NATIONAL FELLOW SCHEME

EFFICIENT DESIGN OF EXPERIMENTS FOR QUALITY AGRICULTURAL RESEARCH

ANNUAL PROGRESS REPORT
2007-2008

Rajender Parsad

INDIAN AGRICULTURAL STATISTICS RESEARCH INSTITUTE
(ICAR)
LIBRARY AVENUE, PUSA, NEW DELHI- 110 012

Contents

Introduction

Agricultural research and statistical methodologies including design of experiments go hand in hand. The quality of agricultural research can be improved through the use of modern sophisticated statistical methodologies and appropriate designed experiments, efficient in terms of cost and precision. The choice of an appropriate design for a given experimental setting is important. The design adopted in the experiment should be able to answer all the questions of the experimenter but, more importantly, it should take care of the variability in the experimental material. If the variability in the experimental material is not taken care of properly through the design, then the experimental error is large and, hence, the CV is large. A large number of experiments have been conducted in National Agricultural Research System (NARS) where in the variability in the experimental material is controlled through formation of blocks. Randomized complete block (RCB) designs have been used extensively by the scientists in NARS. For experiments with large number of treatments (levels of a single factor) or treatment combinations (combinations of levels of multi-factors), the blocks become large and it may not be possible for the experimenter to maintain homogeneity among the experimental units within blocks of a RCB design. As such the primary purpose of forming the blocks to have homogeneous experimental units within a block is defeated. A direct consequence of laying out an experiment in RCB design with large number of treatments is that the coefficient of variation (CV) of the design becomes large. It has been seen that in a large number of these experiments, the replication effect is not significant. In other words, the mean square due to replications is not large in comparison to error mean squares. This raises a serious concern about blocking in field experimentation, i.e., either the blocking is not effective or is faulty. This, at times, leads to the rejection of the experiment, which is a loss of expensive resources and time.

This amount to saying that the error sum of squares is large compared to the sum of squares attributable to the model and hence small treatment differences may not be detected as significant. It also leads to a poor precision of comparisons of the two treatment effects or estimation of any normalized treatment contrast. High CV of the experiments is a very serious problem in agricultural experimentation. Many experiments conducted are rejected due to their high CV values.

Another important issue in the choice of an appropriate design is related to the objectives of the experiment or the problems to be solved, which generally is some specified type of comparisons among the treatments or some given treatment contrasts need to be estimated through the design. For instance, in varietal trials all the possible pairwise treatment comparisons could be of interest to the experimenter. In some other situations, when one particular treatment is on a special footing, also termed as control, as compared to other treatments, called the tests, the tests versus control comparisons are of interest to the experimenter. Similarly, when the treatment structure is factorial in nature, some special treatment contrasts known as factorial effects are of interest. But one design cannot be good for every problem. In other words, one has to pick up a design for the specific problem.

We have so far highlighted some problems that need to be taken care of in the choice of an appropriate design. This is the problem of designing an experiment. The data generated from a designed experiment is analyzed using the technique of analysis of variance. The proper choice of an experimental design reduces the per plot variance leading to a considerable reduction in error mean squares and, hence, the precision of the treatment comparisons. Further, contrast analysis can be used to answer further questions that cannot be answered by the usual analysis of variance. At present, the Indian Agriculture has steered from the era of chronic food deficits to that of self sufficiency and even exports. In fast changing scenario, to make agricultural research competitive, it is essential that sound statistical methodologies be used for the collection and analysis of experimental data.

This scheme aims at providing efficient design of experiments and sound analytical procedures so as to maintain high standards and good quality of agricultural research so as to make agricultural research globally competitive. The main emphasis is being laid on obtaining designs that simultaneously maximize the precision and economize on the experimental resources. These designs will be quite useful for the experiments with limited experimental material particularly the experiments for assessing the impact of soil erosion on crop productivity.

In many agricultural experiments generally the data on more than one character is observed. The situations where several responses are observed on each experimental unit are common in agronomy, plant breeding, agro-forestry, microbiology, soil sciences, experiments conducted to evolve strategies of soil and water conservation, etc. Experiments in which data on several responses are measured from an experimental unit corresponding to the application of a treatment are known as multi-response experiments. Multi-response experiments are of two types viz. complete multi-response experiments (all the response variables are recorded from each experimental unit) and incomplete multi-response experiments (recording of all the responses variables from each experimental unit is not feasible).

During the period under report, the emphasis was made on obtaining economic and efficient and designs for complete and incomplete multi-response experiments and on the development of analytical procedures for the analysis of experimental data related to incomplete multi-response experiments.

For complete multi-response experiments, it has been shown that the designs that are efficient for single response experiments are also efficient for complete multi-response experiments provided that the number of response variables is less than the error degrees of freedom.

A method of construction of designs for incomplete multi-response experiments is also obtained using combination of randomized complete block (RCB) designs and balanced incomplete block (BIB) designs. Here RCB design is treatment-wise design (experimental designs used to allocate the treatments in each set) and BIB design is response-wise design (experimental designs used to allocate the response variables in different sets). The designs obtainable from this method are economical from resource point of view.

A step wise procedure of analysis of incomplete multi-response designs obtained as a combination of RCB design and BIB design has been developed.

β-version of software for generation of nested block designs both for independent errors and correlated errors has been developed.

To disseminate the knowledge available on the combinatorial aspects of designs and analytical procedures acquired to the scientists engaged in research in National Agricultural Research System, the advisory services are pursued rigorously. As a consequence, α-designs have been adopted for initial varietal trials with 30 entries by All India Co-ordinated Research Project on Rapeseed and Mustard. Organized **3 dissemination workshops** on (i) Design Resources Server; (ii) Outliers in Designed Experiments and (iii) Fractional Factorials with special emphasis on Experiments with Scarce Resources and **one discussion Seminar** on Experiments with Mixtures: Theoretical Advances and Applications. A meeting with senior breeders of AICRP on Rapeseed and Mustard to finalize the designs for the initial varietal trials and criteria of promotion of entries was also organized.

Design Resources Server developed has been strengthened in collaboration with National Professor by adding links on α-designs, designs for bioassays, supersaturated designs, modules for generation of randomized layout of square lattice designs, basic designs such as completely randomized designs, RCB designs, Latin square designs and augmented designs. A new page "Analysis of Data" has been launched on Design Resources Server that provides steps of analysis of data generated through designed experiments using SAS and SPSS. A Discussion board has been initiated for sharing research with fellow scientists over the globe or for flagging issues for attention of scientific community. A list of experts in design of experiments over the globe is uploaded which will be useful for establishing linkages. A brief description of the achievements made is given in the sequel.

A) Block Designs for Multi-response Experiments

A.1. Introduction

In many agricultural experiments, generally data on more than one character is observed. For example, in a varietal trial, data on several plant characteristics such as plant height, plant population, days to maturity, number of primary branches per plant, 1000 seed weight, yield and disease resistant characteristics, etc. are observed. Similarly an agronomist may be interested in studying the effect of tillage and nutrient interactions on the growth and yield of rice crop. Besides yield, s/he observes the data on dry weight, root weight, leaf area, nitrogen uptake, etc. to study the plant growth. Situations where several responses are observed on each experimental unit are also common in experiments conducted to evolve strategies for soil and water conservation practices. For example, consider an experiment where the experimenter is interested in studying the effects of tillage practices and crop residue on erosion losses and crop yield. The observations may be made on run-off water as percentage of rainfall, soil loss and yield. Many more similar experimental situations exist in practice.

Experiments in which data on several responses are measured from an experimental unit corresponding to the application of a treatment are known as multi-response experiments. Multi-response experiments are of two types viz. complete multi-response experiments (all the response variables are recorded from each experimental unit) and incomplete multi-response experiments (recording of all the responses variables from each experimental unit is not feasible).

We have reported the results obtained on the analytical aspects of data from multi-response experiments conducted in a block design last year. Besides the analysis of data, the designing of multi-response experiments also needs attention. These experiments are generally conducted using designs that are optimal/ efficient for single response situations. The question, therefore, is, *whether the designs that are optimal/ efficient for single response situations are efficient for multi-response experiments*? To answer this question, we have studied the optimality aspects of the designs for complete multi-response experiments, *i.e*. when the data on each of the variables is collected from each experimental unit.

In many multi-response experiments, due to constraints on resources and time, it may not be feasible to collect the observations on all the response variables from each experimental unit. As a consequence, the data on a subset of response variables is collected from one subset of units; the data on other subset of response variables is collected from other subset of units. Some of the subsets of units/ response variables may be common. Such experiments are quite common in agro-forestry experiments, intercropping experiments, microbial experiments where the interest is in studying rate of growth of a number of bacterial cultures and it is not possible to record all the growth parameters.

For the situations where the collection of data on all the responses is physically impossible, consider that an anthropologist has unearthed a number of skulls, each of which is in a partly mutilated condition. Then the kinds of measurements that one can take on a skull

may differ from one skull to the other. In some experimental situations, it may be inconvenient to measure the responses on each unit due to paucity of time. One such example is given by Srivastava (1968) and is given below:

Let a microbiologist be interested in studying the growth of a number of bacterial cultures (*Rhizobium*, *Azotobacter*, *Escherichia Coli.* etc.). Let there be p growth parameters (population per colony @ 2 days after inoculation, size of the colony, colony shape, growth in different carbon sources, etc.) to be observed. On each of the bacterial cultures, the experimenter could observe all the p responses if the process of measurement is fast during which the experimental conditions remains unchanged. However, if the process of measurement is necessarily slow, one may have to content oneself with fewer ($p_1 < p$) responses on each unit. For more examples on such experimental situations, one may refer to Trawinski and Bargmann (1964), Roy *et al*. (1971), Gupta (1988) and Brzeskwiniewicz (1995).

The experiments in which incomplete set of responses are recorded from each experimental unit are known as incomplete multi-response experiments. Block designs for these experiments are referred as *incomplete multi-response* designs. In incomplete multi-response designs, n experimental units are divided into u (>1) disjoint sets such that l^{th} set has n_l experimental units $\sum_{l=1}^{u} n_l = n$. On each of the n_l units (l =1, 2, …, u) in the l^{th} set, the ($p_l < p$) response variables $V_{l1}, V_{l2}, ..., V_{lp_l}$ are measured, these being selected from the total set of p responses $V_1, V_2, ..., V_p$ according to a rule D_1 which may, henceforth, be called the "*response-wise design*" (design used to allocate the response variables in different sets of units). Furthermore, we envisage that for each l, there is an ordinary "*treatment-wise design*" (design used to allocate the treatments in each set) D_{2l} defined over the n_l units in the l^{th} set. The set of blocks under the design D_{2l} may of course be different for different l. The *incomplete multi-response design* D is then defined to be the total design over all the n units, and is fixed by the (u + 1)-tuplets $D = (D_1, D_{21}, D_{22}, ..., D_{2u})$.

Srivastava (1966, 1968) developed a theory of analysis of incomplete multi-response designs by transforming the data to fit into the general framework of the multivariate analysis of variance. The designs which can be analyzed through the approach of Srivastava (1968) are called regular (the incomplete multi-response designs, in which treatment-wise design from all sets are same). In the literature, this has been done for the situations, when the treatment-wise design in each set is a balanced incomplete block (BIB) design (Srivastava, 1968). Some classes of regular incomplete multi-response designs using randomized complete block (RCB) and BIB designs have been given by Srivastava and McDonald (1970).

Gupta (1988) defined a class of designs, termed as Type A designs, for incomplete multi-response experiments. Type A designs consists of u sets of experimental units S_1, S_2, …, S_u, l^{th} set containing n_l units. Divide these sets into u' groups. The *response-wise design*

D_1 consists of u blocks. To the set S_l a block of response-wise design is allotted, the response variables in this block are recorded on each unit in S_l. *Treatment-wise design* is same in a group. Design D_1 consists u' groups of blocks, the l^{th} group containing m_l blocks. Further in the m_l blocks, irrespective of the group to which they belong,

i) any pair of response variables (h, h`) occurs together in $\lambda_{hh`}$ blocks, ($h \neq h` = 1, 2, \ldots, p$), and
ii) the h^{th} response variable occurs in exactly λ blocks, ($h = 1, 2, \ldots, p$).

Mitra and Saha (1991) gave a method of analysis of incomplete multi-response designs where the treatment-wise designs are variance balanced and/ or efficiency balanced block designs.

Brzeskwiniewicz (1995) presented an idea of using incomplete block designs for multi-response experiments when not all responses can be observed from all experimental units, particularly where the labour requirements in taking measurements or observations hinders the recording of all responses on all experimental units at the same time. He has used a partially balanced incomplete block design as treatment-wise design in each set, but the designs used in different sets are not same.

The designs obtained in the literature were either very large or requires huge amount of resources (cost, labour and time). So there is a need to obtain designs which are efficient and require fewer resources.

In the present investigation, an attempt has been made to answer the question, whether the designs that are efficient for single response situations are efficient for complete multi-response experiments as well? An attempt has also been made to obtain a class of designs for incomplete multi-response experiments which require fewer resources (time and labour).

In section A.2, we shall deal with the designing aspects of the complete multi-response experiments, specifically on the optimality aspects of designs for complete multi-response experiments. It has been shown that the designs that are efficient for single response experiments are also efficient for complete multi-response experiments. The only requirement is that the number of response variables should be less than the error degrees of freedom. Designing aspects of incomplete multi-response experiments have been discussed in section A.3, where we have obtained a method of construction of designs for incomplete multi-response experiments. Further, the analytical procedure developed for incomplete multi-response designs obtained in the present study is given in Section A.4.

A.2. Designs for Complete Multi-response Experiments

Consider an experiment where v treatments are arranged in b blocks such that j^{th} block contains k_j experimental units; $j = 1,2,\ldots,b$ and treatment i is replicated r_i times, $\sum_{j=1}^{b} k_j = \sum_{i=1}^{v} r_i = n$, the total number of experimental units. From each experimental unit p

responses are observed. Let $\mathbf{Y} = [\mathbf{y}_1\ \mathbf{y}_2\ \ldots\ \mathbf{y}_p]$ be $n \times p$ matrix of observations, where $\mathbf{y}_s$ is an $n \times 1$ vector of observations corresponding to the s^{th} response ($s = 1,2, \ldots, p$).

For s^{th} response the model is given by

$$\mathbf{y}_s = \mathbf{X}_s\boldsymbol{\theta}_s + \boldsymbol{\varepsilon}_s \qquad s = 1, 2, \ldots, p \tag{A.2.1}$$

where $\mathbf{X}_s = \left[\boldsymbol{\Delta}_s' \quad \mathbf{1} \quad \mathbf{D}_s'\right]$ is the design matrix for s^{th} response partitioned in conformity with the parameters, $\boldsymbol{\Delta}_s'$ is $(n \times v)$ design matrix of treatments for the s^{th} response, $\mathbf{1}$ is the n dimensional column vector of all elements unity and $\mathbf{D}_s'$ is the design matrix of blocks for the s^{th} response.

$\boldsymbol{\theta}_s = \begin{bmatrix} \boldsymbol{\tau}_s \\ \mu_s \\ \boldsymbol{\beta}_s \end{bmatrix}$ is a $(v + b + 1) \times 1$ vector of parameters, $\boldsymbol{\tau}_s$ is a $v \times 1$ vector of treatment effects, μ_s is general mean and $\boldsymbol{\beta}_s$ is a $b \times 1$ vector of block effects and $\boldsymbol{\varepsilon}_s$ is the residual vector for s^{th} response variable distributed as N ($\mathbf{0}$, $\sigma_{ss}\mathbf{I}_n$).

The model for multi-response experiments in block design set up after rolling down the data pertaining to p-responses one below another is

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{A.2.2}$$

where $\mathbf{Y} = \left(\mathbf{y}_1' \quad \mathbf{y}_2' \quad \cdots \quad \mathbf{y}_p'\right)'$

$$\mathbf{Z} = \left[\begin{bmatrix} \boldsymbol{\Delta}_1' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Delta}_2' & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Delta}_p' \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2' & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}_p' \end{bmatrix}\right] = \bigoplus_{s=1}^{p} \mathbf{X}_s \text{ and } \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\tau} \\ \boldsymbol{\mu} \\ \boldsymbol{\beta} \end{bmatrix}. \tag{A.2.3}$$

Here treatment effects vectors, general mean and block effects vectors for all the response variables are appended one below the other to obtain a single treatment effect vector $\boldsymbol{\tau}$, general mean vector $\boldsymbol{\mu}$ and block effect vector $\boldsymbol{\beta}$, $\boldsymbol{\varepsilon}$ follows p-variate normal distribution with response variables from same observation are correlated but there is no correlation between different observations and $\oplus$ denotes the direct sum of matrices. As a consequence $\boldsymbol{\varepsilon} \sim \mathrm{N}_p(\mathbf{0}, \boldsymbol{\Omega})$, where

$$\boldsymbol{\Omega} = D(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma_{11}\mathbf{I}_n & \sigma_{12}\mathbf{I}_n & \cdots & \sigma_{1p}\mathbf{I}_n \\ \sigma_{21}\mathbf{I}_n & \sigma_{22}\mathbf{I}_n & \cdots & \sigma_{2p}\mathbf{I}_n \\ \vdots & \vdots & & \vdots \\ \sigma_{p1}\mathbf{I}_n & \sigma_{p2}\mathbf{I}_n & \cdots & \sigma_{pp}\mathbf{I}_n \end{bmatrix} = \boldsymbol{\Sigma}_{pp} \otimes \mathbf{I}_n, \tag{A.2.4}$$

σ_{ss} denotes the variance of s^{th} response variable and $\sigma_{ss'}$ is the covariance between s^{th} and s'^{th} response variables when $s \neq s'$. $s, s' = 1, 2, \ldots, p$; $\otimes$ denotes Kronecker product of matrices and D(.) denotes the dispersion matrix.

Using the Generalized Least Squares (GLS) estimation procedure, the normal equations are

$$\left(\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}\right)\boldsymbol{\theta} = \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Y} \qquad \text{(A.2.5)}$$

Now consider that from each experimental unit all responses are observed, i.e., we are in a complete multi-response situation. Then we have,

$\boldsymbol{\Delta}'_s = \boldsymbol{\Delta}'$, $\mathbf{D}'_s = \mathbf{D}'$ $\forall\, s = 1,2,\ldots,p$.

Hence, all $\mathbf{X}_s$ become same, i.e, $\mathbf{X}_s = \mathbf{X} = \left[\boldsymbol{\Delta}' \; \mathbf{1} \; \mathbf{D}'\right]$ and

$$\mathbf{Z} = \oplus\mathbf{X}_s = \oplus\mathbf{X} = \mathbf{I}_p \otimes \mathbf{X} \qquad \text{(A.2.6)}$$

The reduced normal equations for estimating the linear functions of treatment effects are

$$\mathbf{C}^*\boldsymbol{\tau} = \mathbf{Q}^* \qquad \text{(A.2.7)}$$

where

$$\mathbf{C}^* = \boldsymbol{\Sigma}^{-1} \otimes \left(\boldsymbol{\Delta}\boldsymbol{\Delta}' - \boldsymbol{\Delta}\mathbf{D}'\left(\mathbf{D}\mathbf{D}'\right)^{-1}\mathbf{D}\boldsymbol{\Delta}'\right) = \boldsymbol{\Sigma}^{-1} \otimes \mathbf{C}. \qquad \text{(A.2.8)}$$

and

$$\mathbf{Q}^* = \left[\boldsymbol{\Sigma}^{-1} \otimes \left(\boldsymbol{\Delta} - \boldsymbol{\Delta}\mathbf{D}'\left(\mathbf{D}\mathbf{D}'\right)^{-1}\mathbf{D}\right)\right]\mathbf{Y}. \qquad \text{(A.2.9)}$$

Here **C** is the information matrix of the design considered in univariate case. Solving the reduced normal equations (A.2.7) a solution of treatment effects is

$$\hat{\boldsymbol{\tau}} = \mathbf{C}^{*-}\mathbf{Q}^*. \qquad \text{(A.2.10)}$$

Optimality Aspects

In comparative experiments, interest centers around estimation and testing the significance of treatment contrasts. In a given class of designs, a design should be chosen which is "good" according to some well defined statistical criterion. These statistical criteria, in the literature on Design of Experiments are termed as optimality criteria. The optimality criteria are defined as meaningful functions of the elements of the variance-covariance matrix of BLUE of treatment contrasts of interest. Let the problem of inference may be defined as

$$R\text{: } \boldsymbol{\eta} = \mathbf{P}\boldsymbol{\tau} \qquad \text{(A.2.11)}$$

where **P** is a $mp \times vp$ matrix with $\mathbf{PJ} = \mathbf{0}$ and $\boldsymbol{\tau}$ is a $vp \times 1$ vector of treatment effects and **J** is a matrix of unities.

$\mathbf{P\tau}$ in (A.2.11) can be written as

$$\mathbf{P\tau} = \left(\mathbf{I}_p \otimes \mathbf{L}_{m\times v}\right)\begin{pmatrix}\boldsymbol{\tau}_1\\ \boldsymbol{\tau}_2\\ \vdots\\ \boldsymbol{\tau}_p\end{pmatrix}$$

$$= \begin{bmatrix}\mathbf{L}_{m\times v} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{L}_{m\times v} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{L}_{m\times v}\end{bmatrix}\begin{bmatrix}\boldsymbol{\tau}_1\\ \boldsymbol{\tau}_2\\ \vdots\\ \boldsymbol{\tau}_p\end{bmatrix}$$

$$= \begin{bmatrix}\mathbf{L}\boldsymbol{\tau}_1\\ \mathbf{L}\boldsymbol{\tau}_2\\ \vdots\\ \mathbf{L}\boldsymbol{\tau}_p\end{bmatrix}.$$

With reference to *R*, a design *d* is said to be acceptable iff all components of $\boldsymbol{\eta}$ are estimable under *d*. Let $\boldsymbol{D}_R$ (*v*, *b*, *n*; $p \geq 1$) denote the class of all acceptable designs with reference to *R*. We say the problem *R* to be a *non-singular full rank problem* if and only if:

$$\text{Rank } (\mathbf{P}) = pm = p\,(v-1). \qquad \text{(A.2.12)}$$

Consequently, $\boldsymbol{D}_R$ consists of only connected block designs for complete multi-response experiments. We consider only non-singular full rank problems. For any design $d \in \boldsymbol{D}_R$, let $\mathbf{V}_d$ denote the dispersion matrix of estimate of $\hat{\boldsymbol{\eta}}$ using *d*. It is then reasonable to define an optimality criterion as a meaningful function of $\mathbf{V}_d$. The following definitions relate to three important optimality criteria.

Definition A.2.1: *A*-optimality

A design $d^* \in \boldsymbol{D}_R$ is said to be *A*-optimal in $\boldsymbol{D}_R$ if

$$\text{trace}(\mathbf{V}_{d^*}) \leq \text{trace}(\mathbf{V}_d) \text{ for any other design } d \in \boldsymbol{D}_R. \qquad \text{(A.2.13)}$$

The *A*-optimality criterion chooses that design for which the average variance of the estimates of $\boldsymbol{\eta}$ is minimum.

Definition A.2.2: *D*-optimality

A design $d^* \in \boldsymbol{D}_R$ is said to be *D*-optimal in $\boldsymbol{D}_R$ if

$$\det(\mathbf{V}_{d^*}) \leq \det(\mathbf{V}_d) \qquad \text{for any other design } d \in \boldsymbol{D}_R. \qquad \text{(A.2.14)}$$

The *D*-optimality criterion chooses the design for which the generalized variance of the estimated parameter vector $\boldsymbol{\eta}$ is minimum.

Definition A.2.3: *E*-optimality

A design $d^* \in \boldsymbol{D}_R$ is said to be *E*-optimal in $\boldsymbol{D}_R$ if

$$\max(\lambda_{d^*}) \le \max(\lambda_d) \quad \text{for any other design } d \in \boldsymbol{D}_R. \tag{A.2.15}$$

where λ_d and λ_{d^*} are the maximum eigenvalues under the design d and d^* respectively. The E-optimality criterion relates to minimization of the maximum variance of estimates of all orthonormalized treatment contrasts.

Let the experimenter be interested in inferring on a complete set of orthonormalized treatment contrasts for $\boldsymbol{\tau}$ be given by $\mathbf{P}\boldsymbol{\tau}$, where $\mathbf{P} = \mathbf{I}_p \otimes \mathbf{L}_{v-1\times v}$ and $\mathbf{L}$ is such that

$$\mathbf{L}\mathbf{L}' = \mathbf{I}_{v-1};\ \mathbf{L}'\mathbf{L} = \mathbf{I}_v - \frac{1}{v}\mathbf{1}\mathbf{1}',\ \mathbf{P}\mathbf{P}' = \mathbf{I}\otimes\mathbf{L}\mathbf{L}' = \mathbf{I}_p \otimes \mathbf{I}_{v-1} \text{ and}$$

$$\mathbf{P}'\mathbf{P} = \mathbf{I}\otimes\mathbf{L}'\mathbf{L} = \mathbf{I}_p \otimes \left(\mathbf{I}_v - \frac{1}{v}\mathbf{1}\mathbf{1}'\right).$$

Under the complete multi-response setup dispersion matrix $\mathbf{P}\hat{\boldsymbol{\tau}}$ denoted by $\mathbf{V}_d$ is

$$\begin{aligned}
\mathrm{D}(\mathbf{P}\hat{\boldsymbol{\tau}}) = \mathbf{V}_d &= \left(\mathbf{P}\mathbf{C}^*\mathbf{P}'\right)^{-1}\mathbf{P}\mathbf{C}^*\mathbf{C}^{*-}\mathbf{C}^*\mathbf{C}^{*-}\mathbf{C}^*\mathbf{P}'\left(\mathbf{P}\mathbf{C}^*\mathbf{P}'\right)^{-1} \quad [\because D(\mathbf{Q}^*) = \mathbf{C}^*] \\
&= \left[\mathbf{P}\mathbf{C}^*\mathbf{P}'\right]^{-1} \qquad [\because \mathbf{C}^*\mathbf{C}^{*-}\mathbf{C}^* = \mathbf{C}^*] \\
&= \left[\left(\mathbf{I}_p \otimes \mathbf{L}\right)\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{C}\right)\left(\mathbf{I}_p \otimes \mathbf{L}'\right)\right]^{-1} \\
&= \left(\boldsymbol{\Sigma}\otimes(\mathbf{L}\mathbf{C}\mathbf{L}')^{-1}\right).
\end{aligned}$$

$[\because (\mathbf{A}_1\otimes\mathbf{B}_1)(\mathbf{A}_2\otimes\mathbf{B}_2) = \mathbf{A}_1\mathbf{A}_2\otimes\mathbf{B}_1\mathbf{B}_2$ and $(\mathbf{A}\otimes\mathbf{B})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$, if the inverse exist]

The function of $\mathbf{V}_d$ will be in terms of eigenvalues of $\mathbf{L}\mathbf{C}\mathbf{L}'$. It is easier to work with eigenvalues of $\mathbf{C}$ as compared to those of $\mathbf{L}\mathbf{C}\mathbf{L}'$. Therefore, we prove the following result.

Lemma A.2.1: Non-zero eigenvalues of $\mathbf{C}$ and $\mathbf{L}\mathbf{C}\mathbf{L}'$ are same.
Proof: Let $\mathbf{A}$ be a orthogonal matrix given by,

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{v}}1 \\ \mathbf{L} \end{bmatrix}.$$

Then using $\mathbf{A}'\mathbf{A} = \mathbf{I}_v = \mathbf{A}\mathbf{A}'$ we have $\mathbf{L}\mathbf{L}' = \mathbf{I}_{v-1}$.

Let $\mathbf{C}$ be a $v \times v$ symmetric matrix of rank $(v-1)$ such that $\mathbf{C1} = \mathbf{0}$; then

$$\det(\mathbf{C} - \lambda\mathbf{I}_v) = \lambda f(\lambda) = \det\left[\mathbf{A}(\mathbf{C} - \lambda\mathbf{I}_v)\mathbf{A}'\right] = \det(\mathbf{A}\mathbf{C}\mathbf{A}' - \lambda\mathbf{I}_v).$$

$$\mathbf{ACA}' = \begin{bmatrix} \frac{1}{\sqrt{v}}1 \\ \mathbf{L} \end{bmatrix} \mathbf{C} \begin{bmatrix} \frac{1}{\sqrt{v}}1' & \mathbf{L}' \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{v}\mathbf{1C1}' & \frac{1}{\sqrt{v}}\mathbf{1'CL}' \\ \frac{1}{\sqrt{v}}\mathbf{LC1} & \mathbf{LCL}' \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{LCL}' \end{bmatrix}$$

Thus $\mathbf{ACA}' - \lambda \mathbf{I}_v = \begin{bmatrix} -\lambda \mathbf{I}_v & \mathbf{0}' \\ \mathbf{0}' & \mathbf{LCL}' - \lambda \mathbf{I}_{v-1} \end{bmatrix}$.

Now $\det(\mathbf{ACA}' - \lambda \mathbf{I}_v) = -\lambda \det(\mathbf{LCL}' - \lambda \mathbf{I}_{v-1})$.

Thus non-zero eigenvalues of $\mathbf{C}$ and $\mathbf{LCL}'$ are same. Let the non-zero eigenvalues of $\mathbf{C} = \mathbf{R} - \mathbf{NK}^{-1}\mathbf{N}'$ be $\theta_1, \theta_2, \ldots, \theta_{v-1}$ such that, $0 \le \theta_{(1)} \le \theta_{(2)} \le \ldots \le \theta_{(v-1)}$, where $\theta_{(1)}$ is the minimum non-zero eigenvalue of $\mathbf{C}$. Similarly let the eigenvalues of $\mathbf{\Sigma}$ be $\mu_1, \mu_2, \ldots, \mu_p$ which can be arranged in ascending order as $\mu_{(1)} \le \mu_{(2)} \le \ldots \le \mu_{(p)}$, where $\mu_{(p)}$ is the maximum eigenvalue of $\mathbf{\Sigma}$. Also let $\mathbf{\Sigma}$ be positive definite known and fixed, then $\mu_{(1)} \le \mu_{(2)} \le \ldots \le \mu_{(p)}$ are fixed. Using the fact that the eigenvalues of $\mathbf{A} \otimes \mathbf{B}$ are product of eigenvalues of $\mathbf{A}$ with those of $\mathbf{B}$, the eigenvalues of $\mathrm{D}(\mathbf{P}\hat{\boldsymbol{\tau}}) = \mathbf{V}_d = \left(\mathbf{\Sigma} \otimes (\mathbf{LCL}')^{-1}\right)$ are $\Phi_{is} = \mu_s / \theta_i$ (Using Lemma 4.2.1) $s = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, v-1$.

If $\mathbf{\Sigma}$ is fixed, then the eigenvalues of $\mathrm{D}(\mathbf{P}\hat{\boldsymbol{\tau}}) = \left(\mathbf{\Sigma} \otimes (\mathbf{LCL}')^{-1}\right)$ depend only on the eigenvalues of $\mathbf{C}$. Therefore, if the design for single response experiments is optimal then the design for multi-response experiments would also be optimal. To be specific, results obtained on *A*-, *D*- and *E*-optimality of block designs for complete multi-response experiments are given in the sequel.

According to Definition A.2.1, for obtaining an ***A*-optimal design**, for inferring on complete set of orthonormalized treatment contrasts we have to minimize the trace of $\mathbf{V}_d$, *i.e.* the sum of eigenvalues of dispersion matrix $\mathbf{V}_d$. In other words, we have to minimize

$$\text{trace } (\mathbf{V}_d) = \sum_{s=1}^{p}\sum_{i=1}^{v-1} \Phi_{is} = \sum_{s=1}^{p}\sum_{i=1}^{v-1} \mu_s / \theta_i = \sum_{s=1}^{p} \mu_s \sum_{i=1}^{v-1} \theta_i^{-1}.$$

For fixed $\mu_1, \mu_2, \ldots, \mu_p$, minimum of trace ($\mathbf{V}_d$) is achieved when $\sum_i \theta_i^{-1}$ is minimum. Therefore, if a design is *A*-optimal for single response experiments, it is also *A*-optimal for complete multi-response experiments.

According to Definition A.2.2, for obtaining a ***D*-optimal design**, for inferring on complete set of orthonormalized treatment contrasts we have to minimize the determinant of $\mathrm{D}(\mathbf{P}\hat{\boldsymbol{\tau}})$ = $\mathbf{V}_d$. Det($\mathbf{V}_d$) is the product of eigenvalues of $\mathbf{V}_d$. Therefore, we have to minimize

$$\text{Det}\,(\mathbf{V}_d) = \prod_{s=1}^{p}\prod_{i=1}^{v-1} \Phi_{is} = \prod_{s=1}^{p}\prod_{i=1}^{v-1} \mu_s / \theta_i = \left(\prod_{s=1}^{p} \mu_s\right)^{v-1} \left(\prod_{i=1}^{v-1} \theta_i\right)^{-p} .$$

For fixed $\mu_1, \mu_2, \ldots, \mu_p$, minimum of det($\mathbf{V}_d$) is achieved when $\left(\prod_{i=1}^{v-1} \theta_i\right)^{-p}$ is minimum. Therefore, if a design is *D*-optimal for single response experiments, it is also *D*-optimal for multi-response experiments.

According to Definition A.2.3, for obtaining a ***E*-optimal design**, for inferring on complete set of orthonormalized treatment contrasts we have to minimize the maximum eigenvalue of $\mathbf{V}_d$. Maximum eigenvalue of $\mathbf{V}_d$ is $\mu_{(p)} / \theta_{(1)}$. Therefore, to obtain the minimum of maximum eigenvalue of $\mathbf{V}_d$, one has to minimize $\mu_{(p)} / \theta_{(1)}$. For fixed $\mu_{(p)}$, minimum of maximum eigenvalues is obtained when $\theta_{(1)}^{-1}$ is minimum. Therefore, if a design is *E*-optimal for single response experiments, it is also *E*-optimal for multi-response experiments.

A.3. Designs for Incomplete Multi-response Experiments

Incomplete multi-response designs available in literature require huge amount of resources in terms of cost, labour and time. Hence, it is required to obtain incomplete multi-response designs which require fewer resources. In this section efforts have been made to obtain block designs for incomplete multi-response experiments that save on resources in terms of cost of labour and time of collecting data. The method of construction along with the analysis procedure for the proposed class of designs is explained in sequel.

Method A.3.1: Consider a randomized complete block (RCB) design with *v* treatments each treatment replicated *r* times and a BIB design with parameters $v^* = p, b^* = r, r^*, k^*, \lambda^* \geq 2$.

Now take BIB design as *response-wise* design (D_1) and RCB design as *treatment-wise* design (D_2). Let us also consider that block contents of BIB design corresponds to the response variables to be observed from each replication of RCB design. Then we define incomplete multi-response design *D* in *v* treatments and *p* response variables as $D = (D_1, D_2)$.

Example A.3.1: Consider an experimental situation where the experimenter is interested to study the effect of nutrient management on growth parameters and yield attributes of a crop. Let the experimenter has 8 different treatments to be tested on $n = 32$ experimental units. Four response variables to be observed are: (1) plant height 30 days after sowing, (2) dry matter at 30 days after sowing, (3) seed per siliqua and (4) harvest index. Due to constraint on availability of skilled manpower, the data on 3 of these response variables can only be collected from each experimental unit. Therefore, experiment has to be conducted using an incomplete multi-response design. The design for this experimental situation may be obtained using Method A.3.1 as follows:

Consider a RCB design with $v = 8$, $r = 4$. The block contents of this design after randomization are

Block1: (1, 2, 3, 4, 5, 6, 7, 8); Block2: (2, 5, 7, 1, 3, 8, 4, 6)
Block3: (3, 4, 8, 7, 6, 5, 2, 1); Block4: (4, 8, 5, 2, 3, 6, 7, 1).

A BIB design with parameters ($v = 4$, $b = 4$, $r = 3$, $k = 3$, $\lambda = 2$) exists with block contents

(1, 2, 3); (1, 2, 4); (1, 3, 4); (2, 3, 4).

Using the procedure of Method 4.3.1, the incomplete multi-response design obtained is

Replication	Responses observed	Treatment-wise design							
Rep-1	(1, 2, 3)	1	2	3	4	5	6	7	8
Rep-2	(1, 2, 4)	2	5	7	1	3	8	4	6
Rep-3	(1, 3, 4)	3	4	8	7	6	5	2	1
Rep-4	(2, 3, 4)	4	8	5	2	3	6	7	1

Hence, instead of taking 124 observations, the experiment can be carried out by taking only 96 observations from 32 experimental units.

Remark A.3.1: The designs for more number of responses than those in Method A.3.1 can be obtained by taking response-wise design as singular group divisible (GD) design, instead of a BIB design. We know, that a singular GD design in $v^* = nv$ treatments can be obtained from a BIB design in v treatments by replacing each of the treatments of BIB design with a group of n new treatments, *i.e.* by replacing the i^{th} treatment with i, $v + i$, …, $(n - 1)v + i$ new treatments, $\forall\ i = 1, 2, \ldots, v$. Now take this singular GD as response-wise design (D_1) and following the procedure of Method A.3.1, we get an IM design $D = (D_1, D_2)$, where D_2 is the treatment-wise design.

Example A.3.2: Consider the experimental situation explained in Example A.3.1 where data on 4 response variables are collected. If the experimenter is interested in collecting 8 responses then the following procedure may be used to get an incomplete multi-response design.

A BIB design with parameters ($v = 4$, $b = 4$, $r = 3$, $k = 3$, $\lambda = 2$) exists with block contents

(1, 2, 3); (1, 2, 4); (1, 3, 4); (2, 3, 4).

The response-wise design in 8 symbols, is obtained by replacing response variable 1 by response variables (1, 5), response variable 2 by response variables (2, 6), response

variables 3 by response variables (3, 7) and response variable 4 by response variables (4, 8).

So the incomplete multi-response design can be shown as follows:

Replication	Responses observed	Treatment-wise design							
Rep-1	(1, 5, 2, 6, 3, 7)	1	2	3	4	5	6	7	8
Rep-2	(1, 5, 2, 6, 4, 8)	2	5	7	1	3	8	4	6
Rep-3	(1, 5, 3, 6, 4, 8)	3	4	8	7	6	5	2	1
Rep-4	(2, 6, 3, 6, 4, 8)	4	8	5	2	3	6	7	1

A.4. Analysis Procedure for Incomplete Multi-response Designs

In this section, we have developed the procedure of analysis of data generated through designs for incomplete multi-response experiments obtainable through Method A.3.1 and Remark A.3.1. Treatment effects matrix is given by

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{11} & \tau_{12} & \cdots & \tau_{1p} \\ \tau_{21} & \tau_{22} & \cdots & \tau_{2p} \\ \vdots & \vdots & & \vdots \\ \tau_{v1} & \tau_{v2} & \cdots & \tau_{vp} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\tau}_1 & \boldsymbol{\tau}_2 & \cdots & \boldsymbol{\tau}_p \end{bmatrix}. \qquad \text{(A.4.1)}$$

In each replication of treatment-wise design, we have a block of response-wise design as discussed in Method A.3.1. So each set is a minimally connected design.

Now, let $\mathbf{T}_{ls}$, $\mathbf{B}_{ls}$ and $\mathbf{Q}_{ls}$ be column vectors of treatment totals, block totals and adjusted treatment totals for l^{th} set and s^{th} response. The intra-block reduced normal equations for the s^{th} response for design in l^{th} set may then be written as,

$$\mathrm{E}(\mathbf{Q}_{ls}) = \mathbf{C}_l \boldsymbol{\tau}_s$$

$$\mathrm{D}(\mathbf{Q}_{ls}) = \mathbf{C}_l \sigma_{ss}$$

$$\mathrm{cov}(\mathbf{Q}_{ls}, \mathbf{Q}_{ls'}) = \mathbf{C}_l \sigma_{ss'} \qquad s \neq s' = 1, 2, \ldots, p_l$$

where

$$\mathbf{C}_l = \mathbf{R}_l^{-1} - \mathbf{N}_l \mathbf{K}_l^{-1} \mathbf{N}_l', \qquad l = 1, 2, \ldots, u$$

$$\mathbf{Q}_{ls} = \mathbf{T}_{ls} - \mathbf{N}_l \mathbf{K}_l^{-1} \mathbf{B}_{ls} \qquad l = 1, 2, \ldots, u;\ s = 1, 2, \ldots, p_l$$

$\mathbf{N}_l$: $v \times b$ treatments vs blocks incidence matrix for l^{th} set.

$\mathbf{K}_l$: diagonal matrix of block sizes for l^{th} set.

$\mathbf{R}_l$: diagonal matrix of replications for l^{th} set.

σ_{ss} is the variance of s^{th} response variable to be estimated from the sets where s^{th} response variable is present and $\sigma_{ss'}$ is the covariance between s^{th} and s'^{th} response variables to be estimated from the sets in which both s^{th} and s'^{th} response variables are recorded.

For the designs obtained in Method A.3.1

$$\mathbf{C}_1 = \mathbf{C}_2 = \cdots = \mathbf{C}_u = \mathbf{I}_v - \frac{1}{v}\mathbf{1}\mathbf{1}' = \mathbf{C}(\text{say}).$$

Then we have reduced normal equations

$$\begin{aligned} \mathrm{E}(\mathbf{Q}_l) &= \mathbf{C}_l \boldsymbol{\tau}^{(l)} \qquad l = 1, 2, \ldots, u \\ &= \mathbf{C}_l \boldsymbol{\tau} \mathbf{M}^{(l)} \end{aligned} \tag{A.4.3}$$

where $\mathbf{M}^{(l)}$ is a $p \times p_l$ matrix with diagonal elements unity if the response variable is observed from the l^{th} set and obtained from an identity matrix of order p by deleting the column corresponding to the response variable absent in the l^{th} set.

Here we can decompose $\mathbf{C}_l$ as follows

$$\mathbf{C}_l = \alpha_{l1}\mathbf{F}_1 + \alpha_{l2}\mathbf{F}_2 \tag{A.4.4}$$

where $\alpha_{l1} = 1 - \frac{1}{v}$, $\alpha_{l2} = -\frac{1}{v}$

$$\mathbf{F}_1 = \mathbf{I}, \ \mathbf{F}_2 = \mathbf{J} - \mathbf{I} \text{ and } \mathbf{F}_1 + \mathbf{F}_2 = \mathbf{J}. \tag{A.4.5}$$

So analysis of the above incomplete multi-response experiment is possible.

Now combining from all the sets using (A.4.2) and (A.4.3) we get

$$\begin{aligned} \mathrm{E}[\mathbf{Q}_1 \quad \mathbf{Q}_2 \quad \cdots \quad \mathbf{Q}_u] &= \left[\mathbf{C}_1\boldsymbol{\tau}\mathbf{M}^{(1)} \quad \mathbf{C}_2\boldsymbol{\tau}\mathbf{M}^{(2)} \quad \cdots \quad \mathbf{C}_u\boldsymbol{\tau}\mathbf{M}^{(u)}\right] \\ &= \mathbf{C}\boldsymbol{\tau}\left[\mathbf{M}^{(1)} \quad \mathbf{M}^{(2)} \quad \cdots \quad \mathbf{M}^{(u)}\right] \\ &= \mathbf{C}\boldsymbol{\tau}\,\mathbf{L} \end{aligned} \tag{A.4.6}$$

$$\mathrm{E}(\mathbf{Q}) = \mathbf{C}\boldsymbol{\tau}\,\mathbf{L} \tag{A.4.7}$$

$$\mathbf{L} = \left[\mathbf{M}^{(1)} \quad \mathbf{M}^{(2)} \quad \cdots \quad \mathbf{M}^{(u)}\right] \tag{A.4.8}$$

$$\mathbf{L}\mathbf{L}' = \sum_l \mathbf{M}^{(l)}\mathbf{M}'^{(l)} = \mathrm{diag}\left(\sum_l \delta_{ls}\right) \tag{A.4.9}$$

$\delta_{ls} = 1$, if s^{th} ($l = 1, 2, \ldots, p$) response is present in the l^{th} ($l = 1, 2, \ldots, u$) set.

It can be seen that

$$\mathbf{L}\mathbf{L}' = r^*\mathbf{I}_v \tag{A.4.10}$$

where r^* is the number of replication of response-wise design.

So, $\mathbf{L}\mathbf{L}'$ is non-singular.

$$\mathrm{E}\left(\mathbf{Q}\mathbf{L}'(\mathbf{L}\mathbf{L}')^{-1}\right) = \mathbf{C}\boldsymbol{\tau} \tag{A.4.11}$$

Q_{isl} = adjusted treatment total for i^{th} treatment in the l^{th} set for the s^{th} response.

Incorporating the result of (A.4.10) into (A.4.11) we have

$$\left(\mathbf{QL}'(\mathbf{LL}')^{-1}\right) = \frac{1}{r^*}\mathbf{QL}' = \frac{1}{r^*}\mathbf{Z} \qquad [\mathbf{LL}' = r^*\mathbf{I}_v] \tag{A.4.12}$$

$$\mathrm{E}\left(\frac{1}{r^*}\mathbf{Z}\right) = \mathbf{C}\boldsymbol{\tau}.$$

$$\therefore \hat{\boldsymbol{\tau}} = \frac{1}{r^*}\mathbf{C}^{-}\mathbf{Z}$$

where

$$\hat{\boldsymbol{\tau}}_s = \frac{1}{r^*}\mathbf{C}^{-}\sum_{l=1}^{u}\mathbf{Q}_{ls}\delta_{ls}$$

$\delta_{ls} = 1$, if s^{th} ($s = 1, 2, \ldots, p$) response is present in the l^{th} ($l = 1, 2, \ldots, u$) set.

$$\hat{\boldsymbol{\tau}}_1 = \frac{1}{r^*}\mathbf{C}^{-}\begin{bmatrix} Q_{11s} + \ldots + Q_{1us} \\ Q_{21s} + \ldots + Q_{2us} \\ \ldots \\ Q_{v1s} + \ldots + Q_{vup} \end{bmatrix}. \tag{A.4.13}$$

Estimate of σ_{ss} and $\sigma_{ss'}$ can be obtained by using the following expression

$$\hat{\sigma}_{ss} = \sum_{l=1}^{u}\left(\mathbf{Y}_{ls}'\mathbf{Y}_{ls} - \mathbf{Q}_{ls}'\mathbf{C}^{-}\mathbf{Q}_{ls}\right)\delta_{ls}/(n-v)$$

$$\hat{\sigma}_{ss'} = \sum_{l=1}^{u}\left(\mathbf{Y}_{ls}'\mathbf{Y}_{ls'} - \mathbf{Q}_{ls}'\mathbf{C}^{-}\mathbf{Q}_{ls'}\right)\delta_{ls}\delta_{ls'}/(n-v)$$

where $\delta_{ls} = 1$, if s^{th} ($s = 1, 2, \ldots, p$) response is present in the l^{th} ($l = 1, 2, \ldots, u$) set.

$\mathbf{Y}_{ls}$ is the observation vector of l^{th} set for the s^{th} response variable.

A.4.1. Multivariate Treatment Contrast Analysis

For this purpose, let the experimenter is interested in testing the significance of treatment contrast of the form $\mathbf{P}'\boldsymbol{\tau}$, where $\mathbf{P}'_{p\times vp} = \left(\mathbf{I}_p \otimes \mathbf{p}'_{1\times v}\right)$. We write the treatment effects matrix in the form of a column vector by arranging treatment effects vector for each response variable one below the other. So we have

$$\hat{\boldsymbol{\tau}}_{vp\times 1} = \begin{bmatrix} \hat{\boldsymbol{\tau}}_1 \\ \hat{\boldsymbol{\tau}}_2 \\ \vdots \\ \hat{\boldsymbol{\tau}}_p \end{bmatrix} \tag{A.4.14}$$

$\mathbf{P}'\hat{\boldsymbol{\tau}}$ is the BLUE of $\mathbf{P}'\boldsymbol{\tau}$. Now we can obtain the dispersion matrix for the estimated treatment contrasts $\mathbf{P}'\hat{\boldsymbol{\tau}}$ as

$$\mathrm{D}(\mathbf{P}'\hat{\boldsymbol{\tau}}) = \begin{bmatrix} \mathrm{D}(\mathbf{p}_1'\hat{\boldsymbol{\tau}}_1) & \mathrm{D}(\mathbf{p}_1'\hat{\boldsymbol{\tau}}_1, \mathbf{p}_2'\hat{\boldsymbol{\tau}}_2) & \cdots & \mathrm{D}(\mathbf{p}_1'\hat{\boldsymbol{\tau}}_1, \mathbf{p}_p'\hat{\boldsymbol{\tau}}_p) \\ \mathrm{D}(\mathbf{p}_2'\hat{\boldsymbol{\tau}}_2, \mathbf{p}_1'\hat{\boldsymbol{\tau}}_1) & \mathrm{D}(\mathbf{p}_2'\hat{\boldsymbol{\tau}}_2) & \cdots & \mathrm{D}(\mathbf{p}_2'\hat{\boldsymbol{\tau}}_2, \mathbf{p}_p'\hat{\boldsymbol{\tau}}_p) \\ \vdots & \vdots & & \vdots \\ \mathrm{D}(\mathbf{p}_p'\hat{\boldsymbol{\tau}}_p, \mathbf{p}_1'\hat{\boldsymbol{\tau}}_1) & \mathrm{D}(\mathbf{p}_p'\hat{\boldsymbol{\tau}}_p, \mathbf{p}_2'\hat{\boldsymbol{\tau}}_2) & \cdots & \mathrm{D}(\mathbf{p}_p'\hat{\boldsymbol{\tau}}_p) \end{bmatrix} \quad \text{(A.4.15)}$$

$$\begin{aligned} \mathrm{D}(\mathbf{p}_s'\hat{\boldsymbol{\tau}}_s) &= \mathbf{p}_s'\mathrm{D}(\hat{\boldsymbol{\tau}}_s)\mathbf{p}_s \\ &= \mathbf{p}_s'\mathrm{D}\left(\frac{1}{r^*}\mathbf{C}^-\left(\sum_{l=1}^{u}\mathbf{Q}_{ls}\delta_{ls}\right)\right)\mathbf{p}_s \\ &= \mathbf{p}_s'\left(\frac{1}{r^{*2}}\mathbf{C}^-\left(\sum_{l=1}^{u}\mathbf{D}(\mathbf{Q}_{ls})\delta_{ls}^2\right)\mathbf{C}^-\right)\mathbf{p}_s \\ &= \frac{\sigma_{ss}}{r^{*2}}\mathbf{p}_s'\left(\mathbf{C}^-\left(\sum_{l=1}^{u}\mathbf{C}^-\delta_{ls}^2\right)\mathbf{C}^-\right)\mathbf{p}_s \\ &= \frac{\sigma_{ss}}{r^{*2}}\mathbf{p}_s'\mathbf{C}^-\mathbf{p}_s\sum_{l=1}^{u}\delta_{ls}^2 \\ &= \frac{\sigma_{ss}}{r^{*2}}\mathbf{p}_s'\mathbf{C}^-\mathbf{p}_s r^* \\ &= \frac{1}{r^*}\mathbf{p}_s'\mathbf{C}^-\mathbf{p}_s\sigma_{ss} \end{aligned} \quad \text{(A.4.16)}$$

$$\begin{aligned} \mathrm{D}(\mathbf{p}_s'\hat{\boldsymbol{\tau}}_s, \mathbf{p}_{s'}'\hat{\boldsymbol{\tau}}_{s'}) &= \mathbf{p}_s'\mathrm{D}(\hat{\boldsymbol{\tau}}_s, \hat{\boldsymbol{\tau}}_{s'})\mathbf{p}_{s'} \\ &= \mathbf{p}_s'\left(\frac{1}{r^{*4}}\mathbf{C}^-\left(\sum_{l=1}^{u}\mathbf{D}(\mathbf{Q}_{ls}, \mathbf{Q}_{ls'})\delta_{ls}\delta_{ls'}\right)\mathbf{C}^-\right)\mathbf{p}_{s'} \\ &= \left(\frac{1}{r^{*4}}\mathbf{p}_s'\mathbf{C}^-\mathbf{p}_{s'}\left(\sigma_{ss'}\sum_{l=1}^{u}\delta_{ls}\delta_{ls'}\right)\right) \\ &= \frac{\lambda^*}{r^{*4}}\mathbf{p}_s'\mathbf{C}^-\mathbf{p}_{s'}\sigma_{ss'} \end{aligned} \quad \text{(A.4.17)}$$

As mentioned earlier, the estimate of $\sigma_{ss'}$ can be obtained from those sets where data on s^{th} and s'^{th} response variables are collected.

For testing the null hypothesis H_0: $\mathbf{P}'\boldsymbol{\tau} = 0$, against H_1: $\mathbf{P}'\boldsymbol{\tau} \neq 0$, let us take the estimate of the treatment contrasts as $\mathbf{P}'\hat{\boldsymbol{\tau}}$. Now the test statistic can be given by

$$TC = \left(\mathbf{P}'\hat{\boldsymbol{\tau}}\right)'\left[\mathrm{D}(\mathbf{P}'\hat{\boldsymbol{\tau}})\right]^{-1}\left(\mathbf{P}'\hat{\boldsymbol{\tau}}\right) \quad \text{(A.4.18)}$$

where D(.) is dispersion matrix of estimated treatment contrasts.

Test statistic given in (A.4.18) is asymptotically χ^2- distribution with one degree of freedom. Therefore, reject the null hypothesis at α% level of significance if $TC > \chi^2_{1-\alpha,1}$ and conclude that the treatment effects are significantly different. The above procedure is illustrated with the help of following example.

Example A.4.1: Consider an experiment on identification of the best seed treating insecticide of soybean based on the following response variables: Pods/plant (P1), Seeds/pod (P2), Stem fly infestation (%) (P3), 100 Seed weight (P4). Following are the 12 treatments tested in this experiment:

T1: Pusa 16 + Thiamethoxam 4g/kg	T7: Pusa 16 + Imidacloprid 5ml/kg
T2: Pusa 24 + Thiamethoxam 4g/kg	T8: Pusa 24 + Imidacloprid 5ml/kg
T3: JS 335 + Thiamethoxam 4g/ kg	T9: JS 335 + Imidacloprid 5ml/kg
T4: Pusa 16 + Thiamethoxam 8g/kg	T10: Pusa 16 + Imidacloprid 10ml/ kg
T5: Pusa 24 + Thiamethoxam 8g/kg	T11: Pusa 24 + Imidacloprid 10ml/ kg
T6: JS 335 + Thiamethoxam 8g/kg	T12: JS 335 + Imidacloprid 10ml/ kg

The experiment was conducted using RCB design with 12 treatments and each treatment was replicated 4 times. Due to labour constraints, data on a subset of the response variables are collected from each replication of the treatment-wise design. The response-wise design is BIB design with block contents:

Block1: (1, 2, 3); Block2: (1, 2, 4); Block3: (1, 3, 4); Block4: (2, 3, 4);

Replication of treatment-wise design	Responses observed
Rep-1	P1, P2, P3
Rep-2	P1, P2, P4
Rep-3	P1, P3, P4
Rep-4	P2, P3, P4

Replication	treatment-wise design	Responses observed
Rep-1	(1,2,3,4,5,6,7, 8,9,10,11,12)	P1: (49.2,47.4, 52.2, 49.8, 45.6, 70.6, 43.6, 72.2, 66.4, 52.4, 60.6, 96.9) P2: (2.1, 2.8, 2.5, 2.0, 2.6, 2.6, 2.0, 2.0, 2.0, 2.0, 2.0, 2.3) P3: (22.98, 35.43, 28.15, 25.43, 37.72, 31.59, 31.13, 39.26, 29.70, 19.75, 13.70, 12.97)
Rep-2	(1,2,3,4,5,6,7, 8,9,10,11,12)	P1:(53.0, 38.8, 57.8, 45.3, 40.2, 77.8, 46.7, 36.6, 61.6, 54.4, 51.0, 72.4) P2:(2.3, 2.3, 2.4, 2.0, 2.5, 2.6, 2.4, 2.5, 2.3, 2.0, 2.0, 2.4) P4: (8.59, 9.35, 11.08, 9.37, 10.70, 9.44, 8.21, 10.17, 9.18, 8.51, 9.77, 10.62)
Rep-3	(1,2,3,4,5,6,7, 8,9,10,11,12)	P1:(50.4, 44.2, 54.5, 51.1, 44.5, 69.7, 44.8, 68.7, 64.0, 57.2, 62.7, 88.8) P3: (40.6, 43.07, 41.03, 37.84, 35.26, 41.00, 48.66, 41.37, 39.90, 67.24, 20.62, 34.26) P4: (9.48, 10.04, 10.20, 9.22, 10.56, 10.51, 9.70, 10.49, 12.30, 9.20, 10.25, 0.00)
Rep-4	(1,2,3,4,5,6,7, 8,9,10,11,12)	P2: (2.4, 2.1, 2.0, 2.3, 2.2, 2.1, 2.0, 2.2, 2.1, 2.1, 2.1, 2.4) P3: (45.60, 49.07, 42.03, 33.84, 39.26, 45.00, 41.66, 49.37, 35.90, 63.24, 28.62, 32.26) P4: (9.12, 11.04, 13.20, 10.22, 11.56, 11.51, 10.70, 11.49, 11.30, 8.90, 10.40, 9.00)

In each replication of treatment-wise design we have a set of response wise design. So each set is a connected design. For the l^{th} set we can obtain information matrix of the form

$$\mathbf{C}_i = \mathbf{I}_{12} - \frac{1}{12}\mathbf{1}\mathbf{1}'$$

Here, $\mathbf{M}^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{M}^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{M}^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{M}^{(4)} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Treatment totals for the 4 sets are

T1			T2			T3			T4		
49.2	2.1	22.98	53	2.3	8.59	50.4	40.60	9.48	2.4	45.60	9.12
47.4	2.8	35.43	38.8	2.3	9.35	44.2	43.07	10.04	2.1	49.07	11.04
52.2	2.5	28.15	57.8	2.4	11.08	54.5	41.03	10.20	2.0	42.03	13.20
49.8	2.0	25.43	45.3	2.0	9.37	51.1	37.84	9.22	2.3	33.84	10.22
45.6	2.6	37.72	40.2	2.5	10.70	44.5	35.26	10.56	2.2	39.26	11.56
70.6	2.6	31.59	77.8	2.6	9.44	69.7	41.00	10.51	2.1	45.00	11.51
43.6	2.0	31.13	46.7	2.4	8.21	44.8	48.66	9.70	2.0	41.66	10.70
72.2	2.0	39.26	36.6	2.5	10.17	68.7	41.37	10.49	2.2	49.37	11.49
66.4	2.0	29.70	61.6	2.3	9.18	64.0	39.90	12.30	2.1	35.90	11.30
52.4	2.0	19.75	54.4	2.0	8.51	57.2	67.24	9.20	2.1	63.24	8.90
60.6	2.0	13.70	51.0	2.0	9.77	62.7	20.62	10.25	2.1	28.62	10.40
96.6	2.3	12.97	72.4	2.4	10.62	88.8	34.26	12.00	2.4	32.26	9.00

Combining from all the set adjusted treatment total is

Q1			Q2			Q3			Q4		
-9.68	-0.14	-4.34	0.03	-0.01	-0.99	-7.98	-0.30	0.15	0.23	3.45	-1.58
-11.48	0.56	8.11	-14.17	-0.01	-0.23	-14.18	2.17	0.71	-0.07	6.92	0.34
-6.68	0.26	0.83	4.83	0.09	1.50	-3.88	0.13	0.87	-0.17	-0.12	2.50
-9.08	-0.24	-1.89	-7.67	-0.31	-0.21	-7.28	-3.06	-0.11	0.13	-8.31	-0.48
-13.28	0.36	10.40	-12.77	0.19	1.12	-13.88	-5.64	1.23	0.03	-2.89	0.86
11.72	0.36	4.27	24.83	0.29	-0.14	11.32	0.10	1.18	-0.07	2.85	0.81
-15.28	-0.24	3.81	-6.27	0.09	-1.37	-13.58	7.76	0.37	-0.17	-0.49	0.00
13.32	-0.24	11.94	-16.37	0.19	0.59	10.32	0.47	1.16	0.03	7.22	0.79
7.52	-0.24	2.38	8.63	-0.01	-0.40	5.62	-1.00	2.97	-0.07	-6.25	0.60
-6.48	-0.24	-7.57	1.43	-0.31	-1.07	-1.18	26.34	-0.13	-0.07	21.09	-1.80
1.72	-0.24	-13.62	-1.97	-0.31	0.19	4.32	-20.28	0.92	-0.07	-13.53	-0.30
37.72	0.06	-14.35	19.43	0.09	1.04	30.42	-6.64	-9.33	0.23	-9.89	-1.70

The estimates of variance of the response variables are obtained by taking replications of treatment wise design where the response variable is present and for covariance those replications are considered where two response variables occur together. The estimates of variance and covariance are as follows.

$\hat{\sigma}_{11}$ = 1058.03, $\hat{\sigma}_{12}$ = – 5.32, $\hat{\sigma}_{13}$ = 91.90, $\hat{\sigma}_{14}$ = –70.52, $\hat{\sigma}_{22}$ = 1.66, $\hat{\sigma}_{23}$ =4.6325,

$\hat{\sigma}_{24}$ = –0.9897, $\hat{\sigma}_{33}$ = 1094.68, $\hat{\sigma}_{34}$ = –10.006, $\hat{\sigma}_{44}$ = 72.16.

Now let the experimenter is interested in testing the equality of effects of treatment T1 and treatment T6 and treatment contrast is $\mathbf{p}'$ = (1 0 0 0 0 –1 0 0 0 0 0). To test the significance of the treatment contrast $\mathbf{P}'\boldsymbol{\tau}$ the null hypothesis is

H_0: $\mathbf{P}'\boldsymbol{\tau} = 0$ against H_1: $\mathbf{P}'\boldsymbol{\tau} \neq 0$,

where $\mathbf{P}' = \mathbf{I}_4 \otimes \mathbf{p}'$ and $\hat{\boldsymbol{\tau}}$ is the vector of estimates of the treatment effects arranged one below another.

Estimate of treatment contrast $\mathbf{P}'\boldsymbol{\tau}$ is obtained by computing the estimates of the treatment contrast separately for all the response variables and arranging one below the another and is

$$\mathbf{P}'\hat{\boldsymbol{\tau}} = \begin{bmatrix} -65.50 \\ -0.50 \\ -8.41 \\ -4.27 \end{bmatrix}.$$

Dispersion matrix of treatment contrast is

$$\mathrm{D}(\mathbf{P}'\hat{\boldsymbol{\tau}}) = \begin{bmatrix} 235.12 & -0.13 & 2.27 & -1.74 \\ -0.13 & 0.21 & 0.11 & -0.02 \\ 2.27 & 0.11 & 243.26 & -0.25 \\ -1.74 & -0.02 & -0.25 & 16.04 \end{bmatrix}$$

Now using the test statistic for testing the null hypothesis explained above we get

$TC = (\mathbf{P}'\hat{\boldsymbol{\tau}})'[\mathrm{D}(\mathbf{P}'\hat{\boldsymbol{\tau}})]^{-1}(\mathbf{P}'\hat{\boldsymbol{\tau}}) =$ **21.31**

The calculated value of the χ^2 - statistic is **21.31** which is more than the tabulated value of χ^2 at 1 degree of freedom (3.84). So it can be concluded that the treatment T1 (Pusa 16 + Thiamethoxam 4g/kg) is significantly different from treatment T6 (JS 335 + Thiamethoxam 8g/kg). Similarly, one can test the significance of any other treatment contrast of interest.

AppendiX

```
ods listing;
ods html file="C:\IMD_4Resp.xls" style=none;
/*one can give the location of the output file required*/

proc iml;
/*enter data for each of the sets, trt, block, response variables on which data is collected in that
particular set*/
s1 =
{1      1      49.2    2.1     22.98,
2      1      47.4    2.8     35.43,
3      1      52.2    2.5     28.15,
4      1      49.8    2       25.43,
```

```
5     1     45.6   2.6    37.72,
6     1     70.6   2.6    31.59,
7     1     43.6   2      31.13,
8     1     72.2   2      39.26,
9     1     66.4   2      29.7,
10    1     52.4   2      19.75,
11    1     60.6   2      13.7,
12    1     96.6   2.3    12.97};

s2 =
{1    2     53     2.3        8.59,
2     2     38.8   2.3        9.35,
3     2     57.8   2.4        11.08,
4     2     45.3   2          9.37,
5     2     40.2   2.5        10.7,
6     2     77.8   2.6        9.44,
7     2     46.7   2.4        8.21,
8     2     36.6   2.5        10.17,
9     2     61.6   2.3        9.18,
10    2     54.4   2          8.51,
11    2     51     2          9.77,
12    2     72.4   2.4        10.62};

s3=
{1    3     50.4   40.6   9.48,
2     3     44.2   43.07  10.04,
3     3     54.5   41.03  10.2,
4     3     51.1   37.84  9.22,
5     3     44.5   35.26  10.56,
6     3     69.7   41     10.51,
7     3     44.8   48.66  9.7,
8     3     68.7   41.37  10.49,
9     3     64     39.9   12.3,
10    3     57.2   67.24  9.2,
11    3     62.7   20.62  10.25,
12    3     88.8   34.26  0};

s4=
{1    4          2.4    45.6   9.12,
2     4          2.1    49.07  11.04,
3     4          2      42.03  13.2,
4     4          2.3    33.84  10.22,
5     4          2.2    39.26  11.56,
6     4          2.1    45     11.51,
7     4          2      41.66  10.7,
8     4          2.2    49.37  11.49,
9     4          2.1    35.9   11.3,
10    4          2.1    63.24  8.9,
11    4          2.1    28.62  10.4,
12    4          2.4    32.26  9.0};
```

```
RD =
{1 2 3,
1 2 4,
1 3 4,
2 3 4};

/*RD is response-wise design*/

v = 12;
b = 1;
k = 12;
r = 1;
rr = 3;
nn = 36;

y1 = s1[,3:5];
y2 = s2[,3:5];
y3 = s3[,3:5];
y4 = s4[,3:5];
/*take ys from 3 to 5 column of ss*/
C = I(v) - (1/v)*j(v,v,1);
N = j(v,b,1);

T1 = s1[,3:5];
T2 = s2[,3:5];
T3 = s3[,3:5];
T4 = s4[,3:5];

print t1 t2 t3 t4;

B1 = s1[+,3:5];
B2 = s2[+,3:5];
B3 = s3[+,3:5];
B4 = s4[+,3:5];

Q1 = T1 - N*inv(k)*B1;
Q2 = T2 - N*inv(k)*B2;
Q3 = T3 - N*inv(k)*B3;
Q4 = T4 - N*inv(k)*B4;

print q1 q2 q3 q4;

/*Q11 =  (Q1[,1]+Q2[,1])/rr;*/
/*Q22 =  (Q1[,2]+Q3[,1])/rr;*/
/*Q33 =  (Q2[,2]+Q3[,2])/rr;*/

Q11 =  (Q1[,1]+Q2[,1] + Q3[,1]);
Q22 =  (Q1[,2]+Q2[,2] + Q4[,1]);
Q33 =  (Q1[,3]+Q3[,2] + Q4[,2]);
Q44 =  (Q2[,3]+Q3[,3] + Q4[,3]);
```

```
print q11 q22 q33 q44;

tau1 = ginv(C)*Q11;
tau2 = ginv(C)*Q22;
tau3 = ginv(C)*Q33;
tau4 = ginv(C)*Q44;

tau = tau1//tau2//tau3//tau4;
print tau;

/*sigma11 = (y1`*y1 - q1`*ginv(c)*q1)#(1/(nn - v));*/
/*sigma22 = (y2`*y2 - q2`*ginv(c)*q2)#(1/(nn - v));*/
/*sigma33 = (y3`*y3 - q3`*ginv(c)*q3)#(1/(nn - v));*/

/*s11 = (sigma11[1,1] + sigma22[1,1])/2;*/
/*s12 =  sigma11[1,2];*/
/*s22 = (sigma11[2,2] + sigma33[1,1])/2;*/
/*s13 =  sigma22[1,2];*/
/*s33 = (sigma22[2,2] + sigma33[2,2])/2;*/
/*s23 =  sigma33[1,2];*/
/*one can obtain thezse values and substitute*/
s11 = 1058.03;
s12 = -5.32;
s13 = 91.90;
s14 = -70.52;
s22 = 0.9394;
s23 = 4.6325;
s24 = -0.9897;
s33 = 1094.68;
s34 = -10.006;
s44 = 72.16;

sigma = (s11||s12||s13||s14)//(s12||s22||s23||s24)//(s13||s23||s33||s34)//(s14||s24||s34||s44);

print sigma;

*p1 = {1 -1 0 0 0 0 0 0 0 0 0 0};    /*NS*/
*p1 = {1  0 -1  0 0 0 0 0 0 0 0 0};  /*NS*/
*p1 = {1  0 0  -1 0 0 0 0 0 0 0 0};  /*NS*/
*p1 = {1  0 0  0 -1 0 0 0 0 0 0 0};  /*NS*/
*p1 =  {1 0 0 0 0 -1 0 0 0 0 0 0};    /*S*/
*p1 =  {1 0 0 0 0 0 -1 0 0 0 0 0};    /*NS*/
*p1 =  {1 0 0 0 0 0 0 -1 0 0 0 0};    /*NS*/
*p1 =  {1 0 0 0 0 0 0 0 -1 0 0 0};    /*NS*/
*p1 = {1 0 0 0 0 0 0 0 0 0 0 -1};    /*S*/
*p1 = {1 0 0 0 0 0 0 0 0 0 -1 0};    /*S*/
*p1 = {1 0 0 0 0 0 0 0 0 -1 0 0};    /*NS*/
*p1 = {1 0 0 0 0 0 0 0 -1 0 0 0};    /*NS*/
*p1 = {1 0 0 0 -1 0 0 0 0 0 0 0};    /*NS*/
```

```
*Chi square;
/*these are to be tested individually*/
d_p1t1 = (p1*ginv(c)*t(p1)*s11)/(rr**2);
d_p2t2 = (p1*ginv(c)*t(p1)*s22)/(rr**2);
d_p3t3 = (p1*ginv(c)*t(p1)*s33)/(rr**2);
d_p4t4 = (p1*ginv(c)*t(p1)*s44)/(rr**2);

P = I(4)@p1;
Ptau = P*tau;

d_p1t1p2t2 = p1*ginv(c)*t(p1)*s12/(rr**4);
d_p1t1p3t3 = p1*ginv(c)*t(p1)*s13/(rr**4);
d_p1t1p4t4 = p1*ginv(c)*t(p1)*s14/(rr**4);
d_p2t2p3t3 = p1*ginv(c)*t(p1)*s23/(rr**4);
d_p2t2p4t4 = p1*ginv(c)*t(p1)*s24/(rr**4);
d_p3t3p4t4 = p1*ginv(c)*t(p1)*s34/(rr**4);

d_ptau =
(d_p1t1||d_p1t1p2t2||d_p1t1p3t3||d_p1t1p4t4)//(d_p1t1p2t2||d_p2t2||d_p2t2p3t3||d_p2t2p4t4)

        //(d_p1t1p3t3||d_p2t2p3t3||d_p3t3||d_p3t3p4t4)//(d_p1t1p4t4||d_p2t2p4t4||d_p3t3p4t4||d_p4
t4);

TC = t(ptau)*inv(d_ptau)*ptau;


print ptau  d_ptau tc;

run;
quit;
ods html close;
```

A.5. Discussion

The above approach of obtaining the designs for incomplete multi-response designs assumes that the number of error degrees of freedom for each set is more than the number of response variables collected from that particular set. This allows estimations of error variances and covariances. For practical considerations, one may consider a minimally connected design for each set provided that for response wise design $\lambda_{ss'} \geq 2$, where $\lambda_{ss'}$ denote the concurrences of responses s and s'; $\forall s, s' = 1,2,\ldots, p$. If $\lambda_{ss'}$ =1, then the covariance between the response variables s and s' cannot be estimated. A single replication of RCB design as taken in method A.3.1 is also a minimally connected design and satisfies the property of variance balance structure. Therefore, in method A.3.1, we have taken response wise design as BIB design with $\lambda \geq 2$.

Taking, single replication of RCB design as treatment wise design gives a complete block design as treatment wise design. In some experimental situations it may not be possible to have as many homogenous experimental units as the number of treatments. In such

situations one has to make use of incomplete block designs. Here, in minimally connected incomplete block designs may be used for efficient minimally connected designs, a reference may be made to Mandal, Shah and Sinha (1991), Bapat and Dey (1991), Dey, Shah and Das (1995) and Das, Dean and Notz (1998). The minimally connected designs can be obtained in various ways. The question therefore is which minimally connected design for an individual set should be used? Further, one may also think of different minimally connected designs as treatment wise designs for different sets, such that the overall treatment wise design is a BIB design, GD design, standard reinforced resolvable BIB designs etc. Standard reinforced resolvable BIB designs may be quite useful for these situations.

References

Bapat, R.B. and Dey, A. (1991). Optimal block designs with minimal number of observations. *Statist. Prob. Letters*, **11**, 399-402.

Brzeskwiniewicz, H. (1995). On the PBB designs for multi-response experiments. *Biometrical Journal*, **37**(2), 191-198.

Das, A., Dean, A.M. and Notz., W.I. (1998). Designs with nearly minimal number of observations and flexible blocking. *Jour. Statist.Plann. Inf.*, **72**, 133-147.

Dey, A., Shah, K.R. and Das, A. (1995). Optimal block designs with minimal and nearly minimal number of units. *Statistica Sinica*, **5**, 547-558.

Gupta, S.C. (1988). Type A incomplete multi-response designs. *Comm. Statist:Theory & Meth.*, **17**(**9**), 3089-3105.

Mandal, N.K., Shah, K.R. and Sinha, B.K.(1991). Uncertain resources and optimal designs: Problems and perspectives. *Calcutta Statist. Assoc.Bull.*, **40**, 267-282.

Mitra, R.K. and Saha, G.M. (1991). On the analysis of balanced incomplete multi-response designs. *J. Ind. Soc. Agril. Stat*., **43**(**1**), 1-8.

Roy, S. N. (1957). *Some Aspects of Multivariate Analysis.* John Wiley & Sons, New York.

Roy, S.N., Gnanadesikan, R. and Srivastava, J. N. (1971). *Analysis and Design of Certain Quantitative Multiresponse Experiments*, Pregamon Press, Oxford.

SAS Institute Inc. (1989). *SAS/STAT User's Guide*, Vol.2, version 6, 4th ed. Cary, NC.

Srivastava, J.N. (1966). Incomplete multiresponse designs. *Sankhya*, **A 28**, 377-388.

Srivastava, J.N. (1968). On a general class of designs for multi-response experiments. *Ann. Inst. Statist. Math.*, **39**,1825-1843.

Srivastava, J.N. and McDonald, L.L. (1969). On the costwise optimality of hierarchical mltiresponse randomized block designs under the trace criterion. *Ann. Inst. Statist. Math*., **21**, 507-514.

Srivastava, J.N. and McDonald, L.L. (1970). On large class of incomplete imultivariate models, which can be transformed to made manova applicable. *Metron*, **28**, 241-252.

Trawinksi, I.M. and Bargmann, R.E. (1964). Maximum likelihood estimation with incomplete multivariate data. *Ann. Math. Statist*., **35**, 647-657.

B) Software for Nested Block Designs

Nested block designs are the designs in which one system of blocks is nested within another system of blocks. Here the blocks with larger size are called bigger blocks and sub-blocks are nested within these bigger blocks. Nested block designs are useful in many situations. For example, consider a field experiment conducted using a block design where harvesting is done block-wise. The harvested samples are to be analyzed for their contents (quality indicators viz. protein content, etc.) in the laboratory by different technicians at same time or by a technician over different periods of time. The variation due to technicians or time periods may be controlled by another blocking system. Technicians or time periods form a system of blocks called sub-blocks that are nested within blocks. Preece (1967) introduced nested balanced incomplete block (NBIB) designs and gave methods of construction of NBIB designs. Morgan *et al.* (2001) gave a comprehensive review on systematic methods of construction of NBIB designs and catalogued all NBIB designs for v (number of treatments) ≤ 16, r (replication number) ≤ 30 are catalogued by Morgan *et al.* (2001). For more details on NBIB designs, a reference may be made to Morgan *et al.* and references cited therein. An NBIB design may not always exist or even if it exists may require a large number of replications, which the experimenter may not be able to afford. To deal with such situations, Homel and Robinson (1975) defined nested partially balanced incomplete block (NPBIB) designs. For more details on NPBIB designs and their catalogues, one may refer to Satpati and Parsad (2004) and therefirences cited therein.

NBIB and NPBIB designs are for the situtations when observations within sub-blocks, between sub-blocks within blocks are independent. There, however, do occur experimental situations in which the assumption of independence of observations gets violated; observations within sub-blocks or between sub-blocks within blocks or both may be correlated. The observations between any two blocks are generally assumed as independent. To make the exposition clear about the correlation structure for nested block designs: consider an experimental situation, where v treatments are to be compared via n experimental units, arranged in a nested block set up involving b_1 blocks; there are q mutually exclusive sub-blocks nested within each block, so that $qb_1 = b_2$ is the total number of sub-blocks. Let k_1 be the bigger-block size and $k_2 = k_1/q$ be the sub-block size so that total number of observations is $n = b_1k_1 = b_2k_2$. Let the observations be arranged in such a way that the first k_2 observations come from first sub-block nested within first bigger block, second k_2 observations come from second sub-block nested within the first bigger block, likewise last k_2 observations come from the last sub-block of the last bigger block. All the observations are arranged according to the positions of occurrences within a sub-block. The correlation structure that may exist among the observations is of Kronecker product type and is of the following form $\mathbf{I}_{b_1} \otimes \mathbf{\Phi}_q \otimes \mathbf{\Omega}_{k_2}$, the first part denotes that observations from any two different bigger blocks are independent, second part denotes the correlation structure of the observations between sub-blocks nested within a bigger block and the last part represents the correlation structure among the observations within a sub-block. Further, it is assumed that the correlation structure along with correlation values is same for all the sub-blocks and blocks. $\mathbf{\Phi}_q$ and $\mathbf{\Omega}_{k_2}$ may be any one of the nearest

neighbour (NN), autoregressive of order 1 (AR(1)), equi-correlated or no correlation structures. When $\boldsymbol{\Phi}_q = \mathbf{I}_q$ and $\boldsymbol{\Omega}_{k_2} = \mathbf{I}_{k_2}$ then it reduces to the set up with uncorrelated error structure. Let $y_{tl(j)}$ be the observation pertaining to the t^{th} position within l^{th} sub-block nested within the j^{th} block, then the different types of correlation structures can be summarized as follows

The correlation between $y_{tl(j)}$ and $y_{t'l'(j')}$ is of the following form:

$$\text{Corr}(y_{tl(j)}, y_{t'l'(j')}) = \begin{cases} \rho_1^{f_1(h)} \rho_2^{f_2(g)} & \text{if } |t-t'| = g, |l-l'| = h \text{ and } j = j' \\ 0 & \text{elsewhere} \end{cases} \qquad \text{(F.1.1)}$$

where $|\rho_1| \le 1$ and $|\rho_2| \le 1$ and ρ_1 is the correlation between observations coming from the experimental units of two neighbouring sub-blocks and ρ_2 is the correlation between two experimental units that are neighbour within a sub-block ignoring other types of correlations and Corr(.) denotes correlation coefficient. Different choices of functions $f_2(g)$ and $f_1(h)$ defined give rise to different types of Kronecker Product type correlation structure $\mathbf{I}_{b_1} \otimes \boldsymbol{\Phi}_q \otimes \boldsymbol{\Omega}_{k_2}$. If

$$f_1(h) = \begin{cases} 0 & \text{for } h = 0 \\ 1 & \text{for } h = 1 \\ -\infty & \text{for } h > 1 \end{cases} \quad \text{and } \rho_1 \neq 0$$

then the correlation structure between the sub-blocks nested within a bigger-block is of NN type. If $f_1(h) = h$, $\forall\ h = 0, 1, 2, \ldots, q-1$ and $\rho_1 \neq 0$, then this correlation structure is of AR(1) type. The correlation structure is called equi-correlated structure if

$$f_1(h) = \begin{cases} 0 & \text{for } h = 0 \\ 1 & \text{for } h > 0 \end{cases} \quad \text{and } \rho_1 \neq 0$$

Defining the identical function $f_2(g)$ as $f_1(h)$, we get similar correlation structure between the observations pertaining to the same sub-block. If

$$f_1(g) = \begin{cases} 0 & \text{for } g = 0 \\ 1 & \text{for } g = 1 \\ -\infty & \text{for } g > 1 \end{cases} \quad \text{and } \rho_2 \neq 0$$

then the correlation structure between the observations within a sub-block is of NN type. If $f_2(g) = g$, $\forall\ g = 1, 2, \ldots, k_2 - 1$ and $\rho_2 \neq 0$, then this correlation structure is of AR(1) type. The correlation structure is called equi-correlated structure if

$$f_1(g) = \begin{cases} 0 & \text{for } g = 0 \\ 1 & \text{for } g > 0 \end{cases} \quad \text{and } \rho_2 \neq 0.$$

The generalized Kronecker product type of correlation structure defined above can produce $\mathbf{NN} \otimes \mathbf{NN}$, $\mathbf{AR(1)} \otimes \mathbf{AR(1)}$, $\mathbf{NN} \otimes \mathbf{AR(1)}$, $\mathbf{AR(1)} \otimes \mathbf{NN}$, $\mathbf{I}_q \otimes \mathbf{AR(1)}$, $\mathbf{I}_q \otimes \mathbf{NN}$, $\mathbf{AR(1)} \otimes \mathbf{I}_{k_2}$, $\mathbf{NN} \otimes \mathbf{I}_{k_2}$ and $\mathbf{I}_q \otimes \mathbf{I}_{k_2} = \mathbf{I}_{k_1}$.

For these experimental situations with correlated observations, it is useful to have efficient nested block designs that provide protection against the effects of correlated observations or potentially unknown trends, which are highly correlated with positions of experimental

units within blocks. In the example described earlier, nested block designs are used for laboratory analysis whereas the field experiment is conducted using a block design and harvested samples are to be analyzed for their contents in the laboratory by different technicians at the same time. The variation due to technicians or time periods is controlled through forming sub-blocks within blocks. In this experiment, experimenter also records yield in the field besides the characters based on laboratory analysis. The experimenter may want to compare the treatments based on their yield performance and as well as on the character like protein contents recorded in the laboratory based on the analysis of harvested samples. The yield is obtained from the field itself and has not been subjected to the laboratory analysis. Therefore, the analysis of experimental data has to be carried out as per design adopted for field experimentation *i.e.* a block design. Therefore, we have to choose a design which is efficient both under nested block design setup and block design setup ignoring sub-block classification.

The available methods can be used for obtaining efficient nested block designs for independent observations. However, using these methods of construction one may not be able to obtain designs for all parametric combinations. It seems that very little attention has been paid to obtain efficient nested block designs when observations are correlated. Satpati *et al.* (2006) made a computer-aided search has been made of efficient nested block designs for given parametric combinations both under uncorrelated/ correlated error structure. The algorithm for obtaining efficient nested block designs for dependent and independent observations was reported last year.

The algorithm developed has been used for development of a β-version of the software for generation of efficient nested block designs has been developed. It consists of 3 modules viz. (i) nested block designs with independent observations; (ii) nested block designs when the observations within a sub-block have a nearest neighbour correlation structure and (iii) nested block designs when the observations within a sub-block have a nearest autoregressive correlation structure. Once the number of treatments (v), number of bigger blocks (b), bigger block size (k), number of sub-blocks nested within a block (q) and the value of correlation coefficient (ρ) is entered (in case of correlated case), the design gets generated. This software can usefully be employed for generation of efficient resolvable block designs by taking $k_1 = v$. A screen shot of the software developed is given on the left page.

References

Homel, R.J. and Robinson, J. (1975). Nested partially balanced incomplete block designs. *Sankhy* $\bar{a}$, **B37**, 201-210.

Morgan, J.P., Preece, D.A. and Rees, D.H. (2001). Nested balanced incomplete block designs. *Discrete Mathematics*, **231**, 351-389.

Preece, D.A.(1967).Nested balanced incomplete block designs. *Biometrika*, **54**, 479-486.

Satpati, S.K. and Parsad, R. (2004). Constructions and cataloguing of the Nested partially balanced incomplete block designs. *Ars Combinatoria*, **73**, 299-309.

Satpati, S.K., Parsad, R., Gupta, V.K. and Nigam, A.K. (2006). Computer aided search of efficient nested incomplete block designs for correlated observations. *Journal of Combinatorics, Information and System Sciences (Special issue in Honour of Prof. Bhu Dev Sharma),* **31(1-4)**, 163-186.

C) Design Resources Server

For disseminating research in Design of Experiments among the scientists of NARS in particular and researchers all over globe in general, an initiative was taken under the present National Fellow Scheme to create a Design Resources Server in 2005. This server is hosted at the home page of IASRI (www.iasri.res.in/design) so as to make available design theory and actual layout of the designs through web. The Design Resources Server is further being strengthened with new research findings under the National Professor Scheme entitled "Designs for Single Factor and Multi-Factor Experiments and Their Applications in Agricultural Systems Research". It was thought that the fundamentals of design of experiments and other related issues available in the literature and the new research findings of the National Fellow and National Professor schemes as well as research in newer emerging areas of design of experiments being conducted elsewhere should be made available at one place for the benefit of researchers.

The basic purpose of this amalgamation and integration is to provide a platform to share the literature available on design of experiments, research findings of National Fellow and National Professor Schemes and other recent advances made in the literature at International level. An interesting feature of Design Resource Server is that it is open to everyone and anybody from all over the globe can join this and add information to the site to strengthen it further

The goal of this server is to reach experimenters in agricultural, biological and social sciences, industry, etc. and help them in planning and designing their experiments and then to analyze the data gathered so as to enable them to improve the quality of their research and make it globally competitive. The aim is to spread advances in theoretical, analytical and applicational aspects of design of experiments among mathematicians and statisticians both in academia and also involved in advisory and consultancy services. Most importantly, this server targets to popularize the research in design of experiments at a global level and also to create a virtual mobile library on design of experiments. This is also an attempt to illustrate the usefulness of statistics in agricultural research.

During the period under report, Design Resources Server has been strengthened in collaboration with National Professor. The following new additions were made

C.1. Modules for Generation of Randomized Layout of Basic Designs

Online modules for generation of randomized layout of basic designs such as completely randomized designs, randomized complete block designs and Latin square designs have been uploaded on Design Resources Server. Randomized layouts of basic designs (completely randomized designs, randomized complete block designs, Latin square designs, etc.) can be obtained at (www.iasri.res.in/design/Basic%20Designs/generate_designs.htm).

C.2. Randomized Layout of Augmented Designs

Augmented designs are very popular among experimenters. A question generally asked by the experimenters is concerning the replication of the control treatments in each block. An algebraic treatment of this problem was done and optimum replication of controls in each block has been worked out by Parsad and Gupta (2000). This optimum replication is a function of number of tests, number of controls and the number of blocks. Another

problem with the use of augmented designs is that generally the randomization done is faulty. In view of all these difficulties, an online generation of randomized layout of an augmented design is available at the server. This gives optimum replication of controls in each block so as to maximize the efficiency per observation for making tests treatment-control treatment comparisons. The experimenter has also the flexibility of choosing replication of controls of his own choice. Further, one can also form blocks with unequal sizes, depending upon the resources available and the variability in the experimental material. Randomized layout of augmented randomized complete block designs for given number of test treatments, control treatments and number of blocks can be generated using (www.iasri.res.in/design/Augmented%20Designs/home.htm).

C.3. Square Lattice Designs

Square lattice designs form an important class of incomplete resolvable block designs which provide flexibility to suit various experimental situations. Square lattice designs require that the number of treatments must be a perfect square, i.e. $v = s^2$. Simple lattice (square lattice designs with 2 replications) and triple lattice (square lattice designs with 3 replications) can be constructed for $v = s^2$, where s is a positive integer. Existence of square lattice designs with number of replications greater than 3 dependds on the existence of mutually orthogonal Latin squares (MOLS) of order *s*. If *s* is a prime or prime power, the complete set of $s-1$ (MOLS) of order *s* exist. Therefore, a balanced lattice design (square lattice design with number of replications as $s+1$ exists only when *s* is a prime or prime power. As triple lattice designs can be constructed for any $v = s^2$, where s is a positive integer, therefore, online software for generation of square lattice designs with 3 replications has been prepared and uploaded on the Server (www.iasri.res.in/WebHadamard/square%20lattice.htm). It is being extended for generation of balanced lattice designs.

C.4. α-Designs

α-designs are essentially resolvable block designs in which the blocks can be grouped such that in each group, every treatment appears exactly once. On the face of it, these designs look like randomized complete block designs and the purpose of demonstration of variety effect in the field to the inspection team can be made as all the varieties will be appearing on the adjacent piece of land. These designs are very useful for initial varietal trials conducted under the crop improvement programmes. In these designs, the between blocks within replication variation helps in reducing the experimental error, increasing thereby the precision of estimation of treatment contrasts of interest. These designs are also useful for the field trials with large number of treatments/crop varieties which may not always be laid out in a single location or a single season. Therefore, it is desired that variation due to location or time periods may also be controlled along with controlling within location or time period variation. By taking locations or time periods as replications and the variation within a location or a time period as blocks, these situations can also be handled using resolvable block designs. α-designs have been recommended for varietal trials conducted by AICRP on Rapeseed and Mustard, National Bureau of Plant Genetic Resources and Central Institute of Cotton Research. In many of the trials conducted using α-designs the blocks nested within replication differences were found to be significant and low CV in

comparison to a RCB design. This corroborates the hypothesis that these designs are quite useful for such trials. There is a feeling that it may not be possible for the experimenters to lay their hands on α-designs. For the benefit of experimenters and practicing statisticians, a webpage of α-designs has been linked to Design Resources Server and is available at www.iasri.res.in/design/Alpha/Home.htm. One can generate α-designs along with the randomized layout for number of treatments lying between 6 and 150, replications less than or equal to 5 and of block sizes lying between 3 to 10 along with block contents. All concepts and details on α-designs are also available on this URL. The designs given on this URL are essentially the same as catalogued by Parsad *et al*. (2007).

C.5. Designs for Bioassays

Designs for biological assays help in the estimation of the *relative potency* of the *test preparation* with respect to *standard* one. The material uploaded on Design Resources Server includes contrasts of interest in parallel line assays and slope ratio assays. The catalogue of efficient block designs for parallel line assays is given. One can generate the block contents of the required design by clicking on the design in catalogue. (www.iasri.res.in/design/BioAssays/bioassay.html). The designs given on this page are same as those listed in Srivastava *et al*. (2006).

C.6. Supersaturated Designs

Supersaturated designs are fractional factorial designs in which the degrees of freedom for all its main effects and the intercept term exceed the total number of distinct factor level combinations of the design. These designs are useful when the experimenter is interested in identifying the active factors through the experiment and experimental resources are scarce. Definition of supersaturated designs, experimental situations in which supersaturated designs are useful, efficiency criteria for evaluation of supersaturated designs, catalogue of supersaturated designs for 2-level factorial experiments and asymmetrical factorial experiments and bibliography on supersaturated designs has been uploaded on the Server. The complete details of the runs can be obtained by clicking on the required design in the catalogue. (www.iasri.res.in/design/Supersaturated_Design/Supersaturated.html). The supersaturated designs for asymmetrical factorial experiments are essentially those catalogued in Gupta et al. (2008).

C.7. Analysis of Data

A new page "Analysis of Data" has been launched on Design Resources Server. The purpose of this web page is to provide steps of analysis of data generated from designed experiments by using statistical packages like SAS, SPSS, MINITAB, SYSTAT, MS-EXCEL, etc. At present steps for analysis of data generated from randomized complete block designs, incomplete block designs, resolvable incomplete block designs, Latin square designs, factorial experiments conducted using a randomized complete block design, partially confounded factorial experiments, balanced confounded factorial experiments with extra treatment, response surface designs, correlation and regression and test of significance using SAS and SPSS have been uploaded on this page. The steps for performing multiple comparison procedures and treatment contrast analysis are also given both using SAS and SPSS. The data files and result files can also be downloaded.

A SAS macro for performing diagnostics (normality and homogeneity of errors) in experimental data generated through randomized complete block designs and then applying remedial measures such as Box-Cox transformation is also made available on this server.

C.8. Discussion Board

A Discussion Board has also been initiated for providing as a platform for sharing any useful piece of research or idea with any other scientist over the globe. The user can use this board for learning and disseminating information after registering on the discussion board. The information can be viewed by anybody over the globe. If there are some querries or some researchable issues then other peers can also respond to these querries. This helps in creating a network of scientists. Number of registered participants so far is 38 (18: Agricultural Research Statisticians; 19: Experimenters and One Vice-Chancellor). (www.iasri.res.in/design/MessageBoard/MessageBoard.asp).

C.9. Who-is-where

Addresses of important contributors in Design of Experiments including their E-mail addresses were linked to Design Resources Server. The list includes experts from USA, Canada, Australia, UK, China, Japan, Mexico, New Zealand, Oman, Syria, Taiwan, Vietnam and India. This information is useful for all the researchers in Design of Experiments in establishing linkages with their counterparts over the globe.

C.10. Feedback/ Comments

Feedback/ comments received on the Design Resources Server were also linked with Design Resources Server. We have received feedback from 18 researchers (6: Design Experts from India; 7: Experts from abroad; 3: Experimenters and 2: Agricultural Research Statisticians). The first feedback was received from Dr K Rameash, Entomologist working at ICAR Research Complex for NEH Region, Sikkim Centre, Tadong, Gangtok.

C.11. Some Information on the Usage of the Server

- A hit counter was put on the server on November 20, 2007. Since then there have been 1710 hits on this site.
- External links of the server are also available at:
  http://en.wikipedia.org/wiki/Design_of_experiments
  http://en.wikipedia.org/wiki/Hadamard_matrix

- The server has been cited at:

https://dspace.ist.utl.pt/bitstream/2295/145675/1/licao_21.pdf for lecture presentation on Unitary operators.

INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

C.12. Future Directions

The design resources server is dynamic in nature and new research findings and new links would be added to the server on a continuing basis. In the near future, the server is intended to take the following form so that both the statisticians and the stake holders can take utmost benefit from this resource of knowledge.

The Design Resource Server would be redesigned and converted into two parts. Part-I would generally cater to the needs of the stakeholders (scientists of NARS in particular). Part-II would by and large be devoted to statisticians engaged in conducting research in Design of Experiments. However, this divide would only be for convenience, but the information contained in both the parts would be useful for both the groups.

Once the split described above is achieved, an effort would be made to split the material according to the Divisions of the ICAR. This would help the subject matter divisions to look into the contents useful in their research. But as mentioned earlier, the information contained in this would cut across the divisions.

Most importantly, the network of scientists would be expanded and strengthened so as to make this server more useful to the stake holders and strengthen the advisory service.

Part – I would comprise of:

- E-advisory and E-consultancy
  - Electronic books
  - Ask a question
- Network of scientists
  - Discussion/ Advisory board
- Analysis of data
  - Steps of analysis of data using SAS, SPSS, SPAR2, MINITAB, SPAD, SPFE, etc.
  - Contrast analysis and covariance analysis (both intra class and inter class)

Note: Online availability of software, whenever possible and in what so ever form possible, would depend upon the availability of resources like a lot of server space, memory and speed and dedicated high speed lease line.

- Resolvable block designs useful for varietal trials
  - Resolvable BIB designs
  - Resolvable PBIB designs
  - Square and Rectangular Lattice
  - Alpha designs
- Block designs with nested structures (other than those listed above)
- Augmented designs
- Block designs for making test treatments vs control treatment(s) comparisons
- Regression designs
  - Response surface methodology
  - Experiments with mixtures methodology

- Designs for factorial experiments
  - Symmetrical factorials
  - Asymmetrical factorials
  - Fractional factorials
  - Supersaturated designs

Part – II would comprise of:

- Designs for one blocking system
- Designs for two blocking systems
- Designs with correlated error structures
- Designs with nested structures (incomplete bock)
  - Single nesting
  - Nesting in two directions
- Designs for factorial experiments
  - Symmetrical factorials
  - Asymmetrical factorials
  - Fractional factorials
  - Orthogonal main effect plans
  - Supersaturated designs
- Designs for biological assays (symmetrical and asymmetrical)
  - Parallel line assays
  - Slope ratio assays
- Regression designs
  - Designs for fitting response surfaces
  - Designs for experiments with mixtures
- Other considerations
  - Optimality aspects
  - Robustness of designs against
    - Loss of data
    - Outliers
    - Interchange and exchange of treatments
    - Model misspecification

- Online generation of designs
  - Hadamard matrices
  - Orthogonal arrays
  - Nested orthogonal arrays
  - Fractional factorials
  - Supersaturated designs
  - Designs for Multi-response experiments
  - Mutually orthogonal Latin squares
  - Microarray experiments
  - Designs for computer experiments

For all the items listed in Part – II, the following aspects would be covered:

- Methods of construction

- Catalogue of designs along with randomized layout
- Efficiency of designs
- Bibliography

References

1. Gupta, V.K., Parsad, R., Bhar, L.M. and Kole, B. (2008). Supersaturated designs for asymmetrical factorial experiments. *Journal of Statistical Theory and Practice,* **2(1),** 95-108.
2. Parsad, R. and Gupta, V.K. (2000). A note on augmented designs. *Indian Journal of Plant Genetic Resources,* **13(1)**, 53-58.
3. Parsad, R., Gupta, V.K., Batra, P.K., Satpati, S.K. and Biswas, P. (2007). *Monograph on α-Designs.* IASRI, New Delhi.
4. Srivastava, R., Gupta, V.K. and Parsad, R. (2006). *Studies on Block Designs for Biological Assays*. IASRI, New Delhi.

D) Transfer of Technology/ Advisory Services

Advisory services were pursued rigorously. Details of some of them are given below:

National Bureau of Plant Genetic Resources, New Delhi

- Sh. Narendra Panwar was advised on the combined analysis of data pertaining to the experiment conducted to study the genetic variability and stability analysis in *Ocium species* at four locations using a randomized complete block design in 30 cultivars each replicated thrice. There were 16 characters on which data were collected.

National Bureau of Soil Survey and Land Use Planning, Delhi Centre

- Dr. Dharam Singh, Senior Scientist was advised on the combined analysis of data pertaining to the experiment conducted to evaluate 4 cropping systems viz., Rice-Fallow, Rice-Pea, Rice-Potato and Rice-Mustard conducted in 2004 and 2005. The data were analyzed as rice equivalent yield and potato equivalent yield.

Indian Agricultural Research Institute, New Delhi

- Dr. Yashbir Shivay, Senior Scientist from Division of Agronomy was advised on the analysis of data pertaining to the experiments 'Effect of zinc-enriched urea on productivity, zinc uptake and efficiency of aromatic rice-wheat cropping system' conducted to study the effect of various concentrations of zinc-enriched urea on productivity of aromatic rice-wheat cropping system and on zinc concentrations of aromatic rice and wheat grain and straw and their uptake by the rice-wheat cropping system. He was advised on the use of Tukey's Honest Significant Difference and Duncan's Multiple Range Tests.
- Sh. Nishant Kumar Sinha, M.Sc. student of Agricultural Physics was advised on the development of soil index using the data on pH, electrical conductivity, bulk density, porosity, organic carbon, hydraulic conductivity, microbial biomass, dehydrogenase activity (active population of microbes), Nitrate 'N' (nitrogen in nitrate form), NH_3 'N' (nitrogen in ammonia form), aggregate stability (mean weight diameter of soil particle which is important for plant growth), available water (difference between field capacity water and wilting point water) from soil in rice-wheat and maize-wheat cropping systems using principal component analysis. In most of the cases first principal component explained more than 75% of the variation and first principal component score was used as soil quality index for different management practices. He was also advised on the identification of treatment with highest soil quality index and testing the significance of differences in treatment effects. The experiment was conducted during 2006-07.
- Dr. Jagmail Singh, Principal Scientist from Division of Genetics was advised on the analysis of data from complete diallel cross experiment pertaining to cotton (*Gossypium hirsutum* L.) conducted using randomized complete block designs. Three different sets of diallels were made to study genetics of fibre quality, especially fibre strength and its association, if any, with other characters. At Delhi, diallels were made involving 6 genotypes with medium staple (medium fibre length); at Dharwad, crosses were made involving 6 genotypes with long fibre length and at Nagpur, the crosses were made with 8 genotypes for seed oil contents as there is good scope for genetic improvement of

seed oil content in cotton. The complete diallel crosses system (p parents, $p(p-1)/2$ crosses and $p(p-1)/2$ reciprocals) were tried at all the three locations to understand the influence of environment and cytoplasm on expression of these characters. Two traits selected for the study are at present very important in the context of suitability for high speed spinning by textile industry and for increasing oil content in our cultivars. The data was analyzed for individual location separately and combined over all the three locations for all the 3 sets of crosses.

Nagaland

- Sh. W. Rungsung Asemosem from Nagaland was advised on the analysis of data pertaining to an experiment conducted on 50 varieties of rice bean using a RCB design with 3 replications conducted for 2004 and 2005 at three locations Patkai, SASRD and Kohima. He was advised on the analysis of data (both individual environment and combined over 6 environments) pertaining to 13 characters viz. flowering, maturity, plant height, pods per cluster, cluster per plant, pods per plant, pod length, seeds per pod, biomass per plant, protein content, carbohydrate content, 100-seed weight, seed yield per plant.

National Research Centre on Rapeseed and Mustard

- In 2006-07, 24 initial varietal trials were conducted using α-designs. The parameters of these designs are: (i) $v = 12$, $b = 12$, $r = 3$, $k = 3$, A-efficiency = 0.9241; (ii) $v = 15$, $b = 15$, $r = 3$; $k = 3$, A-efficiency = 0.9067; (iii) $v = 24$, $b = 12$, $r = 3$, $k = 6$, A-efficiency = 0.9699; (iv) $v = 28$, $b = 12$, $r = 3$, $k = 7$, A-efficiency = 0.9603 and (v) $v = 36$, $b = 12$, $r = 2$, $k = 6$, A-efficiency = 0.9074. The data were analyzed. It was observed that in 8 trials there were transcription errors such as exchange of treatments. In 3 trials blocks within replication differences were found to be significant and in other 10 of the 24 trials, CV% reduced in comparison to a RCB design.
- The data from initial varietal trials conducted during 2006-07 by National Research Centre on Rapeseed and Mustard, Bharatpur was analyzed and it has been seen that in all the trials location × genotype interactions were highly significant. Therefore, a single strain cannot be promoted for all the locations in a given zone. Hence, the strains promoted may be different for different subset of locations. Further, the entries giving 10% higher yields than the best performing check may not be significantly different from the best performing check. Therefore, it was suggested that an entry should be promoted only when it is statistically significant from the best performing check. It has also been shown that variability in plant population also plays a role in the variability in yield. If analysis of covariance is performed using plant population as covariate, there may be an improvement in the precision of treatment comparisons.
- Following α-designs for 30 entries were recommended for initial varietal trials to be conducted by AICRP on Rapeseed and Mustard at 22 research centres. Separate randomized layouts for all the centres were provided. Centres may choose a design depending upon the nature of variability in the experimental units.
  1. v = 30, b = 15, r = 3, k = 6, A-efficiency = 0.9536, D-efficiency = 0.9778;
  2. v = 30, b = 18, r = 3, k = 5, A-efficiency = 0.9478, D-efficiency = 0.9745;
  3. v = 30, b = 9, r = 3, k = 10, A-efficiency = 0.9683, D-efficiency = 0.9857.

➢ Publications

Research Papers: 18 (07 published, 05 Accepted for Publication and 06 Communicated/Under Revision); Monograph: 01.

(A) Research Papers

Published

1. Jitendra Kumar, Gynendra Singh, R.K. Palta, Suresh Walia, **Rajender Parsad** and Balraj S. Parmar (2006). Field appraisal of controlled release formulations of carbofuran against the rice leaf folder (*cnaphalocrocis medinalis*). *Indian Journal Agricultural Sciences*, **76(12)**, 732-735.
2. A.K. Joshi, G.Ortiz-Ferrara, J. Crossa, G. Singh, R.C. Sharma, R. Chand and **Rajender Parsad** (2007). Combining superior agronomic performance and terminal heat tolerance with resistance to spot blotch (*Bipolaris sorokiniana*) of wheat in the warm humid Gangetic Plains of South Asia. *Field Crops Research*, **103**, 53-61.
3. Subrata Kumar Satpati, **Rajender Parsad**, V.K. Gupta and A.K. Nigam (2006). Computer aided search of efficient nested incomplete block designs for correlated observations. *Journal of Combinatorics, Information and System Sciences (Special issue in Honour of Prof. Bhu Dev Sharma),* **31(1-4)**, 163-186.
4. Subrata Kumar Satpati, **Rajender Parsad** and V.K. Gupta (2007). Efficient block designs for dependent observations: A computer aided search. *Communications-in-Statistics: Theory & Methods,* **36(6)**, 1187-1223.
5. R. Srivastava, **Rajender Parsad**, Amitava Dey and V.K. Gupta (2007). A-efficient block designs for symmetric parallel line assays. *Utilitas Mathematica,* **73**, 239-253.
6. Jitendra Kumar, K. Nisar, M.B. Arun Kumar, Suresh Walia, N.A. Shakil, **Rajender Parsad** and Balraj S. Parmar (2007). Development of polymeric seed coats for seed quality enhancement of soybean (*Glycine max)*. *Indian Journal of Agricultural Sciences,* **77**(11), 738-743.
7. **Rajender Parsad**, Sanpei Kageyama and V.K. Gupta (2007). Use of complementary property of block designs in PBIB designs. *Ars Combinatoria*, **85**, 173-182.

Accepted for Publication

8. V.K. Gupta, **Rajender Parsad**, Lal Mohan Bhar and Basudev Kole. Supersaturated designs for asymmetrical factorial experiments. *Journal of Statistical Theory and Practice.*
9. Ananta Sarkar, **Rajender Parsad**, Abhishek Rathore and V.K. Gupta. Efficient block designs for microarray experiments. *Journal of Indian Society of Agricultural Statistics*.
10. B.N. Mandal, **Rajender Parsad** and V.K. Gupta. Computer-aided construction of Balanced Sampling Plans Excluding Contiguous Units. *Journal of Statistics and Applications*.
11. B.N. Mandal, **Rajender Parsad** and V.K. Gupta. IPPS Sampling Plans Excluding Adjacent Units. *Communications-in-Statistics: Theory and Methods.*
12. **Rajender Parsad**, Abhishek Ratore and V.K. Gupta. Computer aided construction of efficient designs for making treatment-treatment and treatment-control comparisons. *Special Issue of American Journal of Mathematical and Management Sciences in memory of special volume on Bechhofer-Gupta-Sobel.*

Communicated/Under Revision

13. B.N. Mandal, Rajender Parsad and V.K. Gupta. Construction of Doubly Nested Partially Balanced Incomplete Block Designs. *ARS Combinatoria*.

14. P. K. Nandi, **Rajender Parsad**, L. M. Bhar and V. K. Gupta. Outliers in Multi-response Experiments. *Metron.*
15. P. K. Nandi, **Rajender Parsad** and V. K. Gupta. Simultaneous Optimization of Incomplete Multi-response Experiments. *Sankhya*.
16. B.N. Mandal, Rajender Parsad, V.K. Gupta and U.C. Sud. Nested stratified sampling. *Model Assisted Statistics and Applications.*
17. Krishan Lal, **Rajender Parsad** and V.K. Gupta. Trend-free nested balanced incomplete block designs and designs for diallel cross experiments. *Calcutta Statistical Association Bulletein.*
18. V.K. Gupta, Kishore Sinha and **Rajender Parsad**. Some constructions of orthogonal Arrays. *Journal of Statistical Theory and Practice.*

(B) Monographs

1. V.K. Gupta, A. Dhandapani and **Rajender Parsad** (2007). *Hadamard Matrices*. IASRI, New Delhi.

➢ Awards and Recognitions

Awards

- Received the NAAS-Associateship on June 04, 2007 in the 14th Annual General Body Meeting of NAAS. The Associateship was conferred by Dr. M.S. Swaminathan.

Recognitions

- **Member Secretary**, of the committee constituted to assess the requirement and potential use of a multi-site and multi-user package under NAIP.
- **Judge for adjudication of best poster presentation** during the 10th Annual Conference of Society of Statistics, Computer and Applications held at St. Thomas College, Pala during 16-18, November 2007.
- **Chairman,** Contributed Paper presentations during the 61st Annual Conference of Indian Society of Agricultural Statistics held at Birsa Agricultural University, Kanke, Ranchi during November 30 – December 02, 2007.
- **Reviewer** for Journal of Statistical Planning and Inference, Journal of Statistical Theory and Practice, ARS Combinatoria, Journal of Combinatorics, Information Systems and Sciences and Pusa AgriScience
- Member, Editorial Board for सांख्यिकी विमर्श 2007-2008.
- संस्थान में 01 से 30 सितम्बर, 2008 के दौरान आयोजित किये गये हिन्दी चेतना मास की निम्न प्रतियोगिताओं के लिए गठित समितियों की सदस्यता एवं संचालनः
  1. सदस्य मूल्यांकन समिति–प्रभागीय चल–शील्ड निर्धारण
  2. संचालक उप–समिति, वाद–विवाद प्रतियोगिता
  3. सदस्य उप–समिति हिन्दी शोध–पत्र–प्रदर्शन प्रतियोगिता
  4. मूल्यांकन, निबन्ध लेखन प्रतियोगिता
- वर्ष 2008 की विदेशो वैज्ञानिक पत्रिकाओं को पुस्तकालय में मंगाने हेतु तथा निविदाये खोलने हेतु समिति के मनोनीत अध्यक्ष।

➢ Conferences/ Workshops/ Scientific Meetings organized

➢ Dissemination workshop on ***Design Resources Server*** on March 26, 2008 at I.A.S.R.I., New Delhi was organized by Dr. Rajender Parsad, National Fellow jointly with Dr. V.K. Gupta, National Professor. The workshop was inaugurated by the Dr. S.P. Tiwari, Deputy Director General (Education). Dr. B.S. Bisht, Aaaistant Director General (HRD-I) and Dr. H.S. Gaur, Dean and Joint Director (Education), I.A.R.I., New Delhi also participated along with other delegates. The participants of the workshop included the experimenters from IISS, Bhopal; ANGRAU, Hyderabad; NDRI, Karnal, NBPGR, New Delhi; Divisions of Soil Science, Agronomy, Post Harvest Technology, Agricultural Chemicals, Agronomy and Genetics of IARI, New Delhi. During the Introductory Session, Dr. Rajender Parsad made a presentation on **Genesis and Main Features of Design Resources Server**. Dr. S.P. Tiwari was quite appreciative of the efforts made and said that such workshops should be organized in different regions of the country. The practicing statisticians and experimenters should use this server rigorously and send their comments for further improvements. Introductory Session was followed by two Technical Sessions, in which indepth discussions were held on the Design Resources Server. The Technical Sessions were chaired by Dr. B.M. Prasanna, National Fellow, IARI, New Delhi and Dr. R.C. Agrawal, Principal Scientist, NBPGR, New Delhi. The concluding session was chaired by Dr. S.D. Sharma, Director, IASRI, New Delhi. All the participants were highly impressed with the efforts made. After long deliberations, following points were emerged for wider dissemination and further improvement in the server:

Dissemination of Design Resources Server

1. Dissemination workshops should be conducted in different regions of the country.
2. Faculty Seminars may be delivered at IARI, New Delhi.
3. Efforts should be made to make one presentation on Design Resources Server in the foundation course of Agricultural Research Scientists at NAARM, Hyderabad.
4. The Server may be publicized through ICAR News and Reporter and by providing a link on ICAR Home Page and home page of other ICAR institutes.
5. E-mails may be sent to all organizations in National Agricultural Research System.
6. A letter giving information about the usefulness of the server may be sent to all SAUs and ICAR Institutes through DDG (Education).

Suggestions for Improvements in the Server

1. The material on Server may be divided into two parts. Part I would be primarily useful to scientists in NARS in particular and to stakeholders throughout the globe in general. Part II would be useful for statisticians undertaking research in Design of Experiments. The scientists, however, can use either of the parts or parts of their choice. Part I would include E-learning and E-advisory, Network of scientists, Analysis of data, Designs for single factor experiments, Designs with nested classifications, Designs for multiple factor experiments including fractional factorials, Designs for bioassays, Multi-response experiments, Regression designs like response surface, experiments with mixtures, etc. Part II would mainly comprise of methods of construction of the above said designs, catalogues of the designs along with their efficiency and bibliography on several topics. It would also have online generation of Hadamard matrices, Orthogonal arrays, Fractional factorials, Supersaturated designs, Mutually orthogonal Latin squares, Micro-array experiments and Computer experiments.
2. A disclaimer may be put on the Design Resources Server.
3. Pages of the Design Resources Server may be given a Water Mark.
4. Proper citation of Design Resources Server may also be included.
5. Name of the Server may be changed to Experimental Design Resources Server.

6. A mechanism should be developed for moderation of Discussion Board.
7. Reports for all the designs for which randomized layouts are provided by the Server should be made available automatically. These report should include parameters of the design selected by the user, randomized layout of the design and EXCEL worksheet containing columns of Design parameters for all the classification so that user can enter the data generated.
8. In the analysis of data there should be some provision made for conversion factors and data entry as per field work book records.
9. Titles and abstracts of all the theses on Design of Experiments of IASRI may be provided on the Server.
10. The material on two Electronic books may be clubbed into one avoiding similarity and duplicacy.
11. On every page, date on which last updated may be provided.
12. Efforts should be made to explore the possibility of giving an IP address to the design resources server and also putting it exclusively on a separate and a mechanism be developed for its maintenance.

- Dissemination workshop on **Outliers in Designed Experiments** on 26 July, 2007 at I.A.S.R.I., New Delhi. The workshop was inaugurated by Dr. N.N. Goswami, former Vice-Chancellor Chandra Sekhar Azad University of Agriculture and Technology, Kanpur. Dr. V.K. Gupta gave a brief outline on outliers and there presence in the experimental data. He also gave the motivation for taking up the project. Dr. Rajender Parsad made a presentation on Diagnostics in Designed Experiments. Through two real life examples, he has shown that the assumptions of normality and homogeneity of error variances may be violated due to presence of outliers in the experimental data. Dr. L.M. Bhar gave comprehensive presentation on the findings of the project. Both the technical sessions were chaired by Professor Aloke Dey, Indian Statistical Institute, Delhi. An exposure of Design Resources Server was also given to the participants by Dr. Rajender Parsad and Dr. V.K. Gupta. Dr. Rajendra Prasad, ICAR-National Professor chaired the session on Design Resources Server.

- A meeting with Senior Plant Breeders of All India Co-ordinated Research Project on Rapeseed and Mustard on October 03, 2007 to finalize the designs for the initial varietal trials (IVTs) and criteria of promotion of entries. The meeting was chaired by Dr. S.D. Sharma, Director, I.A.S.R.I., New Delhi. Dr. V.D. Patil, ADG (O & P) and Dr Arvind Kumar, Director, National Research Centre for Rapeseed and Mustard, Bharatpur also participated in this meeting. Dr. Rajender Parsad made a presentation on **Statistical Issues in Rapeseed-Mustard Varietal trials**. Based on the analysis of data generated from α-designs, he has shown that the coefficient of variation reduced in 14 trials out of 27 trials. After long deliberations, following decisions were taken:

- The promotion of entries in IVTs should be done on the basis of combined analysis of data and at 10% level of significance and for advanced varietal trials at 5% level of significance and not on the basis of 10% higher yields than the best performing check.
- If within a zone location*genotype interaction is significant, then promotion should be for a subset of locations
- All IVT with 30 or more entries should be conducted using alpha designs only. The randomized layout will be given by I.A.S.R.I., New Delhi for all the 22 centres.
- A meeting with Professor Bikas K. Sinha, Member, National Statistical Commission was organized Under the Guidance of Dr. V.K. Gupta, ICAR-National Professor on October 31, 2007 at I.A.S.R.I., New Delhi. During this meeting, following presentations were made:

S. No.	Speaker	Title
1.	Krishan Lal	Trend Free Nested Balanced Incomplete Block Designs with Applications in Designs for Diallel Cross Experiments
2.	R. Srivastava	Robustness of Designs for Biological Assays
3.	Lal Mohan Bhar	Robust Methods of Analysis of Experimental Data
4.	Susheel Sarkar	Computer Added Search for Linear Trend-free Factorial Experiments
5.	Cini Varghese	On Some Aspects of Change Over Designs
6.	B.K. Sinha	Some Thoughts on Experiments with Mixtures and Optimal Designs under Covariates in the model

From this meeting, following points were emerged:

- The conditions should be obtained for a nested block design to be trend free both at block and sub-block level. For doing this, two polynomials, one for block and another for sub-block positions may be defined. Efforts may also be made when the trend is different in different blocks/sub-blocks.
- The conditions for robustness of block designs for biological assays against missing blocks may be obtained. The possibility of obtaining/identifying robust designs against any two block missing may be explored.
- Suppose we have a BIB design in *b* blocks. In the beginning of the experiment it is known that resources/funding is available only for $b^* < b$ blocks and it is funding/resources for remaining b-b* blocks is expected with probability α. For this situation develop and algorithm for identification of b* blocks out of b blocks such that the efficiency per observation is maximized. b* may be just enough to have the design connected. The following paper may be helpful in this regard.

Mandal, N. K.; Shah, K. R.; Sinha, B. K. Uncertain resources and optimal designs: problems and perspectives. *Calcutta Statist. Assoc. Bull.* 40 (1990/91), no. 157-160, 267-282.

- While handling outliers in the experimental data, whether the trend in residuals/studentized residuals and Cook-statistic is same or different may be explored. When we make use of robust methods of estimations. Expressions for the variance for the estimated treatment contrasts may be obtained after application of robust methods of estimation. Whether the tests of significance in case of robust estimation remain approximately valid in case of designed experiments like in case of regression. The problem of outliers in case of Change Over designs/ designs for bioequivalence trials may also be attempted.
- The possibility of obtaining error functions in the context of Change over designs may be explored.

- A Discussion Seminar on ***Experiments with Mixtures: Theoretical Advances and Applications*** under the leadership of Dr. VK Gupta, National Professor at I.A.S.R.I., New Delhi on March 17-18, 2008. In this discussion Seminar presentations on **Optimality Aspects of Mixture Experiments: New Directions** were made by Professor Bikas K. Sinha, Member National Statistical Commission and Professor, Indian Statistical Institute, Kolkata, Dr. N.K. Mandal and Dr. Manisha Pal from Calcutta University, Kolkata and Dr. Premadhish Das, Kalyani University, Kalyani. Dr. Rajender Parsad made a presentation on **Experiments with Mixtures: Some Applications**. Dr. Krishan Lal Kalra presented some results on **Experiments with Mixtures**.

- A Dissemination Workshop on ***Fractional Factorials with Special Emphasis on Experiments with Scarce Resources*** under the leadership of Dr. V.K. Gupta, National Professor at I.A.S.R.I., New Delhi on March 18, 2008. During this workshop following presentations were made

V.K. Gupta*, Rajender Parsad and L.M. Bhar	Concepts, Applications and Evaluation criteria of Supersaturated designs
V.K. Gupta* and Rajender Parsad	Orthogonal Arrays and their Applications
V.K. Gupta, Rajender Parsad, L.M. Bhar* and Basudev Kole	Construction of Supersaturated designs for Asymmetrical Factorials
V.K. Gupta, Rajender Parsad*, L.M. Bhar and Basudev Kole	Columnwise Co-ordinate Exchange Algorithm for Generation of Two-levels Supersaturated designs
Krishan Lal	Linear Trend Free Designs for 2-level Fractional Factorial Experiments

- *denotes the author who presented the paper

➢ Teaching and Research Guidance

A) Teaching

Year	Trimester	Course	Taught Jointly with	Number of Lectures Taken
2006-07	Trimester III	AS 163: Statistical Inference (4L + 1P)	Dr. L.M. Bhar	32 (26L+6P)
2006-07	Trimester III	AS 370: Recent Advances in the Field of Specialization (1L+0P)	Dr. V.K. Gupta	5 (5L+0P)
2007-08	Trimester - I	AS 200: Design of Experiments-II (1L+1P)	Dr. Cini Varghese	19 (11L +8P)

B) Research Guidance (P.G. Students Guided)

- **Chairman Advisory Committee**:

- Ph.D. (Agricultural Statistics): 2 students completed their degree.

1. Sh. B.N. Mandal, Roll No. 9070 completed his Ph.D. degree in Agricultural Statistics on August 09, 2007. The topic of his Ph.D. thesis was **Combinatorics and its Applications with Special Reference to Sample Surveys.**

The salient achievements of the above research investigation are:

Sampling plans for the following two different situations viz. (i) when the population units are ordered in time or space and contiguous or adjacent units provide similar observations and ii) when there is heterogeneity in the population due to more than one factor; one nested within the another have been developed. Sampling plans for the first type of problems have been obtained by making use of (i) the combinatorial properties of incomplete block designs and (ii) linear programming approach by minimization of probabilities of samples containing adjacent units. To deal with problem of sample selection from the populations having two factors (one nested within another) causing heterogeneity, theory of nested stratified sampling has been developed.

Balanced sampling plans excluding adjacent units {BSA (*m*) plans} are useful for sampling from populations in which the nearer units provide similar observations due to natural ordering of the units in time or space. For BSA (*m*) plans, first order inclusion probabilities of all units are same and second order inclusion probabilities for pairs of adjacent units are zero and constant for other pairs of units. An important series of incomplete block designs called polygonal designs are useful for obtaining BSA (*m*) plans. Considering the blocks of polygonal designs as samples and the treatments as units, a BSA (*m*) plan can be obtained by assigning equal probability of selection to the blocks. A computer algorithm has been developed to obtain polygonal designs. Computer aided search of polygonal designs was made in the parametric range of $v \le 40, b \le 400, k \le 7 \; and \; m \le 4$. Computer aided search gave all the existent polygonal designs in the literature along with 75 new designs in this parametric range. Number of new designs obtained for *m* = 1, 2, 3 and 4 respectively is 14, 35, 18 and 8. The algorithm has the limitation that generation of polygonal designs for large *v* and *k* may take prohibitively longer time. It is important to note that all polygonal designs are BSA (*m*) plans but vice versa may not be true. BSA (*m*) plans may be obtained by assigning unequal probabilities of selections to the samples in the support. Hence, the computer algorithm developed for generation of polygonal designs has been modified to obtain BSA (*m*) plans directly comparatively in a short time. The plans generated by the modified algorithm may, however, have larger support sizes compared to the plans generated through polygonal designs. The linear programming has also been exploited to obtain both circular and linear BSA (*m*) plans by

minimizing the probabilities of selection of samples containing adjacent units. This method is quite useful for obtaining smaller sampling plans and takes very less time.

BSA (*m*) plans suffers from the drawback that the unbiased estimation of variance of Horvitz-Thompson estimator of population mean is not possible. To tackle this problem, a family of distance balanced sampling plans (DBSP) with the property that the second order inclusion probabilities are non-decreasing function of distance between the two concerned units is developed. Unbiased estimation of variance of Horvitz-Thompson estimator of population mean for DBSP is considered. The conditions for DBSP to be more efficient than simple random sampling without replacement (SRSWOR) and BSA (*m*) plans have been obtained. DBSP can give rise to a large number of sampling plans depending on choice of the distance function. Three particular members of the family of DBSP namely two points, three points and $\left[\frac{N}{2}\right]$-points DBSP were investigated for suitable choice of distance function. It has been shown that if the assumption of decreasing correlation between the units as the distance between the units increases holds then the proposed plans are more efficient than other alternative sampling plans such as SRSWOR and BSA (*m*) plans. A class of incomplete block designs, called as distance balanced incomplete block (DSBIB) designs is introduced. The blocks of a DSBIB design act as a support of the DBSP. Results on existence and construction of DSBIB designs have also been obtained.

IPPS plans excluding adjacent (IPPSEA plans) units have been developed for sampling from populations where there is natural ordering of the units along with variability in sizes of the units. IPPSEA plans have been obtained by making use of binary, proper and unequireplicated block designs and linear programming approach. The performance of proposed IPPSEA plans using Horvitz-Thompson estimator was compared with other alternative sampling plans such as SRSWOR, BSA (*m*) plans, probability proportional to size with replacement, Hartley and Rao' strategy, Rao, Hartley and Cochran's strategy and Sampfords' IPPS plan using the a real life population. It was seen that the proposed plan performs better incomparison to alternative plans. Unbiased estimation of Horvitz-Thompson estimator of population total is not possible in these types of plans because some of the second order inclusion probabilities are zero. To resolve this problem, one approximate variance estimation technique has been suggested.

Stratified sampling is often used to draw a representative sample from a population. But stratification based on geographical contiguity may yield strata which are heterogeneous due to some other factors affecting characteristic under study. To deal with this, nested stratified sampling in which there are secondary strata within each primary stratum has been developed for the estimation of finite population total and its variance. Optimum allocation of sample size in a particular primary stratum to secondary strata within that primary stratum and optimum allocation of total sample size to primary strata in the population was discussed in detail. A condition for nested stratified sampling with arbitrary allocation to be more efficient than usual stratified sampling with arbitrary allocation was obtained. It was shown that nested stratified sampling with optimum, Neyman and proportional allocation is more efficient than usual stratified sampling with corresponding allocations. Conditions for nested stratified sampling to be more efficient than direct stratification under different allocations were obtained. A real life data set was used to study the performance of the proposed sampling plan. Feasibility of application of doubly nested incomplete block designs to nested stratified sampling was discussed in this context. Application of doubly nested block designs was also shown to obtain designs for tetrallel crosses.

2. Sh. Pradip Kumar Nandi, Roll No. 8871 completed his Ph.D. degree in Agricultural Statistics on December 24, 2007. The topic of his Ph.D. thesis was **Design and Analysis for Multi-response Experiments.**

Experiments in which data on several responses are measured from an experimental unit corresponding to the application of a treatment are known as multi-response experiments. Multi-response experiments are of two types viz. complete multi-response experiments (all response variables are recorded from each experimental unit) and incomplete multi-response experiments (recording of all responses variables from each experimental unit is not feasible). A stepwise procedure of performing multivariate analysis of variance (MANOVA) of data from complete multi-response experiments conducted in block designs has been developed. If the treatments are found to be significantly different through MANOVA, then the experimenter is interested in testing the hypothesis regarding some treatment contrasts, particularly making all the possible pair wise treatment comparisons. To answer this question, the procedure of carrying out multivariate treatment contrast analysis has been developed. A method based on Euclidean distance from null vector and J-plot based on the singular value decomposition (SVD) of the treatment effects matrix (treatment means/ adjusted treatment means for all the response variables) have been recommended for identification of best treatment. In a designed experiment, when one experimental plot is heavily infested with pests, disease and/or weeds, the response observed from this plot would be markedly different from the response from all other plots. This response may be abnormally high or abnormally low. Such responses are termed as outlier(s). To tackle the problem of outlier(s) in multi-response experiments, a test statistic has been developed for detection of a single outlier vector in complete multi-response experiments run in a block design.

Besides the analysis of data, the designing of multi-response experiments is also of paramount importance. It has been shown that designs that are efficient for single response experiments are also efficient for complete multi-response experiments. The only requirement is that the number of response variables should be less than the error degrees of freedom. A method of construction of designs for incomplete multi-response experiments is also obtained using a combination of randomized complete block designs and balanced incomplete block designs or singular group divisible design. The designs obtainable from this method are economical from resource point of view.

The above relates to the discussion on the problems associated with comparative experiments. A large number of experiments are conducted for establishing a relationship between the levels of input factors and several response variables using response surface designs to perform simultaneous optimization of several responses. Simultaneous optimization procedures of several responses for both complete and incomplete multi-response experiments have been considered. Two situations of incomplete multi-response experiments are considered. The procedure of estimation of parameters from linear multi-response models for incomplete multi-response experiments has been developed for both the situations. It has been shown that the parameter estimates are consistent and asymptotically unbiased. Using these parameter estimates, simultaneous optimization of incomplete multi-response experiments is attempted following the generalized distance criterion.

- **Co-Chairman Advisory Committee:**
  Sh. Kaustav Adityam, Roll No. 4493 completed his M.Sc. degree in Agricultural Statistics
- **Member Advisory Committee:** 15 students
  5 Ph.D. (2 in Agricultural Engineering, 1 each in Horticulture, Post Harvest Technology and Seed Technology) and **10** M.Sc. (3 in Computer Application; 2 in Horticulture, 2 in Agronomy; 2 in Agricultural Economics and 1 in Seed Technology) students completed their respective degrees.

➢ Participation/ Presentations in Conferences/ Symposia/ Workshop etc.

- Presented **9** research papers (5 Invited papers by self, 4 invited papers by co-authors)

No.	Name of the Conference/ Sumposia/ Workshop	Organizing Institution, Venue and Duration	Papers Presented
1.	XIV Annual Group Meeting of AICRP-Rapeseed and Mustard	SKUAST, Jammu during August 02-04, 2007	- Statistical Issues in Rapeseed- Mustard Varietal Trials (**Rajender Parsad*** and V.K. Gupta)
2.	Conference in the Honour of Professor Aloke Dey	I.S.I. Delhi Centre during September 05-06, 2007	- Efficient Designs for Microarray Experiments under fixed/ mixed effects model. (**Rajender Parsad***, Ananta Sarkar, V.K. Gupta and Abhishek Rathore) - Some thoughts on fractional factorial plans. (**V.K. Gupta*,** Rajender Parsad and L.M. Bhar)
3.	Workshop of All India Coordinated Research Project on STCR	Indian Institute of Soil Science during September 28-29, 2007	- Design and Analysis of Experiments Under AICRP on STCR(V.K. Gupta, Aloke Lahiri* and **Rajender Parsad**)
4.	10th Annual Conference of Society of Statistics, Computer and Applications	St. Thomas College, Pala during November 16-18, 2007	- Computer aided search of efficient designs for 2-colour microarray experiments. In the workshop on Statistical Computing and Bioinformatics held during the conference. (**Rajender Parsad***, Ananta Sarkar, V.K. Gupta and Abhishek Rathore) - Some considerations on fractional factorial designs. (**V.K. Gupta***, Rajender Parsad and Lalmohan Bhar.) - A robust method for experimental data analysis. (**Lalmohan Bhar***, Rajender Parsad and V.K. Gupta.)
5.	XV National Conference of Agricultural Research Statisticians	Birsa Agricultural University, Kanke, Ranchi during December 03-04, 2007	- Agricultural Statistics Research: Current Status and Future Challenges. **Technical Session** on Priorities in Agricultural Statistics Research: Current Status and Future Challenges (**Rajender Parsad*** and V.K. Gupta.)
6.	IUPAC sponsored First International Conference on Agrochemicals Protecting Crop, Health and Natural Environment	Division of Agricultural Chemicals, I.A.R.I., New Delhi during January 08-11, 2008.	- Significance of Statistical Tools for Agrochemical Research (**Rajender Parsad***, V.K. Gupta, Jitendra Kumar and N.A. Shakil)

(* represents the author who presented the paper)

Besides the above, I also attended the following

- 61st Annual Confernce of Indian Society ofAgricultural Statistics held at Birsa Agricultural University, Kanke, Ranchi during November 30 to December 02, 2007.
- Brain Storming Session to bring out a status report on **State of Indian Agriculture** organized by National Academy of Agricultural Sciences, New Delhi on January 07, 2008.

➢ Special Lectures Delivered

A) **Lectures Delivered in Training Programmes at IASRI, New Delhi**

- 21 Lectures were delivered on the following topics during various training programmes held at I.A.S.R.I., New Delhi.

• Training programme on **Agribusiness and Market Intelligence** (October 06-15, 2007)	**2 Lectures**
- Minitab: An Overview	1 Lecture
- Multivariate Techniques: An Overview	1 Lecture
• Training programme **Biometrics in Agricultural Research** (October 03 – December 01, 2007)	**5 Lectures**
- SPBD	1 Lectures
- SPFE	1 Lecture
- SAS: An Overview	2 Lectures
- Statistical Package on Augmented Designs	1 Lecture
Training programme on Advances in Quantitative Techniques for Policy Analysis in Agricultural Economics CAS **(**December 06-26, 2007)	**4 Lectures**
- SAS: An Overview	2 lectures
- Multivariate Analytical Techniques: An Overview	2 lectures
• Training programme on Research Methodology for Scientific personnel of Indian Council of Forestry Research and Education, Dehradun (January 07 - 18, 2008)	**2 Lectures**
- SAS: An Overview	1 Lecture
- Augmented Designs	1 Lecture
• Winter School on **Sample Survey Techniques in Agricultural Research (**January 16-February 05, 2008)	**1 Lecture**
- **SAS: An Overview**	1 Lecture
• Training programme on **Advances in Biometrical Techniques** (February 08-28, 2008)	**7 Lecture**
- SAS: An Overview	2 Lecture
- Applications of Multivariate Techniques: An Overview	2 Lecture
- Overview of SPFE 1.0 and SPAD	1 Lecture
- Design and Analysis of Microarray Experiments	1 Lecture
- BiPlot	1 Lecture

B) Invited Lectures Delivered

- 13 Invited Lectures are delivered at Department of Social Work, University of Delhi, New Delhi; Human Resource Development Centre, Council of Scientific & Industrial Research, Ghaziabad.

- 2 lectures on **SPSS: an Overview** at Department of Social Work, University of Delhi, Delhi. (August 11, 2007).
- 10 lectures on the topics Statistics: Introduction and Concepts; Statistical Packages: An Overview; Analysis of Variance; Correlation and Regression; Response Surface Methodology and Experiments with mixtures to the participants of the Training Programme on **Research Methodology and Statistical Methods** organized by Human Resource Development Centre, Council of Scientific & Industrial Research, Ghaziabad. (September 17-20, 2007).
- An Invited Talk on **Statistics: Career and Prospects** at Kalindi College, Delhi University, Delhi on November 23, 2007.

➢ Executive Summary

- Experiments in which data on several responses are measured from an experimental unit corresponding to the application of a treatment are known as multi-response experiments. Multi-response experiments are of two types viz. complete multi-response experiments (all the response variables are recorded from each experimental unit) and incomplete multi-response experiments (recording of all the responses variables from each experimental unit is not feasible). For complete multi-response experiments, it has been shown that the designs that are efficient for single response experiments are also efficient for complete multi-response experiments provided that the response variables are less than the error degrees of freedom.
- Obtained a method of construction of designs for incomplete multi-response experiments using combination of randomized complete block (RCB) designs as treatment-wise design and balanced incomplete block (BIB) designs as response-wise design. The designs obtainable from this method are economical from resource point of view.
- Developed a step wise procedure of analysis of incomplete multi-response designs obtained as a combination of RCB design and BIB design.
- Developed β-version of software for generation of nested block designs both for independent errors and correlated errors.
- Developed an algorithm for generation of efficient supersaturated designs for two-level factorial experiments and obtained several efficient super saturated designs. Developed a software for detection and handling of outlier(s) in the experimental data.
- Efficient designs for asymmetric parallel line assays and slope ratio assays have been obtained by deleting all observations corresponding to a dose (of standard or test preparation) from efficient block designs for symmetric parallel line assays and slope ratio assays.
- To disseminate the knowledge available on combinatorial aspects of designs and analytical procedures acquired to scientists engaged in research in National Agricultural Research System, advisory services were pursued rigorously. Design Resources Server has been strengthened in collaboration with National Professor by adding links on α-designs, designs for bioassays, supersaturated designs, modules for generation of randomized layout of square lattice designs, completely randomized designs, RCB designs, Latin square designs and augmented designs. To provide steps of analysis of data generated through designed experiments using SAS and SPSS, a new page "Analysis of Data" has been launched on Design Resources Server. Discussion Board has been initiated for sharing research with fellow scientists over the globe or for flagging issues for attention of scientific community. A list of experts in design of experiments over the globe is uploaded which will be useful for establishing linkages.
- Taught **3** courses to M.Sc. and Ph.D. students and guided **2 Ph.D.** students as **Chairman**; **1** M.Sc. student as **Co-Chairman** and **10** M.Sc. and **5** Ph.D. students as member advisory committee.
- Published **7** research papers and **2** monographs. **5** papers have been accepted for publication and communicated **7** papers. **9** papers are presented in International/National Conferences. Delivered **21** lectures in ad-hoc training programmes organized at IASRI and **13 invited lectures** in other academic organizations.
- Organized **3 dissemination workshops** on (i) Design Resources Server; (ii) Outliers in Designed Experiments and (iii) Fractional Factorials with special emphasis on Experiments with Scarce Resources and **one discussion Seminar** on Experiments with Mixtures: Theoretical Advances and Applications. A meeting with senior breeders of AICRP on Rapeseed and Mustard to finalize the designs for the initial varietal trials and criteria of promotion of entries was also organized.
- As a Guest Editor, brought out proceedings of invited papers presented during the International Conference on Statistics and Informatics in Agricultural Research as a special volume of Journal of Indian Society of Agricultural Statistics.
- Received **NAAS-Associateship** in 14th Annual General Body Meeting of National Academy of Agricultural Sciences.

25/4/08

Signature of Head of the Institute

(Rajender Parsad) 25/04/08

Principal Investigator