mathbf{N}^{(2)} = \begin{pmatrix} 1\,1\,1\,1\,0\,0\,0\,0\,1\,0\,0\,0\,1\,1 \\ 0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1\,1\,1\,1 \end{pmatrix}$. So in Step 3, we compute $\mathbf{k}^{(2)} = (1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 2\ 1\ 1\ 1\ 2\ 2)'$ which gives $\mathbf{w} = (1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 0.5\ 1\ 1\ 1\ 0.5\ 0.5)'$. To allocate treatment 3, we solve the following formulation:

$$\text{Maximize } \phi = x_1 + x_2 + ... + x_8 + 0.5x_9 + x_{10} + x_{11} + x_{12} + 0.5x_{13} + 0.5x_{14}$$
$$\text{subject to } x_1 + x_2 + ... + x_{14} = 7$$
$$x_1 + x_2 + x_3 + x_4 + x_9 + x_{13} + x_{14} = 3$$
$$x_8 + x_9 + x_{10} + x_{11} + x_{12} + x_{13} + x_{14} = 3$$
$$x_1 \le 3 \quad x_2 \le 3 \quad x_3 \le 3 \quad x_4 \le 3$$
$$x_5 \le 4 \quad x_6 \le 4 \quad x_7 \le 4 \quad x_8 \le 3$$
$$x_9 \le 2 \quad x_{10} \le 3 \quad x_{11} \le 3 \quad x_{12} \le 3$$
$$x_{13} \le 2 \quad x_{14} \le 2$$

An optimal solution to this formulation is (1 1 0 1 0 0 1 0 0 1 1 1 0 0)′ and thus

$$\mathbf{N}^{(3)} = \begin{pmatrix} 1\,1\,1\,1\,0\,0\,0\,0\,1\,0\,0\,0\,1\,1 \\ 0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,0\,1\,0\,0\,1\,0\,0\,1\,1\,1\,0\,0 \end{pmatrix}.$$

This process continues till treatments 4 to 8 are allocated. After obtaining optimal solution in all the 8 steps, $\mathbf{N}^{(8)}$ matrix is obtained as

$$\mathbf{N}(8) = \begin{pmatrix} 1\,1\,1\,1\,0\,0\,0\,0\,1\,0\,0\,0\,1\,1 \\ 0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,0\,1\,0\,0\,1\,0\,0\,1\,1\,1\,0\,0 \\ 1\,0\,1\,0\,0\,1\,1\,1\,0\,0\,0\,1\,0\,1 \\ 1\,0\,1\,0\,1\,1\,0\,0\,0\,1\,1\,0\,1\,0 \\ 0\,1\,1\,0\,1\,0\,1\,1\,1\,1\,0\,0\,0\,0 \\ 0\,0\,0\,1\,1\,1\,1\,0\,1\,0\,1\,0\,0\,1 \\ 0\,1\,0\,1\,1\,1\,0\,1\,0\,0\,0\,1\,1\,0 \end{pmatrix}.$$

The block contents of the design is shown in Table 1.

2.1.2 Construction of regular graph design

A regular graph design is an incomplete block design with replications $r_i = r, \forall i$ and con currences as λ and $\lambda + 1$ where $\lambda = \lfloor r(k-1)/(v-1) \rfloor$ (John and Mitchell, 1977; Cheng and Wu, 1981; Mandal *et al.*, 2014). Consider construction of a regular graph design with $v = 6, b = 9, r = 6, k = 4$ and concurrences 3 and 4. A solution of such a design is given in Table 2.

Similarly, binary block designs belonging to various classes such as partially balanced 6 incomplete block (PBIB) designs, semi-regular graph designs etc. can be easily obtained using the proposed approach.

Remark 3. *There is a possibility that even when a design exists, the method may not give a solution. This is due to the fact that before some ith step, the method may continue to give a* $\mathbf{N}^{(i-1)}$ *which does not lead to final* **N** *matrix even after substep 2. It is difficult to tell which row or rows is responsible for this and thus, substep 2 may not always be effective. However, we found that the algorithm is very effective to obtain binary incomplete block designs for* $v \leq 30$, $k \leq 10$. *For larger v and k, the proposed approach may not be practical to use.*

Table 1. A BIB design with $v = 8, b = 14, r = 7, k = 4, \lambda = 3$

Block-1	1	3	4	5
Block-2	1	3	6	8
Block-3	1	4	5	6
Block-4	1	3	7	8
Block-5	5	6	7	8
Block-6	4	5	7	8
Block-7	3	4	6	7
Block-8	2	4	6	8
Block-9	1	2	6	7
Block-10	2	3	5	6
Block-11	2	3	5	7
Block-12	2	3	4	8
Block-13	1	2	5	8
Block-14	1	2	4	7

Table 2. A regular graph design with $v = 6, b = 9, r = 6, k = 4$

Block-1	2	3	5	6
Block-2	1	4	5	6
Block-3	1	3	4	6
Block-4	2	3	4	6
Block-5	1	3	4	5
Block-6	2	3	4	5
Block-7	1	2	5	6
Block-8	1	2	4	5
Block-9	1	2	3	6

3. APPLICATIONS IN SAMPLE SURVEYS

In sample surveys, controlled sampling is often used when it is not advisable to adopt simple random sampling without replacement (SRSWOR) due to administrative and other reasons. Avadhani and Sukhatme (1967) presented the following layout of $N = 7$ villages from which $n = 3$ villages are to be sampled.

```
*   2   *   1   *
7   *   5   *   4
*   6   *   3
```

Table 3 gives the list of preferred and non-preferred samples from the above population.

A controlled sampling plan should minimize probability of non-preferred samples and also should retain desirable features of SRSWOR. Solution to this problem was given by Avadhani and Sukhatme (1967) using a BIB design. Rao and Nigam (1990) presented

Table 3. List of preferred and non-preferred samples

Preferred samples			Non-preferred samples	
1 2 4	1 5 7	3 4 6	1 2 3	2 3 4
1 2 5	2 3 5	3 5 6	1 2 6	2 3 6
1 2 7	2 4 5	3 5 7	1 3 6	2 3 7
1 3 4	2 5 6	3 6 7	1 3 7	2 4 6
1 3 5	2 5 7	4 5 6	1 4 6	2 4 7
1 4 5	2 6 7	4 5 7	1 4 7	3 4 7
1 5 6	3 4 5	5 6 7	1 6 7	4 6 7

a solution using a linear programming formulation to this problem. We present a LIP formulation to obtain a controlled sampling plan for this problem. For this, let S_1 and S_2 denote respectively, the set of preferred and non-preferred samples and let the total number of samples be t. Define a vector $\mathbf{a} = (a_1, a_2, ..., a_t)'$ such that if jth sample belong to S_2 then $a_j = 1$, otherwise, $a_j = 0$. Also, for a given pair of units $i \neq i'$, define a vector $\boldsymbol{c}^{(ii')} = \left(c_1^{(ii')}, c_2^{(ii')}, ..., c_t^{(ii')}\right)'$ such that $c_j^{(ii')} = 1$ if jth sample contains units i, i'. Now a LIP formulation with integer decision variables to obtain controlled sampling plan can be devised as:

Minimize $\phi = \mathbf{a}'\mathbf{x}$

subject to $\mathbf{1}_t'\mathbf{x} = b$

$$\mathbf{C}\mathbf{x} = b\boldsymbol{\pi}^{(2)} \quad (5)$$

where b is a suitably chosen positive integer, $\mathbf{C} = (\mathbf{c}^{(12)} \vdots \mathbf{c}^{(13)} \vdots \cdots \vdots \mathbf{c}^{(N-1,N)})'$ and $\boldsymbol{\pi}^{(2)}$ is the vector of second order inclusion probabilities. For a controlled sampling plan with second order inclusion probabilities as that of SRSWOR,

$\boldsymbol{\pi}^{(2)} = \frac{n(n-1)}{N(N-1)}\mathbf{1}_{N(N-1)/2}$ where $\mathbf{1}_m$ denote a $m \times 1$ column vector of ones.

Remark 4. *In formulation (5), for a given controlled sampling problem, the objective function weights, coefficient matrix and right side of constraints are known except the value of b. To have an idea of possible value of b, note that elements of* **C** *are 0s and 1s and* **x** *is a vector of non-negative integers. Hence, the product* **Cx** *must be a vector of non-negative integers. This implies that b must be chosen in such a way that* $b\boldsymbol{\pi}^{(2)}$ *becomes a vector of non-negative integers.*

For the problem of Avadhani and Sukhatme (1967), $t = {}^N C_n = 35$,

$$\mathbf{a} = (10010001101100110110110000010000010)',$$

For this example, $\boldsymbol{\pi}^{(2)} = \frac{1}{7}\mathbf{1}_{21}$. Using $b = 7$ and solving (5) gives an optimal solution $\mathbf{x}_{opt}$ = (010000001000100010000000010100000100)$'$. Samples corresponding to nonzero values of $\mathbf{x}_{opt}$ gives the support of the sampling plan as in Table 4 where probability of selection of samples, denoted as $p(s)$, for each sample is 1/7.

Table 4. A controlled sampling plan for the problem of Avadhani and Sukhatme (1967)

Support			*p(s)*
1	2	4	1/7
1	3	7	1/7
1	5	6	1/7
2	3	5	1/7
2	6	7	1/7
3	4	6	1/7
4	5	7	1/7

Remark 5. *It may be mentioned here that a proper block design with v = N treatments arranged in b blocks each of size k = n can be used to get a sampling plan. Considering v treatments as N population units, if each block of the design is given probability of selection* 1/*b then first order inclusion probability* $\pi_i = r_i/b$ *where* r_i *is the number of replications of the ith* (i = 1, 2, ..., v) *treatment. With equireplicated designs, this reduces to* $\pi_i = r/b = k/v = n/N$. *Unequal probability of selection may also be given to the blocks. Let* **p** *be the vector of probability of selection of the blocks of the design. Then it is easy to see that* $\mathbf{Np} = \boldsymbol{\pi}$ *where* **N** *is treatment-block incidence matrix of a binary incomplete block design and* $\boldsymbol{\pi}$ *is the vector of first order inclusion probabilities of the units. Then* $\mathbf{p} = \mathbf{N}'(\mathbf{NN}')^{-1}\boldsymbol{\pi}$. *Thus, one can get a sampling design with specified first order inclusion probabilities by selecting blocks of the design according to probability of selection* p *given above.*

A special type of sampling plans called balanced sampling plans excluding contiguous units were introduced by Hedayat *et al.* (1988) for sampling from naturally ordered populations in which nearby units give similar observations. Stufken *et al.* (1999) extended the concept to balanced sampling plans excluding adjacent units (BSA plans). Here two units are said to be adjacent when their distance is less than or equal to m. Under these plans, all units have same first order inclusion probabilities, all non-adjacent pairs of units have constant second order inclusion probabilities and all adjacent pairs have second order inclusion probability as 0. In other words, π_{ij} = 0 whenever i, j are adjacent and $= \frac{n(n-1)}{N(N-2m-1)}$ under circular ordering of units. Algebraic, algorithmic as well as linear programming based methods for obtaining BSA plans are available in literature (Stufken *et al.* (1999); Stufken and Wright (2001); Wright and Stufken (2008); Mandal *et al.* (2008). BSA plans can be obtained using linear integer programming approach also. Let the total number of samples which do not contain any adjacent pairs of units be t. For a given nonadjacent pair of units $i \neq i'$, define a vector $\boldsymbol{c}^{(ii')} = \left(c_1^{(ii')}, c_2^{(ii')}, ..., c_t^{(ii')}\right)'$ such that $c_j^{(ii')} = 1$ if jth sample contains units i, i'. Let $\mathbf{C}^* = (\mathbf{c}^{(1,m+1)} \vdots \mathbf{c}^{(1,m+2)} \ldots \vdots \mathbf{c}^{(Nm1,N)})'$. A BSA plan, if exists, can be obtained by solving the linear integer programming formulation

Minimize $\phi = \mathbf{1}_t'\mathbf{x}$

subject to $\mathbf{C}^*\mathbf{x} = b\boldsymbol{\pi}^{(2)}$ (6)

with b suitably chosen positive integer and $\boldsymbol{\pi}^{(2)}$ is the vector of second order inclusion probabilities for non-adjacent pairs of units. Choice of b may be made according to Remark 4.

To illustrate the above procedure, consider $N = 15$, $n = 4$, $m = 1$. Then it can be checked that total numbers of possible samples without containing any adjacent pairs of units is $t = 450$, $\boldsymbol{\pi}^{(2)} = \frac{n(n-1)}{N(N-2m-1)}\mathbf{1}_{90} = \frac{1}{15}\mathbf{1}_{90}$ and $\mathbf{C}^*$ is a 90 × 450 matrix of 0s and 1s.Using $b = 15$ and solving (6), we get an optimal solution as $\mathbf{x}_{opt}$ with exactly 15 unities in the positions 11, 54, 83, 102, 131, 174, 203, 222, 250, 304, 316, 333, 378, 388 and 422, respectively and in the rest of the positions with 0s. The 15 samples corresponding to unities in the set of all possible samples without adjacent pairs of units gives the support of a BSA plans with $N = 15$, $n = 4$, $m = 1$ as shown in the columns of Table 5. Each sample in the support is given probability of selection $p(s) = 1/15$. One can check that for this plan, π_{ij} = for all adjacent pairs of units and $= \frac{1}{15}$ for all other pairs.

Thus, one can use the above approach to get a BSA plans with specified N, n and m.

Since BSA plans does not permit variance estimation of Narain-Horvitz-Thompson estimator, hence distance balanced sampling plans (DBSPs) were introduced in literature by Mandal *et al.* (2009). Under DBSPs, π_i's are same for all units and any two units with same distance have same π_{ij} 's and $\pi_{ij} \geq \pi_{i'j'}$ if distance between $i, j \geq$ distance between i', j'. Second order inclusion probabilities depend on the choice of distance function between two units. To be more specific, $\pi_{ij} = \frac{n(n-1)}{N}\frac{f_{ij}}{\sum_{j=i} f_{ij}}$ where f_{ij} is a non-decreasing function of distance between unit i and j. Under the choice, $f_{ij} = \delta(i, j)$ where $\delta(i, j) = \min\{|i - j|, N - |i - j|\}$, second order inclusion probabilities are given by $\pi_{ij} = \frac{4n(n-1)}{N^3}\delta(i, j)$ for even N and $= \frac{4n(n-1)}{N^3 - N}\delta(i, j)$ for odd N.

Table 5. Support of BSA plan with $N = 15$, $n = 4$, $m = 1$

1	1	1	1	2	2	2	2	3	3	3	4	4	5	6
3	4	5	7	4	5	6	8	5	7	9	6	10	7	8
6	8	11	9	7	9	12	10	8	13	11	9	12	10	11
10	14	13	12	11	15	14	13	12	15	14	13	15	14	15

DBSPs can be obtained by algorithmic approaches and algebraic approaches as suggested by Mandal *et al.* (2009) and Mandal *et al.* (2010). Linear integer programming approach can also be effectively used to obtain such plans. Define t and $\mathbf{C}$ as in formulation (5). Then a DBSP can be obtained by solving the following formulation:

Minimize $\phi = \mathbf{1}_t'\mathbf{x}$

subject to $\mathbf{C}\mathbf{x} = b\boldsymbol{\pi}^{(2)}$ (7)

with $\boldsymbol{\pi}^{(2)}$ being the vector of second order inclusion probabilities with elements as given in the previous paragraph and b being a suitably chosen positive integer.

For illustration, consider $N = 9$, $n = 3$. Here $t = 84$, $\mathbf{C}$ is a 36×84 matrix with 0s and 1s containing elements 1 rth row and cth column whenever rth pair of units is contained in cth sample. Here, $\boldsymbol{\pi}^{(2)} = \frac{1}{30}$ (1, 2, 3, 4, 4, 3, 2, 1, 1, 2, 3, 4, 4, 3, 2, 1, 2, 3, 4, 4, 3, 1, 2, 3, 4, 4, 1, 2, 3, 4, 1, 2, 3, 1, 2, 1)$'$. Choosing $b = 30$ and solving (7) gives the solution $\boldsymbol{x}_{opt}$ = (000100000020000012012010 000000001000020000121200000001000200011200 0010020010000000000)$'$ which leads to the plan given in Table 6.

4. DISCUSSION

It is shown in this article that linear integer programming is useful in construction of block designs which deals with one-way elimination of heterogeneity. This is possible because the treatment-block incidence matrix uniquely determines the design. But in elimination of two and more heterogeneity settings, there are two or more incidence matrices and so the determination of such designs uniquely becomes very difficult. Research attempts may be made to solve this problem. Apart from block designs, linear integer programming are also useful in obtaining fractional factorial designs (Fontana, 2013), supersaturated designs (Mandal and Koukouvinos, 2014), cross over designs (Mandal *et al.*, 2016), among others.

Table 6. A DBSP with $N = 9$, $n = 3$

Support			$p(s)$	Support			$p(s)$
1	2	6	1/30	2	6	7	1/30
1	3	6	2/30	2	6	8	2/30
1	4	7	1/30	3	4	8	1/30
1	4	8	2/30	3	5	8	2/30
1	5	6	1/30	3	6	9	1/30
1	5	7	2/30	3	7	8	1/30
1	5	9	1/30	3	7	9	2/30
2	3	7	1/30	4	5	9	1/30
2	4	7	2/30	4	6	9	2/30
2	5	8	1/30	4	8	9	1/30
2	5	9	2/30				

Sampling designs with appealing properties can be obtained by making use of linear integer programming approach. In this article, we have seen how various controlled sampling designs can be obtained through their application. Linear integer programming can also be used to obtain two-dimensional controlled sampling designs, for example, one can see Tiwari and Nigam (1998, 2010). Further, unequal probability sampling designs including inclusion probability proportional to size (IPPS) sampling designs can also be obtained using linear integer programming formulations with suitable constraints.

ACKNOWLEDGEMENTS

The authors are thankful to the reviewer for valuable suggestions which led to the improvement in the presentation of the article.

REFERENCES

Avadhani, M.S. and Sukhatme, B.V. (1967). Controlled sampling with varying probabilities with and without replacement. *Australian Journal of Statistics*, **9**(1), 8-15.

Bose, R.C. (1939). On the construction of balanced incomplete block designs. *Annals of Eugenics*, **9**(4), 353-399.

Chen, D.-S., Batson, R.G., and Dang, Y. (2011). *Applied Integer Programming: Modeling and Solution*. John Wiley & Sons.

Cheng, C.-S. and Wu, C.-F. (1981). Nearly balanced incomplete block designs. *Biometrika*, **68**(2), 493-500.

Conforti, M., Cornu´ejols, G., Zambelli, G., *et al.* (2014). *Integer Programming*, volume 271. Springer.

Fontana, R. (2013). Algebraic generation of minimum size orthogonal fractional factorial designs: an approach based on integer linear programming. *Computational Statistics*, **28**, 241-253.

Genova, K. and Guliashki, V. (2011). Linear integer programming methods and approaches– a survey. *Journal of Cybernetics and Information Technologies*, **11**(1), 1-25.

Hedayat, A.S., Rao, C.R., and Stufken, J. (1988). Sampling plans excluding contiguous units. *Journal of Statistical Planning and Inference*, **19**(2), 159-170.

John, J.A. and Mitchell, T.J. (1977). Optimal incomplete block designs. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**(1), 39-43.

Mandal, B.N., Gupta, V.K., and Parsad, R. (2014). Efficient incomplete block designs through linear integer programming. *American Journal of Mathematical and Management Sciences*, **33**(2), 110-124.

Mandal, B.N. and Koukouvinos, C. (2014). Optimal multi-level supersaturated designs through integer programming. *Statistics & Probability Letters*, **84**, 183-191.

Mandal, B.N., Parsad, R., and Gupta, V.K. (2008). Computer-aided construction of balanced sampling plans excluding contiguous units. *Journal of Statistics and Applications*, **3**(1-2).

Mandal, B.N., Parsad, R., and Gupta, V.K. (2010). Linear integer programming approach to construct distance balanced sampling plans. *Journal of Indian Society of Agricultural Statistics*, **64**(2), 303-312.

Mandal, B.N., Parsad, R., and Gupta, V.K. (2016). Cyclic circular balanced and strongly balanced crossover designs through integer programming. *Communications in Statistics Theory and Methods*, **45**(4), 859-871.

Mandal, B.N., Parsad, R., Gupta, V.K., and Sud, U.C. (2009). A family of distance balanced sampling plans. *Journal of Statistical Planning and Inference*, **139**(3), 860-874.

Rao, J.N.K. and Nigam, A.K. (1990). Optimal controlled sampling designs. *Biometrika*, **77**(4), 807-814.

Schrijver, A. (1998). *Theory of Linear and Integer Programming*. John Wiley & Sons.

Stufken, J., Song, S.Y., See, K., and Driessel, K.R. (1999). Polygonal designs: Some existence and non-existence results. *Journal of Statistical Planning and Inference*, **77**(1), 155-166.

Stufken, J. and Wright, J.H. (2001). Polygonal designs with blocks of size $k \leq 10$. *Metrika*, **54**(2), 179-184.

Tiwari, N. and Nigam, A. (1998). On two-dimensional optimal controlled selection. *Journal of Statistical Planning and Inference*, **69**(1), 89-100.

Tiwari, N. and Nigam, A. (2010). On two-dimensional optimal controlled nearest propor tional to size sampling designs. *Statistical Methodology*, **7**(6), 601-613.

Wright, J.H. and Stufken, J. (2008). New balanced sampling plans excluding adjacent units. *Journal of Statistical Planning and Inference*, **138**(11), 3326-3335.