Design Issues Related to Allocation of Experimental Units with Known Covariates into Two Treatment Groups

Technical Report No. ASU/2013/6

Dated: 3 December 2013

Samrat Hore
Department of Statistics, Tripura University

Anup Dewanji
Applied Statistics Unit, Indian Statistical Institute, Kolkata

Aditya Chatterjee
Department of Statistics, University of Calcutta

Design Issues Related to Allocation of Experimental Units with Known Covariates into Two Treatment Groups

Samrat Hore, Department of Statistics, Tripura University
Anup Dewanji, Applied Statistics Unit, Indian Statistical Institute, Kolkata
Aditya Chatterjee, Department of Statistics, University of Calcutta

Abstract

An experimental situation corresponding to an ANCOVA model involving several known covariates is of interest in many practical situations related to various disciplines. The design problem here is to determine an optimal allocation rule for the experimental units with known covariate(s) into two treatments such that both treatment and covariate effects, or only treatment effect, are efficiently estimated in terms of the commonly used $D-$ or $A-$optimalities. The present work deals with development of a new allocation algorithm for this purpose. The algorithm generates efficient allocation in comparison to random allocation and allocation designs already proposed in the literature. With large number of experimental units and several covariates, the exact optimal allocation is computationally intractable. In such situation, we have proposed an algorithm for finding a near-optimal allocation design that is motivated by computationally tractable meta-heuristic optimization technique. Through simulation experiments and real life data analysis, we have demonstrated the efficiency of the proposed algorithm.

Keywords: $D-$optimality, $A-$optimality, D_s-optimality, A_s-optimality, Allocation design, Near-optimal design, Variable neighborhood search, Robust allocation design.

1. Introduction

The problem of comparing several treatments in presence of a single or multiple covariates has been of interest in many application areas including chemical industry, agriculture, geology, biology and medicine. Generally, in such problems, we consider a linear model such as in $ANCOVA$. The usual design issue is to find the optimum values of the covariates with respect to the well-known $D-$ or $A-$optimalities where, respectively, the determinant or trace of the inverse of information matrix are minimized (Keifer, 1959). This optimal setting of the covariate values (called design points) along with the number of such values are required to be considered under linear models (Das *et.al.*, 2003; Dey and Mukerjee, 2006). The second type of design issue is concerned with the allocation of experimental units with known covariates into several treatments, so that the treatment effect or the covariate effect

or both can be efficiently estimated with respect to some desired optimality criteria (like $D-$ or $A-$optimalities) under various model assumptions. This problem may again be subdivided into two classes depending on if the allocation is done sequentially or all at one time (at the beginning of the study) on a fixed sample basis. The sequential allocation scheme is by and large implemented using the knowledge of the previous responses and allocations in conjunction with the covariate values of the available experimental units. This is generally known as adaptive allocation scheme (Atkinson, 1982,2002; Bandyopadhyay and Biswas, 2001) in the context of clinical trials. The other problem is concerned with allocation of a fixed number of experimental units with known covariates into different treatments. For example, as reported in Saville and Wood (1991), an experiment was carried out at commercial goat farms to determine whether the standard worm drenching program was adequate. In this experiment with forty goats, twenty were chosen completely at random to be allocated to the standard drenching program, while the remaining twenty were drenched more frequently (i.e, the intensive drenching program). The goats (the experimental units) were individually tagged and weighed at the start and end of the year-long study. The primary interest was to study the significance of live weight gains between the two drenching programs using the initial live weight as the single known covariate. An optimum allocation design for the forty experimental units, instead of the randomized allocation rule as had been used in the experiment, towards efficient comparison of two programs might be of interest (See Section 7). More examples were described in Harville (1974) concerning allocation of experimental units to different treatment groups in the area of biology, aircraft engineering and chemical technology. In the present paper we address this allocation design problem in which, as also mentioned in Harville (1974), owing to the large number of possible allocations, finding the exact $D-$ or $A-$optimal allocation design is generally not practical. Moreover, the problem becomes computationally intractable with large number of experimental units. For example, with n experimental units to be allocated into two treatments, one needs to examine $(2^{n-1}-1)$ such non-trivial allocation schemes for the optimal solution making the procedure to be exponentially complex. However, as we shall explain in later sections, the actual search may be judiciously confined to a smaller number of schemes. With the increase in the number of experimental units and the number of treatments under consideration, the problem becomes immensely difficult as well as computationally intractable for deriving the exact solution. Therefore, an approximate and intelligent search algorithm is the only way out that leads to optimal or near-optimal solutions.

With the knowledge of the covariate values and without the information regarding the responses, corresponding to fixed sample experimental units, an efficient allocation algorithm through a multistage procedure has been proposed by Harville (1974) that allocates one or

more units at each stage and improvement in the allocation scheme is achieved through an iterative algorithm that induces small changes in the design at each iteration. Cook and Nachstsheim (1989) considered an extension of Harville's problem by incorporating a block structure (with pre-specified block sizes) in the design and obtained nearly $D-$optimal design by carrying out essentially the same method as advocated by Harville (1974).

In the present paper, the two allocation schemes as mentioned in Harville (1974) have been critically examined with reference to several allocation designs *viz.* the randomized allocation or some optimal or nearly optimal allocation schemes introduced here. When the total number of experimental units is small, all the schemes may be compared with the actual optimal allocation scheme obtained through total search method. However, with large number of experimental units, Harville's initial as well as fine tuned methods take large amount of computer time; while the methods introduced in this paper require relatively less amount of computation. As an example, let us consider the problem of allocating n experimental units with known single covariate into two treatments. While implementing Harville's initial method, any 3 out of n units are chosen randomly and then allocated into two treatments optimally. Then, one unit out of the $(n-3)$ non-allocated units is allocated to either of the two treatments by optimizing the objective function, which requires $2 \times (n-3)$ searches. The process continues with one unit being allocated in each step until all the units are exhausted. In this way, we have to consider a total of $[2(n-3) + 2(n-4) + 2(n-5) + \cdots + 2] = (n-3) \times (n-2)$ searches. The computational burden thus increases quadratically with n. Harville (1974) has also suggested some fine tuning on the initial algorithm by switching the allocation of one experimental unit from one treatment to another treatment, or by exchanging the allocation of experimental units from one treatment to the other. However, it has been demonstrated through simulation experiments (in Section 7) that such fine tuning may also fail to achieve the global optimal. In this work, an initial allocation design is suggested which is non-iterative in nature and requires little computation. Then, an algorithm is proposed to improve upon this initial design. Often the suggested initial design itself performs better than Harville's designs, as indicated in our simulation study.

The proposed algorithm arrives at a near-optimal allocation design by a discrete search method, exploring the neighborhoods of the current design. It is noteworthy to mention that combinatorial optimization techniques and meta-heuristics are very popular search algorithms (*e.g.*, Simulated Annealing) to tackle computationally complex problems (See Kirkpatrick *et.al.*, 1983). Such search algorithms are found to be more effective to suggest an acceptably good solution in a fixed amount of time that is usually negligible as compared to time required for exhaustive enumeration procedure. An intelligent transition from one neighbor to another through a randomized local search algorithm yields a simple but ef-

fective meta-heuristic technique for combinatorial and global optimization and is popularly known as *Variable Neighborhood Search (VNS)* algorithm. The basic scheme in this regard may be easily implemented using any local search algorithm as a subroutine. An exhaustive review of such methods along with its effectiveness by solving several classical combinatorial or global optimization problems has been provided in Hansen and Mladenovic (2001).

The paper has been structured into several sections. In Section 2, the problem has been formulated in the $ANCOVA$ framework. A quick method has been proposed in Section 3 to obtain an initial design. This has been found to be nearly fully efficient for single covariate when the number of experimental units is small. The proposed algorithm has been described in Section 4 and the case of categorical covariates has been taken up in Section 5. In this context, it has been analytically shown that ensuring $D-$optimality necessarily imply balanced allocation with regard to various categories of a single covariate. A compromise between $D-$ and $A-$optimalities has been proposed in Section 6 which has been termed as *robust* optimality and the proposed algorithm has been extended in this context as well. A detailed simulation study under various univariate and multivariate probability models for the covariates are considered in Section 7. The efficacy of the proposed algorithm in comparison to completely randomized allocation and other available schemes, under various setups, has been established through such simulation study. Three real life examples have been considered with data from animal health laboratory study (Saville and Wood, 1991), dairy cow diet study (Wattiaux *et.al.*,1994) and clinical trial on patients with severe aplastic anemia (Kalbfleisch and Prentice, 2002), and it has been demonstrated that the proposed allocation scheme in comparison to the implemented random allocation procedure, or the procedures proposed by Harville (1974), is more efficient. Section 8 ends with some concluding remarks and scope of further research.

2. Problem formulation

The analysis of covariance is a commonly used technique for analyzing comparative experiments. Let us consider the one-way ANCOVA model with two treatments and p covariates whose values are assumed to be known for the n experimental units. Specifically, the linear model used for the analysis is given by

$$y_{li} \;=\; \mu_l + \sum_{j=1}^{p} \beta_j x_{lij} + \epsilon_{li}, \text{ for } \quad i = 1(1)n_l, \quad l = 1, 2, \tag{2.1}$$

where n_l (≥ 1) denotes the number of times the l^{th} treatment is replicated and $\sum_{l=1}^{2} n_l = n$. Here,

y_{li} = value of the response from i^{th} experimental unit receiving the l^{th} treatment,

$\mu_l = l^{th}$ treatment effect,
β_j = regression coefficient for the j^{th} covariate,
x_{lij} = value of the j^{th} covariate corresponding to i^{th} experimental unit receiving the l^{th} treatment, and
ϵ_{li} = observational error assumed to be *iid* following $N(0, \sigma^2)$ distribution.

With reference to the model (2.1), the parameters of interest are (μ_1, μ_2) and $(\beta_1, \beta_2, ...\beta_p)$. Let us define $\theta = (\mu_1, \mu_2, \beta_1, \beta_2, ...\beta_p)$. For the estimability of θ it is necessary that $n \geq (p+2)$. The optimality criteria depends on the information matrix $\sigma^{-2}I$ with

$$I = \begin{pmatrix} n_1 & 0 & \mathbf{x}_{\cdot\ 1}^T \\ 0 & n_2 & \mathbf{x}_{\cdot\ 2}^T \\ \mathbf{x}_{\cdot\ 1} & \mathbf{x}_{\cdot\ 2} & \mathbf{x}^T\mathbf{x} \end{pmatrix},$$

where $\mathbf{x}_{\cdot\ l}^T = 1_{n_l}^T \mathbf{x}_l$ with 1_{n_l} being the $(n_l \times 1)$ vector of all ones, $\mathbf{x}_l$ the $(n_l \times p)$ matrix given by $\mathbf{x}_l = ((x_{lij}))$ for $l = 1, 2$ and $\mathbf{x}$ is the $(n \times p)$ matrix given by $\mathbf{x}^T = (\mathbf{x}_1^T, \mathbf{x}_2^T)$. For $D-$ and $A-$optimalities (for efficient estimation of both treatment and covariate effects), we need to minimize $det(I^{-1})$ and $trace(I^{-1})$, respectively. The information matrix (I) may be partitioned as

$$I = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix},$$

where $A = Diag(n_1, n_2)$; $B^T = (\mathbf{x}_{\cdot\ 1}, \mathbf{x}_{\cdot\ 2})$ and $D = \mathbf{x^T x}$. Following Cook and Nachtsheim (1989), D_s- and A_s-optimal designs correspond to minimization of determinant or trace, respectively, of the covariance matrix of the least squares estimates of a selected subset of parameters under investigation. We investigate D_s- (A_s-)optimal designs that minimize the determinant (trace) of the covariance matrix of the least square estimates of only the treatment effect. Here, we need to consider the $(1,1)$ block entry of the inverse of the partitioned matrix (I). By following Rao and Bhimasankaran (2000, pp 138), this entry is given by

$$I^{*-1} = A^{-1} + A^{-1}B(D - B^TA^{-1}B)^{-1}B^TA^{-1}. \tag{2.2}$$

Consequently, for D_s- and A_s-optimalities with regard to only the treatment effect, we need to minimize $det(I^{*-1})$ and $trace(I^{*-1})$, respectively. Defining $E = (\mathbf{x^T x} - \sum_{\mathbf{l=1}}^{\mathbf{2}} \frac{\mathbf{x_{\cdot\ l}^T x_{\cdot\ l}}}{\mathbf{n_l}})$ and $F = (\frac{\mathbf{x_{\cdot\ 1}}}{n_1}, \frac{\mathbf{x_{\cdot\ 2}}}{n_2})$, one can obtain

$$\det(I^{-1}) = (n_1 n_2 \det(E))^{-1} \tag{2.3}$$

$$\text{trace}(I^{-1}) = \text{trace}\left(\text{diag}(\frac{1}{n_1}, \frac{1}{n_2}) + (F^TEF) + E\right) \tag{2.4}$$

and

$$\det(I^{*-1}) = \det\left(\text{diag}(\frac{1}{n_1},\frac{1}{n_2}) + (F^TEF)\right) \tag{2.5}$$

$$\text{trace}(I^{*-1}) = \text{trace}\left(\text{diag}(\frac{1}{n_1},\frac{1}{n_2}) + (F^TEF)\right). \tag{2.6}$$

For the sake of exposition, let us consider the case of a single covariate. Let us denote x_{li} to be the value of the single covariate corresponding to i^{th} experimental unit in the l^{th} treatment group, for $i = 1, 2, ..., n_l, l = 1, 2$. The corresponding linear model is given by

$$y_{li} = \mu_l + \beta x_{li} + \epsilon_{li}; \quad i = 1, 2, ...n_l, \quad l = 1, 2. \tag{2.7}$$

Here n $(= n_1 + n_2)$ experimental units with known values of the single covariate given by $x = (x_{11}, \cdots, x_{1n_1}, x_{21}, \cdots, x_{2n_2})$ are to be assigned to two treatments. Denoting $x_{\cdot l} = \sum_{i=1}^{n_l} x_{il}/n_l$, for $l = 1, 2$, it can be checked that

$$det(I^{-1}) = (\prod_{l=1}^{2} n_l)^{-1}(\sum_{l=1}^{2}\sum_{i=1}^{n_l}(x_{li} - x_{\cdot l})^2)^{-1} \tag{2.8}$$

$$trace(I^{-}) = \sum_{l=1}^{2} 1/n_l + \left[(\sum_{l=1}^{2}(x_{\cdot l})^2 + 1)/(\sum_{l=1}^{2}\sum_{i=1}^{n_l}(x_{li} - x_{\cdot l})^2)\right] \tag{2.9}$$

and

$$\begin{aligned} det(I^{*-1}) &= (\prod_{l=1}^{2} 1/n_l)\left[1 + (\sum_{l=1}^{2} n_l {x_{\cdot l}}^2)/(\sum_{l=1}^{2}\sum_{i=1}^{n_l}(x_{li} - x_{\cdot l})^2)\right] \\ &= (\prod_{l=1}^{2} 1/n_l)\left[1 + (\sum_{l=1}^{2} n_l {x_{\cdot l}}^2)/(\sum_{l=1}^{2}\sum_{i=1}^{n_l} x_{li}^2 - \sum_{l=1}^{2} n_l {x_{\cdot l}}^2)\right] \end{aligned} \tag{2.10}$$

$$trace(I^{*-1}) = \sum_{l=1}^{2} 1/n_l + \left[(\sum_{l=1}^{2}(x_{\cdot l})^2)/(\sum_{l=1}^{2}\sum_{i=1}^{n_l}(x_{li} - x_{\cdot l})^2)\right]. \tag{2.11}$$

Under this set up, our problem is to develop an allocation scheme for the n experimental units with known covariates into two treatments such that $D-$ or $A-$ (or D_s- or A_s-)optimalities are achieved. In what follows, we state a quick method in this regard mainly to find a reasonable initial solution. The iterative procedure of finding the optimum allocation scheme for the units starts from this initial solution. Note that the determinant and trace of I^{-1} remain the same if the units allocated to the first treatment are given the second treatment and vice versa. Hence, the $D-$ and $A-$optimality remain invariant if the allocation of units is reversed. Our objective is to propose an algorithm to obtain a near-optimal solution that minimizes the computational burden, so that $(2^{n-1} - 1)$ allocation schemes need not be examined for the exact optimum solution through direct search method.

3. Initial allocation of experimental units

This section proposes a quick method to obtain an initial solution. This method is intuitively very simple and seems to achieve good efficiency in most cases. The following two subsections describe the method for a single covariate and multiple covariates, respectively, for n experimental units. We first assume that all the covariates are quantitative. Categorical covariates will be discussed in Section 5.

3.1 Single Covariate

Note that, by looking at (2.8) to (2.11), the general idea is to choose an allocation so that the within sum of squares of the covariate values in the treatment groups is maximum. In order to achieve that, let us first sort the n covariate values $x_1, \cdots, x_n$ as $x_{(1)} < \cdots < x_{(n)}$. We have considered the following four cases by considering the value of n to be $4t$, $4t+1$, $4t+2$ and $4t+3$, respectively, where $t = 0, 1, \cdots$. Let $m = [n/2]$, where $[x]$ is the greatest integer contained in x.

Case I (n = 4t , m = 2t): Allocate the experimental units corresponding to $(x_{(1)}, x_{(n)})$ to the first treatment and those corresponding to $(x_{(2)}, x_{(n-1)})$ to the second treatment. Then, allocate those corresponding to $(x_{(3)}, x_{(n-2)})$ to the first and those corresponding to $(x_{(4)}, x_{(n-3)})$ to the second, and so on. In this process, all the units will be exhausted with each treatment getting $m = 2t$ units. So, the ultimate allocation is given by
$\{(1), (3), ..., (n-2), (n)\} \in I_1$ and $\{(2), (4), ..., (n-3), (n-1)\} \in I_2$, where I_1 and I_2 are the two sets of experimental units allocated to the two treatments and (i) denotes the label of the experimental unit corresponding to $x_{(i)}$. As remarked before, the labeling of the first or second treatment does not make a difference in optimality.

Case II (n = 4t+1 , m = 2t): The scheme is same as that given in Case I. Here, at the end, only the experimental unit $(m+1)$ corresponding to $x_{(m+1)}$ remains to be allocated. Let I_1 and I_2 denote the two sets of already allocated units corresponding to the two treatments. This unit will be allocated to the first treatment corresponding to I_1 if

$$\sum_{i' \in \{I_1, (m+1)\}} (x_{i'} - x_{\cdot 1})^2 + \sum_{k' \in I_2} (x_{k'} - x_{\cdot 2})^2 > \sum_{i'' \in I_1} (x_{i''} - x_{\cdot 1})^2 + \sum_{k'' \in \{I_2, (m+1)\}} (x_{k''} - x_{\cdot 2})^2.$$

This is equivalent to the condition

$$(\sum_{i \in I_1} x_i)^2 - 2m(\sum_{i \in I_1} x_i x_{(m+1)}) > (\sum_{k \in I_2} x_k)^2 - 2m(\sum_{k \in I_2} x_k x_{(m+1)}).$$

Otherwise, the unit $(m+1)$ is allocated to the second treatment. In case of equality, random allocation of the unit $(m+1)$ to either of the two treatments is suggested. This proposed

allocation of the unit $(m+1)$ ensures better optimality or lower value of the objective function.

Case III (n = 4t+2 , m = 2t+1): The scheme is same as that in the previous two cases. However, at the end of the allocation process, experimental units (m) and $(m+1)$ corresponding to the covariate values $x_{(m)}$ and $x_{(m+1)}$ remain to be allocated, with I_1 and I_2 denoting the two sets of already allocated units, as before. The unit (m) is allocated to the first treatment and $(m+1)$ to the second treatment if

$$\sum_{i' \in \{I_1, m\}} (x_{i'} - x_{\cdot 1})^2 + \sum_{k' \in \{I_2, (m+1)\}} (x_{k'} - x_{\cdot 2})^2 > \sum_{i'' \in \{I_1, (m+1)\}} (x_{i''} - x_{\cdot 1})^2 + \sum_{k'' \in \{I_2, m\}} (x_{k''} - x_{\cdot 2})^2,$$

which is equivalent to

$$\sum_{i \in I_1} x_i > \sum_{k \in I_2} x_k.$$

The allocation scheme is reversed otherwise. In case of equality, random allocation of the two units to the two treatments is recommended.

Case IV (n = 4t+3 , m = 2t+1): The scheme is as before and, at the end of the process, three experimental units (m), $(m+1)$ and $(m+2)$ corresponding to covariate values $x_{(m)}$, $x_{(m+1)}$ and $x_{(m+2)}$ remain to be allocated. The unit (m) is allocated to the first treatment and the unit $(m+1)$ to the second if

$$\sum_{i \in I_1} x_i > \sum_{k \in I_2} x_k,$$

and the allocation is reversed otherwise. The remaining unit $(m+2)$ is allocated to the first treatment if

$$(\sum_{i \in I_1'} x_i)^2 - 2m(\sum_{i \in I_1'} x_i x_{(m+2)}) > (\sum_{k \in I_2'} x_k)^2 - 2m(\sum_{k \in I_2'} x_k x_{(m+2)}),$$

where I_1' and I_2' are the new sets obtained from the previous step satisfying the corresponding inequality. Otherwise, the unit $(m+2)$ is allocated to the second treatment. In case of equality, the unit $(m+2)$ may be allocated randomly to any of the two treatments.

3.2 Multiple Covariates

In case of multiple covariates $(p \geq 2)$, construction of the initial allocation design is somewhat different. First, taking each covariate at a time and treating it to be the only single covariate, the initial allocation is obtained using the method of Section 3.1, resulting in p allocation designs. We choose that allocation (out of these p) as the initial design, for which the value of the objective function with all the covariates taken together is minimum.

4. Proposed Algorithm for Near-Optimal Allocation

To obtain the optimal or near-optimal allocation design (as the case may be) for fixed (known) covariate(s), an initial design is needed. After selecting the initial design through the quick method, as described in the previous section, a neighborhood search algorithm is proposed to arrive at the optimal or near-optimal allocation.

Define $\alpha = (\alpha_1, \cdots, \alpha_n)$ to be an n-dimensional vector with each α_i being binary (0 or 1) depending on whether the i^{th} experimental unit is allocated to the first or second treatment, respectively. Therefore, the vector α represents an allocation design. Let us define a neighborhood of an allocation design α consisting of the allocation designs given by $\{\alpha_{(1)}, \cdots, \alpha_{(n)}\}$, where $\alpha_{(i)}$ is the vector obtained from α by replacing α_i by $1 - \alpha_i$, for $i = 1, 2, \cdots, n$. Let us denote this neighborhood of α by $\mathcal{N}_\alpha$. Note that, if an allocation design $\alpha' \in \mathcal{N}_\alpha$, then $\alpha \in \mathcal{N}_{\alpha'}$. Also note that all the $(2^{n-1} - 1)$ allocation designs are connected through this concept of neighborhood, in the sense that any allocation design can be accessed from another allocation design through successive neighborhoods. Let us denote the value of the objective function (to be minimized) for an allocation design α by $V(\alpha)$, as defined in Section 2; see, for example (2.3) to (2.6) and (2.8) to (2.11). The proposed algorithm for obtaining a near-optimal allocation design is as follows:

Step 1: Start with the initial allocation, denoted by the vector $\alpha^{(0)}$, obtained by using the method stated in Section 3.

Step 2: Consider all the allocations in $\mathcal{N}_{\alpha^{(0)}}$ and compute $V(\alpha')$, for all $\alpha' \in \mathcal{N}_{\alpha^{(0)}}$.

Step 3: If $\min\{V(\alpha'), \alpha' \in \mathcal{N}_{\alpha^{(0)}}\} < V(\alpha^{(0)})$, choose the next improved allocation to be $\alpha^{(1)} = arg\ min\{V(\alpha'), \alpha' \in \mathcal{N}_{\alpha^{(0)}}\}$. Otherwise, choose $\alpha^{(1)}$ randomly from among the allocations in $\{\alpha^{(0)}, \mathcal{N}_{\alpha^{(0)}}\}$ with probabilities given by $p_0 = \frac{V(\alpha^{(0)})}{D(\alpha^{(0)})}$, $p_{\alpha'} = \frac{V(\alpha')}{D(\alpha^{(0)})}$, for $\alpha' \in \mathcal{N}_{\alpha^{(0)}}$, where $D(\alpha^{(0)}) = V(\alpha^{(0)}) + \sum\limits_{\alpha' \in \mathcal{N}_{\alpha^{(0)}}} V(\alpha')$.

Step 4: Replace $\alpha^{(0)}$ by $\alpha^{(1)}$ and repeat steps 2 and 3 until convergence with some desired accuracy.

The accuracy may be achieved in several ways, for example, by limiting the number of returns to an allocation design. In this paper, we suggest a probabilistic method to achieve this as described in the following. At the i^{th} return ($i \geq 1$) to a design α, denoted by $\alpha^{(0)}$, the probabilities p_0 and $p_{\alpha'}$ for choosing $\alpha^{(0)}$ and α' in $\mathcal{N}_{\alpha^{(0)}}$, respectively, are modified as $p_0^{(i)} = \frac{V^{(i)}(\alpha^{(0)})}{D^{(i)}(\alpha^{(0)})}$, $p_{\alpha'}^{(i)} = \frac{V^{(i)}(\alpha')}{D^{(i)}(\alpha^{(0)})}$ with $V^{(i)}(\alpha^{(0)}) = V(\alpha^{(0)}) + (D(\alpha^{(0)}) - V(\alpha^{(0)})) \times \frac{i}{n}$, $V^{(i)}(\alpha') = max\ \{(V(\alpha') - (D(\alpha^{(0)}) - V(\alpha^{(0)})) \times \frac{i}{n^2}, 0), \alpha' \in \mathcal{N}_{\alpha^{(0)}}\}$ and $D^{(i)}(\alpha^{(0)}) = V^{(i)}(\alpha^{(0)}) + \sum\limits_{\alpha' \in \mathcal{N}_{\alpha^{(0)}}} V^{(i)}(\alpha')$. This increases the probability of choosing the same design $\alpha^{(0)}$ at successive returns. We suggest stopping the algorithm when this probability exceeds a value close to unity, say 0.95 or 0.99.

Note that, because of the connectedness property stated above (before the statement of the algorithm) the algorithm has the potential to cover all the $(2^{n-1}-1)$ allocation designs. However, looking at the choice of $\alpha^{(1)}$ in step 3, it is clear that the algorithm may get stuck at a local minimum, if there is any. A remedy in this regard may be achieved by restarting the algorithm with an initial allocation chosen randomly from $\bigcup_{\alpha' \in \mathcal{N}_{\alpha^*}} \mathcal{N}_{\alpha'} - \{\alpha^*\}$, where α^* denotes the current optimal allocation design. This, however, may not entirely solve the problem of local minimum. In this regard, the choice of the initial allocation is very important and the quick method of the previous section provides a suitable solution for this. As will be seen in Section 7, this seems to perform better than even the final solution of Harville (1974).

The proposed algorithm has inherent similarity with the well-known *VNS algorithm* (Hansen and Mladenovic, 2001), in which the neighborhood is dynamically updated in each iteration. Moreover, the algorithm is quite similar to the famous *Metropolis Algorithm* and their extensions (Kirkpatrick *et. al.*, 1983), in which moving towards better direction is done with certainty but moving towards a worse direction is accepted through some random mechanism given by Boltzmann factor. In our case also, we choose the best alternative among various superior alternatives, if any, with certainty, but are prepared to accept an inferior alternative through some stochastic mechanism, as given in step 3 of the algorithm. At the same time, if a particular allocation is obtained repeatedly in subsequent steps, the likelihood of accepting that allocation as the optimum is enhanced through the random mechanism, as explained after step 4 of the algorithm. In some sense, this provides us the annealing schedule with clearly defined stopping rule. The neighborhood, as described above, is easy to interpret both in terms of allocation of units in two treatment groups and the geometry of the design vector. One could possibly think of some other definition of neighborhood and carry out the same algorithm to search for near-optimal allocation. However, the definition given here is simple and seems to be natural.

5. Categorical covariates

In practice, the covariates may often be binary or categorical (nominal or ordinal). In such situations, the problem may be approached with the help of *dummy variables*. Let us, for simplicity, consider a single categorical covariate involving $(K+1)$ levels $A_1, \cdots, A_{K+1}$. The problem is to allocate the n experimental units with known levels of the categorical variable to either of the two treatments in an optimal way as suggested in case of quantitative covariates. We consider the model (2.7) with x_{li} denoting the value of the *dummy variable* for the categorical covariate for the i^{th} experimental unit in the l^{th} treatment group. Since

there are in all $(K+1)$ levels, x_{li} is a K dimensional vector with all 0's and at most one 1. Let n_l be the number of experimental units in the l^{th} treatment group and m_{lj} be the number of experimental units in the l^{th} treatment group belonging to the category A_j, $l = 1, 2$, $j = 1, \cdots, K+1$, with $m_{l\ \overline{K+1}} = n_l - \sum_{j=1}^{K} m_{lj}$. Here $n_1 + n_2 = n$. Write $m_j = \sum_{l=1}^{2} m_{lj}$, for $j = 1, \cdots, K+1$, where m_j denotes the number of experimental units in category A_j. The information matrix may be written as

$$\mathbf{I} = \begin{pmatrix} n_1 & 0 & m_{11} & m_{12} & ..... & m_{1K} \\ 0 & n_2 & m_{21} & m_{22} & ..... & m_{2K} \\ m_{11} & m_{21} & m_1 & 0 & ..... & 0 \\ m_{12} & m_{22} & 0 & m_2 & ..... & 0 \\ \vdots & \vdots & ... & ... & .... & \vdots \\ m_{1K} & m_{2K} & 0 & 0 & ..... & m_K \end{pmatrix}.$$

and the determinant of the above matrix is given by

$$det(I) \;=\; (\prod_{j=1}^{K} m_j)\left[n_1 n_2 - n_1 \sum_{j=1}^{K} \frac{m_{2j}^2}{m_j} - n_2 \sum_{j=1}^{K} \frac{m_{1j}^2}{m_j} + (\sum_{j=1}^{K} \frac{m_{2j}^2}{m_j})(\sum_{j=1}^{K} \frac{m_{1j}^2}{m_j}) - (\sum_{j=1}^{K} \frac{m_{1j}m_{2j}}{m_j})^2\right],$$

For $D-$optimality, we have to maximize $det(I)$ with respect to (n_1, n_2) and (m_{1j}, m_{2j}), for $j = 1, 2, \cdots, K$; subject to the conditions $n = n_1 + n_2$ and $m_j = \sum_{l=1}^{2} m_{lj}$, $j = 1, 2, \cdots, K$. For the sake of mathematical convenience, instead of considering the discrete design variables n_l and m_{lj} indicating numbers, we consider corresponding proportions given by $p = \frac{n_1}{n}, 1-p = \frac{n_2}{n}$, $q_j = \frac{m_{1j}}{m_j}, 1 - q_j = \frac{m_{2j}}{m_j}$, for $j = 1, 2, \cdots, K+1$. Defining $m = (\prod_{j=1}^{K} m_j)$, we can rewrite $det(I)$ in terms of p and q_j, for $j = 1, 2, \cdots, K$, as

$$\begin{aligned} det(I) \;&=\; m\left[n^2 p(1-p) - np\sum_{j=1}^{K} m_j(1-q_j)^2 - n(1-p)\sum_{j=1}^{K} m_j q_j^2\right] \\ &\quad + m\left[(\sum_{j=1}^{K} m_j q_j^2)(\sum_{j=1}^{K} m_j(1-q_j)^2) - (\sum_{j=1}^{K} m_j q_j(1-q_j))^2\right]. \end{aligned}$$

Differentiating $det(I)$ partially with respect to p and equating to zero, we get

$$mn\left[n(1-2p) - \sum_{j=1}^{K} m_j(1-q_j)^2 + \sum_{j=1}^{K} m_j q_j^2\right] \;=\; 0. \tag{5.1}$$

Observing $m_j > 0, j = 1, 2, \cdots, K$; and hence $m > 0$, we get after some routine algebra

$$n(1-2p) - \sum_{j=1}^{K} m_j + 2\sum_{j=1}^{K} m_j q_j \;=\; 0. \tag{5.2}$$

Now, differentiating $det(I)$ partially with respect to q_j and equating to zero, for $j = 1, 2, \cdots, K$, we get

$$\begin{aligned} 0 &= 2mm_j \left[np(1-q_j) - n(1-p)q_j + q_j \sum_{j=1}^{K} m_j (1-q_j)^2 \right] \\ &\quad - 2mm_j \left[(1-q_j) \sum_{j=1}^{K} m_j q_j^2 + (1-2q_j)(\sum_{j=1}^{K} (m_j q_j - m_j q_j^2)) \right]. \end{aligned} \tag{5.3}$$

Again observing $m_j > 0,$ for $j = 1, 2, \cdots, K$, $m > 0$, and $m_{K+1} = n - \sum_{j=1}^{K} m_j$, we get

$$np - q_j m_{K+1} - \sum_{j=1}^{K} m_j q_j = 0. \tag{5.4}$$

Multiplying the equation (5.4) by m_j and summing over j, for $j = 1(1)K$, we have

$$p = \frac{\sum_{j=1}^{K} m_j q_j}{\sum_{j=1}^{K} m_j}. \tag{5.5}$$

Using the value of p in equation (5.4) and after some simplifications, we obtain

$$q_j = \frac{\sum_{j=1}^{K} m_j q_j}{\sum_{j=1}^{K} m_j}, \quad j = 1(1)K. \tag{5.6}$$

Comparing equations (5.5) and (5.6), it follows that, the p and $q_j, j = 1(1)K$, have the same solution. So, in the equation (5.2), replacing q_j by p we have

$$(1-2p)(n - \sum_{j=1}^{K} m_j) = 0. \tag{5.7}$$

Observing $m_{K+1} > 0$, we get $p = \frac{1}{2}$ which implies $q_j = \frac{1}{2}, j = 1(1)K$. The relation $q_{K+1} = \frac{1}{2}$ follows immediately. Thus for a single categorical covariate, the $D-$optimal allocation of the experimental units into two treatments yields a balanced design. The analytical solution corresponding to other optimality criteria as well as dealing with $D-$optimality with more than one categorical covariates turn out to be difficult. However, in our simulation study, the optimum allocation scheme under all these criteria, obtained through the proposed algorithm, yields balanced allocation even for more than one categorical covariates each having several categories.

6. Robust Algorithm

It is to be noted that the objective functions for obtaining allocation schemes through $D-$ and $A-$optimalities are not the same and hence the allocation scheme obtained through one optimality criterion is likely to be different from that obtained through the other. As such, a natural question may be how well a nearly $D-$optimal allocation performs with regard to $A-$optimality and *vice versa*. In this regard, development of some algorithm is required that may not lead to $D-$ or $A-$optimalities individually, but gives an allocation scheme that is somewhat optimal with regard to both $D-$ and $A-$optimalities. We call such allocation to be a *robust* allocation scheme. In the present section, we will first introduce the idea of such allocation scheme through the concept of *max-min* efficiency and then modify the proposed algorithm in deriving such allocation scheme.

The efficiency for any allocation scheme α with respect to some benchmark or target allocation, denoted by α^b, is computed as the inverse ratio of the values of the objective function (usually variance) corresponding to α and α^b. In case the benchmark is the optimal allocation so that the efficiency of any other allocation scheme is less than or equal to unity. However, when the optimal allocation is difficult or impossible to find, especially with large number of experimental units involving several covariates of any type (quantitative, categorical, binary *etc.*), the benchmark might be taken to be an allocation that is easy to derive. In such situation, the computed efficiencies are usually more than or equal to one (since the benchmark allocation is not generally as good as the optimal allocation) and the objective of any algorithm is to maximize this efficiency as much as possible. As a candidate for the benchmark allocation α^b, we suggest using the solution $\alpha^{(0)}$ obtained through the quick method of Section 3. The idea is to consider the minimum of $D-$ and $A-$ efficiency of an allocation scheme α, computed with respect to the benchmark allocation α^b, and then to maximize this minimum over all the $(2^{n-1}-1)$ possible allocation schemes. The algorithm for finding the *max-min robust* allocation scheme with n experimental units with the initial allocation $\alpha^{(0)}$ is as follows:

Step 1 : Compute the efficiency of an allocation scheme α' with respect to α^b, as given by $\frac{V(\alpha^{(b)})}{V(\alpha')}$, for $\alpha' \in \mathcal{N}_{\alpha^{(0)}}$, where $V(\cdot)$ is the objective function as defined in Section 2, for both $D-$ and $A-$optimalities. Let us call them $D-$ and $A-$ efficiencies and denote them by $e_D(\alpha'|\alpha^b)$ and $e_A(\alpha'|\alpha^b)$, respectively.

Step 2: Compute $r(\alpha'|\alpha^b) = min[e_D(\alpha'|\alpha^b), e_A(\alpha'|\alpha^b)]$, for $\alpha' \in \mathcal{N}_{\alpha^{(0)}}$.

Step 3: If $max\{r(\alpha'|\alpha^b), \alpha' \in \mathcal{N}_{\alpha^{(0)}}\} > r(\alpha^{(0)}|\alpha^b)$, choose the next improved allocation to be $\alpha^{(1)} = arg\ max\{r(\alpha'|\alpha^b), \alpha' \in \mathcal{N}_{\alpha^{(0)}}\}$. Otherwise, choose $\alpha^{(1)}$ randomly from among the allocations given by $\{\alpha^{(0)}, \alpha' \in \mathcal{N}_{\alpha^{(0)}}\}$ with probabilities given by $p_0 = \frac{r(\alpha^{(0)}|\alpha^b)}{r(\alpha^{(0)})}$ and

$p_{\alpha'} = \frac{r(\alpha'|\alpha^b)}{r(\alpha^{(0)})}$, for $\alpha' \in \mathcal{N}_{\alpha^{(0)}}$, where $r(\alpha^{(0)}) = r(\alpha^{(0)}|\alpha^b) + \sum_{\alpha' \in \mathcal{N}_{\alpha^{(0)}}} r(\alpha'|\alpha^b)$.

Step 4: Replace $\alpha^{(0)}$ by $\alpha^{(1)}$ and repeat steps 2 and 3 until the convergence with some desired accuracy, which may be achieved in the same way as that in Section 4.

Note that this algorithm is essentially same as that of Section 3 except that the objective function here is to be maximized. It is interesting to note that, although we have considered only $D-$ and $A-$optimalities in the robust algorithm, the concept may be generalized to any number of optimality criteria by generalizing only the steps 1 and 2 of the algorithm by incorporating several efficiency values corresponding to the optimality criteria under consideration.

As an alternative, we may take a convex combination of efficiency values (instead of minimum) corresponding to $D-$ and $A-$optimalities (or any other criteria of interest) in step 2 of the algorithm, by computing

$$r(\alpha'|\alpha^b) = \delta \; e_D(\alpha'|\alpha^b) + (1-\delta) \; e_A(\alpha'|\alpha^b) \quad \text{with} \quad 0 \leq \delta \leq 1.$$

Here δ and $1-\delta$ are the indices of importance to be associated respectively with $D-$ and $A-$optimality efficiency values and, without any prior knowledge, the value of δ may be taken as 0.5. It may be noted that by putting δ to be *unity* or *zero*, we get the corresponding allocation with $D-$ or $A-$optimalities, respectively.

7. Illustrations

In this section, we investigate the performance of the proposed algorithm with regard to both $D-$ and $A-$optimalities and also with regard to the robust optimality. For this, we need to compare the design obtained by the proposed algorithm, or any other method, with a particular baseline design, usually chosen as the exact optimal design, with respect to the particular choice of the optimality criterion. However, as discussed before, the exact optimal design may not always be obtained, especially for large number of experimental units. In such situations, some other suitable choice for the baseline design may be considered. In the following, we consider the problem with a single covariate and multiple covariates (in particular, two covariates) as well and carry out extensive simulations to study the efficiency of our proposed algorithm. We also consider three real life examples to illustrate how the proposed algorithm would have performed.

7.1 Simulation Study with Single Covariate

In order to compare with the exact optimal design, we first consider only $n = 10$ experimental units in our simulation study so that, the exact optimal design can be worked

out. In our simulation, the n values for the single covariate are generated from a particular distribution. We consider four choices for this purpose as given by (1) Uniform $[0, 1]$ denoted by $U[0, 1]$, (2) Normal distribution with mean 0 and variance 10, denoted by $N(0, 10)$, (3) Exponential distribution with mean 25, denoted by $E(0.04)$ and (4) Cauchy distribution with location 0 and scale 1, denoted by $C(0, 1)$ with corresponding density

$$f(x) = \frac{1}{\pi(1 + x^2)} \quad .$$

These four distributions have been chosen to reflect different patterns (flat, symmetric, asymmetric and heavy tailed indicating presence of outliers, respectively) in the covariate values. Once the $n = 10$ values are generated from one of the four distributions, we obtain the following five allocation designs for the $n = 10$ units with these covariate values: (1) the initial design $\alpha^{(0)}$ by the quick method of Section 3, (2) the near-optimal design α^* by our proposed algorithm of Section 4, (3) the design α^R obtained by a random allocation, (4) the design α^{H_1} obtained by using Harville's method with interchange of experimental units, and (5) the design α^{H_2} obtained by using Harville's method with exchange of experimental units. We also obtain the exact optimal design α^e obtained by exhaustive search of all the $(2^{n-1} - 1)$ allocation designs. Note that the efficiency of any of the first five designs with respect to the exact optimal design, is defined by the ratio of the values of the objective functions for the exact optimal design and the particular design, $\frac{V(\alpha^e)}{V(\alpha)}$, with α denoting the particular design. The efficiencies of the five designs, given by $\frac{V(\alpha^e)}{V(\alpha^{(0)})}, \frac{V(\alpha^e)}{V(\alpha^*)}, \frac{V(\alpha^e)}{V(\alpha^R)}$, $\frac{V(\alpha^e)}{V(\alpha^{H_1})}$ and $\frac{V(\alpha^e)}{V(\alpha^{H_2})}$, respectively, are computed for this given set of $n = 10$ covariate values. In our simulation study, this exercise is carried out with both $D-$ and $A-$optimality criteria by considering 'both treatment and covariate effects' (denoted by $D-$ and $A-$, respectively) or 'only treatment effect' (denoted by D_s- and A_s-, respectively) for each of the four distributions mentioned at the beginning of this subsection. This results in $4 \times 5 = 20$ different cases, each with four different efficiency values. For each case, this computation of four efficiency values are repeated 1000 times. The means and ranges (in parentheses) over the 1000 such repetitions are reported in Table 1 for all the four optimality criteria. It has been found that efficiencies of all the designs, except the randomized design α^R, are close to 1 and the proposed design α^* seems to have the highest efficiency followed by the initial design $\alpha^{(0)}$ and then the Harville's two designs α^{H_1} and α^{H_2}. For that matter, even the randomized design α^R is nearly 90% as efficient as the exact optimal design α^e. The range of efficiency values corresponding to α^R is, however, quite large indicating that the randomized design may often turn out to be very inefficient. On the other hand, efficiency of α^* seems to be most stable with the smallest range followed by $\alpha^{(0)}$ and then α^{H_1} and α^{H_2}. We have carried out this excercise for values of n up to 20 (since finding the value of $V(\alpha^e)$ for larger values of n is computationally difficult) and in all cases the findings are more or less similar.

It has been already observed that obtaining exact optimal design with large number of experimental units with several known covariate values is immensely difficult. In such situations, we compare our proposed design α^* with those of Harville (α^{H_1} and α^{H_2}). We have taken the number of experimental units (n) to be 50 and 100 to represent moderately large and large sample sizes. For each simulation, n values of the single covariate are generated from one of the four distributions mentioned at the beginning of this subsection. Then, with these n values, the proposed design α^* and Harville's α^{H_1} and α^{H_2} are obtained and the efficiency of α^* with respect to α^{H_i}, given by $\frac{V(\alpha^{H_i})}{V(\alpha^*)}$, is computed, for $i = 1, 2$. This is repeated 1000 times and the corresponding mean and range (in parentheses) are presented in Table 2. As before, we consider both $D-$ and $A-$ as well as D_s- and A_s-optimality criteria. In each entry, the two rows correspond to comparison with α^{H_1} and α^{H_2}, respectively. Although the mean efficiency is marginally greater than *unity* in all cases, indicating marginal superiority of α^* over α^{H_1} or α^{H_2}, it is to be noted that this efficiency is never less than *unity* and sometimes as high as 2 (in case of C(0,1) distribution, that might be attributed to presence of outliers). This indicates that, besides being uniformly more efficient, α^* can be much superior to both α^{H_1} and α^{H_2} depending on the covariate values, apparently when these are skewed. Since obtaining α^{H_1} and α^{H_2} are computationally more involved, especially for large n, this observation appears to be an important factor in favor of α^*.

We now investigate the performance of the robust design of Section 6, denoted by α^{Ro}. By definition, this will be less efficient than α^* with regard to both $D-$ and $A-$optimality criteria, denoted by α^{*D} and α^{*A}, respectively. It is of interest to study how less efficient α^{Ro} is compared to α^{*D} (α^{*A}) when $D-$optimality ($A-$optimality) is considered. Therefore, the performance of α^{Ro} over α^{*D} (α^{*A}) under $D-$optimality ($A-$optimality) is of interest. In this context, the D_s-optimality (A_s-optimality) is also of interest in which case the robust design is obtained by considering the minimum of D_s- and A_s-efficiencies in step 2 of the algorithm in Section 6. The performance of α^{Ro} over the randomized design α^R has also been studied. Note that, for a particular optimality criterion (say, $D-$optimality) with the objective function $V(\cdot)$, the efficiency of α^{Ro} with respect to the reference design α^{*D} and α^R are given by $\frac{V(\alpha^{*D})}{V(\alpha^{Ro})}$ and $\frac{V(\alpha^R)}{V(\alpha^{Ro})}$, respectively. While considering $A-$,D_s- and A_s-optimalities, the efficiency of α^{Ro} with respect to α^{*A},α^{*D_s} and α^{*A_s}, respectively, (in addition to α^R) are accordingly obtained. The results based on 1000 simulations are presented in Table 3 for *max-min* robustness criterion. Similar findings are observed for *convex combination* criterion also and, hence, are not reported here. In each entry, the two rows correspond to comparison with the proposed and randomized designs, respectively. The mean efficiency is less than unity in the first case, as expected, while in the second case, the mean efficiency is more than unity. In all cases, the stopping probability for the proposed

algorithm as mentioned in step 4 of the algorithm has been taken to be 0.99.

Table 1. Mean efficiency of the designs obtained by different algorithms with respect to the exact optimal design over 1000 simulations (range in square brackets) with $n = 10$

Dist.	Design	$D-$	D_s-	$A-$	A_s-
U [0,1]	α^*	0.9997 [0.9947,1]	0.9999 [0.9945,1]	0.9999 [0.9948,1]	0.9999 [0.9949,1]
	$\alpha^{(0)}$	0.9971 [0.9480,1]	0.9970 [0.9480,1]	0.9972 [0.9520,1]	0.9976 [0.9696,1]
	$\alpha^{(R)}$	0.8847 [0.1684,1]	0.8847 [0.1684,1]	0.8916 [0.1842,1]	0.9399 [0.4141,1]
	$\alpha^{(H_1)}$	0.9519 [0.7975,1]	0.9650 [0.9182,1]	0.9816 [0.9428,1]	0.9555 [0.8789,1]
	$\alpha^{(H_2)}$	0.9332 [0.7020,1]	0.9568 [0.8767,1]	0.9732 [0.9359,1]	0.9265 [0.8304,1]
N (0,10)	α^*	0.9998 [0.9968,1]	0.9998 [0.9959,1]	0.9999 [0.9982,1]	0.9999 [0.9982,1]
	$\alpha^{(0)}$	0.9992 [0.9634,1]	0.9992 [0.9364,1]	0.9975 [0.9660,1]	0.9975 [0.9662,1]
	$\alpha^{(R)}$	0.8933 [0.3325,1]	0.8933 [0.3325,1]	0.9351 [0.4578,1]	0.9353 [0.4581,1]
	$\alpha^{(H_1)}$	0.9665 [0.9013,1]	0.9693 [0.9110,1]	0.9745 [0.9555,1]	0.9746 [0.9555,1]
	$\alpha^{(H_2)}$	0.9584 [0.8444,1]	0.9507 [0.8585,1]	0.9688 [0.9409,1]	0.9712 [0.9409,1]
E (0.04)	α^*	0.9997 [0.9859,1]	0.9998 [0.9927,1]	0.9999 [0.9817,1]	0.9999 [0.9817,1]
	$\alpha^{(0)}$	0.9983 [0.9621,1]	0.9848 [0.9217,1]	0.9988 [0.9504,1]	0.9940 [0.9504,1]
	$\alpha^{(R)}$	0.9145 [0.3823,1]	0.8864 [0.2244,1]	0.9132 [0.2918,1]	0.9132 [0.2920,1]
	$\alpha^{(H_1)}$	0.9681 [0.9372,1]	0.9626 [0.8314,1]	0.9786 [0.9064,1]	0.9725 [0.7969,1]
	$\alpha^{(H_2)}$	0.9489 [0.8901,1]	0.9505 [0.8179,1]	0.9617 [0.9043,1]	0.9563 [0.7437,1]
C (0,1)	α^*	0.9998 [0.9914,1]	0.9998 [0.9957,1]	0.9999 [0.9964,1]	0.9999 [0.9964,1]
	$\alpha^{(0)}$	0.9834 [0.9208,1]	0.9818 [0.9259,1]	0.9908 [0.9539,1]	0.9901 [0.9580,1]
	$\alpha^{(R)}$	0.9046 [0.2773,1]	0.9036 [0.2787,1]	0.9415 [0.4052,1]	0.9430 [0.3603,1]
	$\alpha^{(H_1)}$	0.9566 [0.8989,1]	0.9353 [0.8641,1]	0.9615 [0.9304,1]	0.9664 [0.9367,1]
	$\alpha^{(H_2)}$	0.9666 [0.9265,1]	0.9485 [0.8949,1]	0.9748 [0.9483,1]	0.9727 [0.9492,1]

Table 2. Mean efficiency of the proposed design with respect to those obtained by Harville's interchange algorithm (H_1) and exchange algorithm (H_2), respectively, over 1000 simulations (range in square brackets).

Distribution	n	$D-$	D_s-	$A-$	A_s-
U [0,1]	50	1.019 [1 , 1.089]	1.008 [1 , 1.064]	1.005 [1 , 1.058]	1.002 [1 , 1.054]
	50	1.020 [1 , 1.091]	1.043 [1 , 1.145]	1.0016 [1 , 1.111]	1.011 [1 , 1.091]
U [0,1]	100	1.004 [1 , 1.024]	1.031 [1 , 1.094]	1.006 [1 , 1.041]	1.002 [1 , 1.048]
	100	1.007 [1 , 1.033]	1.036 [1 , 1.105]	1.012 [1 , 1.061]	1.006 [1 , 1.054]
N (0,10)	50	1.016 [1 , 1.112]	1.002 [1 , 1.008]	1.003 [1 , 1.047]	1.003 [1 , 1.047]
	50	1.019 [1 , 1.154]	1.010 [1 , 1.044]	1.011 [1 , 1.068]	1.011 [1 , 1.069]
N (0,10)	100	1.019 [1 , 1.122]	1.006 [1 , 1.037]	1.003 [1 , 1.054]	1.002 [1 , 1.040]
	100	1.026 [1 , 1.223]	1.017 [1 , 1.068]	1.006 [1 , 1.072]	1.006 [1 , 1.057]
E (0.04)	50	1.075 [1 , 1.453]	1.001 [1 , 1.005]	1.025 [1 , 1.283]	1.016 [1 , 1.283]
	50	1.084 [1 , 1.488]	1.016 [1 , 1.104]	1.037 [1 , 1.361]	1.027 [1 , 1.361]
E (0.04)	100	1.034 [1 , 1.109]	1.010 [1 , 1.067]	1.024 [1 , 1.261]	1.054 [1 , 1.698]
	100	1.035 [1 , 1.112]	1.026 [1 , 1.090]	1.030 [1 , 1.276]	1.060 [1 , 1.691]
C (0,1)	50	1.166 [1 , 2.048]	1.007 [1 , 1.030]	1.001 [1 , 1.014]	1.002 [1 , 1.034]
	50	1.177 [1 , 2.083]	1.028 [1 , 1.066]	1.008 [1 , 1.038]	1.010 [1 , 1.051]
C (0,1)	100	1.092 [1 , 1.631]	1.008 [1 , 1.098]	1.001 [1 , 1.002]	1.001 [1 , 1.010]
	100	1.099 [1 , 1.637]	1.017 [1 , 1.057]	1.004 [1 , 1.015]	1.004 [1 , 1.023]

Table 3: Mean efficiency of the robust design (*max-min* criterion) with respect to the proposed design and randomised design over 1000 simulations (range in square brackets).

Distribution	n	$D-$	D_s-	$A-$	A_s-
U [0,1]	50	0.992 [0.897,1]	0.989 [0.754,1]	0.998 [0.724,1]	0.986 [0.595,1]
	50	1.017 [0.998,1.089]	1.001[1,1.002]	1.018 [1,1.144]	1.015[1, 1.152]
U [0,1]	100	0.997 [0.919,1]	0.998 [0.819,1]	0.993 [0.973,1]	0.989 [0.767,1]
	100	1.007 [0.999,1.049]	1.002 [1,1.005]	1.008 [0.999,1.075]	1.009 [1,1.123]
N (0,10)	50	0.995 [0.9303,1]	0.998 [0.9196,1]	0.998 [0.9462,1]	0.997[0.9847,1]
	50	1.017 [0.990,1.093]	1.017 [1,1.086]	1.021 [0.993,1.116]	1.010 [1,1.095]
N (0,10)	100	0.994 [0.9791, 1]	0.992 [0.964,1]	0.996 [0.985, 1]	0.998 [0.972,1]
	100	1.006[0.999,1.028]	1.003[1,1.015]	1.009[1,1.052]	1.005[1,1.022]
E (0.04)	50	0.996 [0.8647, 1]	0.994 [0.8237,1]	0.997 [0.926, 1]	0.998 [0.748,1]
	50	1.018 [0.986,1.201]	1.017 [1,1.099]	1.026 [0.983,1.166]	1.008 [1,1.140]
E (0.04)	100	0.997 [0.821, 1]	0.998[0.710,1]	0.987[0.896,1]	0.994 [0.797,1]
	100	1.017 [1,1.104]	1.006 [1,1.019]	1.009 [0.998,1.063]	1.007 [1,1.100]
C (0,1)	50	0.999 [0.952,1]	0.999 [0.945,1]	0.993 [0.958, 1]	0.998 [0.969,1]
	50	1.022 [0.996,1.083]	1.009 [1,1.036]	1.012 [1,1.062]	1.013 [1,1.066]
C (0,1)	100	0.997 [0.976,1]	0.999 [0.961,1]	0.998 [0.974,1]	0.998 [0.978,1]
	100	1.007 [0.999,1.031]	1.003 [1,1.012]	1.005 [0.991,1.066]	1.007 [1,1.045]

7.2 Simulation Study with Multiple Covariates

We now illustrate the superiority of the proposed design α^* over those of Harville (α^{H_1} and α^{H_2}) for multiple covariates. In particular, for the sake of simplicity, we consider two covariates which may be correlated. We also consider moderately large sample size of $n = 50$. We consider two particular configurations of covariates as described in the following:

Configuration1 : The vector $X = (X_1, X_2)$ of two dependent covariates for the n experimental units are generated from a bivariate normal distribution with parameters $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, denoted by $BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

Configuration2 : The two covariates X_1 and X_2 are independent with X_1 following a parametric distribution (Exponential or Logistic) and X_2 following a Bernoulli distribution with success probability p, denoted by $Ber(p)$. The logistic distribution with location parameter (α) 1.78 and scale parameter (β) 2.17 is a negatively skewed distribution, denoted by $L(1.78, 2.17)$, with the density

$$f(x; \alpha, \beta) = \frac{e^{-(x-\alpha)/\beta}}{\beta(1 + e^{-(x-\alpha)/\beta})^2}.$$

Note that, while the two covariates X_1 and X_2 are continuous and dependent in configuration 1, these two are independent with one being continuous and the other being binary in configuration 2. It is of interest to see how the optimal design allocates the units with respect to the binary covariate (e.g., male and female). Table 4 presents the results on efficiency of the proposed design α^* with respect to Harville's designs α^{H_1} and α^{H_2} based on 1000 simulations. In each entry, the first row gives efficiency of α^* over α^{H_1} and the second row gives the same over α^{H_2}. Comparing these efficiency values with those in Table 2, it appears that, with increasing number of covariates, the proposed design is more efficient than Harville's designs. We have also worked with three covariates (results not reported here) and found similar trend in efficiency values.

Table 4: Mean efficiency of the proposed design with respect to those obtained by Harville's interchange algorithm (H_1) and exchange algorithm (H_2), respectively, over 1000 simulations (range in square brackets) with $n = 50$.

distribution	$D-$	D_s-	$A-$	A_s-
BVN$(10, 5, 4, 5, 2)$	1.036 [1, 1.331]	1.052 [1, 1.289]	1.008 [1, 1.076]	1.011 [1, 1.101]
	1.050 [1, 2.554]	1.053 [1, 1.417]	1.008 [1, 1.076]	1.009 [1, 1.137]
BVN$(1, 10, 4, 5, 2)$	1.035 [1, 1.345]	1.042 [1, 1.200]	1.009 [1, 1.067]	1.008 [1, 1.065]
	1.054 [1, 2.053]	1.076 [1, 1.417]	1.010 [1, 1.064]	1.012 [1, 2.036]
E(0.04),	1.042 [1, 1.448]	1.030 [1, 1.187]	1.009 [1, 1.082]	1.018 [1, 1.095]
Ber(0.4)	1.064 [1, 2.704]	1.044 [1, 1.194]	1.010 [1, 1.086]	1.029 [1, 1.134]
L(1.78,2.17),	1.042 [1,1.546]	1.044 [1, 1.327]	1.019 [1, 1.123]	1.017 [1, 1.101]
Ber(0.35)	1.050 [1,1.722]	1.061 [1, 1.403]	1.020 [1, 1.126]	1.028 [1, 1.137]

7.3 The Analysis of Animal Health Laboratory Data

The data is taken from an experiment carried out at commercial goat farms in the Animal Health Laboratory, Lincoln, New Zealand (Saville and Wood, 1991) with an objective to determine whether the standard worm drenching program was adequate. For this, 40 goats were randomly allocated to either of the standard drenching program or the intensive drenching program, in which the goats were drenched more frequently with 20 goats in each program. The goats were individually tagged and weighed at the start and end of the year-long study. The response variable was live weight gain whereas the initial live weight was the single covariate. The main interest in the experiment was to compare the two programs with respect to live weight gain.

We have computed the efficiencies of the proposed design α^* and Harville's designs α^{H_1} and α^{H_2} with respect to the random allocation design α^R that has been used in the exper-

iment. Considering both treatment and covariate effects, the efficiency of α^*, α^{H_1} and α^{H_2} come out to be 1.0012(1.0002), 0.9757(0.9788) and 0.9379(0.9874), respectively, under $D-$(A−optimality). If our interest is on the treatment effect only, then the efficiency with regard to $D_s-(A_s-)$optimalities of α^*, α^{H_1} and α^{H_2} come out to be 1.0012(1.0002), 0.9759(0.9787) and 0.9405(0.9868), respectively. Consequently, the proposed design would have achieved some gain in efficiency, however small, while the Harville's designs would have been worse. The efficiency of the *robust* design α^{Ro} with respect to the randomized allocation design α^R turns out to be 1.0011(1.0001) and 1.0011(1.0002) for *max-min* procedure with regard to $D-(A-)$optimalities and $D_s-(A_s-)$optimalities, respectively. The small gain in efficiency of the proposed design over the random allocation design may be attributed to the fact that out of the 40 covariate values corresponding to the experimental units only 13 are distinct. The gain in efficiency would have been more had the covariate values corresponding to experimental units been distinct.

7.4 The Analysis of Dairy Cow Diets Data

This experiment (Wattiaux *et.al.*, 1994) involved 60 cows that were randomly assigned to one of 6 diets and followed for a number of weeks. Diets were begun after the third week, allowing the animals some initial time to adjust to their new environment. Interest focused on the effect of diet on the average dry matter intake (dmi), the amount of food eaten by each cow, a surrogate measure of milk production. As the covariate, the dmi values at 3^{rd} week were used. For the sake of demonstration, we have selected 'diet 4' and 'diet 6' as the two treatments, applied to 20 experimental units.

We have computed the efficiencies of the proposed algorithm α^*, Harville's interchange algorithm α^{H_1}, Harville's exchange algorithm α^{H_2} with respect to the random allocation design α^R. Considering both treatment and covariate effects as well as only treatment effect with regard to $D-(A-)$optimalities respectively, the efficiencies of α^*, α^{H_1} and α^{H_2} come out to be 1.0420(1.0410), 1.0090(1.0070), 1.0070(1.0050). At the same time the efficiency values by considering $D_s-(A_s-)$optimalities for the same designs are respectively 1.0416(1.0408), 1.0120(1.0140), 1.0120(1.0090). Thus the proposed design is more efficient than any of Harville's methods, in comparison to the randomization procedure.

As the sample size is relatively small here, we have also computed the exact optimal allocation scheme through the time consuming direct computational approach, and found the loss in efficiency of the proposed algorithm as well as Harville's algorithm. Considering the two cases of estimation of both the treatment and covariate effects and only the treatment effect with regard to their optimalities. For $D-(A-)$optimalities, the efficiencies of α^* , α^{H_1} and α^{H_2} with respect to the exact optimal design α^e respectively come out

to be 1.0000(0.9999), 0.9987(0.9977), 0.9957(0.9926). At the same time similar efficiencies with regard to $D_s-(A_s-)$optimalities have been obtained as 0.9999(0.9999), 0.9955(0.9975), 0.9985(0.9984) respectively. It is apparent from the results that the proposed algorithm sometimes produce the exact optimal allocation design and performs consistently better than Harville's allocation scheme in all the cases considered.

The robust algorithm for *max-min* procedure is also applied on the given data to evaluate the efficiency with respect to the given randomized allocation. The mean efficiency of the robust algorithm α^{Ro} with regard to $D-(A-)$optimalities as well as $D_s-(A_s-)$optimalities are 1.0420(1.0410) and 1.0415(1.0407), respectively.

7.5 The Analysis of Randomized Clinical Trial for Severe Aplastic Anemia

The data from a randomized clinical trial on 64 patients with severe aplastic anemia has been reported in Table 1.2 of Kalbfleisch and Prentice (2002). Before the trial started, all the patients were treated with high-dose cyclophosphamide followed by an infusion of bone marrow from an HLA-identical family. The patients were assigned then to each of the two treatment groups: either the combined therapy of cyclosporine and methotrexate ($CSP + MTX$) or methotrexate alone (MTX). For each patient the information on patient's age (in years) at the time of transplant and an indicator of whether or not the patient was assigned to laminar airflow (LAF) isolation room. In the present study, these two will be used as known covariates. The endpoint of interest was the time from assignment until the diagnosis of a life-threatening stage of acute graft versus host disease. In the present set up we assume log(time) to follow a normal distribution such that the existing distributional assumption is met. Our objective is to obtain the gain in efficiency with respect to estimation of the treatment effect along with covariate effect, or only the treatment effect, had the experimenter used the present allocation scheme or Harville's schemes (H_1 and H_2) in place of the random allocation scheme that was used in the study. The efficiencies of α^*, α^{H_1} and α^{H_2} with regard to $D-(A-)$optimalities respectively come out to be 1.0133(1.0075), 1.0001(1.0002), 1.0003(1.0001). Similarly, the respective efficiencies with regard to $D_s-(A_s-)$optimalities have been obtained as 1.0133(1.0106), 1.0002(1.0002), 1.0002(1.0001). The efficiency of the robust algorithm α^{Ro} for *max-min* procedure with respect to the given allocation design for $D-(A-)$optimalities as well as $D_s-(A_s-)$optimalities are also obtained as 1.0074(1.0034) and 1.0074(1.0034), respectively.

7.6 Discussion

The first two real life examples are concerned with single continuous covariate while the last one deals with a mixture of binary as well as continuous covariate. The final allocation through the proposed allocation scheme in the last example also preserves balancing at least with the binary covariate. In the three real life examples, it is evident that although the proposed algorithm allocates treatments in an efficient way to the available experimental units with known covariate(s), the gain in efficiency is marginal compared to the randomized allocation rule or Harville's exchange or interchange algorithms. However, in several simulation studies, it has been aptly demonstrated that there is considerable gain in efficiency in the proposed algorithm in comparison to the other allocation algorithms, provided the covariate values are highly skewed instead of being symmetrical, or there are several outliers in the covariates even if they are coming from a symmetrical distribution. As for example, in table 2, the efficiency of the proposed algorithm in comparison to Harville's (α^{H_1} or α^{H_2}) algorithms with regard to $D-$optimality is sometimes increased by two times in case of covariates coming from a Cauchy distribution, which is heavy tailed and thus, although symmetric, consists of several outlying observations. As such, we have tried to assess the gain in efficiency of the proposed algorithm with respect to other algorithms for the real life data by randomly replacing a small portion of the continuous covariate values with skewed or outlying values thus making the overall covariate values to some extent skewed, or forcing the presence of outliers in the data. To accomplish this *what if* type analysis, we have drawn a random sample of size m experimental units from each of the three real life data and, for the selected experimental units, the continuous covariate values are replaced by random values drawn from $Uniform\ (1.5x_{(n)}, 2x_{(n)})$ distribution, where $x_{(n)}$ stands for the largest continuous covariate value for the given data.

We elaborate the procedure for the analysis of animal health laboratory data. Out of 40 experimental units, a sample of size $m = 10$ is drawn randomly. Then the covariate values of the selected units are replaced by 10 values randomly selected from $Uniform\ (45, 60)$ distribution, as the value of the highest observation in the given data is 30. In the process we have introduced some outliers in the covariate values with the assumption that there are few goats which are giants comparable to others with very high initial live weight. The process of such random removal and replacement by high covariate values, is repeated 1000 times and, for each repetition, the efficiencies of the proposed algorithm as well as other allocation algorithms over random allocation scheme are computed . Similar exercise has been carried out for the other two examples also. For the second example, 5 experimental units are randomly selected and their covariate values are replaced by random samples from $Uniform\ (41, 54)$ distribution. For the third example, 15 experimental units are randomly

selected and their *age* covariate is replaced by random samples from $Uniform\ (63, 84)$ distribution, keeping the other binary covariate value corresponding to LMF fixed. The mean efficiencies over 1000 such repetitions for the proposed algorithm (α^*) as well as Harville's algorithms (α^{H_1} and α^{H_2}) in comparison to the random allocation scheme with regard to $D-(A-)$ as well as $D_s-(A_s-)$optimalities for the three examples are reported in Table 5. The corresponding values for the robust algorithm α^{Ro} with *max-min* criterion in comparison to random allocation scheme are also reported in Table 5.

Table 5. Mean efficiency (range in square brackets) of the designs obtained by different algorithms with respect to the random allocation design over 1000 simulations for the forced skewed data of real life examples.

Source	m	Design	$D-$	D_s-	$A-$	A_s-
Animal		α^*	1.029[1.004,1.265]	1.029[1.004,1.265]	1.027[1.004,1.203]	1.028[1.004,1.205]
Health		$\alpha^{(Ro)}$	1.029[1.004,1.265]	1.029[1.004,1.265]	1.027[1.004,1.203]	1.028[1.004,1.205]
Lab	10	$\alpha^{(H_1)}$	1.009[1.001,1.034]	1.010[1.001,1.033]	1.004[1.000,1.015]	1.004[1.000,1.014]
Data		$\alpha^{(H_2)}$	1.007[1.000,1.031]	1.008[1.000,1.032]	1.005[1.000,1.019]	1.005[1.000,1.019]
Dairy		α^*	1.111[1.014,1.671]	1.111[1.014,1.672]	1.086[1.009,1.469]	1.079[1.008,1.439]
Cow		$\alpha^{(Ro)}$	1.108[1.012,1.664]	1.108[1.012,1.664]	1.082[1.009,1.432]	1.076[1.008,1.431]
Diet	5	$\alpha^{(H_1)}$	1.026[1.003,1.063]	1.027[1.003,1.061]	1.018[1.002,1.031]	1.016[1.002,1.028]
Data		$\alpha^{(H_2)}$	1.023[1.003,1.057]	1.024[1.003,1.055]	1.017[1.002,1.030]	1.016[1.002,1.031]
Severe		α^*	1.228[1.031,1.723]	1.229[1.031,1.722]	1.132[1.023,1.481]	1.130[1.023,1.483]
Aplastic		$\alpha^{(Ro)}$	1.228[1.030,1.722]	1.228[1.031,1.722]	1.132[1.023,1.479]	1.131[1.022,1.480]
Anemia	15	$\alpha^{(H_1)}$	1.051[1.007,1.084]	1.049[1.007,1.084]	1.038[1.006,1.063]	1.036[1.005,1.064]
Data		$\alpha^{(H_2)}$	1.053[1.008,1.087]	1.050[1.007,1.075]	1.037[1.005,1.063]	1.036[1.005,1.062]

It is evident from Table 5 that there is considerable gain in average efficiency of the proposed algorithm for the skewed data in comparison to Harville's algorithms. In some situations, even the robust algorithm gives more efficiency than the Harville's allocation schemes. The 70% gain in maximum efficiency value by following the proposed algorithm over the random allocation design demonstrates that the present algorithm is best suited when the covariate values are skewed, or there are outliers in the covariate values.

8. Concluding Remarks

The present paper deals with optimal allocation of several experimental units to two treatments so as to attain optimality through various criteria in the context of efficient estimation of only treatment effect or both treatment and covariate effects, when the covariate(s)

values are known for all the experimental units. A quick method, mainly intended for initial solution is introduced, while an algorithm similar to VNS method has been proposed to obtain a better design. It has been demonstrated through various simulation studies that the proposed algorithm has maximum gain in efficiency in comparison to either completely random allocation or the Harville's methods. A robust method has also been explored to formulate a compromise solution between the objectives of $D-$ and $A-$optimalities by modifying the proposed algorithm. The proposed allocation rule is computationally convenient and gives comparatively better design especially with large number of experimental units and several covariates. One may be interested in efficient estimation of only the covariate effect. The corresponding optimal allocation design can be similarly obtained by considering the corresponding submatrix of I^{-1}

Our quick algorithm is used to maximize the within sum of squares of covariate values over the groups which is equivalent to minimizing the between sum of squares among the given covariates. As such this procedure helps us to restore homogeneity among the covariate values over various groups as much as possible. One significant contribution of this work is the judicious choice of the initial design which often has almost full efficiency and is better than those of Harville but requiring much less computation. The algorithm with a randomization step increases the possibility of attaining the globally optimal design. Harville one-at-a-time approach with an initial randomization not only requires heavy computations, may also drift away from the optimal design. With the refinement suggested in that work, the performance does not seem to improve much. While choosing ρ (= 2 or 3) units at a time, there may be little gain in efficiency, but at the cost of much computing burden.

Several interesting extensions and generalizations may be suggested in the context of the present work. The first in line is how the proposed algorithm might be implemented if more than two treatments are to be considered for allocation of several experimental units with known covariate(s) values. The problem may be translated in terms of finding all the neighbors of a particular allocation scheme through some computationally convenient way and then developing the corresponding VNS algorithm in line of the present work. The analytical solution of several categorical covariates with multiple labels needs to be explored more exhaustively for two as well as more than two treatment cases. There may be situations when a block structure is to be maintained in the optimal allocation. Cook and Nachstsheim (1989) generalized Harville's allocation scheme in this context. The proposed algorithm may be generalized by considering a suitable model incorporating block effect.

The design problems considered here and also in several other citations are concerned with efficient estimation of effects, be it treatment or covariate or both. Another class of design problem may be formulated in terms of efficient allocation of experimental units with

known covariates to the treatments while the significance of the treatment effect is to be tested. Works in this direction are being carried out by the present authors and will be communicated separately.

Acknowledgment

The first author greatfully acknowledges the financial support of the University Grant Commission (Ref. No. UGC/176/Jr. Fellow(Sc)),Govt. of India .

References:

1. Atkinson, A.C. (1982): Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika* **69** (1), 61 - 67.

2. Atkinson, A.C. (2002): The comparison of designs for sequential clinical trials with covariate information. *Journal of Royal Statistical Society, Series A* **165** (2), 349 - 373.

3. Bandyopadhyay, U. and Biswas, A. (2001): Adaptive designs for normal responses with prognostic factors. *Biometrika* **88** (2), 409 - 419.

4. Cook,R.D. and Nachstsheim,C.J. (1989): Computer-aided blocking of factorial and response-surface designs. *Technometrics* **31** (3), 332 - 346.

5. Das, K., Mondal, N.K. and Sinha, B.K. (2003): Optimal experimental designs for models with covariates. *Journal of Statistical Planning and Inference* **115** (1), 273 - 285.

6. Dey, A. and Mukerjee, R. (2006): D-optimal designs for covariate models. *Statistics* **40** (4), 297 - 305.

7. Hansen, P. and Mladenovic, N. (2001): Variable Neighborhood search: Principles and Applications. *European Journal of Operational Research* **130** (3), 449 - 467.

8. Harville, D. A. (1974): Nearly optimal allocation of experimental units using observed covariate values. *Technometrics* **16** (4), 589 - 599.

9. Kalbfleisch, J. D., and Prentice, R. L. (2002) : The Statistical Analysis of Failure Time Data: Second Edition, *Ch. 1.1.2*, New York: John Wiley and Sons Inc.

10. Kiefer, J. (1959): Optimum Experimental Designs. *Journal of Royal Statistical Society, Series B* **21** (2), 272 - 319.

11. Kirkpatrick, S., Gelatt, C.D. and Vecci, M.P. (1983): Optimization by Simulated Annealing. *Science* **220** (4598), 671 - 680.

12. Rao, A.R. and Bhimasankaran, P. (2000): Linear Algebra, Second edition, *Ch. 3*, New Delhi: Hindustan Book Agency.

13. Saville, D. J. and Wood, G. R. (1991): Statistical Methods: The Geometric Approach. *Ch. 17*, New York: Springer.

14. Wattiaux MA, Combs DK and Shaver RD (1994): Lactational responses to ruminally undegradable protein by dairy cows fed diets based on alfalfa silage. *Journal of Dairy Science* **77** (6), 1604-1617.