Chapter 1

Introduction

1.1 GENERAL INTRODUCTION

The modern concepts of experimental designs are primarily introduced by Ronald A. Fisher in the 1920s and 1930s at Rothamasted Experimental Station, an agricultural research station of London. In Fisher's first book on design of experiments he showed how valid conclusions could be drawn efficiently from experiments with natural fluctuation such as temperature, soil conditions and rainfall, that is, in the presence of nuisance variables. The known nuisance variables usually cause systematic biases in groups of results (e.g. batch-to-batch variables). The unknown nuisance variables usually cause random variability in the results and are called inherent variability or noise.

The experimental design was first used in an agricultural context; the method has been applied successfully in the military and in industry since the 1940s. Besse Day, while working at U.S Naval Experimentation Laboratory, used experimental designs to solve problems such as finding the cause of bad welds at the naval shipyards during World War II. W. Edwards Deming taught statistical methods, including experimental designs, to Japanese scientist and engineers in the early 1950s at a time when "made in Japan" meant poor quality. Genichi Taguchi, the most well known of this group of Japanese scientists is famous for his equality improvement methods who used factorial designs and orthogonal arrays to improve the quality of the product. One of the companies where Taguchi first applied his methods was Toyota. Since the late 1970s, U.S. industry become interested again in quality improvement initiatives, now known as “Total Quality” & “Six sigma” programs. Design of Experiments is considered an advanced method in the six sigma programs, which were pioneered at Motorola & GE.

The fundamental principles in Design of Experiments are the solutions to the problems in experimentation posed by the two types of nuisance factors and serve to improve the efficiency of experiments. Those fundamental principles are

- Randomization
- Replication
- Local control

Based on these fundamental principles we have certain design whose analysis is based on the theory of least squares this gives the best estimates of the treatments effects and was initiated by Fisher (1926) followed by Yates (1936), Bose & Nair (1939) and Rao (1976).

The simplest of all the design is completely randomized designs (CRD) which is applied in the case when the experimental material is homogeneous, CRD is based on two principles i.e. randomization and replication. In this, the treatments are allocated randomly to the experimental units and each treatment can be repeated any number of times. But when the experimental units are heterogeneous, then CRD fails. To overcome this difficulty we have the simplest and the most commonly used block design is a randomized complete block (RCB) design. This is also referred as a Randomized Block designs (RBD).

When the Experimental units are heterogeneous, a part of the variability can be accounted for by grouping the experimental units in such a way that those experimental units within each group are as homogeneous as possible. The treatments are then allotted randomly to the experimental units within each group (or block). This results in an increase in precision of estimates of the treatment contrasts, due to the fact that error variance that is a function of comparisons within blocks is smaller because of homogeneous blocks.

Latin square designs (LSD) are normally used in experiments where it is required to remove the heterogeneity of experimental material in two directions. These designs require that the number of rows and columns are equal to the number of treatments or varieties. LSD is based on three principles randomization, replication and local control. A Latin square is an arrangement of v treatments in v^2 cells arranged in v-rows and v-columns, such that every treatment occurs precisely once in each row and precisely once in each column. The term v is known as the order of the Latin square design. But these two designs RBD and LSD are helpful for the limited number of treatments. However, when the number of treatments in an experiment increases, the blocks of a RBD and the row-columns of LS design become large; and it may not be possible to maintain the homogeneity. In such situations incomplete block designs can be advantageously used as they required smaller block sizes. Since these designs are non-orthogonal and hence their efficiency factor compared to RCB design is less than one under the assumption of equal intra-block variances. But this is offset by the reduced intra-block variance in incomplete block designs. Consequently, the paired comparisons are estimated more precisely. In the class of incomplete block designs, balanced incomplete block (BIB) design, given by Yates (1936), is the simplest one. These designs estimate all possible paired comparisons with same variance and hence are variance balanced.

BIB design is a binary, proper, equireplicate, connected, and balanced design. BIB designs are not available for every parametric combination. Also, even if a BIB design exists for a given number of treatments (*v*) and block size (*k*), it may require too many replications. To overcome this problem, Bose and Nair (1939) introduced a class of binary, equi-replicate and proper designs, which are called partially balanced incomplete block (PBIB) designs with *m*-associate classes. In these designs, the variance of every estimated elementary contrast among treatment effects is not the same. However two-associate class PBIB designs have been extensively studied in the literature. A catalogue of different PBIB(2) designs is available in Clatworthy (1973).

The partially balanced incomplete block designs are available with smaller number of replications for any number of treatments. The number of replication of pair of treatments in these designs is not constant. For defining a PBIB designs, the concept of association scheme on *v* treatments is needed. The original definition of PBIB designs was modified by Nair and Rao (1942).

For $m = 2$, Bose and Shimamoto (1952) classified the known PBIB designs into

(1) Group Divisible (GD),

(2) Simple (S.I),

(3) Triangular (T),

(4) Latin Square Type (L_i) and

(5) Cyclic Designs.

The PBIB designs with two-associate classes are the most important among m-associate class PBIB designs. Designs with more than two classes have also been studied quite extensively in the literature. Vartak (1955) introduced three-associate-class association scheme which is known as a rectangular association scheme. Further properties of the rectangular association scheme was given by Shah (1964) and Vartak (1955, 1959) and Raghavarao (1970).

One of the main aims of genetical investigation involving crops is to develop improved crop varieties using genetic architecture. When there is number of strain of a crop, one type of investigation consists of crossing between such strains to evolve new varieties. Through this type of investigation the combining capacities of the strains can be studied. Normally highly inbred parental lines are considered for such investigation. Various form of diallel crosses as mating designs are used in plant to study the genetic properties and potential of inbred lines or individuals.

Let p denote the number of lines and let a cross between lines i and j be denoted by $(i \times j) = (j \times i)$ $i \neq j = 1, 2, \ldots, p$ and let v be number of crosses. Among the four types of diallel discussed by Griffing (1956), method 4 is the most commonly used diallel in plant breeding. This type of diallel crossing includes the genotypes of one set of F1's, but neither the parents nor reciprocals i.e. $v = p\ (p-1)/2$ crosses.We shall refer to it as a complete diallel cross (CDC). The common practice with CDC is to evaluate the crosses in a completely randomized design or randomized complete block design as the environment designs. Due to limitation of homogeneous experimental units in a block to accommodate all the chosen crosses, the estimates of genetic parameters would not be precise enough if a complete block design was adopted for the large number of crosses. Let there be v lines with ${}^{v}C_2$ crosses tried in an experiment adopting a randomized block design with r replications. There will be ${}^{rv}C_2$ observations in this experiment.

Diallel cross experiments consisting of all possible single crosses among a group of parental lines, are conducted in plant and animal breeding programmers to estimate combining abilities of lines and variance components needed for certain genetical investigations. When the number of lines v is large, the number of crosses in a complete diallel cross experiment may become prohibitively large and in such a case, one might use a "sample" of the crosses only, leading to what is known as partial diallel crosses. (PDC) which is reported (Kempthorne and Curnow, 1961) to have several advantages over complete diallel cross (CDC).

Residual treatment effects designs in which the effect of the treatment continues beyond the period of its application. Residual treatment effects designs have been extensively used for many years in broad spectrum of research areas, including agriculture experiments, dairy husbandry, clinical trials, psychological experiments, medical applications, and industrial settings.

Williams (1945, 1950) introduced balanced residual treatment effects designs. In experiments involving residual effects when a balanced residual treatment effects design is not available, a useful compromise is to use a partially balanced design with respect to direct and residual effects. Blaisdell and Raghavarao (1980, 1985) systematically developed and formally defined the concept of partially balanced changeover designs based on m-associate class PBIB designs [PBCOD(m)] and also obtained the maximum efficiencies for the estimated direct and residual elementary treatment contrasts. Biswas and Raghavarao (1998) constructed partially balanced residual treatment effects designs, where the treatments have rectangular and group divisible association schemes.

In many experimental situations, interest is focused on the simultaneous comparisons of several test treatments to a control treatment rather than on all pairwise comparisons. Pigeon (1984) and Pigeon and Raghavarao (1987) proposed a class of control balanced residual treatment effects designs and investigated their relative efficiencies. We have the definition due to Pigeon and Raghavarao (1987).

Let us consider a residual treatment effects design of first order in which v treatments are compared with a control, which is a standard treatment, using N experimental units and the experiment lasts for p periods and let $D(v+1, N, p)$ denote the class of all such residual treatment effects designs. The linear model for the analysis is

$$y_{ij} = \mu + \alpha_i + \beta_j + \tau_{d(i,j)} + \rho_{d(i-1,j)} + e_{ij}$$

Where y_{ij} $(1 \leq i \leq p;\ 1 \leq j \leq N)$ is the response observed in the ith period on the jth unit, μ the general mean, α_i the ith period effect, β_j the jth unit effect, $\tau_{d(i,j)}$ the direct effect of the treatment $d(i, j)$, $\rho_{d(i-1,j)}$the first-order residual effect of the treatment $d(i, j)$ with $\rho_{d(0, j)} = 0$ for all j, e_{ij} are random errors assumed to be normally and independently distributed with mean zero and variance σ^2. All the parameters are assumed to be fixed.

1.2 BASIC TEMINOLOGY AND DEFINITIONS

1.2.1 Fields:

Let F be a set of elements and let there be two binary operations, denoted by addition (+) and multiplication (.) signs. The system $\mathcal{F} = \langle F,+,\cdot \rangle$ is a field if it satisfies the following axioms:

(1) $\langle F,+,\cdot \rangle$ is an Abelian group whose identity will be denoted by 0 and inverse of an element $x \in F$ will be denoted by $-x$.

(2) $\langle F_0,+,\cdot \rangle$ is an Abelian group, where $F_0 = \{x \in F / x \neq 0\}$. The identity $x \in F_0$ will be denoted by x^{-1}.

(3) The operation multiplication is distributive over addition, i.e. $a\ .\ (b + c) = a.\ b + a\ .\ c$ for all $a, b, c \in F$.

1.2.2 Galois Field:

If the set F has a finite number of elements, the field $\mathcal{F}$ is called a finite field or a Galois field. If $\mathcal{F} = \langle F_0,+,\cdot \rangle$ is a Galois field, then $|F| = p^n$ where p is a prime number and n (≥ 1) is a positive integer.

1.2.3 Incidence Matrix for Binary Designs:

Associated with any design D, an incidence matrix N is defined as $N = (n_{ij})$, $(i = 1, 2,.., v;$ $j = 1, 2,\ldots b)$ where n_{ij} denotes the number of time the ith treatment occurs in the jth block. In case of binary designs,

$$n_{ij} = \begin{cases} 1, & \text{if } i\text{th treatment occurs in the } j\text{th block} \\ 0, & \text{Otherwise} \end{cases}$$

1.2.4 Juxtaposition :

A juxtaposition is a union of some designs, whose incidence matrices are N_h, $h = 1, 2,\ldots, a$, with the common number of treatments. This union is described by the incidence matrix $[N_1 : N_2 :\ldots: N_a]$.

1.2.5 Orthogonal Latin Squares:

If in two Latin squares of the same order, when super imposed on one another, every ordered pair of symbols occurs exactly once, the two Latin squares are said to be orthogonal Latin Squares.

1.2.6 Mutually Orthogonal Latin Square Design:

If in a set of Latin squares every pair of orthogonal, the set is called a set of mutually orthogonal Latin squares (MOLS).

1.2.7 Balanced Incomplete Block Designs:

Balanced incomplete block design is an arrangement of v treatments in b blocks of k ($<v$) plots each, if

1. Each treatment occurs in r blocks,
2. Each pair of treatments occur together in λ blocks,

v, b, r, k and λ are the parameters of BIB design and satisfy the following relations.

(i). $vr = bk$

(ii). $\lambda(v-1) = r(k-1)$

(iii). $b \geq v$

1.2.8 Resolvable BIB Design:

An incomplete block design is said to be resolvable if the blocks can be divided into sets such that each set is a complete replication for a (v, b, r, k) resolvable design. It follows easily that $v = tk$, where t is some positive integer.

1.2.9 Association Scheme:

Given v treatments denoted by 1, 2, 3,…, v a relation satisfying the following conditions is said to be an association scheme with m classes.

(i). Any two treatments α and β are either 1^{st}, 2^{nd}….. m^{th} associate. The relation of association being symmetrical; that is, if the treatment α is the i^{th} associate of the treatment β, then β is the i-th associate of α.

(ii). Each treatment α has n_i i^{th} associates, the number n_i being independent of α.

(iii). If any two treatments α and β are the i^{th} associates, then the number of treatments that are j^{th} associates of α and k^{th} associates of β is p^i_{jk} and is independent of the pair of i^{th} associates α and β.

1.2.10 Partially Balanced Incomplete Block Designs

PBIB design with *m*-associate classes if the *v*-treatments are arranged into b sets of size k (< v) such that

(i). Every treatment occurs at most once in a set.

(ii). Every treatment occurs in exactly r sets.

(iii). If two treatments α and β are *i*th associates, then they occur together in λ_i sets, the number λ_i being independent of the particular pair of *i*th associates α and β. The numbers v, b, r, k, and λ_i (i = 1, 2 … m) are called the parameters of this designs.

The following relations between parameters of the association scheme and of the PBIB designs are known

(i). $vr = bk$

(ii). $\sum_{i=1}^{m} n_i p^i_{jk} = n_i - \delta_{ij}$ where δ_{ij} is the Kronecker's delta, i.e. $\delta_{ij} = 1$ if $i = j$ and is zero otherwise.

(iii). $\sum_{i=1}^{m} n_i = v - 1$

(iv). $\sum_{i=1}^{m} n_i \lambda_i = r(k-1)$

(v). $n_i p_{jk}^{i} = n_j p_{ik}^{j} = n_k p_{ij}^{k}$

1.2.11 Group Divisible Designs

In this category the $v = mn$ treatments are divided into m groups of n each such that any two treatments of the same group are the first associates and two treatments from different groups are the second associates. The association scheme can be displayed by arranging the treatment numbers in a rectangular arrangement of m rows and n columns where each row of n treatments gives a group. Evidently, $n_1 = n - 1$, $n_2 = n(m-1)$. The secondary parameters are

$$p_{jk}^{1} = \begin{pmatrix} n-2 & 0 \\ 0 & n(m-1) \end{pmatrix}$$

$$p_{jk}^{2} = \begin{pmatrix} 0 & n-1 \\ n-1 & n(m-2) \end{pmatrix}$$

Furthermore group divisible designs are classified as:

(1) Singular GD design if $r - \lambda_1 = 0$;

(2) Semi-regular GD design if $r - \lambda_1 > 0$ and $rk - v\lambda_2 = 0$;

(3) Regular GD design if $r - \lambda_1 > 0$, $rk - v\lambda_2 > 0$

1.2.12 Rectangular Design

Let there be $v=mn$ treatments arranged in a rectangle of m rows and n columns. With respect to each treatment, the first associates are the other (n-1) treatments of the same row, the second associates are the other (m-1) treatments of the same column, and the remaining (m-1)(n-1) treatments are third associates. For this association scheme we have $n_1 = n - 1$, $n_2 = m - 1$, $n_3 = (m-1)(n-1)$. The secondary parameters are

$$p_{jk}^1 = \begin{pmatrix} n-2 & 0 & 0 \\ 0 & 0 & m-1 \\ 0 & m-1 & (m-1)(n-2) \end{pmatrix}$$

$$p_{jk}^2 = \begin{pmatrix} 0 & 0 & n-1 \\ 0 & m-2 & 0 \\ n-1 & 0 & (m-2)(n-1) \end{pmatrix}$$

$$p_{jk}^3 = \begin{pmatrix} 0 & 1 & n-2 \\ 1 & 0 & m-2 \\ n-2 & m-2 & (n-2)(m-2) \end{pmatrix}$$

1.2.13 Partially Balanced Residual treatment Effects Designs

A partially balanced residual treatment effects design (PBRTED) based on an m-associate class PBIB design is an arrangement of v treatments in k rows and N columns such that

1. Every treatment occurs at most once in a column.
2. Every treatment occurs t times in each row.
3. Every pair of treatments (θ, Φ) occurs together in μ_i columns if θ and Φ are the i^{th} associates.

4. Deleting the last row of the design, every pair of treatments (θ, Φ) occur together in v_i columns if θ and Φ are the i[th] associates.

5. Every ordered pair of treatments (θ, Φ) occurs together in successive periods in λ_i columns if θ and Φ are the i[th] associates.

6. For every pair of treatments (θ, Φ), the number of columns in which θ occurs with Φ in the last row is the same as the number of columns in which Φ occurs with θ in the last row.

The parameters of a PBRTED are $v, k, b, t, \lambda_i, \mu_i, v_i$ (i= 1, 2, 3…., m) and satisfy the following parametric relations :

(i). $b = vt$

(ii). $\Sigma n_i \mu_i = tk(k-1)$

(iii). $v\Sigma \lambda_i n_i = b(k-1)$

(iv). $\Sigma v_i n_i = t(k-1)(k-2)$

Where, n_i denotes the number of i^{th} associates of any treatment.

1.2.14 Control Balanced residual Treatment Effects Designs

Let $v + 1$ treatments be represented as x, 0, 1, 2,…., $v - 1$ such that the control treatment has a treatment x and the v test treatments have treatments 0, 1, 2,…., $v - 1$. Then arrangement of the $v + 1$ treatments in p periods ($p \leq v + 1$) and N units such that every unit receives a treatment in each period is said to be a control balanced residual treatment effects design (CBRTED) if:

(i). No treatment occurs more than once in an experimental unit.

(ii). The control treatment occurs t_0 times in each period and each test treatment occurs t_1 times in each period.

(iii). The control treatment occurs with each test treatment in λ_0 units and each test treatment occurs with every test treatment in λ_1 units.

(iv). Excluding the last period, the control treatment occurs with each test treatment in μ_0 units and each test treatment occurs with every other test treatment in μ_1 units.

(v). The ordered treatment pairs (x, i) and (i, x) $(i = 0, 1, 2, \ldots., v-1)$ occurs in successive periods in γ_0 units and the ordered treatment pair (i, j) $(i \neq j\ i, j = 0, 1, 2, \ldots., v-1)$ occurs in successive periods in γ_1 units.

(vi). For every pair of distinct treatments, Φ and ψ, the number of units in which Φ occurs with ψ in the last period is the same as the number of units in which ψ occurs with Φ in the last period.

It is desirable that the control treatment be replicated at least as often as each of the test treatments and therefore we will consider only designs for which $t_0 \geq t_1$.

The numbers $v, p, N, t_0, t_1, \lambda_0, \lambda_1, \mu_0, \mu_1, \gamma_0$ and γ_1 are called parameters of the design and they satisfy the following relations:

(i). $N = t_0 + v\, t_1$

(ii). $(p - 1)t_0 \leq vt_1$

(iii). $p(p - 1)t_0 = v\lambda_0$

(iv). $p(p - 1)(vt_1 - t_0) = v(v - 1)\, \lambda_1$

(v). $(p - 1)(p - 2)t_0 = v\, \mu_0$

(vi). $(p - 1)(p - 2)\, (vt_1 - t_0) = v(v - 1)\mu_1$

(vii). $(p - 1)t_0 = v\gamma_0$

(viii). $(p - 1)\, (vt_1 - t_0) = v(v - 1)\, \gamma_1$

1.3 LITERATURE REVIEW

Dey (1986) in corollary 5.3 obtained a method to construct a regular GD design where s is a prime or prime power. They obtained a new incidence matrix $N = \begin{bmatrix} N_1 & N_2 \\ N_2 & N_1 \end{bmatrix}$, where N_1 and N_2 are obtained by hidden trial method. The above design yields a GD design.

Kageyama and Mohan (1985) developed a method to construct GD designs considering a BIB designs as a basic design they define a new incidence matrix based on Kronecker product. Sharma (1994) gave the simple method to construct a semi-regular GD designs with a group size two from a Hadamard matrices.

Kageyama and Sinha (2003) described some new patterned methods of constructing rectangular designs using balanced incomplete block designs and nested balanced incomplete block designs and at the end of the paper based on these methods of construction they gave the table of rectangular designs in the range of $r, k \leq 10$.

Singh *et.al.* (2011) described some methods to construct rectangular PBIB design using incidence matrix of known PBIB design and Hadamard Matrix. They also said if the dual of these rectangular designs generates the semi regular GD designs. They have also constructed rectangular PBIB design using $v = n^2$ treatments arranged in $n \times n$ square array and the dual of these designs will also be the rectangular design only.

Blaisdell and Raghavarao (1980) have developed and define the concept of partially change-over designs based on *m*-associate class PBIB designs. And this definition is the formal mathematical approach to the problem of developing residual effects designs and balanced residual effects designs form a special case of this definition. They also introduced four new series for these designs of which two series are based on rectangular association scheme and the rest two are based on triangular association scheme. Raghavarao and Blaisdell (1985) has determined the maximum efficiencies for the estimated direct and residual elementary treatment contrast of these designs mentioned in Blaisdell and Raghavarao (1980) and the *A*-optimality of these designs is also explored.

Aggarwal and Jha (2006) has obtained some new series of partially balanced residual treatment effects designs using GD designs, the Latin sqaure association scheme with two constraints *i.e.* L_2 association scheme, the rectangular association scheme and two associate cyclic designs. They have also estimated the efficiencies for the estimated direct and residual elementary treatment contrast of these designs.

Pigeon and Raghavarao (1987) introduced control balanced residual treatment effects designs for the situation where one treatment is a control or standard and is compared with the v test treatments, and they have also given methods to construction of control balanced residual treatment effects designs and have also investigated their efficiencies. Aggarwal, Deng and Jha (2004) have developed some new families of control balanced residual treatment effects designs using BIB and unreduced BIB designs. They also showed that they all are Schur-optimal. Some more construction have been developed by Aggarwal and Jha (2009) using BIB designs where treatment be a prime or prime power.

Schellenberg, Van Rees and Vanstone (1977) considered balanced tournament design using BIB designs and other combinatorial structure.

1.4 OVERVIEW OF THESIS

The thesis comprises of six chapters.

First Chapter, is the *introduction* in which we have discussed the concept, review and literature of the block designs.

Second Chapter, is *the construction of group divisible designs.* In this chapter we have proposed the alternative methods for constructing the singular, semi regular and regular group divisible designs. Here, we gave three approaches to obtain GD design. Here in the first approach the group divisible designs carried out Dey (1986) who had chosen N_1 and N_2 by trail method but in this chapter a specified method is considered to construct the GD designs. In the second approach we obtain a GD design from a resolvable BIB design and in the third approach we make use of juxtaposition for obtaining GD design.

Third Chapter is *the construction of three-associate-class partially balanced incomplete block designs*. In this chapter, we have given five approaches to construct three-associate-class PBIB design called as rectangular PBIB designs. Finally at the end of the chapter a table has been provided based on these five approaches and the cases are considered where $r, k \leq 20$.

Fourth Chapter is *the construction of partially balanced residual treatment effects designs*. In this chapter some new series of partially balanced residual treatment effects designs have been constructed using group divisible (GD) designs, the mutually orthogonal Latin square designs, group and set of blocks constituting replications based on group divisible designs.

Fifth Chapter is *the construction of partially balanced residual treatment effects designs for comparing test treatment with a control*. In literature the work is only done on the residual treatment effects designs for comparing test treatment with a control using balanced designs but we in this chapter made the same comparison for the partially balanced designs. The concept of this chapter is totally new. So we gave the definition of control partially balanced residual treatment effects designs (CPBRTED) and the construction.

Six Chapter is the final chapter which comprises of the application of *complete diallel crosses in the tournament designs*. In the end, a detailed bibliography on the literature of design of experiments is also appended.

So, the main objective of the thesis is to give some simple and easy approaches to construct PBIB designs which are available, some new methods to construct residual treatment effects designs, the application of complete diallel crosses and also to introduce the some new but quite efficient plans so that the new scope of PBIB is enhanced.