V K Gupta
Rajender Parsad
Lal Mohan Bhar
Baidya Nath Mandal

Statistical Analysis of Agricultural Experiments

Part - I: Single Factor Experiments

Statistical Analysis
of Agricultural Experiments

Important Web Resources to Visit

Design Resources Server
www.iasri.res.in/design

Sample Survey Resources Server
http://sample.iasri.res.in/ssrs/

Indian NARS Statistical Computing Portal
http://stat.iasri.res.in/sscnarsportal

Statistical Analysis of Agricultural Experiments

Part - I : Single Factor Experiments

V K Gupta
Rajender Parsad
Lal Mohan Bhar
Baidya Nath Mandal

Preface

Design of Experiments is an integral component of every scientific endeavour, particularly the agricultural sciences where the designing of an experiment is an inevitable component of research. This is due to the fact that advancement of any scientific discipline happens through new knowledge and technology developments and indeed, a scientifically designed experiment is a valuable tool in this process. The doyen of agricultural sciences Professor M.S. Swaminathan rightly said, "It is the effective use of the tools of statistical design of experiments that paved the way for the green revolution."

A carefully designed experiment is able to answer all the queries of a researcher with accuracy and reliability with efficient use of available resources of the experimenters. Thus, for successful experimentation, it is highly desirable that scientists and researchers of scientific disciplines, including agricultural sciences, understand the basic principles of designing an experiment and analysis of resultant data from the completed experiment. It may be emphasized that a researcher should always consult a statistician before, during and after experimentation, if he is not convinced enough about using a design for his experiment or an analysis technique for this data. In any designed experiments, there are certain factors which are beyond the control of the experimenter. In order to reduce the influence of these factors, the experiment has to be suitably planned and executed scientifically with immense care. In this regard, interactions with a statistician before the experiment begins is expected to be highly beneficial to the experimenter because he can devise a suitable design keeping in view the questions the experimenter has in mind, the factors he wants to consider and the amount of available resources he can afford.

Though, a number of statistical books on designing and analysis of experiments are available, most of these books are intended for statisticians because the subject is treated through rigorous algebraic theories. Researchers of applied scientific disciplines are often not interested in the mathematical derivation of the statistical formulae. They would rather be happy with basic ideas behind using a design for a particular experimental situation and subsequently analysis steps with necessary precautions while doing the analysis.

The authors have conducted a number of training programmes on designing and analysis of experiments for the research scientists of Indian National Agricultural Research and Education System. Through the interaction with the trainee scientists, the authors felt that there is a strong need for preparing a book on design and analysis of experiments covering basic concepts, enough examples of situations for

use of various designs, and easy to follow analysis steps. This book is an attempt in that direction. In this book, the subject of design and analysis of experiments has been treated in simple language giving basic concepts of various designs and essential steps of analysis of data from designed experiments. Moreover, examples of analysis of data have been provided for each of the designs covered in the book. The additional strength of the book is that for each of the designs covered in this book, SAS and R codes for analysis have been provided for performing the analysis. It is hoped that this will enable the readers to directly analyze their data from their experiments.

Since the topic of design and analysis of experiments is vast, the book has been divided into two parts; first part covering fundamentals of design and analysis of experiments related to, by and large, single factor experiments and the second part focusing mostly on multifactor experiments and advanced topics. The composition of the first part of the book is as follows. It starts with an introductory Chapter on design of experiments which gives an introduction to the terminologies and basic concepts of experimental designs along with a brief history of the subject. No claim is being made for this Chapter to be completely exhaustive. In Chapter 2, introduction is made to some most commonly used designs namely completely randomized designs, randomized complete block designs and Latin square designs. Chapter 3 discusses contrast analysis, a very useful tool for answering many questions of the experimenters through designed experiments. Chapter 4 deals with analysis of covariance, a useful technique for controlling variability when one or more covariates linearly related to response variable are available. Chapters 5, 6 and 7 are devoted to study of various important classes of block designs. The topics such as outliers in experimental designs, groups of experiments are studied in the subsequent Chapters. One Chapter has been devoted to multivariate analysis of variance for simultaneously analyzing multiple response variables. The book has four Annexures. Annexure I gives introduction to SAS and Annexure II introduces the reader to the R software, currently trending open source statistical software. Annexure III is devoted to multiple comparisons while Annexure IV describes web resource "Design Resources Server." The readers would find Annexure I and II very useful in having a good understanding of the codes that have been used throughout the book. Annexure III and IV will help in a better understanding of the contents of the book. Part I of the book ends with a bibliography on the topics covered in this book. Although an attempt has been made to give an elaborate bibliography, no claim is made to this being exhaustive.

Help has been received from many places and persons in preparing the manuscript of this book. The ideas received from our teachers Dr. Arun Nigam and Dr. Aloke Dey have been very helpful in giving a shape to this book. In fact the authors are indebted to both Dr. Nigam and Dr. Dey for introducing us to this very important area of research in Statistics which is intertwined with agricultural sciences and helps improving the quality of agricultural research by making use

of appropriate and sophisticated designs of experiments and analysis of data. We convey our gratitude towards Dr. Bikas Sinha and Dr. Rahul Mukerjee who have always motivated us in this endeavour. They have always been a deep source of inspiration for us.

A special mention may be made about the invaluable help and support received from Mrs. Jyoti Gangwani, a technical officer working in the National Professor scheme who has very meticulously checked the results of all the experimental designs given in this book by reanalyzing the data from every single designed experiment reported in this book. She has also helped with the editing of the manuscript and then reading very carefully the galley proofs. Her contribution in the development of the manuscript of this book has been monumental in many ways. The authors would like to thank her whole heartedly for being always present with a helping hand, whenever asked for.

The authors are thankful to the Director, IASRI for his encouragement throughout the preparation of this manuscript. Thanks are also due to Director, IASRI for providing all facilities required for handling the project of writing this book.

The authors, in particular the first author, would like to place on record the strength received from the Education Division of the ICAR for providing the financial support in the National Professor Scheme for undertaking this assignment of writing this book and getting this published. The authors are beholden to the Deputy Director General (Education), ICAR for his constant encouragement throughout the preparation of this manuscript. The funds received from the Division of Education, ICAR, for publication of scientific manuscripts is highly appreciated and immensely acknowledged.

The authors fervently expect that the book would be enormously useful to the agricultural scientists in planning and designing their experiments and analyzing the data generated from their experiments. The authors would look forward to getting suggestions from all around to help improve the contents of the next edition of the book so that the readers can benefit from the contents. The errors and omissions are likely to be present in this book. The authors would exceedingly feel grateful if these are brought to their notice.

V K Gupta
Rajender Parsad
Lal Mohan Bhar
Baidya Nath Mandal

Place: New Delhi

Date: February, 2016

Contents

1

Introduction to Design of Experiments

1.1 Introduction

It may be emphasized in the beginning itself that experimental design is first about agriculture, animal science, biology, chemistry, industry, education, etc. and then about Statistics and Mathematics. In fact, experimental design forms the backbone of agricultural sciences; it is an integral component of every research endeavour in agricultural sciences. To design a good experiment the researcher first needs to outline questions to be answered or needs one or more well defined hypotheses. Some examples of typical questions or hypotheses are

(i)   How does the feed formulation affect the body weight of animals?

(ii)   Which variety of crop species would be good for particular region?

(iii)   Does the date of sowing affect the crop yield?

(iv)   How does the water availability and its quality influence the crop yield?

(v)   How does greenhouse gases emission influence the global warming?

(vi)   Does the use of pesticide in crops affect the health of farmers as well as the people consuming the produce?

(vii)   Do the micronutrients and minerals influence the productivity of crops?

(viii) Are the resource conservation technologies counterproductive?

(ix)   How do altering manure management strategies at livestock operations or animal feeding practices control the methane emission?

It is hard to define a design of experiment, because it is a form of art along with the science. It may be borne in mind that no experiment could be the ultimate one. A good experiment would be one that allows testing what the researcher wants to test and exercises control over everything else. In that sense, a good experiment is one that estimates the effects that the researcher is interested in and simultaneously minimizes controls or eliminates confounding factor(s). A confounding factor is also at times called the nuisance factor. It potentially distorts the data. This factor is sitting hidden in a model and affects the variable being studied, but is not known or acknowledged.

An example would be a study of nutrients like nitrogen, phosphorous, potash and sulphur on the yield of wheat. If the minerals like zinc and manganese in the soil are likely to be present along with the nutrients, and the study measures only nutrients but not the minerals, the study may find that the nutrients do affect the yield of wheat which may or may not be true. The presence of minerals in the soil might also be affecting the yield. If this confounding factor is

identified early enough, adjustments can be made so that the confounding does not destroy the results or introduces bias in results.

Another example could be a study of feed formulation on body weight of animals. If the initial body weights of the animals are likely to be markedly different and the study measures only the periodic body weights of the animals after giving them feeds, the study may find that the feeds do affect the body weight of animals. But this may or may not be true. The initial body weights of animals might also be affecting the final bodyweight of the animals. If the experiment does not take care of this confounding factor, it may influence the results.

In planning any experiment, the experimenter needs to decide

(a)   What conditions to study or what are the treatments, e.g., feed formulation, nutrients, irrigations, pesticides, varieties of a crop, resource conservation technologies, dates of sowing, etc.?

(b)   What is the experimental material on which the experiment is to be conducted, e.g., animals, human beings, plots in a field, pots in a glasshouse, birds in a pen, tissues in a laboratory, trees, branches of a tree, leaf position on a tree, etc.? To be more specific, experimental material is actually a collection of subjects or units, or plots, etc. and is termed as experimental units or simply units.

(c)   What measurements to make or what are the responses and how to measure these accurately and correctly, e.g., yield of crop, body weight of animal, milk yield, number of eggs layed, percentage of plants infected by disease, etc.? Response also denotes the measurable outcome as a result of application of treatments on the experimental units.

In any planned experiment, there are four major sources of variability. These are

(a)   Variability due to the conditions under study or the treatments. This variability is desirable and is in fact a deliberate attempt of the researcher to create this variability.

(b)   Variability in the experimental units. This variability is unwanted and undesirable but needs to be accounted for. Generally this variability is overlooked by the researchers.

(c)   Variability in the measurement process or measuring the response. This part of the variability is unwanted and undesirable. We shall assume throughout that this variability is not present.

(d)   Variability absolutely unaccounted for, unwanted and undesirable. The reason for this part of the variability is unknown to the experimenter.

Since it will be assumed that the variability in the measurement process or measuring the response is absent and the response obtained is true, accurate and correct response of the treatment applied, there will be in fact three major sources of variability to reckon with, *viz.*, (a), (b) and (d).

Looking at the requirements of planning an experiment and the various sources of variability in the planned experiment, the thinking with respect to subject matter (agricultural, biological, industrial etc.) and statistical thinking is needed to reach for a good experimental design. In order to give a concrete form to this thinking, a strong interaction between the researcher and the statistician is absolutely essential for planning and executing an experiment. We begin with three important principles of a designed experiment.

1.2 Principles of design of experiments

There are three basic principles of designing an experiment namely *randomization, replication and local control (blocking)*. These techniques are discussed briefly in the sequel.

1.2.1 Randomization

Randomization means random assignment of conditions to study or treatments to the subjects or experimental material (in fact experimental units), without an obvious plan, prior to start of the experiment. Randomization converts unplanned, systematic variability into planned, chance-like variability. An analytical reason in support of randomization is that essentially it ensures observations generated to be independent and hence the statistical tools used for analysis of observations gathered become applicable. This is more important for the use of test statistic like Snedecor's *F* and Student's *t* in hypothesis testing, wherein a pre-condition is that the observations are independent and are identically distributed as normal variate. This is the major concern of randomization.

Randomization also serves the following purposes:

A random assignment of conditions to study or treatments to experimental units ensures that no experimental unit or no treatment received any favour in the beginning of the experiment. Randomization prevents systematic and subjective biases from being introduced into the experiment by the experimenter. In other words, randomization controls the experimenter bias. Lack of a random assignment of experimental units or subjects leaves the experimental procedure open to experimenter bias. It ensures that subjects or experimental units that are favoured or are adversely affected by unknown sources of variation are those "selected using chance device or random permutation" and not systematically selected. For example, in an initial varietal trial of a crop improvement programme, a breeder may assign his or her new strain of experimental crop to the parts of the field that look the most fertile to promote his or her strain; or a nutritionist may assign newly developed feed formulation to healthy and well growing animals to promote a favourite feed. The preferred variety or formulation may then appear to give better results no matter how good or bad it actually is.

Lack of random assignment can also leave the procedure open to systematic biases. Presence of systematic errors in an experiment makes the comparisons among treatments biased, no matter how precise measurements are or how many experimental units are used.

Consider an experiment involving response of four feed formulations to influence the growth in terms of body weight of animals. Suppose that the four feeds, *viz.*, A, B, C, D are given

to 12 animals. Each feed is observed on 3 animals. Without randomization experimenter would take 3 observations on feed one administered to three animals; then on feed two; then on feed three; and then on feed four, i.e., the order of the feeds given to 12 animals are A, A, A, B, B, B, C, C, C, D, D, D. This order might be perfectly satisfactory but could equally well prove to be disastrous. It may be possible that the first three observations on feed A arise from animals that have no disease, the next three observations from feed B arise from animals that acquired foot and mouth disease recently, while the next three observations from feed C arise from animals that are suffering from bloat and the last three observations on feed D arise from animals that are suffering from mastitis. Obviously, the response to feed A would be more pronounced than that to feeds B, C or D, whereas in fact feed A may actually not be better than any of these feeds. Similarly the response to feed C may be more pronounced than that of feeds B or D. It is quite likely that the experimental conditions might favour a particular feed. Order A, B, C, D, A, B, C, D, A, B, C, D, A, B, C, D might help to solve the problem, but it does not eliminate it completely.

Consider this experiment to study the influence of four feed compositions on growth of animals. The total number of ways in which 12 animals can be assigned to 4 feeds so that 3 animals are assigned to each feed is

$$\frac{12!}{3!3!3!3!} = 369,600.$$

A random assignment can lead to order A, A, A, B, B, B, C, C, C, D, D, D or A, B, C, D, A, B, C, D, A, B, C, D, A, B, C, D with probability 1/369,600. The probability is indeed infinitesimal, almost zero. Even though such arrangements can happen in a proper randomization, but to avoid such a thing happening purposefully, one must resort to a proper randomization.

Having said this, randomization has at times its limitations also because randomized experiments violate ethical standards and so cannot be adopted in practice in some situations. To make the exposition clear, an example is considered from clinical trials on human beings.

Suppose that a researcher wants to investigate the abortion–breast cancer hypothesis, which postulates a causal link between induced abortion and the incidence of breast cancer. A hypothetical controlled experiment starts with subjects (pregnant women) and divides them randomly into treatment group (receiving induced abortions) and control group (bearing children). Regular cancer screenings are conducted for women from both groups. Such an experiment would always run counter to common ethical principles. It would also suffer from various confounds and sources of bias, *e.g.*, it would be impossible to conduct it as a blind experiment.

The published studies investigating the abortion–breast cancer hypothesis generally start with a group of women who already have received abortions. Membership in this "treated" group is not controlled by the investigator: the group is formed after the "treatment" has been assigned.

Consider another study in which a researcher wants to compare some phenotypic traits among animals of three different breeds. In this case the experimenter starts with a number of animals, some animals belong to breed one, some belong to breed two and rest of them

belong to breed three. The researcher collects observations on the phenotypic traits of interest. In this experiment, it is not possible to assign breed to an animal at random because the animals already are of a particular breed. In other words, the assignment of breeds to animals is not under the control of the experimenter.

In view of this, design of experiments or experimental design is the design of all information-gathering exercises where variation is present, whether under the full control of the experimenter or not. The latter situation is usually called an observational study, and would be beyond the scope of this book. We shall, henceforth, focus on randomized experiments.

1.2.2 Replication

Replication is the repetition of the conditions of study or treatments under investigation to different experimental units, be it animals or pots or plots in a field, or position of leaf on a plant. Replication intends to increase the size of the experiment.

Replication enables the experimenter to obtain a valid estimate of the experimental error. Estimate of experimental error permits statistical inference; for example, performing tests of significance or obtaining confidence interval, etc. If there is no replication, then the researcher would not be able to estimate the experimental error. And as will be seen in the later Chapters, it is against this estimated experimental error the null hypotheses are tested.

Consider an example where two levels of Nitrogen as A = 30 kg/ha and B = 60 kg/ha are applied to wheat crop. The interest of study is to see how nitrogen influences the yield of wheat. In experiment 1, there are four plots available and each level of nitrogen is applied to two plots randomly. The plots receiving the same level of nitrogen are expected to give the same response. The difference gives the experimental error. In experiment 2, there are six plots and each level of nitrogen is applied to three plots randomly. The yield in kg per plot is given in bracket.

| Experiment 1 | A (31.5) | B (30.6) | B (28.2) | A (32.8) | | |
| Experiment 2 | B (26.8) | A (31.7) | A (33.4) | B (28.6) | A (32.9) | B (27.9) |

In experiment 1, the experimental error can be estimated as

$$\frac{(31.5-32.8)^2 + (30.6-28.2)^2}{2} = \frac{7.45}{2} = 3.725 \quad .$$

This can also be estimated as

$$\left(31.50-32.15\right)^2 + (32.80-32.15)^2 + (30.60-29.4)^2 + (28.20-29.40)^2$$
$$= 0.4225 + 0.4225 + 1.4400 + 1.4400 = 3.725$$

Here the average yield from A is $\dfrac{31.5+32.8}{2} = 32.15$ and the average yield from B is

$$\frac{30.6 + 28.2}{2} = 29.40 \ .$$

In this experiment, the experimental error is 3.725.

In experiment 2, the experimental error can be estimated as

$$\frac{(31.7 - 33.4)^2 + (31.7 - 32.9)^2 + (33.4 - 32.9)^2 + (26.8 - 28.6)^2 + (26.8 - 27.9)^2 + (28.6 - 27.9)^2}{3}$$

$$= \frac{9.52}{3} = 3.173$$

This can also be estimated as

$$(31.700 - 32.667)^2 + (33.400 - 32.667)^2 + (32.900 - 32.667)^2 + (26.800 - 27.767)^2$$
$$+ (28.600 - 27.767)^2 + (27.900 - 27.767)^2 = 3.173$$

Here the average yield from A is $\dfrac{31.7 + 33.4 + 32.9}{3} = 32.667$ and the average yield from B is

$\dfrac{26.8 + 28.6 + 27.9}{3} = 27.767$

In this experiment, the experimental error is 3.173.

Increasing the size of the experiment or increasing the replication also helps to increase the precision of estimating the pairwise differences among the treatment effects. This is so because with the increase in the size of the experiment, the experimental error reduces. As can be seen from the example above, the experimental error in experiment 2 reduces from that in experiment 1.

It may be emphasized here that replication is different from repeated measurements. Suppose that the four animals are each assigned to a feed and a measurement is taken on each animal. The result is four independent observations on the feed. This is replication. On the other hand, if one animal is assigned to a feed and then measurements are taken four times on that animal, the measurements are not independent. We call them repeated measurements. The variation recorded in repeated measurements taken at the same time reflects the variation in the measurement process, while variation recorded in repeated measurements taken over a time interval reflects the variation in the single animal's responses to the feed over time. Neither reflects the variation in independent animal's responses to feed. We need to know about the latter variation in order to generalize any conclusion about the feed so that it is relevant to all similar animals.

Generally speaking, all the treatments should be replicated same number of times. In that case the total number of experimental units is a scalar multiple of the number of treatments. In case the total number of experimental units is not a scalar multiple of the number of treatments, then the replication of treatments should be as equal as possible. In other words, the replications of treatments should not differ by more than one. For instance, if the number of treatments

is 7 and the number of experimental units is 24, then 4 treatments may be replicated three times and three treatments may be replicated 4 times. But there might occur some experimental situations where some treatments may need to be replicated more number of times than the other treatments and the difference in replications is more than one. In fact, there do occur experimental situations where some treatments are not replicated because not enough material is available for replication. There are other experimental situations where even a single complete replication is not experimented because of the resource constraint and economy. There will be occasions to refer to all such situations later in the book (both parts I and II).

1.2.3 Local control or blocking

Experimental conditions under which an experiment is run should be representative of those to which the conclusions of the experiment are to be applied. For inferences to be broad in scope, experimental conditions should be rather varied. Unfortunate consequence of increasing scope of experiment is an increase in variability of response. Blocking is a technique that is often used to help deal with this problem

As mentioned earlier, one source of variability is the experimental material or experimental units. Local control or blocking is a technique to account for the variability in response because of the variability in the experimental units. To block an experiment is to divide the experimental units into groups or blocks of similar units in such a way that the observations in each block are collected under relatively similar experimental conditions. If blocking is done well, the comparisons of two or more treatments are made with more precision than similar comparisons from an unblocked design.

It may be mentioned that the blocking is advantageous if the variability within the groups or blocks is as small as possible and between groups or blocks is as large as possible.

In feeding trials litters of the same animal can form natural blocks. Similarly, animals with similar body weights can also form blocks; animals with genetic similarity can also form blocks; animals with same age can also be a criterion for forming the blocks; animals with same lactation number or stage can be another consideration for forming blocks. Fertility gradient in field experiments can be a way of forming blocks. In this case, the blocks are formed perpendicular to the fertility gradient. Salinity levels in the field could also be a criterion for forming blocks in field experiments. Age of the trees in horticultural experiments could be a source of variability and trees of same age can form natural blocks. The soil depth may be another criterion of blocking. In hilly areas, terraces may be taken as natural blocks.

From practical considerations, the contiguous experimental units should form blocks. But sometimes it may so happen that the homogeneous experimental units may not be contiguous. In that case blocks formed are irregular in shape. It is indeed possible that the blocks may not have same number of experimental units. If we force the blocks to be of same size, then again variability may creep in and the purpose of blocking is defeated.

There can be more than one source of variability in the experimental material. If there are two sources of variability in the experimental units, then recourse is made to forming blocks in two directions, called rows and columns. The conditions under study or the treatments are applied to the cells at the intersection of rows and columns. Row-column designs are also useful for the situations, wherein, the fertility gradient is along the diagonal in the field. Sometimes, the two blocking systems may be nested one within another. There may be larger blocks and within each larger block there are smaller blocks, called sub-blocks. The treatments are applied to the sub-blocks within larger blocks. There may be another type of experimental situation where within the larger blocks, rows and columns are formed. The treatments are applied to the cells within each larger block.

When there is blocking, then the randomization of conditions to study or treatments to experimental units changes. The exact randomization will be described in the respective chapters.

1.3 Brief history of design of experiments

The statistical principles underlying design of experiments were pioneered by R. A. Fisher in the 1920s and 1930s at Rothamsted Experimental Station, an agricultural research station around forty kilometres north of London. Fisher had shown the way on how to draw valid conclusions from field experiments where nuisance variables such as temperature, soil conditions, and rainfall are present. He had shown that the known nuisance variables usually cause systematic biases in results of experiments and the unknown nuisance variables usually cause random variability in the results and are called inherent variability or noise. He introduced the concept of analysis of variance (ANOVA) for partitioning the variation present in data (a) due to attributable factors, and (b) due to chance factors. The methodologies he and his colleague Frank Yates developed are now widely used. Their methodologies have a profound impact on agricultural sciences research.

Though the experimental design was initially introduced in an agricultural context, the method has been applied successfully in the industry since the 1940s. George Box and his co-workers developed experimental design procedures for optimizing chemical processes, particularly response surface designs for chemical and process industries. W. Edwards Deming taught experimental designs to Japanese scientists and engineers in the early 1950s at a time when Japanese products were considered to be of poor quality. Genichi Taguchi, a Japanese engineer, suggested a number of techniques using orthogonal arrays. Taguchi coined the concept of robust parameter design and process robustness. Around 1990, Six Sigma, a new way of representing continuous quality improvement came into existence. Six sigma employs a technique that uses statistics to make decisions based on quality and feedback loops and is widely used by many large manufacturing companies. Design of experiments is considered an advanced method in the Six Sigma programs.

Recently, experimental designs are also being used in clinical trials. This evolved in the 1960s when medical advances were previously based on unreliable data. For example, doctors used to examine a few patients and publish papers based on such data. The biases resulting

from these kinds of studies became known. This led to a move toward making the randomized double-blind clinical trial the standard for approval of any new product, medical device, or procedure. The scientific application of the valid designing and analysis following proper statistical methods became very important in clinical trials.

More recently the experimental design techniques have started gaining popularity in the area of computer-aided design and engineering using computer/simulation models including applications in manufacturing industries.

1.4   Some preliminaries

In the context of design of experiments, some widely used terminologies including those discussed earlier are now defined in the sequence.

The term conditions to study or *treatments* is used to denote the different objects, methods or processes among which comparison is made. Some examples of treatments are different kinds of fertilizer in agronomic experiments, different irrigation methods or levels of irrigation, different fungicides in pest management experiments and doses of different drugs or chemicals in laboratory experiments, different varieties of crops, different pesticides, grazing systems for animals, different tree species in agro-forestry experiments, different concentrations of a solute in chemical experiments, etc.

A *control* treatment is a standard treatment that is used as a baseline or basis of comparison for the other treatments. This control treatment might be the treatment which is currently in use, or it might be a no treatment at all. For example, a study of new pesticides could use a standard pesticide as a control treatment, or an experiment involving fertilizers may have one treatment as no fertilizers at all. In clinical trials, a control treatment is generally a placebo.

Experimental units are the subjects or objects on which the treatments are applied. For example, plots of land receiving fertilizer, groups of animals receiving different feeds, or batches of chemicals receiving different temperatures, pots in glasshouse experiments, Petri dishes or tissues to culture bacteria or micro-organisms in laboratory experiments, etc.

Responses are measurable outcomes, which are observed after applying a treatment to an experimental unit. Alternatively, the response is what we measure to find out what happened in the experiment. In an experiment, there may be more than one response. Some examples of responses are grain yield or straw yield, nitrogen content in plants or biomass of plants, quality parameters of the produce, percentage of plants infested by disease, weight gain by animals, etc.

Factors are the variables whose influence on a response variable is being studied in the experiment. If only one factor is being studied in an experiment then such an experiment is called a single factor experiment. If more than one factor is being studied simultaneously in an experiment, then such an experiment is called multi-factor or factorial experiment. The term factor is commonly used in the case of factorial experiments. For example, temperature and concentration of chemicals in a chemical experiment are two factors, Nitrogen, Phosphorus and Potassium fertilizers are three factors in an agronomic experiment, dose and time of application of a chemical formulation are two factors in a laboratory experiment.

The term *factor levels* or a simply *levels* is used to denote the values or settings that a factor takes in a factorial experiment. For example, doses of a nitrogenous fertilizer as 0 kg/ha, 30 kg/ha, 80 kg/ha are three levels of the fertilizer, 10°C, 20°C, 30°C are three levels of temperatures in a chemical experiment, 10%, 20%, 30%, 40% concentration of a solute in a solution are four levels in a laboratory experiment, presence of polythene sheet on the surface of soil or its absence could be two levels of a practice in water management study.

Treatment combination or *level combination*: In factorial experiments, the set of values for all factors in a trial is called treatment combination or level combination. For example, if in a chemical experiment, there are two factors *viz.*, temperature and concentration and both these factors have three levels each as 10°C, 20°C, 30°C and 10%, 20%, 30%, respectively, then total number of treatment combinations is 3 × 3 = 9 and these 9 combinations are (10°C, 10%); (10°C, 20%); (10°C, 30%); (20°C, 10%); (20°C, 20%); (20°C, 30%); (30°C, 10%); (30°C, 20%); (30°C, 30%). These combinations are in fact 9 treatments. We can label the 9 treatments as 1, 2, 3, 4, 5, 6, 7, 8, 9. The association is the following: 1 ~ (10°C 10%); 2 ~ (10°C, 20%); 3 ~ (10°C, 30%); 4 ~ (20°C, 10%); 5 ~ (20°C, 20%); 6 ~ (20°C, 30%); 7 ~ (30°C, 10%); 8 ~ (30°C, 20%); 9 ~ (30°C, 30%).

Conversely, if there are 9 treatments and these 9 treatments can be thought of as combination of levels of two factors, both having 3 levels each, then the same association can be used to convert the treatments into treatment combinations.

Application of a treatment combination to an experimental unit is called a *run* or a *design point* in factorial experiments.

An *observational unit* is a unit on which the response variables are measured. Observational units are often the same as experimental units, but this may not be true always. The mistake of confusing observational unit with experimental unit leads to pseudo-replication as discussed in a paper by Hurlbert (1984). For example, consider an experiment to investigate the effects of ultraviolet (UV) levels on the growth of smolt. The experiment is conducted in two tanks where one tank receives high levels of UV light and the other tank receives no UV light. Fish are placed in each tank and at the end of the experiment growths of the individual fish are measured. In this experiment, the tanks are the experimental units but the observational units are the smolts. The treatments, presence and absence of UV light, are applied to the tanks and not to individual fish but a whole group of fish are simultaneously exposed to the UV radiation. Here any tank effect is completely confounded with the treatment effect and cannot be separated. Another example is that inorganic fertilizers are applied to plots in a field containing some plants. At the time of harvest, all the plants in the plot are not harvested. Only a sample of plants is harvested. In this case once again the plot is the experimental unit to which fertilizers are applied but the observational units are the plants sampled.

A *treatment contrast* or simply a *contrast* is a linear function of treatment effects such that the sum of the coefficients is zero. For instance, if $\tau_1, \tau_2, \cdots, \tau_v$ denote the v treatment effects, then $p_1\tau_1 + p_2\tau_2 + \cdots + p_v\tau_v$ is a contrast if and only if $p_1 + p_2 + \cdots + p_v = 0$. A big advantage of contrast is that one can make all the possible pairwise treatment comparisons. It also enables to make any

other comparison among treatment effects. For details the reader may see Chapter 3. A contrast is said to be elementary contrast if and only if only two of the coefficients are non-zero while all other coefficients are zero. $\tau_1 - \tau_2$, $\tau_i - \tau_l$, etc. are elementary contrasts. Other contrasts could be $2\tau_i - \tau_l - \tau_u$ or $\tau_i - \tau_l - \tau_u + \tau_w$.

1.5 Factorial experiment

There has been a description of treatments in Section 1.4. There has also been a description of factors and treatment combinations. It may be emphasised that the treatments in any experiment may either be unstructured or structured. Unstructured treatments are actually levels of a single factor in a single factor experiment. In these experiments, the interest is in making all the possible pairwise treatment comparisons. At times comparisons between subsets of treatments or among treatments within subgroups also form a part of the hypotheses to be tested. These experiments are generally conducted as unblocked design, block design or a row-column design or a nested design depending upon the problem to be solved and the nature of the experimental material.

On the other hand, the treatments may be structured in the sense that there are several factors and each factor has several levels. The treatments in this case are the level combinations of all the factors. The interest of the researcher is in estimating the factorial effects comprising of main effects and the interaction effects rather than making all the possible pairwise treatment comparisons or subgroups testing. The treatment sum of squares in this case is partitioned into main effects and interaction effects sum of squares. Otherwise, the experiment once again is conducted using an unblocked design, a block design or a row column design or a nested design as one would have used in case of unstructured treatments. There are no special designs for running factorial experiments. However, treatment structure or their fraction may be obtained based on availability of resources and objectives of the experiment. An incomplete block design in factorial experiment may be obtained in such a way that the desired factorial effects are estimated with more precision by sacrificing information on factorial effects of less interest, particularly the higher order interactions.

If there are several factors it is always advantageous to study them simultaneously rather than studying them separately. Suppose that there are two factors A and B, A having three levels and B having four levels. Let the levels of the two factors be denoted by 0, 1, 2 and 0, 1, 2, 3, respectively. The association between the 12 treatment combinations and the treatments is the following:

1 ~ (0, 0); 2 ~ (0, 1); 3 ~ (0, 2); 4 ~ (0, 3); 5 ~ (1, 0); 6 ~ (1, 1); 7 ~ (1, 2); 8 ~ (1, 3); 9 ~ (2, 0); 10 ~ (2, 1); 11 ~ (2, 2); 12 ~ (2, 3). The analysis of 12 treatments run in two replications as a completely randomized design (CRD) or an unblocked design is

Source	DF
Treatments	11
Error	12
Total	23

On the other hand if it is known that the treatments are structured as two factors with three and four levels, respectively, the same design can be analysed as factorial experiment run in a CRD with two replications. In that case the treatments can be partitioned into main effects and interaction effects as explained below.

Source	DF
Treatments	11
Main Effect A	2
Main Effect B	3
Interaction A*B	6
Error	12
Total	23

Another advantage of using a factorial experiment is the following: If one looks carefully at the design in the example, each treatment combination appears twice in the completely randomized design (CRD) because the replication is two. However, if one looks at the replications of levels, then the levels 0, 1, 2 of factor A are replicated eight times each. Similarly the levels 0, 1, 2, 3 of factor B are replicated six times each. This replication of levels within the replication of treatment combinations is known as hidden replication. It is because of this hidden replication that some comparisons are made with higher precision in factorial experiments. These experiments have another advantage that these allow to study the interaction effects. If an experiment is conducted separately for each factor, the interaction effects cannot be estimated. Moreover, to achieve the same precision as in factorial experiment, the replications would have to be large. For example, if factor A with three levels is conducted as CRD, to have 12 degrees of freedom for error, one needs to have 5 replications making a total of 15 observations. Similarly, if the factor B is run as a CRD, then to have 12 degrees of freedom for error, the replication should be four making a total of 16 observations. So the total number of observations becomes 31, but the interaction effect cannot be estimated. On the other hand a factorial experiment requires only 24 observations to have 12 degrees of freedom for error and allows estimation of interaction effect also. This means that running an experiment separately for each factor would result into an increase in the cost of the experimentation and interaction effects would have to be sacrificed. But factorial experiments have an advantage that not only the cost is reduced, the interaction effects are estimable and can be studied. Further the hidden replication in factorial experiments leads to an improved precision of the factorial effects.

1.6 Variability in the experimental data

The data generated through designed experiments exhibit a lot of variability. In Section 1.1 there was a mention of various type of variability in the data generated. The variability may be wanted, desirable, unwanted, undesirable but is controllable in the sense that it can be accounted for. There is also some more variability, unwanted, undesirable and uncontrollable. The reason for its presence is unknown. In an example in Section 1.2.2, it has been seen that even the experimental units (plots) subjected to the same treatment also give rise to different observations, thus creating variability. These plots are expected to give same response, but

actually the responses are different; reasons unknown. The statistical methodologies, in particular the theory of linear estimation and analysis of variance, enable us to partition the total variability in the data into two major components. The first major component comprises of that part of the total variability to which we can assign causes or reasons. The second component comprises of that part of the total variability to which we cannot assign any cause or reason. This variability arises because some factors are unidentified as a source of variation. Even after careful planning of the experiment, this component is always present and is known as *experimental error*. The observations obtained from experimental units identically treated are useful for the estimation of this experimental error. Ideally one should select a design that will give experimental error as small as possible. There is, though, no rule of thumb to describe what amount of experimental error is small and what amount of it can be termed as large. A popular measure of the experimental error is the percent Coefficient of Variation (CV). Generally the researcher desires the CV to be small, though there is no degree of smallness defined.

The explainable part of the total variability again has two major components. One major component is the conditions to study or the treatments. This part of the variability is wanted or desirable. There is always a deliberate attempt on the part of the experimenter to create variability by the application of several treatments. So in every designed experiment treatments are one component that cause variability. The other component of the explainable part of variability is the experimental units. This variability is unwanted and undesirable. The factors that cause this variability are called nuisance factors. This part of the variability is accounted for by using the principle of local control. Before planning the experiment, the experimenter must have a complete knowledge about the experimental units on which the experiment would be conducted and the sources of variability in the experimental units. If this variability is substantial and is not accounted for by proper designing of experiment, then this component would sit in the experimental error and make it unduly large. The end result would be a bad experiment. As mentioned earlier in Section 1.2.3, there could be many ways of accounting for the variability due to experimental units. The remedy will depend upon the sources and nature of the factors causing variability in the experimental units. As a matter of fact, the way to account for the variability in the experimental units will dictate what type of design is to be used. Many a time, depending upon practical constraints, a naive design may be the best design.

As-a-matter-of-factly many designs have been evolved in the literature depending upon how the variability present in the experimental units is taken care of and how the treatments are allocated to the experimental units or how the randomization is done. If the experimental units are homogeneous and do not exhibit considerable variability, then the treatments are applied randomly to all the experimental units assuming that all the experimental units are uniform. Such designs are known as zero-way elimination of heterogeneity designs or completely randomized designs (CRD) and will be dealt with in detail in Chapter 2. On the contrary, if the variability present in the experimental units is sizeable, then forming groups called blocks containing homogeneous experimental units can account for this variability if the variability in the experimental units is due to one nuisance factor only. As opposed to the allotment of treatments randomly to all the experimental units in a CRD, the treatments in this case are allotted randomly to the experimental units within each block. Such designs are termed as one-

way elimination of heterogeneity setting designs or the block designs. The most common block design is the randomized complete block (RCB) design which is also considered in Chapter 2. If there are two sources of variability in the experimental units, then the experimental units are grouped into arrays, called rows and columns, and the intersection of rows and columns, called cells are the experimental units. The treatments are allocated to the cells. For the randomization purpose, first the rows are randomized and then the columns are randomized. There is no randomization of treatments possible within rows and/or within columns. Such designs are called row-column designs or two-way elimination of heterogeneity designs. A special class of these designs is the Latin square designs and will be studied in Chapter 2.

Generally in experimentation, the number of treatments is large. For large number of treatments, the blocks become large if one has to apply all the treatments in a block, as desired by the RCB design. It may then not be possible to maintain homogeneity among plots within a block and the basic purpose of forming blocks is defeated. The intra block variance or the variance per plot becomes large resulting in a large experimental error and thus a high value of coefficient of variation (CV). To overcome this problem, recourse may be made to an incomplete block design. A block design is said to be an incomplete block design if the design has at least one block that does not contain all the treatments. Some common incomplete block designs are balanced incomplete block (BIB) design, partially balanced incomplete block (PBIB) design including Lattice designs – square and rectangular, cyclic design, alpha design, etc. The concept of incomplete block design can also be extended to incomplete row and / or incomplete column designs. An example of incomplete row-column design is a Youden design or a Generalized Youden design or a Pseudo Youden design.

The unexplainable part of the variability, called the experimental error, is always present. But through controlled experimentation, it is always possible to control this component of variability. It is desirable that this component is as small as possible. This part, therefore, can be controlled by proper designing of an experiment. This means that the design should be such that it accounts for all the sources of variability in the experimental units. If the experimenter fails to control the variability in the experimental units through proper designing, then the experimental error can be controlled by a very useful and important statistical technique called analysis of covariance. This would be dealt with in Chapter 4.

1.7 Shape and size of experimental units

In agricultural field experiments, often plots in fields are used as experimental units. One important issue in this context is the shape and size of the plots and their arrangement. Some general considerations for plot arrangements are given in the sequence.

i)   The experimental area should be as uniform as possible. Uneven sites may lead to high error.

ii)   Plots should be either rectangular or square and equal in area.

iii)   The orientation of the plots should be same, for example, the longer side of the rectangular plots should be parallel to each other.

iv) Uniformity trials may be conducted to get optimum shape and size of the plots. Uniformity trial involves growing a particular crop on a field or piece of land with uniform conditions. All sources of variation except that due to native soil differences, are kept constant. At the time of harvest the entire field is divided into smaller units of same size and shape and the produce from each such unit is recorded separately. The smallest the basic units, the more detailed are the measurements of soil heterogeneity.

v) It may not be economically feasible to conduct a uniformity trial. Even time constraint may be prohibitive in the conduct of a uniformity trial. If soil parameters are known, then these can be used for formation of plots and blocks. At times, the residuals obtained from a previous designed experiment conducted at that place may be used as covariate in the analysis of data generated.

1.7.1 Determining optimum size of plots

In the sequel are described some methods for understanding soil fertility variation/plot size.

i) *Fertility contour map*: An approach to describe the heterogeneity of land is to construct the fertility contour map. This is constructed by taking the moving averages of yields of unit plots and demarcating the regions of same fertility by considering those areas, which have yield of same magnitude. This approach of describing the variation in fertility has been adopted by large number of workers in India and abroad. Fertility contour map can also be developed using the soil parameters in the observed samples obtained from the experimental area.

ii) *Maximum Curvature Method*: In this method basic units of uniformity trials are combined to form new units. The new units are formed by combining columns, rows or both. Combination of columns and rows is done in such a way that no columns or rows are left out. For each set of units, the coefficient of variation (CV) is computed. A curve is plotted by taking the plot size (in terms of basic units) on X-axis and the CV values on the Y-axis of graph sheet. The point at which the curve takes a turn, *i.e.*, the point of maximum curvature is located by inspection. The value corresponding to the point of maximum curvature will be optimum plot size.

iii) *Fairfield Smith's Variance Law*: Smith (1938) suggested an empirical relation between variance and plot size. Smith developed an empirical model representing the relationship between plot size and variance of mean per plot. This model is given by the equation

$$v_x = \frac{v_1}{x^b}$$
$$or \; \log v_x = \log v_1 - b \log x$$

where x is the number of basic units in a plot, is the variance of mean per plot of x units, is the variance of mean per plot of one unit and b is the characteristic of soil and measure of correlation among contiguous units. If $b = 1$, $v_x = \frac{v_1}{x}$ and the units making up the plots of x units are not correlated at all. If $b = 0$, $v_x = v_1$ and the units making up the plots of x units are perfectly correlated and hence there is no gain due to larger size of plots.

Generally b lies between 0 and 1. The values of v_1 and b are determined by least squares method.

This law can further be used for arriving at an optimum plot size. Smith recommended the cost function $C = C_1 + xC_2$, where C_1 is overhead cost which is independent of plot size and C_2 is the consideration of cost by a unit increase in the plot size. Optimum value of plot size is the one which minimizes the cost per unit of information *viz.* $(C_1 + xC_2)$. Once b is estimated from uniformity trial data, the optimum size of plot can be obtained using the following formula

$$x_{opt} = \frac{bC_1}{(1-b)C_2}.$$

Here it must be mentioned that the value of x_{opt} is some multiple of the basic plot size. For example, if $x_{opt} = 2$, then it means that the optimum plot size is twice the basic plot size used in the uniformity trial for estimating b.

1.7.2 Shape of the plots

Shape of plots in agricultural field experiments should be decided after taking care of following points:

i)   Crop to be grown

ii) Convenience of planting and harvesting crop

iii) Ability to use machineries (if machineries are going to be used)

iv) Presence or absence of fertility gradient

v)   Variation in soil depth

1.8 Determination of number of replications

A very important question that needs to be answered by the experimenter is about the number of replications to be used in a design. Although the answer largely depends upon the resources available, there are some scientific reasons also that help in determining the optimum replication number. The following points should be kept in mind while determining number of replications of the treatments.

i)   The foremost important consideration in the determination of replication number is that there should be adequate error degrees of freedom. As far as possible, there should be about 12 degrees of freedom for error. The reason is not far to seek. The error mean square sits in the denominator of the test statistic to be used for testing the null hypothesis. If one looks at the tables of Snedecor's F, the value below 12 degrees of freedom is very high and very variable. So small variations in treatment effects will not be detected significant for smaller degrees of freedom for error. On the other hand, the table values of Snedecor's F stabilize after 12 degrees of freedom. So in order to be able to capture small variations, the error degrees of freedom should be at least 12. On the other hand, the error degrees of freedom should not be unduly large. It would be wastage of resources to spend large

degrees of freedom for estimating experimental error. It may be seen that if there are more number of treatments then lesser number of replications are required to ensure same number of error degrees of freedom.

ii)   Availability of resources and precision required: Number of replications should be determined in such a way that the experiment can be conducted with the available resources namely labour, cost, time, experimental material, etc. and it should be able to achieve desired precision of comparisons among treatments. The smaller the differences that are desired to be detected between treatment means or effects, the more is the number of replications needed. Sometimes it may not be possible to obtain desired precision with available resources and there may be a need of a trade-off between available resources and desired precision level either by sacrificing precision or by increasing available resources.

iii) Type of experimental material: Generally homogenous experimental units require less number of replications and heterogeneous experimental units require more number of replications of the treatments.

iv)   Manageability of the experiment: It should also be kept in mind that the experimenter should be able to manage to conduct the experiment well. This entails that number of treatments, their replications and number of experimental units should not be very large, otherwise it may lead to a poorly managed experiment.

   We describe below a method due to Cochran and Cox (1964) to obtain number of replications of treatments. In conducting an experiment, the experimenter may be interested to detect difference of at least, say d, between two treatment effects. Let the two treatment means be $\bar{t}_1$ and $\bar{t}_2$. The significance of difference between two treatment effects is tested using Student's t statistic given by

$$t = \frac{\left|\bar{t}_1 - \bar{t}_2\right|}{\sqrt{2s_e^2 / r}}$$

where s_e^2 is the measure of error variation and r is the number of replications for both the treatments. If the experimenter wants to detect a difference of at least d between the two treatment effects, then the t-statistic should come significant at desired level of significance α and the corresponding t-statistic would be given by

$$t_\alpha = \frac{\left|d\right|}{\sqrt{2s_e^2 / r}}$$

where t_α denotes the critical value of t_α distribution at level of significance α. From the above equation one can get the number of replications as

$$r \le \frac{2s_e^2 t_\alpha^2}{d^2}.$$

It may be noted that the above formula requires the knowledge of s_e^2. This may be known from similar kinds of experiments conducted earlier or it may be estimated from a pilot experiment.

The description of obtaining the replication number is valid for orthogonal equi-replicated designs. In general, the denominator in the test statistic (t_a) would be the estimated variance of the estimated elementary contrast. The expression for replication number will then be obtained accordingly.

1.9 Steps for running an experimental design

The main steps in conducting an experiment are given below:

i)   State the objectives of the study and the hypotheses to be tested.

ii)   Determine the response variable(s) of interest that can be measured.

iii)   Determine the controllable factors of interest that might affect the response variable(s) and the levels of each factor to be used in the experiment. It is better not to pre-judge any factor to be not significant. Such factors should be included in the design.

iv)   Determine the uncontrollable variables that might affect the response variables.

v)   Determine the total number of experimental units and number of replications of the treatments in the experiment, based on available time and resources and if possible, using estimates of variability, precision required, size of effects expected. Keep some resources for unforeseen contingencies.

vi)   Select a suitable design for the experiment. The chosen design should block the known nuisance variables and randomize the experimental units to protect against unknown nuisance variables.

vii)   Conduct a smaller pilot experiment to see if the results make sense and perform a performance analysis with response variables as random variables to check for estimability of the factor effects and precision of the experiment. Review steps i-vi in case of unsatisfactory situation.

viii) Perform the experiment strictly according to the experimental design.

ix)   Analyse the data from the experiment.

x)   Interpret the results and state the conclusions.

xi)   Document the results and conclusions from the experiment.

xii)   The most important thing to remember is that the treatments are always labelled randomly.

1.10 Scope of the present book

The purpose of the present book is to describe the commonly used experimental designs in agricultural research and using the actual experiments conducted by the researchers, give the analysis of data and interpretation of results. Since this book is targeted for agricultural researchers, there will be a bias towards agriculture, horticulture and animal sciences while

describing the examples. But the applications to other sciences and industry are straight forward.

Since the topic of design and analysis of experiments is vast, the book has been divided into two parts; first part covering fundamentals of design and analysis of experiments related to, by and large, single factor experiments and the second part focusing mostly on multifactor experiments and advanced topics. The composition of the first part of the book is as follows. The book begins with an introductory chapter and then proceeds to describe the basic designs like CRD, RCB design and LSD. Contrast analysis and analysis of covariance are two very powerful techniques that help answer almost all the questions of the researcher. So these two Chapters follow the basic designs. The readers may devote time in understanding these important techniques of analysis of data keeping in mind their application and usefulness. The book then portrays the importance and usefulness of incomplete block designs and resolvable block designs in agricultural research. Alpha designs or resolvable block designs are very useful in crop improvement programmes. Augmented designs also form a part of the discussion. Combined analysis of experiments is an important component of this book.

Although Chapter 3 is devoted to contrast analysis, while making multiple comparisons, adjustments need to be made to attain overall significance level of the comparisons. Various methods of multiple comparisons are given in Annexure-III. The authors are advised to read Annexure-III before reading the main chapters of the book.

There are appendices dealing with the SAS commands and R codes. The readers may read Annexure-I on SAS for better understanding of the main chapters of the book. Annexure-IV describes a very important web resource "Design Resources Sever", which is hosted at www.iasri.res.in/design. The book is concluded with a bibliography.

The main aim of this book is to give simpler analytical solutions to all problems related to designed experiments. To achieve this end, each chapter of the book introduces the subject in detail in a very simple language that can be understood by the researchers in agricultural sciences and industry. Then the subject is explained with the help of an example, which is an actual experiment conducted by the researchers in agricultural sciences. The analysis of data is done using SAS. R code is also given for the benefit of readers who have familiarity with R software. These unique features make this book distinct from other books available on design of experiments.

The examples in all the chapters of the book have been solved using SAS software. The SAS commands are given in detail. The reason for using SAS is that in the National Agricultural Research System, SAS is available and is being used. In fact SAS is one of the popular software being used globally for analysis of data. For the benefit of readers, some codes for using R software are also given. The output obtained from the use of R software is not given to avoid duplication, because the results obtained are similar to those obtained using SAS. The readers may read Annexure-II for better understanding of the basics of R codes. Those using SPSS may refer to steps of analysis given in Design Resources Server.

While using SAS commands, the input variables and the class variables have been given abbreviated names, *e.g.*, rep is used for replication, trt is used for treatment, etc. Obviously the

SAS output will also depict these names only. However, the output described in the book does not conform to this SAS output. Firstly, the SAS output would generally be more than what is described in the text. Secondly, the output may not be exactly in the format it is described. Lastly, the class variables reported in the analysis do not have abbreviated names. Instead we give the full names of the variables. The reader may also note that the word 'block' has been used throughout for 'replication', wherever replication is used as a block. The reader should not have any confusion of the two terms.

2

Some Basic Experimental Designs

2.1 Introduction

An experimental design is essentially a rule that determines the assignment of treatments to the experimental units keeping in mind the principles of randomization, replication and local control. The experiments, however, differ from each other greatly in many respects, depending upon the variability in the experimental units and how it is taken care of. Generally the process of randomization of the experimental units depends upon the way the variability present in the experimental units is accounted for. This then dictates what type of design is used in a given experimental situation. In some experimental situations, a naïve design, generated keeping in mind the experimental conditions and practical considerations, helps answering the objectives of the experiment. Nonetheless, there are some basic (or standard) designs that are used frequently by the experimenters because of the ease in running these experiments. The purpose of this Chapter is to describe such designs.

We begin with a Completely Randomized Design (CRD), which uses the principles of randomization and replication. This design is used when there are strong reasons to believe that there is no variability in the experimental units. This will be followed by Randomized Complete Block (RCB) Design and Latin Square Design, in which all the three principles of randomization, replication and local control are applied. RCB design and Latin Square Design are used when there are one and two sources of variability, respectively present in the experimental units. In both these designs, the treatment replications are equal. There is flexibility in the choice of number of replications in RCB design, but in case of Latin square design, the replication number is equal to the number of treatments. As would be seen later in this Chapter and in other Chapters as well, it is simply the change in the randomization procedure of the treatments to the experimental units that gives rise to different designs. There could, however, be more sources of variability present in the experimental units and the randomization would be done accordingly. It will become obvious through different Chapters that the randomization and control of variability through grouping(s) of the experimental units are related to each other and help in controlling the experimental error.

2.2 Completely randomized design

Consider an experimental situation where the experimenter is interested (a) in comparing four grazing systems (treatments), viz., *rotational*, *deferred rotational*, *continuous* and *cut and carry*, and (b) to study the effect of the grazing systems on the body weight of the animals. Suppose that 16 animals are available for conducting the experiment. Suppose further that the choice of 16 animals is such that they do not contribute to the variability in the final body

weights of the animals after being subjected to the grazing systems. In other words, it is assumed that the experimental units (subjects or animals) do not contribute to the variability in the data and the only explainable part of variability present in the data is because of the four different grazing systems. The unexplained part of variability is the experimental error. An easy way of running this experiment is to allocate the 4 grazing systems randomly to the 16 animals such that each grazing system is received by four animals. However, it is not necessary to have equal replication of grazing systems. One can have unequal replication of the treatments as well. But as far as possible, it is better to have equal, or as equal as possible, replications of the treatments. If that be so, then the replications would differ by at most one. This is known as a design for zero-way elimination of heterogeneity.

A zero-way heterogeneity setting design or an unblocked design or a CRD is the simplest design in which only two principles of design of experiments *viz*. randomization and replication are used. There is no use of local control here, since the experimental units are assumed to be homogeneous. The only identifiable cause of variability is the treatments and the remaining part of the variability is the experimental error.

To make the exposition general, suppose that there are v treatments and n homogeneous experimental units. The v treatments are allotted at random to the n experimental units. Let the ith treatment be replicated r_i times ($i = 1,2,\ldots,v$) such that $\sum_{i=1}^{v} r_i = n$. Normally the number of replications for different treatments should be equal as it ensures equal precision of estimates of linear functions of treatment effects. The average replication number is then n/v, which will be a positive integer if v divides n. The actual number of replications of treatments is, however, determined by the availability of experimental resources and the requirement of precision and sensitivity of comparisons. If the experimental material for some treatments is available in limited quantities, the number of replications of these treatments is reduced. If the estimates of certain treatment effects are required with more precision, the number of replications of such treatments is increased.

2.2.1 Randomization

There are several methods of random allocation of treatments to the experimental units. The v treatments are first assigned numbers (or labels) randomly from 1 to v. The n experimental units are also numbered randomly. One method of randomization uses the random number tables. Any column (or columns) of a randomly opened page of a random number table is taken. If v is a one-digit number, then only one column is consulted digit by digit. If v is a two-digit number, then two columns (or two-digit random numbers) are consulted. All numbers greater than v and zero, are ignored.

Let the first number chosen be n_1 ; then the treatment numbered n_1 is allotted to the first unit. If the second number is n_2 which may or may not be equal to n_1 , then the treatment numbered n_2 is allotted to the second unit. This procedure is continued. When the ith treatment number has occurred r_i times, $(i = 1,2,\ldots,v)$ this treatment is ignored subsequently. This process terminates when all the units are exhausted.

One drawback of the above procedure is that sometimes a very large number of random numbers may have to be ignored because they are greater than v. It may even happen that the random number table is exhausted before the allocation is complete. To avoid this difficulty the following procedure is adopted.

Let v be an s-digit number. Choose P as the highest s-digit number divisible by v. For instance, if $v = 13$, then $P = 91$; when $v = 31$, then $P = 93$; when $v = 123$, then $P = 984$. All numbers greater than P and zero are ignored. If a selected random number is less than v, then it is used as such. If it is greater than or equal to v, then it is divided by v and the remainder is taken to be the random number and used for allotting treatment to experimental unit. When a number is divisible by v (*i.e.*, the remainder is zero), then the random number is v. For example, assume that $v = 123$ and the random number drawn is 991. This number would be rejected because this is greater than 984. If the random number drawn is 95, then the treatment labeled 95 is allotted to that experimental unit. Further, if the random number selected is 567, then dividing 567 by 123 would leave the remainder as 75. So treatment labeled 75 is allotted to that experimental unit. Further, if the random number selected is 615, then dividing 615 by 123 would leave the remainder as zero. In this case, treatment labeled as 123 is allotted to that experimental unit.

Alternative methods of random allocation

If random number tables are not available, treatments can be allotted by drawing *lots* as explained in the sequel. However, these procedures may not help generate strictly random numbers. So these procedures need to be adopted with caution.

The number of the ith treatment is written on r_i pieces of papers ($i = 1,2,...v$). The $\sum_{i=1}^{v} r_i = n$ pieces of papers are then folded individually so that the numbers written on them are not visible. These papers are then drawn one by one at random. Before each draw the slips are thoroughly shuffled. The treatment that is drawn at the tth draw is allotted to the tth unit ($t = 1,2,...,n$)

Random allocation is also possible by using a fair coin. Let there be 5 treatments and 20 experimental units. Each treatment is to be replicated four times. Suppose that the experimental units are labeled by numbers from 1 to 20 randomly.

When a coin is tossed, there are two possible outcomes; either head or tail appears. Denote the "head" by H and the "tail" by T. When the coin is tossed twice, there are four possible outcomes; these are HH, HT, TH or TT. Similarly, when the coin is flipped three times, there are eight possible outcomes; HHH, HHT, HTH, HTT, THH, THT, TTH, TTT. This can be easily generalized to n flippings of the coin.

The 5 treatments are now identified not by serial numbers as earlier but by any five of the above eight possible outcomes obtainable by flipping a coin three times. If any of the remaining three outcomes, say THT, TTH and TTT appear, no treatment is selected for allotment and the coin is again flipped thrice.

A coin is now thrown three times and the outcome is noted. If the outcome is any of the five outcomes HHH, HHT, HTH, HTT, THH, the treatment labeled by it is allotted to the first experimental unit. If the event happened is any of the three, THT, TTH, TTT, it is ignored. The coin is again tossed three times and this event is used to select a treatment for the second experimental unit. If the same outcome appears more than once, do not reject it until the number of times it has appeared equals the number of replications of the treatment it represents. This process is continued till all the experimental units are exhausted.

It may be worthwhile mentioning here that the labels are also allotted randomly to all the treatments. This would hold everywhere, whether mentioned or not.

The linear model in this case is

Expected response = general mean + effect of treatments.

Since there is no source of variation in the experimental units, the model does not contain the effect due to experimental units.

This can also be written as

response = general mean + treatments effect + error,

where the errors are independently distributed as normal variate with zero mean and constant variance σ^2. The partitioning of the total variability in this case is

Source of variation
Due to model
Error
Total

The component "due to model" can be partitioned as

Source of variation
Due to model
Due to treatments

2.2.2 Analysis of CRD

This design provides a one-way classified data according to levels of a single factor, the single factor being the treatments with v levels. Since no variability is expected from the experimental units, the only identifiable source of variability is the treatments. We then have the following linear model:

$$y_{ij} = \mu + \tau_i + e_{ij} , \quad i = 1, \cdots, v; \, j = 1, \cdots, r_i$$

$$E(e_{ij}) = 0, \quad Cov(e_{ij}, e_{kl}) = \begin{cases} \sigma^2, & \text{if } i = k, j = l \\ 0, & \text{otherwise} \end{cases}$$

where the random variable y_{ij} is the observation recorded on the jth replicate of the ith treatment, μ is the general mean, τ_i is the fixed effect of the ith treatment and e_{ij} is the random error component associated with the (i, j)th observation, $i = 1,2,..., v$; $j = 1,2,...,r_i$. These are assumed to be distributed independently and normally with zero mean and constant variance σ^2. We also assume that the replication of the ith treatment is r_i, $i = 1,2,\cdots,v$ and $\sum_{i=1}^{v} r_i = n$.

Let us define the following:

Treatment totals, $T_i = \sum_{j=1}^{r_i} y_{ij}$ $\forall i = 1,2,\cdots,v$, and Grand total as $G = \sum_{i=1}^{v} T_i = \sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}$.

The following formulae can be employed for analysis of variance:

Correction factor $(CF) = \dfrac{G^2}{n}$

Total sum of squares $(SS) = \sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}^2 - CF$

Sum of squares due to treatments $(SST) = \sum_{i=1}^{v} \dfrac{T_i^2}{r_i} - CF$

Error sum of squares $(SSE) = \sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}^2 - \sum_{i=1}^{v} \dfrac{T_i^2}{r_i} = \left(\sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}^2 - CF \right) - \left(\sum_{i=1}^{v} \dfrac{T_i^2}{r_i} - CF \right)$

$$= \text{Total } SS - \text{Treatment } SS.$$

The interest of the experimenter is in testing the null hypothesis: $H_0 : \tau_1 = \tau_2 = \cdots = \tau_v = \tau(say)$ against the alternative that $\tau_i \neq \tau_l, l \neq i = 1,2,\cdots,v$ for at least one pair of treatment effects, say τ_i and τ_j. For testing this hypothesis, we set up the analysis of variance Table 2.1.

Table 2.1: ANOVA table in CRD

Source	DF	SS	MS = SS/DF	F
Treatments	$v - 1$	$\text{SST} = \sum_{i=1}^{v} \dfrac{T_i^2}{r_i} - CF$	$\dfrac{SST}{(v-1)} = s_t^2$	$\dfrac{s_t^2}{s_e^2}$
Error	$n - v$	$\text{SSE} = \sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}^2 - \sum_{i=1}^{v} \dfrac{T_i}{r_i}$	$\dfrac{SSE}{(n-v)} = s_e^2$	
Corrected Total	$n - 1$	$\sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}^2 - CF$		

If the calculated value of F is greater than the table value of $F_{\alpha;(v-1),(n-v)}$ at α level of significance and $(v - 1)$, $(n - v)$ degrees of freedom, then the null hypothesis H_0 is rejected at α level of significance and it can be concluded that equality of all the treatment effects does not hold. In that case, the researcher has no knowledge about the treatment effects except that there is at least one pair of treatments that differs significantly from each other. In that case the researcher has to go for the computation of least significant difference (LSD) or other multiple comparison procedures as explained in Annexure-III to make pairwise treatment comparisons.

It may be seen here that the unbiased estimator of σ^2 is $\hat{\sigma}^2 = s_e^2$.

Further, all the elementary treatment contrasts (or treatment contrasts for pairwise treatment comparisons) are estimable through the design. The best linear unbiased estimator (BLUE) of any treatment contrast $d_{il} = \tau_i - \tau_l$ is

$$\hat{d}_{il} = \frac{T_i}{r_i} - \frac{T_l}{r_l}, \forall\, i \neq l = 1, 2, \cdots, v.$$

The variance of $\hat{d}_{il}$ is $\mathrm{Var}(\hat{d}_{il}) = \sigma^2\left(\dfrac{1}{r_i} + \dfrac{1}{r_l}\right)$. The estimated standard error of the estimated difference between the ith and lth treatment effects is $\hat{\mathrm{SE}}(\hat{d}_{il}) = s_e\left(\dfrac{1}{r_i} + \dfrac{1}{r_l}\right)^{1/2}$.

The least significant difference (LSD) is given as $\text{LSD} = \hat{\mathrm{SE}}(\hat{d}_{il}) \times t_{\alpha/2;\,errorDF}$.

Here $t_{\alpha;\,error\,DF}$ denotes the value of Student's t at α level of significance and error degrees of freedom. The treatment means are given by $\overline{T}_i = \dfrac{T_i}{r_i} \,\forall\, i = 1, 2, \cdots, v$. The pairwise comparison of treatment effects can be made by comparing the difference between any two treatment means with the LSD. Any two treatment effects are said to differ significantly if the difference of their means is larger than the LSD.

2.2.3 Example 1

An experiment was conducted in Rabi season on a variety of tomato during 2010-11 with 5 treatments of integrated nutrient management *viz.* Trt1 ~ farmers' practice (2.5 tonnes farmyard manure/ha), Trt2 ~ recommended dose of fertilizers (NPK 120:75:100), Trt3 ~ 50% recommended dose of fertilizers + vermin-compost 5 tonnes/ha, Trt4 ~ 50% recommended dose of fertilizers + vermin-compost 10 tonnes/ha and Trt5 ~ 50% recommended dose of fertilizers + vermin-compost 2.5 tonnes/ha + farmyard manures 5 tonnes/ha. The objective of the experiment was to find out the most appropriate integrated nutrient management system for tomato. The experiment was conducted using a completely randomized design and the dry matter accumulation (gm/plant) was recorded after the experiment was over. Table 2.2 gives the replicated data on dry matter accumulation in g/plant for each treatment:

Table 2.2: Dry matter accumulation in g/plant

Tr1	Tr2	Tr3	Tr4	Tr5
108.2	225.2	176.5	201.3	214.3
112.7	226.4	195.2	183.6	226.2
116.8	135.2	188.4	197.5	215.0
106.8	227.5	190.3	186.1	230.6
117.9	218.2	210.3	188.6	212.6
	229.1	195.1	210.4	230.4
				227.6
				228.3

In the sequel the data are analyzed to identify the best integrated nutrient management system.

2.2.4 Procedure and Calculations

The inference problem being solved here is the testing of the following null hypothesis: H_0: $\tau_1 = \tau_2 = \cdots = \tau_i = \cdots = \tau_v = \tau$ (say) against the alternative hypothesis H_1: at least two of the τ_i's are different. In the example, $v = 5$.

First compute the following totals:

Treatment totals

$T_1 = 108.2 + 112.7 + 116.8 + 106.8 + 117.9 = 562.4$

$T_2 = 225.2 + 226.4 + 135.2 + 227.5 + 218.2 + 229.1 = 1261.6$

$T_3 = 176.5 + 195.2 + 188.4 + 190.3 + 210.3 + 195.1 = 1155.8$

$T_4 = 201.3 + 183.6 + 197.5 + 186.1 + 188.6 + 210.4 = 1167.5$

$T_5 = 214.3 + 226.2 + 215 + 230.6 + 212.6 + 230.4 + 227.6 + 228.3 = 1785.0$

Gross total $(G) = T_1 + T_2 + T_3 + T_4 + T_5 = 562.4 + 1261.6 + 1155.8 + 1167.5 + 1785 = 5932.3$

Correction factor, $CF = G^2/n = (5932.3)^2/31 = 1135231.719$

SS due to treatments, $SST = \sum_{i=1}^{v} \dfrac{T_i^2}{r_i} - CF$

$$= (562.4)^2/5 + (1261.6)^2/6 + (1155.8)^2/6 + (1167.5)^2/6 + (1785.0)^2/8 - 1135231.719$$

$$= 41399.233$$

$$\text{Total SS} = \sum_{i=1}^{v}\sum_{j=1}^{r_i} y_{ij}^2 - CF$$

$$= (108.2)^2 + (112.7)^2 + \ldots + (227.6)^2 + (228.3)^2 - 1135231.719$$

$$= 49891.17$$

Error SS = Total SS – SS due to treatments = SSE

$$= 49891.17 - 41399.233 = 8491.938$$

The following Analysis of Variance Table is then formed.

Table 2.3: ANOVA table for data in Example 1

Source	DF	SS	MS	F-value	Prob > F
Treatments	4	41399.233	10349.808	31.69	<0.0001
Error	26	8491.938	326.613		
Total	30	49891.171			

R-square	CV	RMSE	Yield Mean
0.830	9.444	18.072	191.364

The model used has been able to explain 83 per cent of the total variability in the data. Since calculated F value = 31.69 is greater than the tabulated F at 4 and 26 degrees of freedom at 5% level of significance (= 2.742), the null hypothesis is rejected and at least two treatment effects are significantly different from each other at 5% level of significance. In fact, the probability of obtaining a value of F greater than 31.69 is smaller than 0.0001, meaning thereby that at least two treatments effects differ significantly even at smaller than 0.01% level of significance.

Now, to compare the treatment pairs, we calculate treatment means and LSD values at 5% level of significance. The treatment means are given in Table 2.4.

Table 2.4: Treatment wise mean and standard deviation of dry matter accumulation

Level of treatment	N	Dry matter accumulation	
		Mean	Standard Deviation
1	5	112.480	4.967
2	6	210.267	36.967
3	6	192.633	11.031
4	6	194.583	10.316
5	8	223.125	7.740

We now proceed to test the equality of treatments effects, *i.e.* H_0: $\tau_i - \tau_l = 0$, for all $i \neq l = 1,2,\cdots,v$. This is equivalent to making all the possible pairwise treatment comparisons. Table 2.5 gives |*difference between two treatment means*| and the least significant difference (LSD). Here |x| is the absolute value of x. In other words, it is the value of x ignoring the sign. If the difference of treatment means is larger than the LSD, then the two treatments are significantly different from each other at 5 per cent level of significance.

Table 2.5: Least significant differences of treatment pairs

Treatment Numbers	Difference of treatments means	Least Significant Difference
1, 2	97.79	22.35
1, 3	80.15	22.35
1, 4	82.1	22.35
1, 5	110.65	21.04
2, 3	17.64	21.31
2, 4	15.69	21.30
2, 5	12.86	19.93
3, 4	1.95	21.31
3, 5	30.5	19.93
4, 5	28.55	19.33

Alternatively, arrange the treatment means in ascending or descending order depending upon the character under study. If it is yield, it may be arranged in descending order and if it is disease infestation, it may be arranged in ascending order.

Table 2.6: Treatment means arranged in descending order

Dry matter accumulation	Treatment	Rank
223.125	5	1
210.267	2	2
194.583	4	3
192.633	3	4
112.480	1	5

Take the different between two treatment means with consecutive ranks. In Table 2.6, the difference Trt5 – Trt2 = 223.125 – 210.267 = 12.858. The LSD at 5% for these two treatments is 19.93. Therefore, Trt5 and Trt2 are not significantly different and may be assigned the same letter A. Since treatments ranked 1 and 2 are statistically not significant, therefore, now check the difference between treatment with rank 1 and rank 3, i.e., Trt5 – Trt4 = 223.125 – 194.583 = 28.542. The LSD at 5% for these two treatments is 19.33. Therefore, these are statistically

different, or we can say that Treatment 5 is statistically better than treatment 3 and we may assign a different letter, say B, to treatment rank 3.

Since treatment with rank 1 is statistically significant compared to treatment with rank 3, now significance of treatment with rank 1 need not be tested with treatments with rank 4 and 5. Treatment with rank 1 will automatically be significantly different from treatments with ranks 4 and 5.

The procedure of making pairwise treatment comparisons just explained always holds when the estimated variance of the estimated difference of every possible pair of treatments is same as happens in CRD with equal replication, RCB design, LSD or any other variance balanced design. In case the estimated variance of the estimated difference of every possible pair of treatments is different, then we may need to check the significance of all treatment contrasts.

In this example, however, estimated variances of the estimated difference of all other pairs of treatments is less than that between treatments with rank 1 and rank 3. Therefore, we may stop checking significance of treatment with rank 1 with treatments with ranks 4 and 5. Now, start with treatment with rank 2 and test the significance of difference of treatment effects with rank 2 and 3 as Trt2 – Trt4 = 210.267 – 194.583 = 15.684, which is less than corresponding LSD at 5% level of significance (21.30). Therefore, we assign a second letter to treatment with rank 2 same as that was assigned to treatment with rank 3 earlier, i.e., B. Now, treatment with rank 2 has two symbols A and B depicting that it is not significantly different from treatments with rank 1 and 3. Now, we proceed to test the significance of difference of treatment effects with ranks 2 and 4 as Trt2 – Trt3 = 210.267 – 192.633 = 17.634, which is less than corresponding LSD at 5% level of significance (21.31). Therefore, now treatment 3 with rank 4 may also be assigned the same letter B. Next, proceed to test the significance of difference of treatment effects with ranks 2 and 5 as Trt2 – Trt1 = 210.267 – 112.480 = 97.787, which is more than corresponding LSD at 5% level of significance (22.35). Therefore, now treatment with rank 5 may be assigned a different letter, say C. Next, proceed to test significance of difference of effects of treatments with ranks 3 and 4, *i.e.*, Trt4 – Trt3 = 194.583 – 192.533 = 2.050, which is less than the LSD at 5%. Therefore, these two treatments are statistically at par and already assigned same letter B. We proceed with testing the same way and will get the Table 2.7.

Table 2.7: Treatments with letter display

Dry matter accumulation	Treatment	Rank
223.125A	5	1
210.267A,B	2	2
194.583B	4	3
192.633B	3	4
112.480C	1	5

Another way of presenting this Table is

Tr5	Tr2	Tr4	Tr3	Tr1
223.125	210.267	194.583	192.633	112.480

As per the Table 2.7, one can say that treatment 5 is significantly better than treatments 4, 3 and 1. Treatment 2 although statistically not significant with treatments 5, 4 and 3, is significantly different from treatment 1. Similarly treatments 4 and 3 are significantly different from treatment 1. Therefore, if treatment with highest mean is best, then any one of the treatments Tr5 or Tr2 may be used as they are statistically at par.

It may be noted that LSD controls only individual error rate and should be used only when null hypothesis of equality of treatment effects through ANOVA is rejected. Other commonly used multiple comparison procedure test that controls only individual error rate is Duncan's Multiple range Test. Some tests which control family error rate are Bonferroni correction and Tukey's Honestly Significant Differences (HSD) test and can be used even when null hypothesis through ANOVA is not rejected. More details on multiple comparison procedures may be seen in Annexure-III.

2.2.5 Analysis using SAS

The design is a CRD with $v = 5$ treatments and $n = 31$ observations. The data has been analyzed using SAS software. The commands and the data preparation are given in the sequel.

```
DATA crd;
INPUT trt dma;
/* trt denotes the treatment number and dma denotes the dry matter accumulation
in g/plant*/;
CARDS;
1       108.2
1       112.7
1       116.8
1       106.8
1       117.9
2       225.2
2       226.4
2       135.2
2       227.5
2       218.2
2       229.1
3       176.5
3       195.2
3       188.4
3       190.3
3       210.3
3       195.1
4       201.3
```

4	183.6
4	197.5
4	186.1
4	188.6
4	210.4
5	214.3
5	226.2
5	215.0
5	230.6
5	212.6
5	230.4
5	227.6
5	228.3

```
;
PROC GLM DATA=crd;
CLASS trt;
MODEL dma = trt;
MEANS trt;
LSMEANS trt/PDIFF  LINES;
RUN;
```

Remark 2.1 It may be worthwhile mentioning here that in the INPUT statement, CLASS statement and MODEL statement etc. the terms like trt, rep, etc. have been used to represent treatments, replications, etc. The output of analysis will also be using these notations. But while giving the results of analysis, the abbreviated forms are not used. Instead, the full forms are used for clarity and better understanding.

2.2.6 Output of analysis

The results obtained by the analysis using SAS are given in Table 2.8. This output is same as described earlier. The model with treatment effects only has been able to explain 83 per cent of the total variation. It is seen from the analysis of variance table that the treatment effects are significantly different (p-value < 0.0001).

Table 2.8: SAS output for data in Example 1

ANOVA

Source	SS	DF	MS	F-value	Prob > F
Model	41399.233	4	10349.808	31.69	<0.0001
Error	8491.938	26	326.613		
Total	49891.171	30			

R-square	CV	RMSE	dma Mean
0.830	9.444	18.072	191.364

ANOVA

Source	Type I SS	DF	MS	F value	Prob > F
Treatments	41399.233	4	10349.808	31.69	<0.0001
Error	8491.938	26	326.613		
Total	49891.171	30			

The distribution of the observations for each treatment is given in the Figure 2.1.
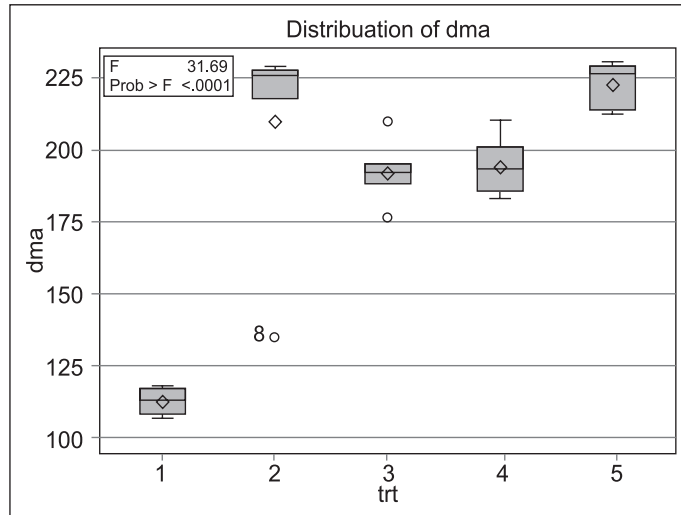
Figure 2.1: Treatment wise Box plot of dry matter accumulation

The mean and standard deviation of the dry matter accumulation for each of the treatments is given in Table 2.9.

Table 2.9: Treatment wise mean and standard deviation of dry matter accumulation

Level of treatment	N	Dry matter accumulation	
		Mean	Standard Deviation
1	5	112.480	4.967
2	6	210.267	36.967
3	6	192.633	11.031
4	6	194.583	10.317
5	8	223.125	7.7434

Table 2.10 gives the *p*-values for making pairwise treatment comparisons. These comparisons are similar to the one made above using LSD table. A *p*-value smaller than 0.05 implies that the pair of treatment effects is significantly different. For example, a *p*-value < 0.0001 indicates that

the treatment effects 1 and 2; 1 and 3; 1 and 4; 1 and 5; are significantly different. A *p*-value of 0.1030 suggests that the treatment effects 2 and 3 are not significantly different at 5 per cent level of significance. Similarly, a *p*-value of 0.1990 suggests that the treatment effects 2 and 5 are not significantly different at 5 per cent level of significance

Table 2.10: *P*-values for pairwise comparison of the treatments

Least Squares Means for effect treatment **Pr > |t| for H0: LSMean(i)=LSMean(j)** **Dependent Variable: Dry Matter Accumulation**					
i/j	**1**	**2**	**3**	**4**	**5**
1		<0.0001	<0.0001	<0.0001	<0.0001
2	<0.0001		0.1030	0.1450	0.1990
3	<0.0001	0.1030		0.8530	0.0040
4	<0.0001	0.1450	0.8530		0.0070
5	<0.0001	0.1990	0.0040	0.0070	

It may be noted from Table 2.11 that in this case the LS means are the same as the unadjusted means. Table 2.11 is another way of explaining the significance of difference of two treatment effects. Treatments with same letter are not significantly different.

Table 2.11: Treatments with letter display

***t* Comparison Lines for Least Squares Means of Treatments**				
LS-means with the same letter are not significantly different				
		DMA LSMEAN	**Treatment**	**Rank of Treatment**
	A	223.125	**5**	5
B	A	210.267	**2**	4
B		194.583	**4**	3
B		192.633	**3**	2
	C	112.480	**1**	1

Since dry matter accumulation is highest for Trt5 and is significantly different from all other treatment effects, except treatment 2, so Trt5 *i.e.*, 50% recommended dose of fertilizers + vermicompost 2.5 tonnes/ha + farmyard manures 5 tonnes/ha is the best integrated nutrient management system, which is at par with 100% NPK (Trt2) so far as dry matter accumulation in tomato is concerned.

The pairwise treatment comparisons can also be made without writing the treatments in descending order of the treatments LS Mean values. The results are given in Table 2.12.

Table 2.12: Treatments with letter display

Treatment	DMA LS Mean	Rank of Treatment
1	112.480^C	5
2	210.267^{A,B}	2
3	192.633^B	4
4	194.583^B	3
5	233.125^A	1
General Mean	191.364	

In Table 2.12, any two treatments whose LS Means have at least one letter common are not statistically significant using LSD at given level of significance. Therefore, it follows that treatment 1 is significantly different from treatments 2, 3, 4, 5. Similarly, treatment 5 is significantly different from treatments 1, 3 and 4 but is not significantly different from treatment 2. On the other hand, treatment 2 is not significantly different from treatments 3, 4 and 5.

It may be worthwhile mentioning here that all comparisons are made at 5 per cent level of significance.

2.2.7 Analysis using R

The purpose of this section is to give the R code for analysis of data generated from a CRD for the benefit of the readers who would like to use R software. It may be mentioned here that the output obtained from R code is not given to avoid repetition.

```
d1=read.table("crd.txt",header=TRUE)
attach(d1)
names(d1)
#Treatment means and standard deviations
aggregate(dma, by=list(trt), mean)
aggregate(dma, by=list(trt), sd)
#Treatment wise box plot of dma
boxplot(dma~trt)
#ANOVA
trt=factor(trt)
crdout<-aov(dma~trt)
summary(crdout)
#Tukey's honest significant difference test is inbuilt part of Base R
TukeyHSD(crdout)
#LSD test, download and install agricolae package
library(agricolae)
lsd.result <- LSD.test(crdout,"trt")
lsd.result
detach(d1)
```

2.3   Randomized complete block design

A CRD assumes that there is no variability in the experimental units. The only source of variability in the data is the treatments. However, the experimental units selected for experimentation may exhibit variability because of many reasons. If the experimental units are the animals and the treatments are the grazing systems, then the initial body weight of the animal may be a major source of variability. Similarly, if the experimental units are plots in a field, and the treatments are the various levels of fertilizers and or irrigation, then the soil fertility may be a source of variability. Similarly with feeding trials in animal experiments, the lactation number may be a source of variability. Litter mates of animals may be source of variability in animal experiments. The salinity patches in the soil may be source of variability in field experiments. The variability in the experimental units needs to be accounted for, otherwise the experimental error will be unduly large and the Coefficient of variation (CV) would be overly large, which may lead to not rejecting the null hypothesis.

The focus of this Section is on designs useful for situations when there is heterogeneity in the experimental units and it is expected that there is only one source of variability in the experimental units. All the three principles of experimentation, viz., *randomization*, *replication* and *local control* are used in these designs. In these designs, the experimental units are partitioned into groups (called blocks) in such a way that experimental units within each block are as homogeneous as possible. As the name itself suggests, a Randomized Complete Block (RCB) design is a complete block design in the sense that each block is a complete replication. In other words, all the treatments in the experiment appear once in each block. Consequently, the block size, or the number of experimental units in each block is equal to the number of treatments. Further, since each block is a complete replication, the number of blocks is also equal to the replication number of treatments.

The randomization procedure in a RCB design is the following: (i) the treatments are randomly allocated the treatment labels, (ii) the treatments are assigned randomly to the experimental units within each block, and (iii) a separate randomization is done in each block.

The linear model in this case is

Expected response = general mean + effect of treatments + effect due to experimental units (*grouped as blocks*)

Since there is only one source of variation in the experimental units, the model can be rewritten as

Expected response = general mean + effect of treatments + effect of blocks (*or replications*).

This can also be written as

response = general mean + treatments effect + block (*or replication*) *effect + error*,

where the errors are independently distributed as normal variate with zero mean and constant variance σ^2. The split of the total variability in this case is

Source of variation
Due to model
Error
Total

The component "due to model" can be split as

Source of variation
Due to model
Due to treatments
Due to blocks (or replications)

2.3.1 Analysis of RCB design

Suppose that an experiment is run in a RCB design with v treatments and b replications (or complete blocks). Suppose that the observation generated on the response variable from the ith treatment in the jth block is represented by y_{ij}, $i = 1,2,....,v; j = 1,2,....,b$. The observations are represented by the following linear, additive model

$$y_{ij} = \mu + \tau_i + \beta_j + e_{ij}$$

where μ is the general mean effect; τ_i is the effect of the ith treatment (fixed); β_j is the effect of the jth block (fixed); e_{ij} is random error associated with y_{ij}, assumed to be mutually independent and distributed identically as normal variable with mean zero and common variance σ^2, i.e.,

$$E(e_{ij}) = 0, \quad Cov(e_{ij}, e_{kl}) = \begin{cases} \sigma^2, \text{ if } i = k, j = l \\ 0, \text{ otherwise} \end{cases}$$

Let the treatment totals and the block totals be denoted as respectively, $T_i = \sum_{j=1}^{b} y_{ij}, \forall i = 1,2,\cdots,v$

and $B_j = \sum_{i=1}^{v} y_{ij}, \forall j = 1,2,\cdots,b$, and grand total as $G = \sum_{i=1}^{v} T_i = \sum_{j=1}^{b} B_j = \sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}$.

The following formulae can be employed for analysis of variance:

Correction factor $(CF) = \dfrac{G^2}{vb}$

Total sum of squares $= \sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}^2 - CF$

Sum of squares due to treatments $(SST) = \sum_{i=1}^{v} \dfrac{T_i^2}{b} - CF$

Sum of squares due to blocks (or replications) $(SSB) = \sum_{i=1}^{v} \dfrac{T_i^2}{b} - CF$

Error sum of squares $(SSE) = \sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}^2 - \sum_{i=1}^{v} \dfrac{T_i^2}{b} - \sum_{j=1}^{b} \dfrac{B_j^2}{v} + CF$

$$= \left(\sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}^2 - CF \right) - \left(\sum_{i=1}^{v} \dfrac{T_i^2}{b} - CF \right) - \left(\sum_{j=1}^{b} \dfrac{B_j^2}{v} - CF \right)$$

= Total *SS* – Treatment *SS* – Block *SS*.

The interest of the experimenter is in testing the null hypothesis: $H_0 : \tau_1 = \tau_2 = \cdots = \tau_v = \tau (say)$ against the alternative that $\tau_i \neq \tau_l, i \neq l = 1,2,\cdots,v$ for at least one pair of treatment effects, say τ_i and τ_l. For testing this hypothesis we set up the analysis of variance Table 2.13.

Table 2.13: ANOVA table for RCB design

Source	DF	SS	MS = SS/DF	F
Treatments	$v-1$	$SST = \sum_{i=1}^{v} \dfrac{T_i^2}{b} - CF$	$\dfrac{SST}{(v-1)} = s_t^2$	$\dfrac{s_t^2}{s_e^2}$
Blocks (or Replications)	$b-1$	$SSB = \sum_{j=1}^{b} \dfrac{B_j^2}{v} - CF$	$\dfrac{SSB}{(b-1)} = s_b^2$	$\dfrac{s_b^2}{s_e^2}$
Error	$(v-1)(b-1)$	$SSE = \sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}^2 - \sum_{i=1}^{v} \dfrac{T_i^2}{b} - \sum_{j=1}^{b} \dfrac{B_j^2}{v} + CF$	$\dfrac{SSE}{(v-1)(b-1)} = s_e^2$	
Total	$vb-1$	$\sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}^2 - CF$		

If the calculated value of F is greater than the table value of $F_{\alpha;(v-1)(v-1)(b-1)}$ at α level of significance and $(v-1), (v-1)(b-1)$ degrees of freedom, then the null hypothesis H_0 is rejected at α level of significance and it can be concluded that the treatment effects are significantly different from one another.

It may be seen here that the unbiased estimator of σ^2 is $\hat{\sigma}^2 = s_e^2$.

Further, all the elementary treatment contrasts are estimable through the design. The Best Linear Unbiased Estimator (BLUE) of any treatment contrast $d_{il} = \tau_i - \tau_l$ is

$$\hat{d}_{il} = \frac{T_i - T_l}{b}, \forall\, i \neq l = 1, 2, \cdots, v\,.$$

The variance of $\hat{d}_{il}$ is $\mathrm{Var}\!\left(\hat{d}_{il}\right) = \dfrac{2\sigma^2}{b}$. The estimated standard error of the estimated difference

between the ith and lth treatment effects is $SE\!\left(\hat{d}_{il}\right) = \left(\dfrac{2s_e^2}{b}\right)^{1/2}$.

The Least Significant Difference (LSD) at α level of significance is given as

$$\mathrm{LSD} = SE\!\left(\hat{d}_{il}\right) \times t_{\alpha/2;\,errorDF}$$

Here $t_{\alpha/2;\,error\,DF}$ denotes the value of Student's t at α level of significance and error degrees

of freedom. The treatment means are given by $\overline{T}_i = \dfrac{T_i}{b} \forall\, i = 1, 2, \cdots, v$. The pairwise comparison of

treatment effects can be made by comparing the difference between any two treatment means with the *LSD*. Any two treatment effects are said to differ significantly if the difference of their means is larger than the *LSD*.

2.3.2 Example 2

An initial varietal trial (Late Sown, irrigated) was conducted to study the performance of 20 new strains of mustard vis-a-vis four checks (Swarna Jyoti: ZC; Vardan: NC; Varuna: NC; and Kranti: NC) using a Randomized Complete Block Design (RCB) design at Bhatinda with 3 replications under the aegis of All India Coordinated Research Project on Rapeseed and Mustard. The seed yield in kg/ha was recorded. The details of the experiment are given in Table 2.14.

In the sequel, the data are analyzed (a) to test whether or not there is any difference among the treatment effects, (b) to make all the possible pairwise treatment comparisons to identify the best treatment *i.e.* the treatment giving highest yield, and (c) to test whether or not the average performance of check varieties (i) Swarna Jyoti (MCN-04-128), (ii) Vardan (MCN-04-129), (iii) Varuna (MCN-04-131), and (iv) Kranti (MCN-04-133) is significantly different from average performance of remaining strains.

Table 2.14: Seed yield (kg/ha) data

Treatment Number	Strain	Code	Replications		
			1	2	3
1	RK-04-3	MCN-04-110	1539.69	1412.35	1319.73
2	RK-04-4	MCN-04-111	1261.85	1065.05	1111.36
3	RGN-124	MCN-04-112	1389.19	1516.54	1203.97
4	HYT-27	MCN-04-113	1192.39	1215.55	1157.66
5	PBR-275	MCN-04-114	1250.27	1203.97	1366.04
6	HUJM-03-03	MCN-04-115	1296.58	1273.43	1308.16
7	RGN-123	MCN-04-116	1227.12	1018.74	937.71
8	BIO-13-01	MCN-04-117	1273.43	1157.66	1088.20
9	RH-0115	MCN-04-118	1180.82	1203.97	1041.90
10	RH-0213	MCN-04-119	1296.58	1458.65	1250.27
11	NRCDR-05	MCN-04-120	1122.93	1065.05	1018.74
12	NRC-323-1	MCN-04-121	1250.27	926.13	1030.32
13	RRN-596	MCN-04-122	1180.82	1053.47	717.75
14	RRN-597	MCN-04-123	1146.09	1180.82	856.67
15	CS-234-2	MCN-04-124	1574.42	1412.35	1597.57
16	RM-109	MCN-04-125	914.55	972.44	659.87
17	BAUSM-2000	MCN-04-126	891.40	937.71	798.79
18	NPJ-99	MCN-04-127	1227.12	1203.97	1389.19
19	**SWARNA JYOTI(ZC)**	**MCN-04-128**	1389.19	1180.82	1273.43
20	**VARDAN(NC)**	**MCN-04-129**	1331.31	1157.66	1180.82
21	PR-2003-27	MCN-04-130	1250.27	1250.27	1296.58
22	**VARUNA(NC)**	**MCN-04-131**	717.75	740.90	578.83
23	PR-2003-30	MCN-04-132	1169.24	1157.66	1111.36
24	**KRANTI-(NC)**	**MCN-04-133**	1203.97	1296.58	1250.27

Note: Strains of mustard in bold are the four checks.

2.3.3 Analysis of data

Treatment Totals

$T1 = 1539.69 + 1412.35 + 1319.73 = 4271.77$

$T2 = 1261.85 + 1065.05 + 1111.36 = 3438.26$

$T3 = 1389.19 + 1516.54 + 1203.97 = 4109.70$

$T4 = 1192.39 + 1215.55 + 1157.66 = 3565.60$

$T5 = 1250.27 + 1203.97 + 1366.04 = 3820.28$

$T6 = 1296.58 + 1273.43 + 1308.16 = 3878.17$

$T7 = 1227.12 + 1018.74 + 937.71 = 3183.57$

$T8 = 1273.43 + 1157.66 + 1088.20 = 3519.29$

$T9 = 1180.82 + 1203.97 + 1041.90 = 3426.69$

$T10 = 1296.58 + 1458.65 + 1250.27 = 4005.50$

$T11 = 1122.93 + 1065.05 + 1018.74 = 3206.72$

$T12 = 1250.27 + 926.13 + 1030.32 = 3206.72$

$T13 = 1180.82 + 1053.47 + 717.75 = 2952.04$

$T14 = 1146.09 + 1180.82 + 856.67 = 3183.58$

$T15 = 1574.42+ 1412.35 + 1597.57 = 4584.34$

$T16 = 914.55 + 972.44 + 659.87 = 2546.86$

$T17 = 891.40 + 937.71 + 798.79 = 2627.90$

$T18 = 1227.12 + 1203.97 + 1389.19 = 3820.28$

$T19 = 1389.19 + 1180.82 + 1273.43 = 3843.44$

$T20 = 1331.31 + 1157.66 + 1180.82 = 3669.79$

$T21 = 1250.27 + 1250.27 + 1296.58 = 3797.12$

$T22 = 717.75 + 740.90 + 578.83 = 2037.48$

$T23 = 1169.24 + 1157.66 + 1111.36 = 3438.26$

$T24 = 1203.97 + 1296.58 + 1250.27 = 3750.82$

Block Totals

$B1 = 1539.69 + 1261.85 + 1389.19 + ... + 717.75 + 1169.24 + 1203.97 = 29277.25$

$B2 = 1412.35 + 1065.05 + 1516.54 + ... + 740.90 + 1157.66 + 1296.58 = 28061.74$

$B3 = 1319.73 + 1111.36 + 1203.97 + ... + 578.83 + 1111.36 + 1250.27 = 26545.19$

Grand Total, $G = 29277.25 + 28061.74 + 26545.19 = 4271.77 + 3438.26 + 4109.70 + ... + 2037.48 + 3438.26 + 3750.82 = 83884.18$

$$\text{Correction Factor, } CF = \frac{(83884.18)^2}{72} = \frac{7036555654}{72} = 97729939.64$$

$$\text{Treatments } SS\ (SST) = \frac{(4271.77)^2 + (3438.26)^2 + \cdots + (3438.26)^2 + (3750.82)^2}{3} - CF$$

$$= 100244098.93 - 97729939.64 = 2514159.29$$

41

$$\text{Block (or Replication) } SS \text{ } (SSB) = \frac{(29277.25)^2 + (28061.74)^2 + (26545.19)^2}{24} - CF$$

$$= 97886072.15 - 97729939.64 = 156132.51$$

$$\text{Total } SS = (1539.69)^2 + (1412.35)^2 + (1319.73)^2 + (1261.85)^2 + (1065.05)^2 + \cdots + (1203.97)^2$$
$$+ (1269.58)^2 + (1250.27)^2 - CF$$

$$= 100863347.59 - 97729939.64 = 3133407.95$$

Error SS (SSE) = Total SS − Treatments SS − Replication (or Block) SS

$$= 3133407.95 - 2514159.29 - 156132.51 = 463116.15$$

We then have the analysis of variance as shown in Table 2.15.

Table 2.15: ANOVA table for the data in Example 2

Source	DF	SS	MS	F-value	Prob > F
Treatments	23	2514159.289	109311.273	10.86	<0.0001
Blocks (or Replications)	2	156132.504	78066.250	7.75	0.0013
Error	46	463116.156	10067.743		
Total	71	3133407.949			

This analysis reveals that the treatment differences are highly significant (p-value < 0.0001). Similarly, the block effects are also highly significant (p-value = 0.0013) meaning thereby that the block formation has proved to be very effective. The blocks formation was genuinely required and blocks formation has been proper.

2.3.4 Analysis using SAS

The design is a RCB design with v = 24 treatments, b = 3 blocks (or replications) and n = 72 observations. The data has been analyzed using SAS software. The commands and the data preparation are given in the sequel.

DATA rbd; /*one can enter any other name for Data*/;
INPUT trt $ 11. trtn rep syield;
*here 11. represents that the value of the variable trt is upto 11 columns;
/*trtn denotes the treatment number, rep the replication number and syield
the seed yield in kg/hectare*/;
CARDS;
MCN-04-110   1   1   1539.69
MCN-04-111   2   1   1261.85
MCN-04-112   3   1   1389.19

MCN-04-113	4	1	1192.39
MCN-04-114	5	1	1250.27
MCN-04-115	6	1	1296.58
MCN-04-116	7	1	1227.12
MCN-04-117	8	1	1273.43
MCN-04-118	9	1	1180.82
MCN-04-119	10	1	1296.58
MCN-04-120	11	1	1122.93
MCN-04-121	12	1	1250.27
MCN-04-122	13	1	1180.82
MCN-04-123	14	1	1146.09
MCN-04-124	15	1	1574.42
MCN-04-125	16	1	914.55
MCN-04-126	17	1	891.40
MCN-04-127	18	1	1227.12
MCN-04-128	19	1	1389.19
MCN-04-129	20	1	1331.31
MCN-04-130	21	1	1250.27
MCN-04-131	22	1	717.75
MCN-04-132	23	1	1169.24
MCN-04-133	24	1	1203.97
MCN-04-110	1	2	1412.35
MCN-04-111	2	2	1065.05
MCN-04-112	3	2	1516.54
MCN-04-113	4	2	1215.55
MCN-04-114	5	2	1203.97
MCN-04-115	6	2	1273.43
MCN-04-116	7	2	1018.74
MCN-04-117	8	2	1157.66
MCN-04-118	9	2	1203.97
MCN-04-119	10	2	1458.65
MCN-04-120	11	2	1065.05
MCN-04-121	12	2	926.13
MCN-04-122	13	2	1053.47
MCN-04-123	14	2	1180.82
MCN-04-124	15	2	1412.35
MCN-04-125	16	2	972.44
MCN-04-126	17	2	937.71
MCN-04-127	18	2	1203.97
MCN-04-128	19	2	1180.82
MCN-04-129	20	2	1157.66
MCN-04-130	21	2	1250.27
MCN-04-131	22	2	740.90

MCN-04-132	23	2	1157.66
MCN-04-133	24	2	1296.58
MCN-04-110	1	3	1319.73
MCN-04-111	2	3	1111.36
MCN-04-112	3	3	1203.97
MCN-04-113	4	3	1157.66
MCN-04-114	5	3	1366.04
MCN-04-115	6	3	1308.16
MCN-04-116	7	3	937.71
MCN-04-117	8	3	1088.20
MCN-04-118	9	3	1041.90
MCN-04-119	10	3	1250.27
MCN-04-120	11	3	1018.74
MCN-04-121	12	3	1030.32
MCN-04-122	13	3	717.75
MCN-04-123	14	3	856.67
MCN-04-124	15	3	1597.57
MCN-04-125	16	3	659.87
MCN-04-126	17	3	798.79
MCN-04-127	18	3	1389.19
MCN-04-128	19	3	1273.43
MCN-04-129	20	3	1180.82
MCN-04-130	21	3	1296.58
MCN-04-131	22	3	578.83
MCN-04-132	23	3	1111.36
MCN-04-133	24	3	1250.27

```
;
RUN;

PROC GLM ;
CLASS trtn rep;
MODEL syield = trtn rep;
LSMEANS trtn/PDIFF LINES;
CONTRAST 'check vs strains' trtn 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 -20 -20 4 -20 4 -20;
RUN;
```

In order to compare the check varieties with the strains, the null hypothesis to be tested is that the average effect of strains is same as the average effect of check varieties. The null hypothesis H_0: $4\tau_1 + 4\tau_2 + ... + 4\tau_{17} + 4\tau_{18} + 4\tau_{21} + 4\tau_{23} - 20\tau_{19} - 20\tau_{20} - 20\tau_{22} - 20\tau_{24} = 0$ is tested against

$$H_1: 4\tau_1 + 4\tau_2 + ... + 4\tau_{17} + 4\tau_{18} + 4\tau_{21} + 4\tau_{23} - 20\tau_{19} - 20\tau_{20} - 20\tau_{22} - 20\tau_{24} \neq 0.$$

For testing the hypothesis H_0, one needs to perform contrast analysis. The problem of contrast analysis has been dealt with in Chapter 3. However, here we have given the SAS steps to perform the contrast analysis.

The output of analysis using SAS is given in Table 2.16.

Table 2.16: Output of analysis using SAS

ANOVA

Source	DF	SS	MS	F-value	Prob > F
Model	25	2670291.793	106811.672	10.61	<0.0001
Error	46	463116.156	10067.743		
Corrected Total	71	3133407.949			

R-square	CV	Root MSE	Yield Mean
0.852	8.612	100.34	1165.06

ANOVA

Source	DF	SS	MS	F-value	Prob > F
Treatments	23	2514159.289	109311.273	10.86	<0.0001
Blocks	2	156132.504	78066.252	7.75	0.0013
Error	46	463116.156	10067.743		
Corrected Total	71	3133407.949			

The model with treatment effects and block effects explains about 85 per cent of the total variability in the data. The treatment effects are highly significant (p-value < 0.0001) meaning thereby that the null hypothesis is rejected. It is interesting to note that the block effects are also highly significant (p-value = 0.0013).

The mean and standard deviation of the treatments are given in Table 2.17.

Table 2.17: Treatment wise mean and standard deviation of seed yield

Level of treatment	N	SYIELD	
		Mean	Standard Deviation
1	3	1423.92	110.44
2	3	1146.09	102.89
3	3	1369.90	157.18
4	3	1188.53	29.14
5	3	1273.43	83.48
6	3	1292.72	17.68
7	3	1061.19	149.30

8	3	1173.10	93.57
9	3	1142.23	87.66
10	3	1335.17	109.42
11	3	1068.91	52.20
12	3	1068.91	165.48
13	3	984.01	239.22
14	3	1061.19	177.97
15	3	1528.11	100.92
16	3	848.95	166.29
17	3	875.97	70.73
18	3	1273.43	100.92
19	3	1281.15	26.74
20	3	1223.26	30.63
21	3	1265.71	104.40
22	3	679.16	94.28
23	3	1146.09	87.66
24	3	1250.27	46.31

The distribution of observations over replications for each treatment is given in Figure 2.2.
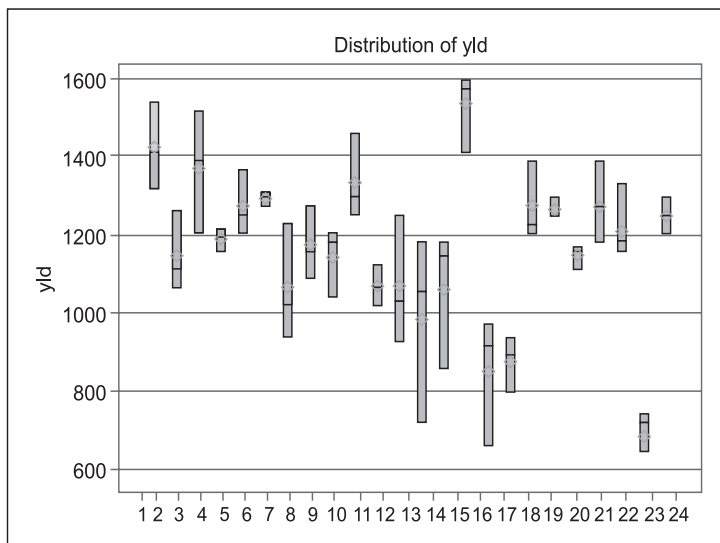
Figure 2.2: Treatment wise Box plot of seed yield

Similarly, the Figure 2.3 gives the plot of observations in each block. It is quite evident from

here that the blocks differ in their effects, *i.e.* mean square between blocks is high as compared to mean square error, a fact supported by the ANOVA as well.
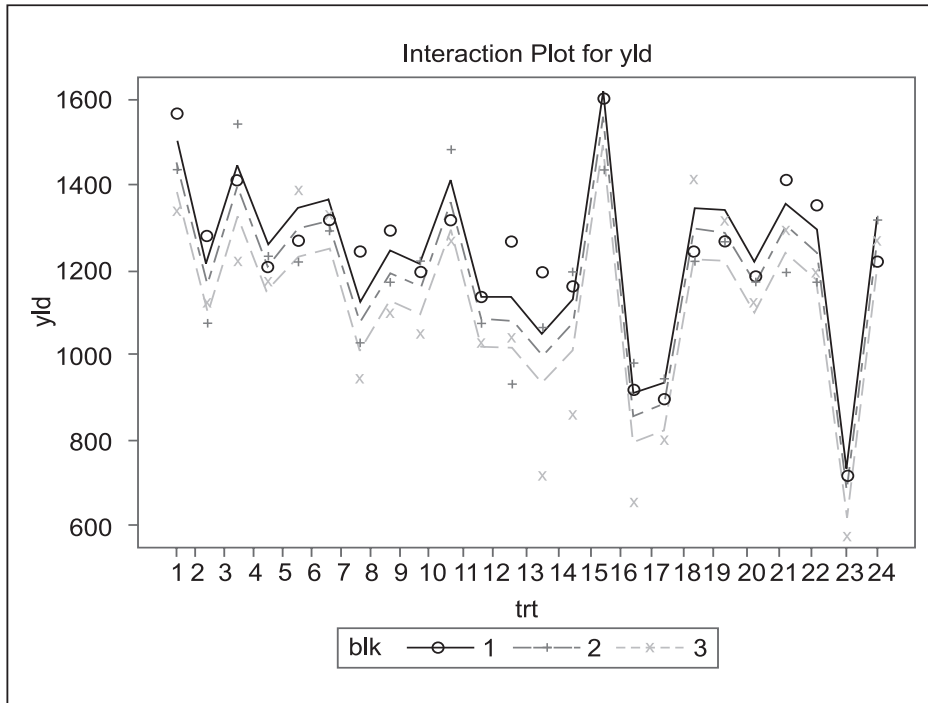
Figure 2.3: Plot of observations in each block

The pairwise comparison of treatment effects is made and is presented in Table 2.18. Treatments having at least one letter common are not significantly different in their effects. The strain CS-234-2 and coded as MCN-04-124 (treatment 15) is the highest seed yielding strain. This strain produces significantly higher seed yield than all other treatments produce except treatment numbers 1 and 3, which produce seed yield statistically at par with the produce of treatment 15. So this strain may be recommended as the best among the lot.

Table 2.18: Treatments in descending order with letter display

Means with the same letter are not significantly different							
Duncan Grouping					Mean	Treatment Number	Replication
A					1528.11	15	3
A	B				1423.92	1	3
A	B	C			1369.90	3	3
	B	C	D		1335.17	10	3
	B	C	D		1292.72	6	3
	B	C	D		1281.15	19	3
	B	C	D		1273.43	18	3
	B	C	D		1273.43	5	3
	B	C	D		1265.71	21	3
	B	C	D	E	1250.27	24	3
		C	D	E	1223.26	20	3
		C	D	E	1188.53	4	3
F			D	E	1173.10	8	3
F			D	E	1146.09	23	3
F			D	E	1146.09	2	3
F			D	E	1142.23	9	3
F				E	1068.91	11	3
F				E	1068.91	12	3
F				E	1061.19	14	3
F				E	1061.19	7	3
F	G				984.01	13	3
	G				875.97	17	3
	G				848.95	16	3
		H			679.16	22	3

From Table 2.18 it is also evident that check variety (Treatment 19: best performing check) is significantly different from strains at Treatment 7, 11, 12, 13, 14, 15, 16, 17. Similarly, check variety (Treatment 20) is significantly different from strains at Treatment 1, 13, 15, 16, 17. Further, check variety (Treatment 24) is significantly different from strains at Treatment 13, 15, 16, 17. The check variety (Treatment 22) is, however, the lowest yielding and is significantly different from all the strains.

The contrast analysis for testing the null hypothesis that the average effect of strains is same as the average effect of check varieties was done and the result is given in Table 2.19.

Table 2.19: Result of contrast analysis

Contrast	DF	Type III SS	MS	F-value	Prob > F
Check vs Strains	1	46126.736	46126.736	4.58	0.0377

It may be noted that the check varieties differ significantly from the strains (p-value = 0.0377).

The pairwise treatment comparisons can also be presented without writing the treatments in descending order of the treatments LS Mean values. The results are given in Table 2.20.

Table 2.20: Treatments with letter display

Treatment Name	LS Mean of Syield	Rank of Treatment	Treatment Name	LS Mean of Syield	Rank of Treatment
1	$1423.92^{A,B}$	2	13	$984.01^{G,H}$	21
2	$1146.09^{E,F,G}$	15	14	$1061.19^{F,G}$	19
3	$1369.90^{A,B,C}$	3	15	1528.11^{A}	1
4	$1188.53^{D,E,F}$	12	16	848.95^{H}	23
5	$1273.43^{B,C,D,E}$	7	17	875.97^{H}	22
6	$1292.72^{B,C,D,E}$	5	18	$1273.43^{B,C,D,E}$	8
7	$1061.19^{F,G}$	20	19	$1281.15^{B,C,D,E}$	6
8	$1173.10^{D,E,F}$	13	20	$1223.26^{C,D,E,F}$	11
9	$1142.23^{E,F,G}$	16	21	$1265.71^{B,C,D,E}$	9
10	$1335.17^{B,C,D}$	4	22	679.16^{I}	24
11	$1068.91^{F,G}$	17	23	$1146.09^{E,F,G}$	14
12	$1068.91^{F,G}$	18	24	$1250.27^{C,D,E}$	10
General Mean			1165.06		
$SE(\hat{d})$			81.927		
LSD at 5%			164.91		

It may be mentioned here that the SAS commands given in Section 2.3.4 do not compute LSD at 5%. This may, therefore, be computed by using the formula given in Section 2.3.1. One may also compute it by adding a SAS command "MEANS trtn/LSD;". In Table 2.20, any two treatments whose LS Means have at least one letter common are not statistically significant using LSD. Therefore, it follows that treatment 15 is the one that produces highest seed yield and is not significantly different from treatments 1 and 3. It is significantly different from all the remaining treatments. Similarly, treatment 22, a control variety, produces the lowest seed yield and has in fact statistically significant lower seed yield from all other strains. The other three control varieties (Treatments 19, 20 and 24) are statistically at par with each other.

2.3.5 Analysis of RCB design using R

R code

```
d2=read.table("rbd.txt",header=TRUE)
attach(d2)
names(d2)
#Treatment means and standard deviations
aggregate(syield, by=list(trt), mean)
aggregate(syield, by=list(trt), sd)
#Treatment wise box plot of yield
boxplot(syield~trtn)
#ANOVA
trtn=factor(trtn)
rep=factor(rep)
aov.out=aov(syield~trtn+rep)
summary(aov.out)
library(lsmeans)
lsm <- lsmeans(aov.out, "trtn")
contrast(lsm, list(con1 = c(4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,-20,-20,4,-20,4,-20)))
#Tukey's honest significant difference test
TukeyHSD(rbdout)
#LSD
library(agricolae)
lsd.result <- LSD.test(aov.out,"trtn")
lsd.result
detach(d2)
```

2.4 Latin square design

A CRD assumes that there is no variability in the experimental units. The only source of variability in the data is the treatments and the remaining variability is the error. On the other hand, an RCB design assumes that other than the treatments, there is one source of variability in the experimental units and this variability in the experimental units is controlled by forming blocks of homogeneous experimental units. In this case, the sources of variability in the data are the treatments and the blocks (or replications) and the remaining part of the variability is the experimental error. This section is devoted to designs which control two sources of variability in the experimental units. When there are two sources of variability in the experimental units, we need to form blocks in two directions, perpendicular to each other. The two blocking systems are cross classified as rows and columns and the intersection of rows and columns is a cell or the experimental unit. Following on the example of four grazing systems and 16 experimental units (animals), one source of variability in the animals could be the initial body weight. The other source of variability could be their physiological behavior. For instance, the calving age or the number of lactations could be another source of variability in the experimental material. The physiological behavior and the initial body weights are the two sources of variability in

the animals and need to be controlled by proper designing of experiment. In the insecticide field trial where the insect migration has a predictable direction that is perpendicular to the dominant fertility gradient of the experimental field, there are two sources of variability in the experimental units. In order to control two-way heterogeneity in the experimental material, we use designs known as *Latin Square* obtained from Latin square arrangement. The following are examples of 4×4 and 5×5 Latin square designs:

$$
\begin{array}{cccc}
A & B & C & D \\
B & C & D & A \\
C & D & A & B \\
D & A & B & C
\end{array}
\; ; \;
\begin{array}{ccccc}
A & B & C & D & E \\
B & C & D & E & A \\
C & D & E & A & B \\
D & E & A & B & C \\
E & A & B & C & D
\end{array}
\; ; \;
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
3 & 4 & 1 & 2 \\
4 & 3 & 2 & 1 \\
2 & 1 & 4 & 3
\end{array}
$$

In such designs two restrictions are imposed by forming blocks in two directions, row-wise and column-wise. A Latin square arrangement is an arrangement of *v* Latin letters in a *v*×*v* square in such a way that each row and each column has all the *v* Latin letters appearing exactly once. A design based upon a Latin square arrangement is called a Latin square design. Ignoring rows and considering columns as blocks gives an RCB design. Similarly, ignoring columns and treating rows as blocks gives an RCB design. So a Latin square design is an RCB design in rows as well as columns. Treatments are allocated in such a way that every treatment occurs once and only once in each row and each column. In this design, the replication number of treatments is same as the number of treatments.

Latin squares have been classified as reduced and standard. The Latin squares have also been classified as squares with normalized or standard and semi-standard form, whereby reduced Latin square is synonym to normalized or standard form and standard Latin square is synonym to semi-standard form. Latin square is considered reduced if its first row and first column contains elements in the numerical (1,2,…,*v*) or lexicographic order (A,B,C,…). On the other hand, it is considered standard if only its first row contains elements in the natural order. In the examples of Latin squares given earlier, the first two Latin squares of order 4 and 5, respectively are the reduced squares (or Latin squares in standard form) while the third Lain square of order 4 is in standard (or semi-standard) form.

The randomization of the *v* treatments over the v^2 experimental units arranged in a *v*×*v* square is difficult. The design obtained after randomization should be a Latin square design. In actual field arrangement during experimentation, first we select a *v*×*v* reduced or normalized or standard Latin square randomly from the Fisher and Yates Tables. Having selected the square, column-wise randomization is carried out first, followed by row-wise randomization. Of course, the treatments labels (or the Latin letters) are randomized separately before starting the actual randomization in the design.

The linear model in this case is

Expected response = general mean + effect of treatments + effect due to experimental units

Since there are two sources of variation in the experimental material, the model can be rewritten as

Expected response = general mean + effect of treatments+ effect of rows + effect of columns

This can also be written as:

Response = general mean + treatments effect + rows effect + columns effect + error,

where the errors are distributed independently as normal variate with zero mean and constant variance σ^2. The partitioning of the total variability in this case is

Source of Variation
Due to model
Error
Total

The component "due to model" can be split as

Source of Variation
Due to model
Due to Treatments
Due to Rows
Due to Columns

2.4.1 Analysis of Latin square design

The v^2 observations generated from a Latin square design of order v are represented by the following linear, additive, fixed effects model:

$$y_{i(j,k)} = \mu + \tau_i + \beta_j + \gamma_k + e_{i(j,k)}; \ \ i = 1,2,\cdots,v; j = 1,2,\cdots,v; k = 1,2,\cdots,v,$$

where $y_{i(j,k)}$ is the observation pertaining to the ith treatment appearing in the (j, k)th cell, μ is the grand mean, τ_i is the ith treatment effect, β_j is the effect of the jth row, γ_k is the effect of the kth column, and $e_{i(j,k)}$ is the random error associated with $y_{i(j,k)}$, assumed to be mutually independent and distributed normally with mean zero and common variance σ^2.

Let the treatment totals, rows totals, column totals be denoted as, respectively,

$$T_i = \sum_{(j,k) \supset i} y_{i(j,k)}, \ \forall\, i = 1,2,\cdots,v;\ \text{sum of observations over cells containing treatment } i;$$

$$R_j = \sum_{i=1}^{v}\sum_{k=1}^{v} y_{i(j,k)}, \ \forall\, j = 1,2,\cdots,v;\ \text{sum of observations in the } j\text{th row;}$$

$C_k = \sum_{i=1}^{v} \sum_{j=1}^{v} y_{i(j,k)}$, $\forall\, k = 1,2,\cdots, v$; sum of observations in the kth column.

The grand total is, $G = \sum_{i=1}^{v} T_i = \sum_{j=1}^{v} R_j = \sum_{k=1}^{v} C_k = \sum_{i=1}^{v} \sum_{j=1}^{v} \sum_{k=1}^{v} y_{i(j,k)}$

The following formulae can be employed for analysis of variance:

Correction factor (CF) $= \dfrac{G^2}{v^2}$

Total sum of squares $(SS) = \sum_{i=1}^{v} \sum_{j=1}^{v} \sum_{k=1}^{v} y_{i(j,k)}^2 - CF$

Sum of Squares due to treatments $(SST) = \dfrac{1}{v} \sum_{i=1}^{v} T_i^2 - CF$

Sum of Squares due to rows $(SSR) = \dfrac{1}{v} \sum_{j=1}^{v} R_j^2 - CF$

Sum of Squares due to columns $(SSC) = \dfrac{1}{v} \sum_{k=1}^{v} C_k^2 - CF$

Error sum of squares (SSE)

$$= \sum_{i=1}^{v} \sum_{j=1}^{v} \sum_{k=1}^{v} y_{i(j,k)}^2 - \frac{1}{v} \sum_{i=1}^{v} T_i^2 - \frac{1}{v} \sum_{j=1}^{v} R_j^2 - \sum_{k=1}^{v} C_k^2 + 2(CF)$$

$$= \left(\sum_{i=1}^{v} \sum_{j=1}^{v} \sum_{k=1}^{v} y_{i(j,k)}^2 - CF \right) - \left(\frac{1}{v} \sum_{i=1}^{v} T_i^2 - CF \right) - \left(\frac{1}{v} \sum_{j=1}^{v} R_j^2 - CF \right) - \left(\frac{1}{v} \sum_{k=1}^{v} C_k^2 - CF \right)$$

= Total SS – Row SS – Column SS – Treatment SS

If the calculated value of F is greater than the table value of $F_{\alpha;(v-1),(v-1)(v-2)}$ at α level of significance and $(v-1),(v-1)(v-2)$ degrees of freedom, then the null hypothesis H_0 is rejected at α level of significance and it can be concluded that the treatment effects are significantly different from one another.

It may be seen here that an unbiased estimator of σ^2 is $\hat{\sigma}^2 = s_e^2$.

Further, all the elementary treatment contrasts are estimable through the design. The Best Linear Unbiased Estimator (BLUE) of any treatment contrast $d_{il} = \tau_i - \tau_l$ is

$$\hat{d}_{il} = \frac{T_i - T_l}{v}, \forall\, i \ne l = 1,2,\cdots, v.$$

Table 2.21: ANOVA table in Latin square design

Source	DF	SS	MS	F
Treatments	$v-1$	$SST = \dfrac{1}{v}\sum\limits_{i=1}^{v} T_i^2 - CF$	$\dfrac{SST}{(v-1)} = s_t^2$	$\dfrac{s_t^2}{s_e^2}$
Rows	$v-1$	$SSR = \dfrac{1}{v}\sum\limits_{j=1}^{v} R_j^2 - CF$	$\dfrac{SSR}{(v-1)} = s_r^2$	$\dfrac{s_r^2}{s_e^2}$
Columns	$v-1$	$SSC = \dfrac{1}{v}\sum\limits_{k=1}^{v} C_k^2 - CF$	$\dfrac{SSC}{(v-1)} = s_c^2$	$\dfrac{s_c^2}{s_e^2}$
Error	$(v-1)(v-2)$	$SSE = \sum\limits_{i=1}^{v}\sum\limits_{j=1}^{v}\sum\limits_{k=1}^{v} y_{i(j,k)}^2 - \dfrac{1}{v}\sum\limits_{i=1}^{v} T_i^2$ $-\dfrac{1}{v}\sum\limits_{j=1}^{v} R_j^2 - \sum\limits_{k=1}^{v} C_k^2 + 2(CF)$	$\dfrac{SSE}{(v-1)(v-2)}$ $= s_e^2$	
Total	v^2-1	$\sum\limits_{i=1}^{v}\sum\limits_{j=1}^{v}\sum\limits_{k=1}^{v} y_{i(j,k)}^2 - CF$		

The variance of $\hat{d}_{il}$ is $\mathrm{Var}\!\left(\hat{d}_{il}\right) = \dfrac{2\sigma^2}{v}$. The estimated standard error of the difference between the estimated ith and lth treatment effects is $SE\!\left(\hat{d}_{il}\right) = \left(\dfrac{2s_e^2}{v}\right)^{1/2}$.

The Least Significant Difference (*LSD*) is given as $LSD = SE\!\left(\hat{d}_{il}\right) \times t_{\alpha;\,errorDF}$.

Here $t_{\alpha/2;\,error\,DF}$ denotes the value of Student's t at $\alpha/2$ level of significance and error degrees of freedom. The treatment means are given by $\bar{T}_i = \dfrac{T_i}{v}\ \forall\, i = 1,2,\cdots,v$. The pairwise comparison of treatment effects can be made by comparing the difference between any two treatment means with the *LSD*. Any two treatment effects are said to differ significantly if the difference of their means is larger than the *LSD*.

2.4.2 Example 3

An experiment was conducted at Agricultural Research Station, Kopurgaon, Maharashtra on Cotton using a Latin Square Design to study the effects of foliar application of urea in combination with insecticidal sprays on the cotton yield. The 6 treatments were {T_1: Control (*i.e.* no N and no insecticides), T_2: 100kg N/ha applied as urea (half at final thinning and half at flowering as top dressing), T_3: 100kg N/ha applied as urea (80 kg N/ha in 4 equal split doses as spray and 20 kg N/ha at final thinning), T_4: 100 kg. N/ha applied as CAN (half at final thinning and half at flowering as top dressing), T_5: T_2+ six insecticidal sprays, T_6: T_4 + six insecticidal

sprays}. There were 6 replications, and the data of Cotton yield in kg per plot is given in Table 2.22.

Table 2.22: Cotton yield data (kg/plot)

T_3 3.10	T_6 5.95	T_1 1.75	T_5 6.40	T_2 3.85	T_4 5.30
T_2 4.80	T_1 2.70	T_3 3.30	T_6 5.95	T_4 3.70	T_5 5.40
T_1 3.00	T_2 2.95	T_5 6.70	T_4 5.95	T_6 7.75	T_3 7.10
T_5 6.40	T_4 5.80	T_2 3.80	T_3 6.55	T_1 4.80	T_6 9.40
T_6 5.20	T_3 4.85	T_4 6.60	T_2 4.60	T_5 7.00	T_1 5.00
T_4 4.25	T_5 6.65	T_6 9.30	T_1 4.95	T_3 9.30	T_2 8.40

In the sequence, the data are analyzed (a) to identify the best treatment, (b) to test whether or not the average effect of T_3 (100kg N/ha applied as urea) and T_4 (100 kg N/ha) is same as the average effect of T_5 (T_2 + six insecticidal sprays) and T_6 (T_4 +six insecticidal sprays).

2.4.3 Analysis of data

We compute the following totals in Table 2.23.

Table 2.23: Treatment, row and column totals

Treatments Totals (T_i)	Rows Totals (R_j)	Columns Totals (C_k)
$T_1 = 1.75 + ... + 4.95 = 22.20$	$R_1 = 3.10 + ... + 5.30 = 26.35$	$C_1 = 3.10 + ... + 4.25 = 26.75$
$T_2 = 3.85 + ... + 8.40 = 28.40$	$R_2 = 4.80 + ... + 5.40 = 25.85$	$C_2 = 5.95 + ... + 6.65 = 28.90$
$T_3 = 3.10 + ... + 9.30 = 34.20$	$R_3 = 3.00 + ... + 7.10 = 33.45$	$C_3 = 1.75 + ... + 9.30 = 31.45$
$T_4 = 5.30 + ... + 4.25 = 31.60$	$R_4 = 6.40 + ... + 9.40 = 36.75$	$C_4 = 6.40 + ... + 4.95 = 34.40$
$T_5 = 6.40 + ... + 6.65 = 38.55$	$R_5 = 5.20 + ... + 5.00 = 33.25$	$C_5 = 3.85 + ... + 9.30 = 36.40$
$T_6 = 5.95 + ... + 9.30 = 43.55$	$R_6 = 4.25 + ... + 8.40 = 42.85$	$C_6 = 5.30 + ... + 8.40 = 40.60$

Grand Total, $G = \sum_{i=1}^{v} T_i = \sum_{j=1}^{v} R_j = \sum_{k=1}^{v} C_k = 198.50$

Correction factor (CF) $= \dfrac{G^2}{v^2} = \dfrac{(198.5)^2}{(6)^2} = \dfrac{39402.25}{36} = 1094.506$

Treatments SS $= \dfrac{1}{v} \sum_{i=1}^{v} T_i^2 - CF$

$= \{(22.20)^2 + ... + (43.55)^2\} / 6 - 1094.506 = 47.211$

Rows SS $= \frac{1}{v} \sum\limits_{j=1}^{v} R_j^2 - CF$

$= \{(22.20)^2 + ... + (43.55)^2\}/6 - 1094.506 = 47.211$

Columns SS $= \frac{1}{v} \sum\limits_{k=1}^{v} C_k^2 - CF$

$= \{(26.75)^2 + ... + (40.6)^2\}/6 - 1094.506 = 21.586$

Total $SS = \sum\limits_{i=1}^{v}\sum\limits_{j=1}^{v}\sum\limits_{k=1}^{v} y_{ijk}^2 - CF$

$= (3.10)^2 + (5.95)^2 + ... + (9.30)^2 + (8.40)^2 - 1094.506 = 128.333$

Error SS = Total SS – Treatments SS – Rows SS – Columns SS

$= 128.333 - 47.210 - 34.442 - 21.585 = 25.094$

We now form the Analysis of Variance Table 2.24.

Table 2.24: ANOVA table for cotton yield data

Source	DF	SS	MS	F-value	Prob > F
Treatments	5	47.211	9.442	7.53	0.0004
Rows	5	34.442	6.883	5.49	0.0024
Columns	5	21.586	4.317	3.44	0.0210
Error	20	25.095	1.255		
Corrected Total	35	128.333			

From Table 2.24, one can easily see that the treatment effects are highly significant (*p*-value = 0.0004) meaning thereby that the null hypothesis of equal treatment effects is rejected. The treatments, therefore, influence the cotton yield. The rows and columns effects are also highly significant with respective *p*-values as 0.0024 and 0.0210. This is an evidence to the fact that the formation of rows and columns have been effective.

2.4.4 Analysis using SAS

The design is a LSD with $v = 6$ treatments and $n = 36$ observations. In this design the number of rows is same as the number of columns, which in turn is same as the number of treatments. So in this design the replication number of treatments is equal to the number of treatments. The data has been analyzed using SAS. The commands and the data preparation are given in the sequel.

DATA lsd;

INPUT row col trt cyield;
/*the first column 'row' denotes the row number; the second column 'col' denotes the column number; the third column "trt' represents the treatment number and the last column 'cyield' represents the cotton yield*/
CARDS;
1 1 3 3.10
1 2 6 5.95
1 3 1 1.75
1 4 5 6.40
1 5 2 3.85
1 6 4 5.30
2 1 2 4.80
2 2 1 2.70
2 3 3 3.30
2 4 6 5.95
2 5 4 3.70
2 6 5 5.40
3 1 1 3.00
3 2 2 2.95
3 3 5 6.70
3 4 4 5.95
3 5 6 7.75
3 6 3 7.10
4 1 5 6.40
4 2 4 5.80
4 3 2 3.80
4 4 3 6.55
4 5 1 4.80
4 6 6 9.40
5 1 6 5.20
5 2 3 4.85
5 3 4 6.60
5 4 2 4.60
5 5 5 7.00
5 6 1 5.00
6 1 4 4.25
6 2 5 6.65
6 3 6 9.30
6 4 1 4.95
6 5 3 9.30
6 6 2 8.40
;

```
PROC GLM data = lsd;
CLASS row col trt;
MODEL cyield = trt row col;
MEANS trt/tukey;
CONTRAST 'T3 T4 vs T5 T6' trt 0 0 1 1 −1 −1;
RUN;
```

2.4.5 Output of analysis

The results obtained from the analysis of data are described in the sequel.

Table 2.25: Output using SAS

ANOVA

Source	DF	SS	MS	F Value	Prob > F
Model	15	103.238	6.882	5.49	0.0003
Error	20	25.095	1.255		
Corrected Total	35	128.333			

R-Square	CV	Root MSE	cyield Mean
0.804	20.315	1.120	5.514

ANOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Treatment	5	47.211	9.442	7.53	0.0004
Row	5	34.442	6.888	5.49	0.0024
Column	5	21.586	4.317	3.44	0.0210
Error	20	25.095	1.255		
Corrected Total	35	128.33			

It is worthwhile noting that the model with treatments effects, row effects and column effects explains about 80 per cent of the total variability in the data. As mentioned earlier also, this analysis of variance table divulges that the treatment effects are highly significant (p-value = 0.0004) meaning thereby that the null hypothesis of equal treatment effects is rejected. The rows and columns effects are also highly significant with respective p-values as 0.0024 and 0.0210. So running this experiment as a row-column design is justified and it is very apparent that there were two sources of variability in the experimental units.

The distribution of observations for each treatment is given Figure in 2.4. It is easily seen from the Figure also that the distribution of observations is very different for different treatments. The Figure clearly reveals that treatment number 3 is most variable and treatment number 5 is least variable.
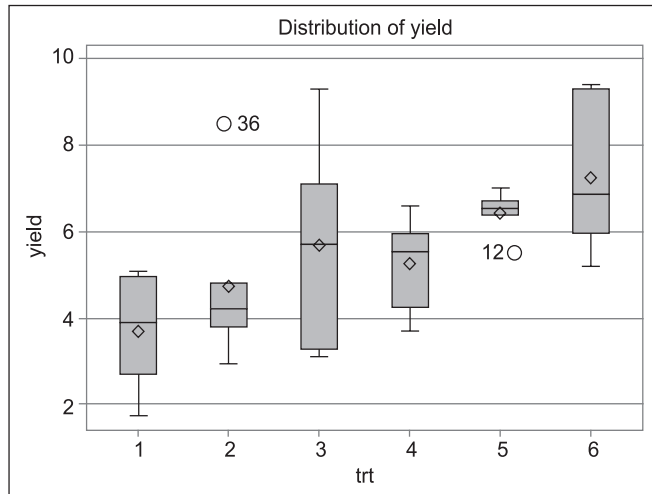
Figure 2.4: Treatment wise Box plot of seed yield

A pairwise comparison of treatment effects is made and the results are summarized in Table 2.26. Treatment 6 is the maximum yielding and is significantly different from treatment 1, which is the lowest yielding. Treatment 6 is, however, statistically at par with treatments 5, 3 and 4. Similarly, treatment 1 is statistically at par with treatments 2, 4 and 3. On the basis of yield, treatment 6 may be recommended as the best for cotton yield.

Table 2.26: Treatments in descending order with letter display

Means with the same letter are not significantly different					
Tukey Grouping			**Mean**	**N**	**Treatment**
	A		7.258	6	6
B	A		6.425	6	5
B	A	C	5.700	6	3
B	A	C	5.267	6	4
B		C	4.733	6	2
		C	3.700	6	1

A comparison of treatments 3 and 4 with treatments 5 and 6 (null hypothesis that the average effect of treatments 3 and 4 is the same as the average effect of treatments 5 and 6) reveals that the difference is significant (p-value = 0.0076).

Table 2.27: Contrast analysis result

Contrast	DF	Contrast SS	Mean Square	F Value	Prob > F
3 4 vs 5 6	1	11.070	11.070	8.82	0.0076

The pairwise treatment comparisons can also be made without writing the treatments in descending order of the treatments LS Mean values. The results are given in Table 2.28.

Table 2.28: Treatments with letter display

Treatment Name	Treatment Description	Mean of 'cyield'	Rank of Treatment
1	Control	3.70^C	6
2	100kg *N*/ha applied as urea (half at final thinning and half at flowering as top dressing)	4.73^{B,C}	5
3	100kg *N*/ha applied as urea (80 kg *N*/ha in 4 equal split doses as spray and 20 kg *N*/ha at final thinning)	5.70^{A,B,C}	3
4	100 kg. *N*/ha applied as CAN (half at final thinning and half at flowering as top dressing)	5.27^{A,B,C}	4
5	$T_5 : T_2$ + six insecticidal sprays	6.43^{A,B}	2
6	T_4 + six insecticidal sprays	7.26^A	1
General Mean		5.51	
$SE(\hat{d})$		0.647	
Tukey HSD at 5%		2.033	

In Table 2.28, any two treatments whose Means have at least one letter common are not statistically significant using Fisher's Least Square Difference. Therefore, it follows that treatment 1 is significantly different from treatments 3, 4, 5 and 6, but is not significantly different from treatment 2. Similarly, treatment 6 is significantly different from treatments 1, 2, 3 and 4, but is not significantly different from treatment 5. Also treatment 2 is significantly different from treatments 5 and 6. Following the Table 2.28, it is also evident that treatment 3 is significantly different from treatments 1 and 6. Treatment 4, however, is not significantly different from treatments 2, 3 and 5.

Further, 100 kg. *N*/ha applied as CAN (half at final thinning and half at flowering as top dressing) along with six insecticidal sprays produces the maximum yield, though it is at par with 100kg *N*/ha applied as urea (half at final thinning and half at flowering as top dressing) coupled with six insecticidal sprays. It may be worthwhile mentioning here that all comparisons are made at 5 per cent level of significance.

2.4.6 Analysis using R

R code
```
d3=read.table("lsd.txt",header=TRUE)
attach(d3)
names(d3)
#Treatment means and standard deviations
aggregate(cyield, by=list(trt), mean)
aggregate(cyield, by=list(trt), sd)
#Treatment wise box plot of yield
boxplot(cyield~trt)
```

#set the contrast coefficients for testing average of treatments 3 and 4 with average of treatments 5 and 6
```
trt=factor(trt)
row=factor(row)
col=factor(col)
aov.out=aov(cyield~trt+row+col)
summary(aov.out)
library(lsmeans)
lsm <- lsmeans(aov.out, "trt")
contrast(lsm, list(con1 = c(0,0,1,1,-1,-1)))
#Tukey's honest significant diffence test
TukeyHSD(aov.out,"trt")
#LSD
library(agricolae)
lsd.result <- LSD.test(aov.out,"trt")
lsd.result
detach(d3)
```

2.5 Conclusion

This Chapter has been devoted to introducing the basic designs like CRD, RCB design and Latin square design. SAS has been used for the analysis of data. The PROC GLM has been the major procedure used for analysis of data. The R code for the analysis of data has also been given.

It has been observed that in many experiments conducted as an RCB design (very few experiments are conducted as Latin square design), the block mean square is not high as compared to mean square error. In other words, block mean square is smaller than the error mean square. This is not a healthy situation. The basic purpose of forming blocks (or two systems of blocks as in Latin square design) is that there was variability in the experimental units. It is expected that the between blocks variability would be large and the within block variability would be small. But if the block effects are not significant, it means that substantial part of variability arising in the experimental units has not been accounted for by forming blocks. Obviously then the CV would also be large.

It may be re-emphasized that the variability in the experimental units is a very disturbing factor and it needs to be taken care of properly so as to enable a proper conduct of experiment.

For the benefit of the experimenters, a utility has been created at the "Design Resources Server" hosted at www.iasri.res.in/design to generate a randomized layout of these basic designs. There is also a provision for generating a data entry sheet based on the randomized plan either in TXT (Text file) or CSV (Comma Separated Values) formats. CSV/TXT files can be opened using any text editor or in MS®-Excel®. The experimenter may use "Datasheet" hyperlink for downloading / opening generated datasheet. Besides randomized layout, an outline of ANOVA is also shown for the benefit of the experimenters. The users may visit http://iasri.res.in/design/ Basic Designs/basicdesign.aspx and take advantage of this utility.

3

Contrast Analysis

3.1 Introduction

The main technique adopted for the analysis and interpretation of the data collected from an experiment is the analysis of variance (ANOVA). This technique essentially consists of partitioning the total variation in an experimental data into components ascribable to different sources of variation due to the controlled factors and error. A standard analysis of variance provides an *F*-test, which is called an *omnibus test*, because it reflects all possible differences between the means of the groups analyzed by the ANOVA. The hypothesis generally tested using the *F*-statistic in analysis of variance is that all the treatment effects are same against an alternative that at least one treatment effect is different from others. The objective of an experiment is often much more specific than merely determining whether or not all of the treatment effects are same and are expected to give rise to similar responses. Precise conclusions can be obtained from *contrast analysis* because a contrast expresses a specific question about the pattern of results of an ANOVA.

When performing a contrast analysis we need to distinguish whether the contrasts are *planned* or *post hoc*. *Planned* or *a priori* contrasts are selected before running the experiment. The design is chosen as per the planned contrasts of interest. In general, they reflect the hypotheses the experimenter wanted to test and there are usually *few* of them. *Post hoc* or *a posteriori* (after the fact) contrasts are decided after the experiment has been run. The goal of *a posteriori* contrasts is to ensure that unexpected results, if any, are reliable.

3.1.1 Examples

Generally, all possible pairwise treatment comparisons need to be made. Similarly it may be of interest to test if the average effect of subgroup of treatments is equal to the average effect of another subgroup of treatments. Some examples are in order. A medical experimenter is concerned with the efficacy of each of several new drugs as compared to a standard drug. A nutrition experiment may be run to compare high fiber diets with low fiber diets. A plant breeder may be interested in comparing exotic collections with indigenous cultivars. An agronomist may be interested in comparing the effects of biofertilisers and chemical fertilizers. A water technologist may be interested in studying the effect of nitrogen with farmyard manure (FYM) over the nitrogen levels without FYM in presence of irrigation. A forestry scientist may be interested in comparing the various tree species in terms of timber volume, the tree species being the ones good from fuel point of view, fodder point of view or timber volume point of view.

In order to answer these types of questions, one has to look beyond analysis of variance. Contrast analysis is an answer to these types of questions. As a matter of fact, almost all the questions of the experimenters can be answered through contrast analysis. The inference problem to be solved is translated in the form of a contrast and then the contrast analysis is done to answer the questions. Before we describe the contrast analysis, we define a contrast.

3.2 Contrasts

Let $\tau_1, \tau_2, \cdots, \tau_v$ denote the v parameters (or treatment effects). Let

$$B = p_1\tau_1 + p_2\tau_2 + \cdots + p_v\tau_v$$

be a linear function of $\tau_1, \tau_2, \cdots, \tau_v$. Here $p_1, p_2, \cdots, p_v$ are arbitrary real numbers such that

$$p_1 + p_2 + \cdots + p_v = 0 .$$

Such a function B is called a *contrast* or a treatment contrast and $p_1, p_2, \cdots, p_v$ are known as the coefficients of the contrast. A contrast is not unique because the choice of $p_1, p_2, \cdots, p_v$ is arbitrary. For example, if there are three treatment effects, then $\tau_1 + \tau_2 - 2\tau_3$ is a contrast because the coefficients of τ_1, τ_2 and τ_3 are 1, 1 and −2, which satisfy 1 + 1 − 2 = 0. Similarly, another example of contrast is $\tau_1 - \tau_3$, yet another is $\tau_1 - \tau_2$.

Let

$$C = q_1\tau_1 + q_2\tau_2 + \cdots + q_v\tau_v$$

be another contrast. As above, once again $q_1, q_2, \cdots, q_v$ are arbitrary real numbers and $q_1 + q_2 + \cdots + q_v = 0$

When performing a planned analysis involving several contrasts, we need to evaluate if these contrasts are mutually orthogonal or not. Two contrasts, B and C, are *orthogonal contrasts* if and only if

$$p_1q_1 + p_2q_2 + \cdots + p_vq_v = 0 .$$

For example, the contrasts $\tau_1 - \tau_2$ and $\tau_1 + \tau_2 - 2\tau_3$ are orthogonal because the coefficients of the treatments in these two contrasts are 1, −1, 0 and 1, 1 and −2 which satisfy $1\times1+(-1)\times1+0\times(-2)=0$. One can obtain a large number of orthogonal contrasts. If there is a set of contrasts such that every pair of contrasts in the set is orthogonal to each other, then the set is said to be a set of mutually orthogonal contrasts. For v parameters (or treatment effects), the maximum number of mutually orthogonal contrasts is v – 1. In other words, the cardinality of the complete set of mutually orthogonal contrasts is v – 1. But the total number of contrasts can be infinite. They need not be mutually orthogonal contrasts.

Let $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_i, \cdots, \mathbf{s}_p$ be the vectors of coefficients of p parametric (treatment) contrasts. This set of p contrasts is said to be linearly independent if and only if the only relationship among these contrasts $l_1\mathbf{s}_1 + l_2\mathbf{s}_2 + \cdots + l_i\mathbf{s}_i + \cdots + l_p\mathbf{s}_p = \mathbf{0}$ is $l_1 = 0, l_2 = 0, \cdots, l_i = 0, \cdots, l_p = 0$, where $l_1, l_2, \cdots, l_i, \cdots, l_p$ are arbitrary scalar constants.

It may be noted here that orthogonality implies linear independence; the converse, however, is not true.

Consider that there are $v = 3$ treatments. The set of contrasts $\tau_1 - \tau_2$ and $\tau_1 - \tau_3$ is linearly independent but not orthogonal. The two vectors of coefficients of the contrasts in this case are $\langle (1, -1, 0) \rangle$ and $\langle (1, 0, -1) \rangle$. Further the two contrasts $\tau_1 - \tau_2$ and $\tau_1 + \tau_2 - 2\tau_3$ are orthogonal. These contrasts are also linearly independent. The two vectors of coefficients of the contrasts in this case are $\langle (1, -1, 0) \rangle$ and $\langle (1, 1, -2) \rangle$.

One way of writing the coefficients of the complete set of mutually orthogonal contrasts with v parameters is the following:

τ_1	τ_2	τ_3	τ_4	τ_5	...	τ_{v-1}	τ_v
1	-1	0	0	0	...	0	0
1	1	-2	0	0	...	0	0
1	1	1	-3	0	...	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
1	1	1	1	1	...	$-(v-2)$	0
1	1	1	1	1	...	1	$-(v-1)$

The matrix above gives the coefficients of the mutually orthogonal contrasts. The actual contrasts are $\tau_1 - \tau_2$; $\tau_1 + \tau_2 - 2\tau_3$; $\tau_1 + \tau_2 + \tau_3 - 3\tau_4$; ...; $\tau_1 + \tau_2 + \tau_3 + \cdots + \tau_{v-2} - (v-2)\tau_{v-1}$; $\tau_1 + \tau_2 + \tau_3 + \cdots + \tau_{v-1} - (v-1)\tau_v$. This way of writing the complete set of mutually orthogonal contrasts gives complete set of linearly independent contrasts.

Remark 3.1 It may be noted that mutually orthogonal contrasts provide a technique for partitioning ANOVA sum of squares due to treatments into sum of squares due to single degrees of freedom contrasts or any contrast among subsets of treatments. If B and C are any two orthogonal treatment contrasts, then the tests for H_0: B = 0 and H_0: C = 0 are independent of one another. In other words, the results of one test (H_0: B = 0) have no effect on the result of the other test (H_0: C = 0). Further, if B_1, B_2, ..., B_{v-1} are mutually orthogonal contrasts obtained from v parameters (or treatment effects), the treatment sum of squares can be partitioned into SS_{treat} = SSB_1 + SSB_2 + ... + SSB_{v-1}. This, however, holds only for orthogonal, equi-replicated designs. For non-orthogonal designs, this kind of partitioning, though possible, is difficult as it would involve some weighing factors. The description of this is beyond the scope of this book. However, the sum of squares for the various subsets of treatments can be obtained in the usual way and the testing of hypothesis can also be done in the usual way.

The set of mutually orthogonal contrasts is not unique. There can be several sets of mutually orthogonal contrasts. For example, if $v = 5$, then the coefficients of the complete set of mutually orthogonal contrasts could be

$$\begin{array}{ccccc} 1 & -1 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 & 0 \\ 1 & 1 & 1 & -3 & 0 \\ 1 & 1 & 1 & 1 & -4 \end{array} \quad \text{or} \quad \begin{array}{ccccc} 3 & 3 & -2 & -2 & -2 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{array} \quad \text{or} \quad \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ -1 & 2 & 0 & -2 & 1 \\ 1 & -4 & 6 & -4 & 1 \end{array}$$

3.2.1 Contrasts in practice

Consider an experiment being conducted with 7 treatments (varieties of maize). For making comparisons of the 7 treatments, *i.e.* for testing $H_0: \tau_1 = \tau_2 = \ldots = \tau_i = \ldots = \tau_7$, we can write 6 degrees of freedom as 6 mutually orthogonal contrasts with coefficient matrix as

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & -3 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & -4 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & -5 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & -6 \end{pmatrix}$$

For making pairwise comparisons of varieties, we may have null hypothesis as $H_0: \tau_1 = \tau_2$ or $H_0: \tau_1 = \tau_3$, or $H_0: \tau_3 = \tau_4$ or $H_0: \tau_6 = \tau_7$, etc. The coefficients of contrasts for testing these problems could be

$H_0: \tau_1 = \tau_2$	1	−1	0	0	0	0	0
$H_0: \tau_1 = \tau_3$	1	0	−1	0	0	0	0
$H_0: \tau_3 = \tau_4$	0	0	1	−1	0	0	0
$H_0: \tau_6 = \tau_7$	0	0	0	0	0	1	−1

Now suppose that the 7 treatments are divided into two disjoint groups, first four are test varieties and the last three are control varieties (local variety, national variety, disease resistant variety). The interest of the experimenter is to make comparisons among tests; among controls; and tests *versus* controls. This can be done by defining contrasts suitably. For comparing the tests (three degrees of freedom), the null hypothesis is $H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4$. The coefficient matrix of the three mutually orthogonal contrasts for testing this null hypothesis is

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & -3 & 0 & 0 & 0 \end{pmatrix}$$

Similarly, for comparisons among controls (two degrees of freedom), the null hypothesis is H_0: $\tau_5 = \tau_6 = \tau_7$. The coefficient matrix of mutually orthogonal contrasts for testing this null hypothesis is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -2 \end{pmatrix}$$

Finally, for making tests *versus* controls comparisons (one degree of freedom), we have H_0: $(\tau_1 + \tau_2 + \tau_3 + \tau_4)/4 = (\tau_5 + \tau_6 + \tau_7)/3$. The coefficients of the treatment effects for this hypothesis testing are

1/4	1/4	1/4	1/4	−1/3	−1/3	−1/3

The inference problem on a contrast is invariant with respect to the choice of coefficients. In other words, the choice of a contrast will not change the inference about the hypothesis to be tested. This is so because the sum of squares due to the contrasts would remain unchanged with a different choice of coefficients of the contrast. We can also choose the coefficients of the treatment effects for testing this null hypothesis as

3	3	3	3	−4	−4	−4

or

6	6	6	6	−8	−8	−8

3.2.2 More about contrasts

Orthonormal contrasts are orthogonal contrasts which satisfy the additional condition that, for each contrast, the sum of squares of the coefficients add up to one. Suppose that $B = p_1\tau_1 + p_2\tau_2 + \cdots + p_v\tau_v$ is a contrast, meaning thereby that $p_1 + p_2 + \cdots + p_v = 0$. Let $c = p_1^2 + p_2^2 + \cdots + p_v^2$. Then

$$B^O = \frac{p_1}{\sqrt{c}}\tau_1 + \frac{p_2}{\sqrt{c}}\tau_2 + \cdots + \frac{p_v}{\sqrt{c}}\tau_v = \alpha_1\tau_1 + \alpha_2\tau_2 + \cdots + \alpha_v\tau_v \text{ is a normalized contrast. Here}$$

$\alpha_1 = \dfrac{p_1}{\sqrt{c}}, \alpha_2 = \dfrac{p_2}{\sqrt{c}}, \cdots, \alpha_v = \dfrac{p_v}{\sqrt{c}}$. Obviously, $\alpha_1 + \alpha_2 + \cdots + \alpha_v = \dfrac{p_1 + p_2 + \cdots + p_v}{\sqrt{c}} = 0$ and

$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_v^2 = \dfrac{p_1^2 + p_2^2 + \cdots + p_v^2}{c} = 1.$

For example with three treatment effects, $\tau_1 + \tau_2 - 2\tau_3$ is a contrast. The coefficients of treatment effects in this contrast are 1, 1, −2 and sum of squares of the coefficients is 6. However,

$$\frac{1}{\sqrt{6}}\tau_1 + \frac{1}{\sqrt{6}}\tau_2 - \frac{2}{\sqrt{6}}\tau_3 \text{ is a normalized contrast because the sum of squares of the coefficients in}$$

this contrast is one.

A set of normalized treatment contrasts, which are pairwise orthogonal, is called a set of orthonormal contrasts. The cardinality of a set of orthonormal contrasts is $v - 1$. For example, if $v = 5$, then the coefficients of the complete set of orthonormal contrasts could be

$$
\begin{array}{ccccc}
3/\sqrt{30} & 3/\sqrt{30} & -2/\sqrt{30} & -2/\sqrt{30} & -2/\sqrt{30} \\
1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 & 0 \\
0 & 0 & 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\
0 & 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2}
\end{array}
$$

There are some advantages of using normalized and orthonormal contrasts. But this is beyond the scope of this book. For our purpose, it would suffice to understand mutually orthogonal contrasts.

3.2.3 Testing the hypothesis related to contrasts

Generally speaking, the testing of hypothesis pertaining to contrasts of interest is a difficult problem. However, using a good statistical package like SAS or R, one can easily test the hypothesis related to contrasts of interest. If the design adopted is a completely randomized design, or a randomized complete block design or a Latin square design, or a factorial experiment conducted in a completely randomized design, or a randomized complete block design or a Latin square design, then obtaining the sum of squares due to contrast (or the hypothesis related to contrast) is relatively easy. Suppose the experimenter wishes to test a null hypothesis about a contrast $B = p_1\tau_1 + p_2\tau_2 + \cdots + p_v\tau_v$ using any of the designs just described. The null hypothesis

is H_0: $B = p_1\tau_1 + p_2\tau_2 + \cdots + p_v\tau_v = 0$. The contrast sum of squares is obtained as $\dfrac{\left(\sum\limits_{i=1}^{v} p_i \overline{T}_i\right)^2}{\sum\limits_{i=1}^{v} \dfrac{p_i^2}{r_i}}$. Here

$\overline{T}_i$ is the mean of the ith treatment obtained from the design and r_i is the replication of the ith treatment, $i = 1,2,\cdots, v$ If the design is equi-replicated, *i.e.*, $r_1 = r_2 = \cdots = r_i = \cdots = r_v = r$,

then the contrast sum of squares is given by $\dfrac{r\left(\sum\limits_{i=1}^{v} p_i \overline{T}_i\right)^2}{\sum\limits_{i=1}^{v} p_i^2}$.

However, for other designs, which are not balanced (or non-orthogonal), like incomplete block designs, nested block designs, incomplete row column designs, etc., the computation of contrast sum of squares is quite involved. General procedure is described in the sequel.

Suppose that the linear model (pertaining to the design used for generation of data) is expressed as $\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{e}$; $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\theta} \Rightarrow E(\mathbf{e}) = \mathbf{0}$ and $D(\mathbf{e}) = \sigma^2 \mathbf{I}_n$. Here $\mathbf{y}$ is an n-component vector of observations, $\mathbf{X}$ is an $n \times p$ design matrix (depends upon the design used for generation of data), $\boldsymbol{\theta}$ is a p-component vector of parameters and $\mathbf{e}$ is an n-component vector of random errors. The vector $\boldsymbol{\theta}$ contains v-component vector of treatment effects $\boldsymbol{\tau}$ and $(p - v)$-component

vector $\boldsymbol{\beta}$ of nuisance parameters.

Suppose that there are $s(<v)$ *testable* hypothesis (a hypothesis that can be expressed in terms of estimable functions)

$$p_1^1 \tau_1 + p_2^1 \tau_2 + \cdots + p_v^1 \tau_v = m_1$$

$$p_1^2 \tau_1 + p_2^2 \tau_2 + \cdots + p_v^2 \tau_v = m_2$$

.

.

.

$$p_1^s \tau_1 + p_2^s \tau_2 + \cdots + p_v^s \tau_v = m_s.$$

A hypothesis is said to be testable if it is estimable through the design. This set of s testable hypotheses can be rewritten in matrix notations as: $\mathbf{P}'\boldsymbol{\tau} = \mathbf{m}$. Here $\mathbf{P} = ((p_i^j))$, $i = 1,2,\cdots, v; j = 1,2,\cdots, s$ and $\mathbf{m} = (m_1, m_2, \cdots, m_s)'$. If the s rows of $\mathbf{P}'$ are linearly independent, then for testing this hypothesis the test statistic is

$$F(H) = \frac{Q/s}{R_0^2/(n-r)} = \frac{Q}{s\hat{\sigma}^2},$$

which under the null hypothesis follows a Snedecor's F distribution with s and $n - r$ degrees of freedom, where r is the rank of $\mathbf{X}$ and n is the total number of observations. If the s rows of $\mathbf{P}'$ are not linearly independent then s may replaced by s^*, the number of linearly independent rows of $\mathbf{P}'$. Here $Q = (\mathbf{P}'\hat{\boldsymbol{\tau}} - \mathbf{m})' [\mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{P}]^{-1} (\mathbf{P}'\hat{\boldsymbol{\tau}} - \mathbf{m})$, $s^2 = \hat{\sigma}^2 = R_0^2/(n-r)$, $R_0^2 = SSE = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}})$ = Error sum of squares in the ANOVA. Further, $(n - r)$ is the error degrees of freedom in the ANOVA.

When $\mathbf{m} = \mathbf{0}$, the testable hypothesis is $\mathbf{P}'\boldsymbol{\tau} = \mathbf{0}$, and then Q becomes

$$Q = \hat{\boldsymbol{\tau}}'\mathbf{P}[\mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{P}]^{-1}\mathbf{P}'\hat{\boldsymbol{\tau}}.$$

The test statistic once again follows Snedecor's F distribution with s and $n - r$ degrees of freedom.

As mentioned earlier in this Chapter, contrast analysis is a very important technique of data analysis and using this technique almost all the questions of the experimenters can be answered. What at all needs to be done is to convert the problem to a contrast or a set of contrasts. Once this is done, the analysis follows immediately. Some practical situations have

been described above through examples. It has been demonstrated as to how the problems can be translated into contrasts. In the sequel some examples and the analysis is described through actual experimental data.

3.3 Example 1

In order to select suitable tree species for Fuel, Fodder and Timber an experiment was conducted in a randomized complete block design with ten different tree species and four replications. The plant height was recorded in centimeter (cm). The details of the experiment are given in Table 3.1.

Table 3.1: Plant height (cms) (Place - Kanpur)

Tree species Number	Tree Species	Spacing	Blocks (or Replications)			
			1	2	3	4
1	A. Indica	4×4	144.44	145.11	104.00	105.44
2	D. Sisso	4×2	113.50	118.61	118.61	123.00
3	A. Procer	4×2	60.88	90.94	80.33	92.00
4	A. Nilotic	4×2	163.44	158.55	158.88	153.11
5	T. Arjuna	4×2	110.11	116.00	119.66	103.22
6	L. Loucoc	4×1	260.05	102.27	256.22	217.80
7	M. Alba	4×2	114.00	115.16	114.88	106.33
8	C. Siamia	4×2	91.94	58.16	76.83	79.50
9	E. Hybrid	4×1	156.11	177.97	148.22	183.17
10	A. Catech	4×2	80.20	108.05	45.18	79.55

In the sequence the data are analyzed using analysis of variance for a two way classified data and then comparisons are made within and between different groups of tree species.

3.3.1 Procedure and Calculations

In this Example, $v = 10$, $b = 4$. We compute the following totals in Table 3.2.

Table 3.2: Treatment and block totals

Treatment Total (T_i)	Treatment mean ($\bar{T}_i = T_i / b$)
T_1= 144.44 +…+ 105.44 = 498.99	$\bar{T}_1$ = 498.99/4 = 124.748
T_2 = 112.50 + … + 123.00 = 473.72	$\bar{T}_2$ = 473.72/4 = 118.430
T_3 = 60.88 + … + 92.00 = 324.15	$\bar{T}_3$ = 324.15/4 = 81.038
T_4 = 163.44 + … + 153.11 = 633.98	$\bar{T}_4$ = 633.98/4 = 158.495

$T_5 = 110.11 + \ldots + 103.22 = 448.99$	$\overline{T}_5 = 448.99/4 = 112.248$
$T_6 = 260.05 + \ldots + 217.8 = 836.34$	$\overline{T}_6 = 836.34/4 = 209.085$
$T_7 = 114.00 + \ldots + 106.33 = 450.37$	$\overline{T}_7 = 450.37/4 = 112.593$
$T_8 = 91.94 + \ldots + 79.50 = 306.43$	$\overline{T}_8 = 306.43/4 = 76.608$
$T_9 = 156.11 + \ldots + 183.17 = 665.47$	$\overline{T}_9 = 665.47/4 = 166.368$
$T_{10} = 80.20 + \ldots + 79.55 = 312.98$	$\overline{T}_{10} = 312.98/4 = 78.245$

Block Total (B_j)	**Block Mean** $(\overline{B}_j = B_j / v)$
$B_1 = 144.44 + \ldots + 80.20 = 1294.67$	$\overline{B}_1 = 1294.67/10 = 129.467$
$B_2 = 145.11 + \ldots + 108.05 = 1190.82$	$\overline{B}_2 = 1190.82/10 = 119.082$
$B_3 = 104.00 + \ldots + 45.18 = 1222.81$	$\overline{B}_3 = 1222.81/10 = 122.281$
$B_4 = 105.44 + \ldots + 79.55 = 1243.12$	$\overline{B}_4 = 1243.12/10 = 124.312$

Grand Total, $G = \sum_{i=1}^{10} T_i = \sum_{j=1}^{4} B_j = 4951.42$.

Correction Factor, $CF = G^2 / vb = (4951.42)^2/40 = 612914.00$

Treatments (Trees) $SS = \sum_{i=1}^{10} \dfrac{T_i^2}{4} - CF$

$= \left((498.99)^2 + \cdots + (312.98)^2\right)/4 - 612914.00 = 66836.355$

Blocks (or Replications) $SS = \sum_{j=1}^{4} \dfrac{B_j^2}{10} - CF$

$\left((1294.67)^2 + \cdots + (1243.12)^2\right)/10 - 6121914.00 = 569.431.$

Total $SS = \sum_{i=1}^{10}\sum_{j=1}^{4} y_{ij}^2 - CF$

$= \left((144.44)^2 + (145.11)^2 + \cdots + (79.55)^2\right) - 6121914.00 = 89101.047.$

Error SS = Total SS – Treatments SS – Blocks SS = 89101.42 – 66836.35 – 569.43 = 21695.262.

We now form the following Analysis of Variance Table 3.3.

Table 3.3: ANOVA table

Source	DF	SS	MS	F-value	Prob > F
Tree Species	9	66836.355	7426.262	9.24	<0.0001
Blocks	3	569.431	189.810	0.24	0.8703
Error	27	21695.262	803.528		
Corrected Total	39	89101.047			

The least significant difference (LSD) between any two treatment means for testing the null hypothesis that two treatment effects are equal, *i.e.*, $\tau_i = \tau_j \ \forall\, i \ne j = 1,2,\cdots,10$

$$= t_{\alpha,\,error\,d.f.} \times \sqrt{2s^2/b} = 2.05 \times \sqrt{(2\times 803.53)/4} = 41.09$$

The analysis of variance Table suggests that the treatment effects are highly significant (*p*-value < 0.0001), but the block effects are not significant or the block mean squares is small as compared to error mean square. On the basis of the LSD we prepare Table 3.4 giving the significance of the difference between two treatments effects:

Table 3.4: Treatments grouping with letter display

							Mean	Tree No. (Treatment)
						A	209.085	6
					B		166.368	9
				C	B		158.495	4
			D	C			124.748	1
	E		D	C			118.430	2
F	E		D				112.593	7
F	E		D				112.248	5
F	E						81.038	3
F	E						78.245	10
F							76.608	8

Treatments with the same letter are not significantly different from each other. It, therefore, follows from the Table that treatment 6 is significantly different from all other treatments. Thus, the tree species L. Loucoc is the best so far as plant height is concerned. Treatment 9 (E. Hybrid) is not significantly different from treatment 4 (A. Nilotica), but is significantly different from all other treatments. Further, treatment 4 (A. Nilotica) is not significantly different from

treatments 1 (A. Indica) and 2 (D. Sisso), but is significantly different from all other treatments. Likewise we can draw conclusions about the pairwise treatment comparisons from all other alphabets. Tree species C. Siamia (Species 8) records the lowest height.

Suppose now that tree species numbers 1, 2, 3, 4, 10 are useful for fuel, fodder and timber and tree species numbers 5, 6, 7, 8, 9 are useful for fuel and fodder only. The interest of the experimenter is to test a null hypothesis that (i) the average effect of trees in the two groups is same. The null hypothesis can be formulated as

$$H_{01}: \frac{\tau_1 + \tau_2 + \tau_3 + \tau_4 + \tau_{10}}{5} = \frac{\tau_5 + \tau_6 + \tau_7 + \tau_8 + \tau_9}{5}, \text{ or}$$

$$H_{01}: \tau_1 + \tau_2 + \tau_3 + \tau_4 - \tau_5 - \tau_6 - \tau_7 - \tau_8 - \tau_9 + \tau_{10} = 0.$$

Similarly, suppose that the experimenter is also interested in testing the following null hypotheses:

(ii) the average effect of tree species numbers 2, 3, . . . , 10 is same as that of tree species number 1;
(iii) the average effect of tree species numbers 1, 2, 3, 4 is same as that of tree species number 9;
(iv) the average effect of tree species numbers 1, 2, 3, 4 is same as that of tree species number 10;
(v) the average effect of tree species numbers 5, 6, 7, 8 is same as that of tree species number 9;
(vi) the average effect of tree species numbers 5, 6, 7, 8 is same as that of tree species number 10.

We can formulate the hypotheses as:

$H_{02} : 9\tau_1 - \tau_2 - \tau_3 - \tau_4 - \tau_5 - \tau_6 - \tau_7 - \tau_8 - \tau_9 - \tau_{10} = 0$
$H_{03}: \tau_1 + \tau_2 + \tau_3 + \tau_4 - 4\tau_9 = 0$
$H_{04} : \tau_1 + \tau_2 + \tau_3 + \tau_4 - 4\tau_{10} = 0$
$H_{05} : \tau_5 + \tau_6 + \tau_7 + \tau_8 - 4\tau_9 = 0$
$H_{06} : \tau_5 + \tau_6 + \tau_7 + \tau_8 - 4\tau_{10} = 0$

The alternative hypothesis in all the cases is that the parametric contrast is not equal to zero. We have Table 3.5 for testing these null hypotheses.

Table 3.5: Testing significance of the contrasts

Hypothesis	DF	Contrast SS	MS	F	Prob > F
H_{01}	1	5377.297	5377.297	6.69	0.0154
H_{02}	1	4.113	4.113	0.01	0.9435
H_{03}	1	6680.244	6680.244	8.31	0.0076
H_{04}	1	5761.655	5761.655	7.17	0.0125
H_{05}	1	4801.126	4801.126	5.98	0.0213
H_{06}	1	7805.398	7805.398	9.71	0.0043

It is evident from Table 3.5 that the average effect of tree species numbers 2, 3, ..., 10 is same as that of tree species number 1 at 5 percent level of significance. All other null hypotheses are rejected at 5 percent level of significance because the p-values are smaller than 0.05.

Suppose now that the interest of the experimenter is to test certain hypothesis concerning the five tree species in the Group 1 (comprising of Tree species Numbers 1, 2, 3, 4, and 10). The null hypothesis now is $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_{10}$. The 4×10 coefficients matrix of the set of linearly independent treatment contrasts (which are actually mutually orthogonal) for testing this null hypothesis are

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & -4 \end{pmatrix}$$

The sum of squares for testing the equality of the five trees effects can be obtained by defining four linearly independent contrasts as

$\tau_1 - \tau_2$; $\tau_1 + \tau_2 - 2\tau_3$; $\tau_1 + \tau_2 + \tau_3 - 3\tau_4$; $\tau_1 + \tau_2 + \tau_3 + \tau_4 - 4\tau_{10}$.

Using these sets of contrasts we get the following:

Hypothesis	DF	SS	MS	F	Prob > F
H_0	4	17854.0908	4463.523	5.55	0.0021

It is again evident that the effect of tree species in this group are significantly different (p-value = 0.0021)

3.3.2 Analysis using SAS

We now describe the analysis of the data in Example 1 by using SAS. The preparation of the data file, the PROC to be adopted and the commands are given below:

DATA Treeheight;
INPUT tree rep height;
/*the first column 'tree' denotes the tree number (or tree species or treatments); the second column 'rep' denotes the replication number; the third column 'height' denotes the plant height*/
CARDS;

```
1    1    144.44
2    1    113.50
3    1     60.88
4    1    163.44
5    1    110.11
6    1    260.05
7    1    114.00
8    1     91.94
```

9	1	156.11
10	1	80.20
1	2	145.11
2	2	118.61
3	2	90.94
4	2	158.55
5	2	116.00
6	2	102.27
7	2	115.16
8	2	58.16
9	2	177.97
10	2	108.05
1	3	104.00
2	3	118.61
3	3	80.33
4	3	158.88
5	3	119.66
6	3	256.22
7	3	114.88
8	3	76.83
9	3	148.22
10	3	45.18
1	4	105.44
2	4	123.00
3	4	92.00
4	4	153.11
5	4	103.22
6	4	217.80
7	4	106.33
8	4	79.50
9	4	183.17
10	4	79.55

```
;
PROC MEANS;
CLASS tree;
VAR height;
PROC MEANS;
CLASS  rep;
VAR height;
PROC GLM;
CLASS tree rep;
MODEL height = tree rep;
LSMEANS tree/PDIFF LINES;
```

CONTRAST '1 2 3 4 10 vs 5 6 7 8 9' tree 1 1 1 1 −1 −1 −1 −1 1 1;
CONTRAST '1 vs 2 3 4 5 6 7 8 9 10' tree 9 −1 −1 −1 −1 −1 −1 −1 −1 −1;
CONTRAST '1 2 3 4 vs 9' tree 1 1 1 1 0 0 0 0 −4 0;
CONTRAST '1 2 3 4 vs 10' tree 1 1 1 1 0 0 0 0 0 −4;
CONTRAST '5 6 7 8 vs 9' tree 0 0 0 0 1 1 1 1 −4 0;
CONTRAST '5 6 7 8 vs 10' tree 0 0 0 0 1 1 1 1 0 −4;
CONTRAST 'within group' tree 1 −1 0 0 0 0 0 0 0 0,
    tree 1 1 −2 0 0 0 0 0 0 0,
    tree 1 1 1 −3 0 0 0 0 0 0,
    tree 1 1 1 1 0 0 0 0 0 −4;
RUN;

When the treatment is a character variable, then the coefficients of treatment effects should be entered with care as the SAS automatically arranges the treatments in lexicographic order. Suppose there are 3 treatments as varieties namely Sonalika, C-306 and PBW343, then if one wants to test H_0: Sonalika + C-306 - 2*PBW343 = 0, the coefficients then would be 1 -2 1 as lexicographic ordering of varieties is C-306, PBW343 and Sonalika.

3.3.3 Output of analysis

In what follows are described the results obtained from the analysis of data. Table 3.6 gives the mean and standard deviation of plant height for each tree species (or levels of treatment). The minimum and maximum values within each group are also given.

Table 3.6: Tree wise mean and standard deviation of plant height

Tree	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
1	4	124.748	23.135	104.000	145.110
2	4	118.430	3.884	113.500	123.000
3	4	81.038	14.434	60.880	92.000
4	4	158.495	4.227	153.110	163.440
5	4	112.248	7.190	103.220	119.660
6	4	209.085	73.721	102.270	260.050
7	4	112.593	4.204	106.330	115.160
8	4	76.608	13.950	58.160	91.940
9	4	166.368	16.847	148.220	183.170
10	4	78.245	25.737	45.180	108.050

Similarly, Table 3.7 gives the mean and standard deviation of plant height for each replication. The minimum and maximum values within each group are also given.

Table 3.7: Replication wise mean and standard deviation of plant height

Replication	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
1	10	129.467	56.321	60.880	260.050
2	10	119.082	34.377	58.160	177.970
3	10	122.281	57.861	45.180	256.220
4	10	124.312	46.207	79.500	217.800

The analysis of variance is then performed and the results obtained are given in Table 3.8.

Table 3.8: Analysis of variance of plant height data

Source	DF	SS	MS	F value	Prob > F
Trees	9	66836.35	7426.26	9.24	<0.0001
Blocks	3	569.43	189.81	0.24	0.8703
Error	27	21695.27	803.53		
Corrected Total	39	89101.05			

R Square	CV	Root MSE	Height Mean
0.76	22.90	28.35	123.79

It is apparent that the model with trees and replications has been able to explain 76 per cent of the total variability in plant height. The CV is high, though (22.90). It may be seen again from this ANOVA that the treatment effects are highly significant (p-value < 0.0001), but the block effects are not significant (this could be a reason for high CV).

Since the design is a randomized complete block design, the unadjusted means of various levels of treatments or tree species are same as the adjusted means or least square means. A pairwise comparison of the tree species is made and the results are given in Table 3.9.

Table 3.9: Pairwise comparison of tree species

t Comparison Lines for Least Squares Means of tree species				
LS-means with the same letter are not significantly different				
			height LSMEAN	LSMEAN Number
	A		209.085	6
	B		166.368	9
C	B		158.495	4
C	D		124.748	1
C	D	E	118.430	2
F	D	E	112.593	7
F	D	E	112.248	5
F		E	81.038	3
F		E	78.245	10
F			76.608	8

Treatments with the same letter are not significantly different from each other. It, therefore, follows from Table 3.8 that treatment 6 stands out in terms of plant height as this is significantly different from all other treatments. Treatment 9 is not significantly different from treatment 4, but is significantly different from all other treatments. Further, treatment 4 is not significantly different from treatments 1 and 2, but is significantly different from all other treatments. Likewise we can draw conclusions about the pairwise treatment comparisons from all other alphabets.

Contrast analysis is also performed and once again it is implicit from the results presented in the Table 3.10 that all the contrasts are significantly different from zero except the contrast defining the equality of the effect of tree species 1 with the average effect of the tree species 2, 3, 4, 5, 6, 7, 8, 9, 10. Similarly, the effect of tree species in the group comprising of trees 1, 2, 3, 4 and 10 are significantly different.

Table 3.10: Result of contrast analysis

Contrast of Trees	DF	Contrast SS	MS	F Value	Prob > F
1, 2, 3, 4, 10 vs 5, 6, 7, 8, 9	1	5377.30	5377.30	6.69	0.0154
1 vs 2, 3, 4, 5, 6, 7, 8, 9, 10	1	4.11	4.11	0.01	0.9435
1, 2, 3, 4, vs 9	1	6680.24	6680.24	8.31	0.0076
1, 2, 3, 4, vs 10	1	5761.65	5761.65	7.17	0.0125
5, 6, 7, 8 vs 9	1	4801.13	4801.13	5.98	0.0213
5, 6, 7, 8, vs 10	1	7805.40	7805.40	9.71	0.0043
Within Group (1, 2, 3, 4, 10)	4	17854.09	4463.52	5.55	0.0021

3.4 Analysis using R

For the benefit of the readers, the analysis using R software is also given in the sequel. The R code is given but the output is not given. The readers may like to use this code for analysis of data.

R code

```
d4=read.table("Treeheight.txt",header=TRUE)
attach(d4)
names(d4)
#Treatment means and standard deviations
aggregate(height, by=list(tree), mean)
aggregate(height, by=list(tree), sd)
#Tree wise box plot of height
boxplot(height~tree)
#Replication wise box plot of height
boxplot(height~rep)
tree=factor(tree)
rep=factor(rep)
```

```
aov.out=aov(height~tree+rep)
summary(aov.out)
#contrast analysis
library(lsmeans)
lsm <- lsmeans(aov.out, "tree")
#For grouping of treatments with letters, you need to install multcompView package and
#run the following code
cld(lsm,Letters="ABCDEF")
contrast(lsm, list(con1 = c(1,1,1,1,-1,-1,-1,-1,-1,1),con2 = c(9,-1,-1,-1,-1,-1,-1,-1,-1,-1),con3
= c(1,1,1,1,0,0,0,0,-4,0),con4 = c(1,1,1,1,0,0,0,0,0,-4),con5 = c(0,0,0,0,1,1,1,1,-4,0),con6 =
c(0,0,0,0,1,1,1,1,0,-4)))
contrast(lsm,list(con7=c(1,-1,0,0,0,0,0,0,0,0),con8=c(1,1,-2,0,0,0,0,0,0,0),con9=c(1,1,1,-
3,0,0,0,0,0,0),con10=c(1,1,1,1,0,0,0,0,0,-4)))
#Through splitting terms in anova without lsmeans package,
#Caution: works for orthogonal contrasts only
contrast.mat=matrix(c(1,-1,0,0,0,0,0,0,0,0,1,1,-2,0,0,0,0,0,0,0,1,1,1,-
3,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,-4),ncol=4)
contrasts(tree)<-contrast.mat
aov.out=aov(height~tree+rep)
summary(aov.out,split=list(tree=list("1vs2"=1,"first2vs3"=2,"first3vs4"=3,"first4vs10"=4)))
detach(d4)
```

Remark 3.2 It may be worthwhile mentioning here that the contrast analysis is a very powerful statistical methodology for answering almost all the questions of the researchers. Generally speaking, the researchers finish their analyses after generating the ANOVA table. But there are many more probing questions that can be answered using contrast analysis. The only effort required is that the researcher should be able to write the problem in terms of a parametric contrast(s). Once this is done, the problem is easy to handle. In this Chapter, one example has been given to describe how contrast analysis helps in solving the problems of the researcher. Contrast analysis has also been done in Chapter 2. It has also been done in many other Chapters, particularly when dealing with the analysis of augmented designs. In Chapter 12 is presented another very interesting application of contrast analysis in solving a difficult problem. The readers may try to grasp that application of contrast analysis because that would be immensely useful to them in their research.

Remark 3.3 In this Chapter it has been demonstrated how contrast analysis helps in making comparisons among subsets of treatments. Once the problem is translated into a contrast, then the analysis can be done very easily for testing a null hypothesis about the contrast. This book is essentially devoted to single factor experiments. In factorial experiments, there are several factors and each factor has several levels. The treatments are all possible combinations of levels of different factors. In case of factorial experiments, like single factor experiments, comparisons can be made among levels of a factor at fixed levels of the other factors. For example, if there are

two factors A and B with 3 and 2 levels each represented as a_0, a_1, a_2 and b_0, b_1, respectively, then the 6 treatment combinations are a_0b_0, a_0b_1, a_1b_0, a_1b_1, a_2b_0, a_2b_1. These 6 treatment combinations are in fact 6 treatments and all type of contrasts can be defined for making subset comparisons among treatment combinations. The contrasts of interest could be $a_0b_0-a_2b_0$; $a_0b_0-2a_1b_0+a_2b_0$; $a_2b_0-a_2b_1$. We may define any other contrast. This would be described in more detail in Part-II of the book.

4

Covariance Analysis

4.1 Introduction

It has been discussed earlier in Chapters 1 and 2 that the data generated from experimental designs exhibits a large variability. The two major components of this variability are, (a) to which some cause can be assigned, (b) to which no cause can be assigned. The second component of variability to which no cause can be assigned is the experimental error. The major concern of any experimental design or statistical analysis of experimental data is to keep the experimental error as small as possible. There is no way the experimental error can be totally eliminated.

The two major components of the explainable part of variability are (a) the treatments, (b) the experimental material or to be specific the experimental units. The variability due to treatments is a deliberate attempt on the part of the experimenter to create variability. However, the experimental units on which the experiment is conducted (to which the treatments are subjected) is a major source of variability to be dealt with through designing the experiment properly. Generally, while designing an experiment the variability in the experimental material gets overlooked. At times this is not properly taken care of. This could be one cause of large experimental error. So in order to control the experimental error, proper designing of experiment is essential. However, there are analytical ways also of controlling the experimental error. One such technique is that of analysis of covariance (ANCOVA). It is expected that a good experiment attempts to incorporate all possible means of minimizing the experimental error.

The analysis of covariance, generally known as ANCOVA, is a technique that combines both the ANOVA (analysis for comparing population means of groups using data collected from a single factor or a multi-factor experiment using some appropriate design and then analyzed using ANOVA) and that of estimating the slope of a straight line between two variables, y and X. In both the cases the response variable y is continuous (measured on interval or ratio scale). In case of the ANOVA, the X variables are generally in nominal or ordinal scale and these generally serve to identify the treatment groups, or the block groups or the sub-block groups within the block groups, or the rows and columns groups, etc. In the regression setting, the X variable is also continuous, like the response variable y.

Of the many uses of ANCOVA, two major uses are, (a) to check if the regression lines for each treatment group are parallel or not (meaning thereby that the regression lines have different intercepts but common slopes) or if the regression lines for each treatment group are coincident or not (meaning thereby that the regression lines have the same slopes and intercepts), (b) to test for the differences in the population means of groups when some of the variation in the response variable can be explained by the covariate. For instance, the effectiveness of different feeds given

to animals can be compared by randomizing the animals to the feeds and measuring the body weights at different points of time during the experiment. However, some of the variation in the body weights of animals after subjecting them to different feeds may be attributable to their initial body weights. So by standardizing all the animals to some common weight can help in detecting the differences among the groups more precisely. In this case the variation caused by different initial body weights of the animals is also taken away from the experimental error leading to a considerable reduction in the experimental error.

Another interesting application of the ANCOVA is the following: In a field experiment the rodents attack the field and as a consequence some of the plots in the experiment are partially damaged. So the observations recorded from the plots damaged by the rodent attack are naturally different from the plots not damaged by rodent attack and, therefore, there is a lot of variability caused in the data by rodent attack. Covariance analysis (ANCOVA), with rodent damage as a covariate, could be useful in adjusting plot yields to the levels that these should have been had there been no damage in any plot due to rodent attack.

ANCOVA requires measurement of the characteristic of primary interest plus the measurement of one or more variables known as *covariates*. It also requires that the functional relationship of the covariates with the character of primary interest is known beforehand. Generally a linear relationship is assumed, though other type of relationships could also be assumed. Further, it is important to note that the covariates used should be such that these are not influenced by the application of treatments. Otherwise, while adjusting the study variable over covariates may take away some part of the variability due to the application of treatments.

Consider the case of a variety trial in which weed incidence is used as a covariate and the grain yield is the characteristic of interest. With a known functional relationship between weed incidence and grain yield, the covariance analysis can adjust grain yield in each plot to a common level of weed incidence. With this adjustment, the variation in yield due to weed incidence is quantified and effectively separated from that due to varietal differences.

ANCOVA can also be used for more than one covariate as well. It can also be applied not only to linear functional relationship between the characteristic of interest and the covariates but also to any type of functional relationship between variables *viz.* quadratic, inverse polynomial, etc. The readers may note that all these are in fact strictly linear models only, because a linear model is one which is linear in parameters. Here we illustrate the use of covariance analysis with the help of a single covariate that is linearly related with the character of primary interest. It is expected that this simplification shall not unduly reduce the applicability of the technique, as a single covariate that is linearly related with the primary variable is adequate for most of the experimental situations in agricultural research.

The covariance analysis can also be used in the analysis of data generated from designed experiments with one or more missing observations. Suppose that an experiment is conducted in a RCB design with 8 treatments and 3 blocks (replications). Suppose during experimentation, the observation pertaining to treatment 5 in block 2 is lost (or missing). For the analysis of data, one can use covariance analysis technique by first assigning a value 0 to the missing observation and then defining an auxiliary variable which takes a value +1 for the missing

observation and 0 for all other observations. The values can also be taken as +1 and −1 instead of +1 and 0. Similarly, if in addition to the lost observation of treatment 5 in block 2, one more observation pertaining to treatment 3 in block 3 is also lost, then one can assign values 0 to the missing observations and then define two pseudo auxiliary variables X_1 and X_2, with X_1 same as defined in case of loss of one observation and X_2 as taking a value +1 for the lost observation pertaining to treatment 3 in block 3 and 0 (or −1) for all other observations. The results from this analysis would be exactly same as one would produce by ignoring the lost data and analyzing the remaining data as a block design (design for one-way elimination of heterogeneity) in the sense that the probability levels of significance of treatment effects and block effects would be the same in both the cases. In fact the results obtained by either using a covariate for each missing observation or by ignoring the lost observation in the analysis of data generated from any designed experiment are same.

As mentioned earlier and above also, it is a realized fact that proper blocking (or grouping) of experimental units helps in reducing the experimental error by maximizing the differences among the blocks and minimizing the differences within the blocks. In experiments conducted in agricultural sciences and other sciences, blocking, however, cannot cope with certain types of variability arising due to spotty soil heterogeneity, soil salinity, unpredictable incidence of pest and diseases, insect manifestation, appearance of weed, etc. In all these instances, heterogeneity among experimental plots will not follow a definite pattern, which causes difficulty in capturing large differences among blocks. Indeed, blocking is ineffective in the case of non-uniform insect incidences because blocking was done prior to the occurrence of insects or pests or disease. Furthermore, even though it is true that a researcher may have some information on the probable path or direction of insect movement, unless the direction of insect movement coincides with the soil fertility gradient, the choice of whether soil heterogeneity or insect incidence should be the criterion for blocking is difficult. The choice is especially difficult if both sources of variation have about the same importance.

Use of covariance analysis should be considered in such experiments where blocking couldn't adequately reduce the experimental error. By measuring an additional variable (*e.g.,* covariate *X*) that is known to be linearly related to the characteristic of interest *y*, the source of variation associated with the covariate can be deducted from experimental error. This adjusts the primary variable *y* linearly upward or downward, depending on the relative size of its respective covariate. The adjustment accomplishes two important improvements:

1. The treatment mean is adjusted to a value that it would have had, had there been no differences in the values of the covariate. The adjustment in treatment mean is generally made using the mean value of the covariate.

2. The experimental error is reduced and the precision for comparing treatment effects is increased.

Although blocking and covariance techniques are both used to reduce experimental error, the differences between the two techniques are such that they are usually not interchangeable. The ANCOVA can be used only when the covariate representing the heterogeneity among the experimental units can be measured quantitatively. However, that is not a necessary condition

for blocking. In addition, because blocking is done before the start of the experiment, it can be used only to cope with sources of variation that are known or predictable. ANCOVA, on the other hand, can take care of unexpected sources of variation that occur during the experiment. Thus, ANCOVA is useful as a supplementary procedure to take care of sources of variation that could not be accounted for by blocking.

When covariance analysis is used for error control and adjustment of treatment means/effects, the covariate must not be affected by the treatments being tested. Otherwise, the adjustment removes both the variation due to experimental error and that due to treatment effects. A good example of covariates that are free of treatment effects are those that are measured before the treatments are applied, such as soil analysis and residual effects of treatments applied in the past experiments. Number of weeds in each plot in a varietal trial is another example of a covariate which is measured after the application of treatments. This cannot be controlled at the designing stage as the intensity of weeds is only known after its emergence, which happens once the experiment has been laid out. In other cases, care must be exercised to ensure that the covariates defined are not affected by the treatments being tested.

A strong assumption while fitting analysis of covariance is that the regression coefficient is common for all the classes of the treatments. But in practice, this assumption may not hold. It may, therefore, be desired to fit a different regression coefficient for each class (or level) of the treatment. This is so because in addition to the usual assumptions on the error variables, the covariance analysis model assumes a linear relationship between the covariate (X) and the mean response with the same slope for each treatment. Therefore, it is essential to test the equality of slopes by comparing the fit of the analysis of covariance model assuming same slope for each treatment with the fit of the corresponding model that uses different slopes for each treatment. In the sequence is given an example that uses SAS commands for testing the equality of slopes for each treatment:

4.2 Example 1

An experiment was conducted with 25 bags of 15 oysters and 5 treatments *viz.* Trt1 as cool-bottom, Trt2 as cool-surface, Trt3 as hot-bottom, Trt4 as hot-surface and Trt5 as control. Each treatment is randomly allocated to the 25 bags so that each treatment receives 5 bags. Each bag of 15 oysters is considered as one experimental unit. The oysters were washed, cleaned, dried and weighed at the beginning of the experiment and then again after 40 days of being subjected to treatments. The purpose of the experiment was, (a) to determine if artificially heated water has any affect on the growth of oysters, and (b) to determine if the position in the water column (surface *vs* bottom) has any affect on the growth of oysters. The initial weight and the final weight of the 25 bags of 15 oysters each for different locations is given in Table 4.1.

Table 4.1: Oyster data

Treatment	1	1	1	1	1	2	2	2	2	2	3	3	3
Replication	1	2	3	4	5	1	2	3	4	5	1	2	3
Initial Weight	26.6	33.1	27.3	31.9	33.4	27.4	29.1	28.1	26.6	26.3	22.6	29.0	26.9
Final Weight	29.9	37.9	31.9	37.1	38.3	31.8	34.2	33.1	31.0	30.2	29.3	35.9	34.5

Treatment	3	3	4	4	4	4	4	5	5	5	5	5	
Replication	4	5	1	2	3	4	5	1	2	3	4	5	
Initial Weight	22.1	27.7	25.7	30.5	21.5	24.8	28.9	20.6	18.0	24.9	25.7	19.2	
Final Weight	29.0	29.3	30.1	37.0	26.9	28.2	32.2	23.8	21.7	27.2	30.8	22.3	

4.2.1 Analysis using SAS

In the sequel are described SAS commands and data structure for performing the analysis of covariance on the data generated.

DATA Oyster;
INPUT trt rep Initialwt Finalwt;
/*the first column (trt) gives the treatment numbers; the second column (rep) gives the replication numbers; the third column (Initialwt) gives the initial weight of the bags; the last column (Finalwt) gives the final weight of the bags taken after 40 days*/
CARDS;

```
1    1    26.6    29.9
1    2    33.1    37.9
1    3    27.3    31.9
1    4    31.9    37.1
1    5    33.4    38.3
2    1    27.4    31.8
2    2    29.1    34.2
2    3    28.1    33.1
2    4    26.6    31.0
2    5    26.3    30.2
3    1    22.6    29.3
3    2    29.0    35.9
3    3    26.9    34.5
3    4    22.1    29.0
3    5    27.7    29.3
4    1    25.7    30.1
4    2    30.5    37.0
4    3    21.5    26.9
4    4    24.8    28.2
4    5    28.9    32.3
5    1    20.6    23.8
```

5	2	18.0	21.7
5	3	24.9	27.2
5	4	25.7	30.8
5	5	19.2	22.3

```
;
ODS RTF FILE = 'OYSTER.RTF';
PROC REG; /*simple overall regression analysis */;
MODEL Finalwt = Initialwt;
RUN;
PROC SORT;
BY trt;
PROC REG; /*simple regression analysis for each treatment*/
MODEL Finalwt = Initialwt;
BY trt;
RUN;
PROC GLM;/*anova for one way classified data*/
CLASS trt;
MODEL Finalwt = trt;
RUN;
PROC GLM;/*ancova for one way classified data*/
CLASS trt;
MODEL Final wt = trt Initialwt / solution;
LSMEANS TRT / STDERR PDIFF ADJUST = TUKEY OUT= adjmeans;
CONTRAST 'Control vs. Treatment' trt  −1 −1 −1 −1 4;
CONTRAST 'Bottom vs. Top' trt −1 1 −1 1 0;
CONTRAST 'Cool vs. Hot' trt −1 −1 1 1 0;
CONTRAST 'Interaction Depth*Temp' trt 1 −1 −1 1 0;
RUN;
PROC PRINT DATA = adjmeans;
RUN;
PROC GLM;/*ancova for homogeneity of slopes*/
CLASS trt;
MODEL Finalwt = trt Initialwt trt*Initialwt;
RUN;
QUIT;
ODS RTF CLOSE;
```

4.2.2 Output of analysis

The results obtained by doing the analysis using the SAS procedures are given in the sequence. The first output of analysis corresponds to the first PROC GLM, which performs a simple linear regression analysis of the final weight on the initial weight over all the 25 observations. From the results it follows that for the overall experiment there is a significant linear relationship between the initial and the final weights. The model is able to explain about 90 per cent of the

total variability in the data (R^2= 0.895; p< 0.0001).

Table 4.2: Output from analysis using SAS

ANOVA

Source	DF	SS	MS	F Value	Prob > F
Initial Weight	1	441.662	441.662	195.27	<0.0001
Error	23	52.020	2.262		
Corrected Total	24	493.682			

R-Square	CV	Root MSE	Finalwt Mean
0.895	4.859	1.504	30.948

Parameter	Estimate	Standard Error	t Value	Prob > |t|
Intercept	2.8194	2.0353	1.39	0.1793
Slope (Initialweight)	1.0689	0.0765	13.97	<0.0001

It may be seen that the regression coefficient (or slope) is also highly significant ($p < 0.0001$), though the intercept is not significantly different from zero. The plot of initial weight against final weight (Figure 4.1) also indicates that the relationship between these two variables is linear.
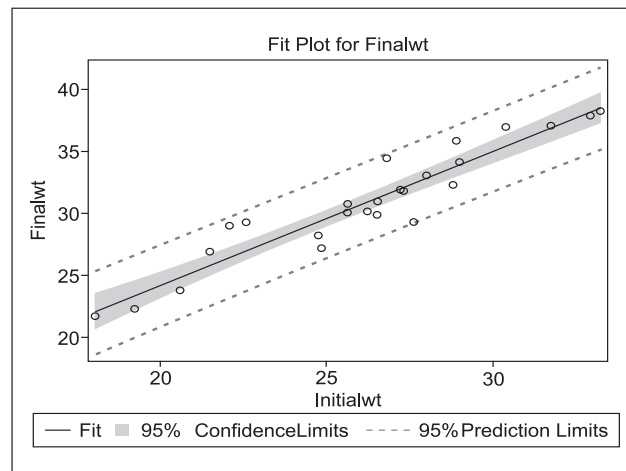
Figure 4.1: Plot of final weight vs initial weight

The highly significant slope in this linear regression analysis implies a strong dependence of final weight on the initial weight. This suggests that the initial weight may prove to be a useful covariate for the analysis.

The second PROC GLM performs a similar linear regression analysis within each treatment group separately. This analysis is based on 5 observations (replications of each treatment). The estimates of the slopes within each treatment group along with the standard errors and the *t* and *p* values are given in Table 4.3.

Table 4.3: Estimates of slopes

Parameter	Slope estimate	Standard error	t value	Prob> |t|
Initialweight (Trt 1 ~ Cool bottom)	1.1713	0.0850	13.78	0.0008
Initialweight (Trt 2 ~ Cool surface)	1.4015	0.0938	14.94	0.0007
Initialweight (Trt 3 ~ Hot bottom)	0.7601	0.4320	1.76	0.1767
Initialweight (Trt 4 ~ Hot surface)	1.0588	0.2158	4.91	0.0162
Initialweight (Trt 5 ~ Control)	1.0664	0.1694	6.30	0.0081

From this analysis it is found that the slope of the linear regression is fairly uniform over all the five treatments, by and large, except for hot-bottom combination for which the slope is slightly low. This fact is important because the analysis of covariance adjusts all the treatment groups by the same slope. The fitted equations for each treatment group are given below:

Trt 1	$y = -0.6574 + 1.1713X$		**Trt 4**	$y = 3.0752 + 1.0588X$
Trt 2	$y = -6.4825 + 1.4015X$		**Trt 5**	$y = 2.0398 + 1.0664X$
Trt 3	$y = 12.097 + 0.7601X$			

The third PROC GLM has only 'trt' (treatments) as the class statement indicating thereby that the only classification variable in the design is the treatment and, therefore, the design is a CRD. The analysis of the data for CRD is given in Table 4.4.

Table 4.4: Analysis of data as a CRD

ANOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Treatment	4	258.730	64.683	5.51	0.0037
Error	20	234.952	11.748		
Corrected Total	24	493.682			

R-Square	CV	Root MSE	Finalwt Mean
0.524	11.075	3.427	30.948

This analysis clearly reveals that the treatment effects are significantly different ($p = 0.0037$) meaning thereby that the location does affect the growth of oysters. The variation explained by fitting this model is merely 52 percent, though. Figure 4.2 gives the distribution of the final weight of oysters for each type of location (treatment).
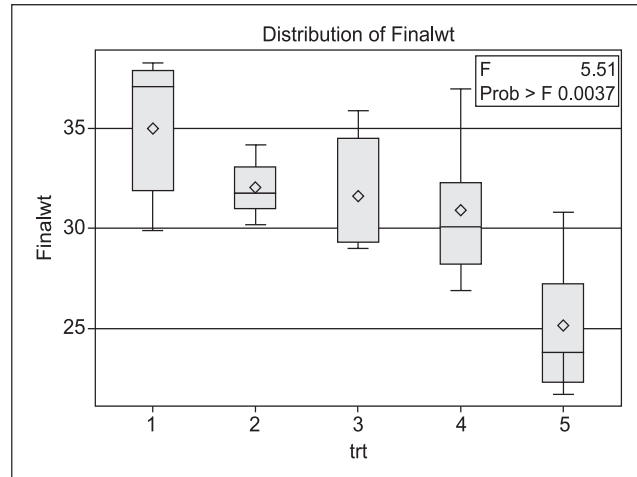
Figure 4.2: Treatment wise Box plot of final weight

The next PROC GLM is related to the ANCOVA, through which we try to answer the question as to whether or not the different locations (treatments) affect the final weights adjusted for differences in the initial weights of 25 bags of oysters. In other words, the ANCOVA answers the question about whether or not the different locations (treatments) affect the final weights if all the twenty five bags of oysters had started with the same initial weight. Once again, it may be noted that the class variable is the 'trt' only. The initial weight is not taken as a class variable, because this variable is designated as a covariate or regression variable. The output of ANCOVA is given in Table 4.5.

Table 4.5: Analysis of covariance result

Source	DF	Type I SS	MS	F Value	Prob > F
Treatment	4	258.730	64.683	32.01	<0.0001
Initial weight	1	196.544	196.544	97.23	<0.0001
Error	19	38.408	2.021		
Corrected Total	24	493.682			

Source	DF	Type III SS	MS	F Value	Prob > F
Treatment	4	13.612	3.403	1.68	0.1953
Initial weight	1	196.544	196.544	97.23	<0.0001
Error	19	38.408	2.021		
Corrected Total	24	493.682			

R-Square	CV	Root MSE	Finalwt Mean
0.922	4.594	1.422	30.948

One glaring thing that is noticeable from this analysis is that the model used explains about 92 per cent of the total variability in the data. On the other hand, the model without covariate (CRD model), explains about 52 percent of the variability in the data.

Since the two factors in the model, the class variable 'treatments (or locations)' and the regression variable 'initial weight' are not orthogonal to each other since all the five levels of treatments do not appear with every level of the variable initial weight, (orthogonality will be explained in detail in Chapter 5) the treatment sum of squares need to be adjusted for all other factors in the model. Similarly, the true test of the significance of the linear components of the relationship between 'initial weight'(X) and final weight (y) needs to use an Initial weight sum of squares adjusted for the effects of locations (or treatments). For this reason, it is preferable to use Type III sum of squares in the ANOVA.

For a better understanding of the importance of Type III SS, the Type I SS is also given above before giving the Type III SS. The Type I treatment SS is 258.730. The treatment effects in this case are highly significant ($p < 0.0001$). The Type I treatment SS is in fact the unadjusted treatment SS and is the same as the one found in the one-way ANOVA. If we subtract this SS from the Total SS, we obtain the error SS for the simple one-way ANOVA (493.682 − 258.730 = 234.952).

On the other hand the Type III SS for treatments is 13.612. Contrary to type I SS, this leads us to a conclusion that the treatment effects do not differ significantly ($p = 0.1953$). This is the adjusted treatment SS and allows us to test the treatment effects, adjusting for all other factors (in this case the initial body weights) included in the model. The reason for adjustments has already been described above. It is, therefore, evident that the covariate has an impact on the inference and treatment effects, which were significantly different in the absence of covariate have become homogeneous in the presence of covariate.

Table 4.6: Parameter estimates with standard errors

Parameter	Estimate	Standard error	t value	Prob > |t|
Intercept	2.7580	2.3592	1.17	0.2568
Trt1	0.7876	1.2865	0.61	0.5477
Trt2	0.8862	1.0865	0.82	0.4248
Trt3	2.3275	0.9912	2.35	0.0299
Trt4	0.9868	1.0203	0.97	0.3456
Trt5	0.0000	.	.	.
Initialweight	1.0333	0.1048	9.86	<0.0001

Note: *The matrix* **X'X** *has been found to be non-singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter B are not uniquely estimable.*

Table 4.6 provides a solution for testing the equality of slopes for each treatment. The last row of the Table "Initial Weight' provides the combined weighted slope of the regressions of final weight on initial weight for each treatment.

To compare the true effects of the locations (treatments), unbiased by differences in initial weights, the treatment means should be adjusted to what their values would have been if all the 25 bags had the same initial weight. The estimated least-squares means followed by their standard errors and the *p*-values for all pairwise tests of treatment differences are given in Table 4.7.

Table 4.7: Standard error and *p*-values for the treatments

Treatment	Finalwt LSMEAN	Standard Error	Prob > |t|	LSMEAN Number
1	30.7380	0.7700	<0.0001	1
2	30.8366	0.6478	<0.0001	2
3	32.2778	0.6395	<0.0001	3
4	30.9372	0.6359	<0.0001	4
5	29.9504	0.8002	<0.0001	5

The pairwise treatment comparisons have also been made and are presented in Table 4.8.

Table 4.8: *p*-values for pairwise treatment comparisons

Least Squares Means for effect Treatment Pr > |t| for H$_0$: LSMean(i)=LSMean(j) Dependent Variable: Finalweight					
i/j	1	2	3	4	5
1		1.0000	0.5781	0.9996	0.9714
2	1.0000		0.5345	1.0000	0.9226
3	0.5781	0.5345		0.5826	0.1731
4	0.9996	1.0000	0.5826		0.8664
5	0.9714	0.9226	0.1731	0.8664	

It may be noted from Table 4.8 that the *p*-values are high indicating that the differences are not significant.

It is also worthwhile noting the differences between the unadjusted to adjusted treatments means for the variable FINAL weight in Table 4.9.

Table 4.9: Treatment wise unadjusted and adjusted least square means

Treatment (or Location)	Unadjusted Means (Final weight)	Adjusted LS Means (Final weight)	Calculations for $\bar{Y}_{adj_i} = \bar{Y}_i - \beta(\bar{X}_i - \bar{X})$
1	35.020	30.7380	35.02 – 1.0333(30.46–26.32)
2	32.060	30.8366	32.06 – 1.0333(27.50–26.32)
3	31.600	32.2778	31.60 – 1.0333(25.66–26.32)
4	30.900	30.9372	30.90 – 1.0333(26.28–26.32)
5	25.160	29.9504	25.16 – 1.0333(21.68–26.32)

The differences in unadjusted and adjusted treatment means are due to the differences in initial weights among the treatment groups (for example, Treatment 5 was assigned much smaller oysters than other treatments). In calculating these adjusted means, the coefficient *β* =

1.0333 is a weighted average of the slopes of the linear regressions for each of the five treatment groups.

Table 4.10: Result of contrast analysis

Contrast	DF	Contrast SS	MS	F Value	Prob > F
Control *vs.* Treatment	1	3.596	3.596	1.78	0.1980
Bottom *vs.* Top	1	1.859	1.859	0.92	0.3496
Cool *vs.* Hot	1	2.700	2.700	1.34	0.2622
Interaction Depth*Temp	1	2.382	2.382	1.18	0.2913

The output indicates that oyster growth is not significantly affected by differences in temperature (cool *vs.* hot) or the depth (bottom *vs* top). Similarly the interaction between the depth and temperature is not significant. Although constructed to be orthogonal, these contrasts are not orthogonal to the covariate; therefore, their sums of squares do not add to the adjusted treatment SS.

The last PROC GLM is used for testing the heterogeneity of regression coefficients of treatment groups (or slopes). This is important because the ANCOVA assumes the homogeneity of slopes or the regression coefficients (or equality of regression coefficients) of the covariate for each treatment group. In other words, since a single regression coefficient (or slope) is used to adjust all observations in the experiment, it is desirable that the regression coefficients are same for each treatment group. This also implies that the estimate of each regression coefficient for each treatment group is an estimate of the same common slope for the entire data. The null hypothesis for testing the heterogeneity of regression coefficients is $H_0 : \beta_1 = \cdots = \beta_i = \cdots = \beta_v$, where β_i is the regression coefficient (or slope) of the regression pertaining to the ith level (or group) of treatment, assuming that there are v groups of treatments. As a matter of fact, the presence of interaction between the treatments (treatment groups) and the covariate is indicative of heterogeneity of regression coefficients. It also means that the regression relationship differs for different treatment groups.

In the sequel is presented the output obtained from this PROC GLM.

Table 4.11: Output from PROC GLM in SAS

ANCOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Treatment	4	5.988	1.497	0.66	0.6277
Initialweight	1	105.143	105.143	46.51	<0.0001
Initialweight*Treatment	4	4.501	1.125	0.50	0.7377
Error	15	33.907	2.260		
Corrected Total	24	493.682			

R-Square	CV	Root MSE	Finalwt Mean
0.931	4.858	1.503	30.948

One obvious thing that is noticeable from this analysis is that this model explains about 93 per cent of the total variability in the data. On the other hand, the model without interaction between treatment levels and the covariate explains about 92 percent of the variability in the data. Evidently, the interaction does not contribute much to the variability as is indicated by the *p*-value = 0.7377. Thus, it can be conclusively established that the null hypothesis of homogeneity of regression coefficients cannot be rejected. Performing an ANCOVA on this data with a common slope is, therefore, justified.

4.2.3 Analysis using R

In the sequence is described the R code for the analysis of covariance and testing the heterogeneity of slopes. Only the code is given for the benefit of readers more familiar with R software. The output of the analysis is not given to save space and repetition.

```
R code
d5=read.table("Oyster.txt",header=TRUE)
attach(d5)
names(d5)
#Treatment means and standard deviations
aggregate(Finalwt, by=list(trt), mean)
aggregate(Finalwt, by=list(trt), sd)
#Treatment wise box plot of Finalwt and Initialwt
boxplot(Finalwt~trt)
boxplot(Initialwt~trt)
#Regression of final weight on initial weight
lm1<-lm(Finalwt~Initialwt)
anova(lm1)
summary(lm1)
#Regression of final weight on initial weight for each treatment
summary(lm(Finalwt~Initialwt,trt==1,data=d5))
summary(lm(Finalwt~Initialwt,trt==2,data=d5))
summary(lm(Finalwt~Initialwt,trt==3,data=d5))
summary(lm(Finalwt~Initialwt,trt==4,data=d5))
summary(lm(Finalwt~Initialwt,trt==5,data=d5))
#onewayanova with treatments only
trt=factor(trt)
lm2=lm(Finalwt~trt)
anova(lm2)
#ancova with class variable treatments and contiuous initial body weight variable
lm3=lm(Finalwt~trt+Initialwt)
anova(lm3)
#To get type III sum of squares, download and install car package
library(car)
Anova(lm3,type="III")
```

summary(lm3)
lsm=lsmeans(lm3,"trt")
lsm
pairs(lsm)
contrast(lsm, list(con1 = c(-1,-1,-1,-1,4),con2=c(-1,1,-1,1,0),con3=c(-1,-1,1,1,0),con4=c(1,- 1,-1,1,0)))
lm4=lm(Finalwt~trt+Initialwt+trt:Initialwt)
Anova(lm4,type="III")
detach(d5)

4.3 Another example of analysis of covariance

In view of the importance of the analysis of covariance in analyzing the data generated from designed experiments and to control the experimental error, it may not be out of place to give one more example of analysis of covariance and to highlight some other important applications of analysis of covariance in analysis of data.

4.3.1 Example 2 (Gomez and Gomez, 1984)

A part of this example has been taken from Gomez and Gomez (1984). In order to study the effect of iron toxicity in soil on rice varieties, an experiment was designed as a RCB design with 15 rice varieties and three replications. The soil in the field where the experiment was conducted had a toxic level of iron. On two sides of each experimental plot, two guard rows of a susceptible check variety were planted to generate the iron toxicity score for a susceptible check variety. Scores for tolerance for iron toxicity were collected from each experimental plot as well as from guard rows. For each experimental plot, the score of susceptible check (averaged over two guard rows) provided the value of the covariate for that plot. Data on the tolerance scores of each variety (Y variable) and the corresponding susceptible check (X variable) are given in the Table 4.12.

In the sequence is done the analysis of covariance by treating the iron toxicity of susceptible check variety as covariate by making adjustments in the iron toxicity of rice varieties against the different values of iron toxicity of susceptible check variety.

Table 4.12: Scores of tolerance for iron toxicity (Y) of 15 rice varieties and those of the corresponding guard rows of a susceptible check variety (X) in a RCB design

Variety Number	Replication-I		Replication-II		Replication-III	
	X	*Y*	*X*	*Y*	*X*	*Y*
1.	15	22	16	13	16	14
2.	16	14	15	23	15	23
3.	15	24	15	24	15	23
4.	16	13	15	23	15	23
5.	17	17	17	16	16	16
6.	16	14	15	23	15	23
7.	16	13	15	23	16	13
8.	16	16	17	17	16	16
9.	17	14	15	23	15	24
10.	17	17	17	17	15	26
11.	16	15	15	24	15	25
12.	16	15	15	23	15	23
13.	15	24	15	24	16	15
14.	15	25	15	24	15	23
15.	15	24	15	25	16	16

4.3.2 Analysis of data

To begin with, the analysis is done without using the auxiliary information (or covariate *X*). The analysis would be same as that of a randomized complete block design as done in Chapter 2. This analysis would clearly highlight the advantage of using the auxiliary information. The SAS commands for analysis are:

DATA irontoxicityinrice;
INPUT trt rep Y X;
/*the first column (trt) gives the treatment numbers; the second column (rep) gives the replication numbers; the third column (Y) gives the tolerance score of iron toxicity in rice; the last column (X) gives the tolerance score of iron toxicity in a susceptible check variety*/
CARDS;

```
1    1    22    15
1    2    13    16
1    3    14    16
2    1    14    16
2    2    23    15
2    3    23    15
3    1    24    15
3    2    24    15
3    3    23    15
```

4	1	13	16
4	2	23	15
4	3	23	15
5	1	17	17
5	2	16	17
5	3	16	16
6	1	14	16
6	2	23	15
6	3	23	15
7	1	13	16
7	2	23	15
7	3	13	16
8	1	16	16
8	2	17	17
8	3	16	16
9	1	14	17
9	2	23	15
9	3	24	15
10	1	17	17
10	2	17	17
10	3	26	15
11	1	15	16
11	2	24	15
11	3	25	15
12	1	15	16
12	2	23	15
12	3	23	15
13	1	24	15
13	2	24	15
13	3	15	16
14	1	25	15
14	2	24	15
14	3	23	15
15	1	24	15
15	2	25	15
15	3	16	16

```
;
ODS RTF FILE = 'OYSTER.RTF';
PROC GLM;/*anova for two way classified data*/
CLASS trt rep;
MODEL Y = trt rep / LSMEANS PDIFF;
RUN;
ODS RTF CLOSE;
```

4.3.3 Output of analysis

The results obtained by using the SAS procedures are described in Table 4.13.

Table 4.13: Results from SAS for iron toxicity data

ANOVA

Source	DF	Type III SS	Mean Square	F Value	Prob > F
Replications	2	104.044	52.022	2.85	0.0745
Treatments	14	265.911	18.994	1.04	0.4448
Error	28	510.622	18.237		
Corrected Total	44	880.578			

R-Square	CV	Root MSE	y Mean
0.420	21.544	4.270	19.822

The fitted model is able to explain only 42 percent of the total variability in the data. It may be seen here that the treatment effects are not significantly different. The replications are also not significantly different at 5 % level of significance, although it appears from the figure below that the replications differ.
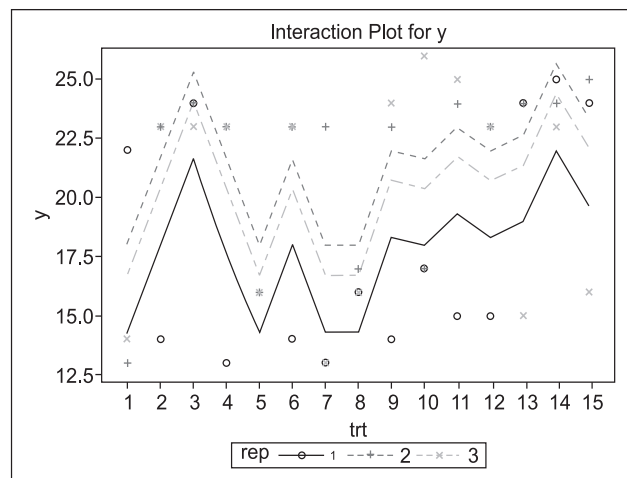
Figure 4.3: Replication wise iron toxicity

Table 4.14: *t* **comparison Lines for Least Squares Means of treatment**

		Y LSMEAN	Treatment	LSMEAN Number
LS-means with the same letter are not significantly different				
	A	24.000	14	14
	A	23.667	3	3
B	A	21.667	15	15
B	A	21.333	11	11
B	A	21.000	13	13
B	A	20.333	12	12
B	A	20.333	9	9
B	A	20.000	2	2
B	A	20.000	10	10
B	A	20.000	6	6
B	A	19.667	4	4
B		16.333	1	1
B		16.333	5	5
B		16.333	8	8
B		16.333	7	7

The analysis of covariance (ANCOVA) is now performed using the covariate, *X*. The analysis is done as follows:

The SAS file has already been prepared while analyzing the data without covariate. To that file, the following SAS commands may be added:

```
DATA irontoxicityinrice;
INPUT trt rep Y X;
/*the first column (trt) gives the treatment numbers; the second column (rep) gives the
replication numbers; the third column (Y) gives the tolerance score of iron toxicity in rice; the
last column (X) gives the tolerance score of iron toxicity in a susceptible check variety*/
CARDS;
. . . . . . . . . . . . . . . . . .; /*insert data here*/
ODS RTF FILE = 'RESULT.RTF';
PROC GLM; /*ancova for two way classified data*/
CLASS trt rep;
MODEL Y = trt rep X;
```

LSMEANS TRT/PDIFF LINES;
RUN;
ODS RTF CLOSE;
The results obtained are given in Table 4.15.

Table 4.15: Output of analysis of covariance

ANOVA

Source	DF	Type III SS	MS	F-Value	Prob> F
Replication	2	22.480	11.240	2.71	0.0844
Treatment	14	152.561	10.897	2.63	0.0151
X (Check Variety as Covariate)	1	398.752	398.752	96.24	<0.0001
Error	27	111.871	4.143		
Corrected Total	44	880.578			

R-Square	CV	Root MSE	y Mean
0.873	10.269	2.035	19.822

A glaring thing to notice is that the total variability explained by the model with covariate has gone up to 87 percent compared to 42 percent with a model without covariate. This clearly indicates that the covariate produces variability and the tolerance score of iron toxicity in rice varieties needs to be adjusted against the variable score of toxicity in the susceptible check variety. Obviously, therefore, the use of covariate has resulted into a considerable reduction in the error mean square and hence the CV has also reduced drastically from 21.544 to 10.269. This has been possible because the effect of the covariate, X, is highly significant (p-value < 0.0001). As a matter of fact, the treatments effects are now significantly different even at 2% level of significance. Without using the auxiliary information, the treatment effects were not significantly different. The use of covariate in the analysis has helped in catching the small differences among the treatment effects as significant. The covariance analysis will thus result into a more precise comparison of treatment effects.

The unadjusted means and the adjusted (LSMEANS) are given in the Table 4.16. It may be seen that the two means differ for all the varieties. This once again indicates that it has been worthwhile using the covariate in the model.

Table 4.16: Treatment wise unadjusted and least square means

Level of Treatment	N	Y	
		Unadjusted Mean	LSMEAN
1	3	16.333	16.875
2	3	20.000	18.512
3	3	23.667	20.149
4	3	19.667	18.178
5	3	16.333	22.963
6	3	20.000	18.512
7	3	16.333	16.875
8	3	16.333	20.934
9	3	20.333	20.875
10	3	20.000	24.600
11	3	21.333	19.845
12	3	20.333	18.845
13	3	21.000	19.512
14	3	24.000	20.483
15	3	21.667	20.178

Table 4.17: t comparison lines for least squares means of treatments after ANCOVA

LS-means with the same letter are not significantly different					
			Y LSMEAN	Treatment	LSMEAN Number
	A		24.600	10	10
B	A		22.963	5	5
B	C		20.934	8	8
B	C		20.875	9	9
B	C		20.482	14	14
B	C	D	20.178	15	15
B	C	D	20.149	3	3
B	C	D	19.845	11	11
B	C	D	19.512	13	13
	C	D	18.845	12	12
	C	D	18.512	6	6

LS-means with the same letter are not significantly different					
			Y LSMEAN	Treatment	LSMEAN Number
	C	D	18.512	2	2
	C	D	18.178	4	4
		D	16.875	7	7
		D	16.875	1	1

It is easily seen from Table 4.17, in comparison to Table 4.14 when the analysis was done without the covariate, that small treatment differences have been caught as significant. The inference emerging from a model without covariate was altogether different from the one obtained using a model with covariate. For instance, treatment 14 was significantly different from treatments 1, 5, 7, 8 in the without covariate model. But in the covariate model, treatment 14 is significantly different from treatments 1 and 7 only. Similarly, in the no covariate model, treatment 10 was statistically at par with all other treatments, whereas in the covariate model, this treatment performs the best and is statistically at par with only treatment 5 and is significantly higher in terms of toxicity compared with all other treatments.

4.3.4 Analysis using R

The purpose of this section is to provide the R code for analysis of covariance. The code is given for the benefit of readers who use R software for analysis. For saving the space and to avoid repetition, the results obtained from analysis using the R code are not given.

R code
```
d6=read.table("irontoxicity.txt",header=TRUE)
attach(d6)
names(d6)
#Treatment means and standard deviations
aggregate(y~trt,data=d6,mean)
aggregate(y~trt,data=d6,sd)
#Treatment wise box plot of y and x
boxplot(y~trt)
boxplot(x~trt)
#Two-way anova with treatments and replications
trt=factor(trt)
rep=factor(rep)
lm1=lm(y~trt+rep)
anova(lm1)
summary(lm1)
library(lsmeans)
```

```
#ancova with class variable treatments and replication and continuous x variable
lm2=lm(y~trt+rep+x)
anova(lm2)
#to get type III sum of squares, install car package
library(car)
Anova(lm2,type="III")
summary(lm2)
library(lsmeans)
lsm=lsmeans(lm2,"trt")
lsm
pairs(lsm)
detach(d6)
```

4.4 ANCOVA in analysis of data with missing observations

Loss of data from a well planned, designed and managed experiment is a common phenomenon. The loss of data may render even a good design to lose its properties. For instance, if the experiment has been run as a randomized complete block (RCB) design or a balanced incomplete block (BIB) design and if there is loss of data, then the properties of the original design are lost. RCB design is an orthogonal design, but if there is a loss of data because of some accident, then the resulting design becomes a non-orthogonal design. Similarly, a BIB design is variance balanced and is most efficient for making all the possible pairwise treatment comparisons. However, with loss of data in a BIB design, the resulting design may not remain a variance balanced design. Further, the loss of data at times, may result in a design which is not treatment connected in the sense that it may not be possible to make all the possible pairwise treatment comparisons.

When there is a loss of data in any designed experiment, the analysis of data can be done in any of the following ways:

(a) Estimate the missing values of observations lost by minimizing the error sum of squares. Plug in the estimated values for the observations lost and analyze the data using the original design. The treatment sum of squares, however, needs to be adjusted in this case.

(b) Use analysis of covariance by defining covariates. The covariates in this case are the pseudo variables. The observations are taken as zero for the missing data. If n is the total number of observations in the original design and if p observations are lost, then the p observations lost are taken as zero and p covariates (pseudo variables) are defined one each for each of the missing observations. A covariate takes $n - 1$ values as -1 (or 0) and the value corresponding to the one missing observation as +1. A covariate is defined in a similar way for each of the p missing observations. The ANCOVA is performed using p covariates.

(c)   Analyze the residual design obtained after losing the data. The observations lost and the treatments corresponding to the lost observations are dropped. If n is the total number of observations in the original design and if p observations are lost, then the residual design in $n - p$ observations is analyzed using ANOVA. Obviously the replication of the treatments and the block sizes change.

The use of ANCOVA in the analysis of data with missing observations is illustrated in the sequel.

4.4.1 Example 3

An experiment was conducted at Jorhat under the aegis of Project Directorate of Cropping Systems Research, Modipuram (now known as Indian Institute of Farming Systems Research, Modipuram) using a balance incomplete block (BIB) design with parameters $v = b = 7$, $r = k = 4$, $\lambda = 2$ (for definition and meaning of the parameters, the reader may refer to Chapter 5). The treatment details, the block structure and the yields from each of the crop sequences converted into calories / hectare are given in the Table 4.18.

Table 4.18: Treatment and block details and yield (calorie/hectare) from crop sequence

Season	Treatments (Crop Sequences)						
	T1	T2	T3	T4	T5	T6	T7
Kharif	Rice	Rice	Rice	Rice	Rice	Rice	Rice
Rabi	-	Boro Rice	Mustard	Brinjal	Tomato	French Bean	Potato
Summer	Rice	-	Rice	Rice	Rice	Rice	Rice

Block 1	T2 (3325060)	T4 (2606200)	T5 (3279420)	T6 (2330180)
Block 2	T1 (2992900)	T2 (−)	T5 (3348780)	T7 (2982000)
Block 3	T1 (3639920)	T4 (2467800)	T6 (2196580)	T7 (−)
Block 4	T3 (2602410)	T5 (3696340)	T6 (2388060)	T7 (2921790)
Block 5	T1 (3055180)	T3 (2653680)	T4 (2501060)	T5 (3594320)
Block 6	T2 (3380420)	T3 (2760690)	T4 (2522100)	T7 (2961270)
Block 7	T1 (2921420)	T2 (3380420)	T3 (2677400)	T6 (2594420)
	T# denotes the treatment number and the figures in brackets are the Kilo calories/hectare			

It may be noted that during the experimentation two observations pertaining to treatment T2 in block 2 and treatment T7 in block 3 are lost. In the sequel the data generated from a BIB design with two missing observations is analyzed using analysis of covariance. Two pseudo variables are defined, one for each of the missing observation.

4.4.2 Analysis using SAS

The analysis of the data obtained from a BIB design with two missing observations is done using analysis of covariance. Two covariates (pseudo variables), X1 and X2, are defined, as described above. The data file for analysis using PROC GLM of SAS is given below.

```
DATA ancova_bibd;
INPUT blk trt cal X1 X2;
/*In the input, blk denotes the block number; trt denotes the treatment number; cal denotes
the total calorie obtained from the crop sequence; X1 denotes the first pseudo variable (or
covariate) and X2 denotes the second pseudo variable (or covariate) */
CARDS;
1       2       3325060         0       0
1       4       2606200         0       0
1       5       3279420         0       0
1       6       2330180         0       0
2       1       2992900         0       0
2       2             0         1       0
2       5       3348780         0       0
2       7       2982000         0       0
3       1       2639980         0       0
3       4       2467800         0       0
3       6       2196580         0       0
3       7             0         0       1
4       3       2602410         0       0
4       5       3696340         0       0
4       6       2388060         0       0
4       7       2921790         0       0
5       1       3055180         0       0
5       3       2653680         0       0
5       4       2501060         0       0
5       5       3594320         0       0
6       2       3380420         0       0
6       3       2760690         0       0
6       4       2522100         0       0
6       7       2961270         0       0
7       1       2921420         0       0
7       2       3380420         0       0
7       3       2677400         0       0
7       6       2594420         0       0
;
ODS RTF FILE = 'RESULT.RTF';
PROC GLM;
CLASS trt blk;
```

MODEL cal = blk trt X1 X2;
MEANS trt;
LSMEANS trt / PDIFF LINES;
RUN;
ODS RTF CLOSE;

4.4.3 Output of analysis

The results obtained from the analysis are given in Table 4.19.

Table 4.19: Output of analysis using SAS

ANCOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks (adjusted)	6	152491189324	25415198221	1.89	0.1588
Treatments (adjusted)	6	3.2181177E12	536352954798	39.81	<0.0001
X1 (Covariate 1)	1	5.9843328E12	7878384853.3	0.58	<0.0001
X2 (Covariate 2)	1	3.9654507E12	8618880	0.00	<0.0001
Error	13	175149529143	13473040703		
Corrected Total	27	4.3125834E12			

R-Square	CV	Root MSE	cal Mean
0.991	4.346	116073.4	2670710

Using the model with two covariates, about 96 per cent of the total variability in the data is explained. Obviously, therefore, the CV obtained from the analysis is 4.025. From the above ANCOVA it may be concluded that the treatment effects differ significantly.

The unadjusted and the adjusted LSMEANS of the treatments are given in the Table 4.20.

Table 4.20: Treatment wise least square mean

Level of Treatment	N	cal	
		Unadjusted Mean	LSMEAN
1	4	2902370.00	2696102.86
2	4	2521475.00	3122858.57
3	4	2673545.00	2396834.29
4	4	2524290.00	2341277.14
5	4	3479715.00	3245591.43
6	4	2377310.00	2190904.29
7	4	2216265.00	2701401.43

It may be seen from Table 4.20 that the unadjusted means differ substantially from the LS means. For any comparisons, therefore, LS means should be used. The pairwise treatment comparisons can be made using the Table 4.21.

Table 4.21: *t* comparison lines for least squares means of trt

		cal LSMEAN	Treatment	LSMEAN Number
	LS-means with the same letter are not significantly different

		cal LSMEAN	Treatment	LSMEAN Number
	A	3245591.43	5	5
	A	3122858.57	2	2
	B	2701401.43	7	7
	B	2696102.86	1	1
	C	2396834.29	3	3
D	C	2341277.14	4	4
D		2190904.29	6	6

4.4.4 Analysis of data after deleting the missing observations

The analysis of the data is done again after deleting the two missing observations corresponding to treatment 2 in block 2 and treatment 7 in block 3, as suggested above. In other words, instead of analyzing 28 observations, we now analyze 26 observations. The original design was balanced incomplete block design with parameters $v = 7$, $b = 7$, $r = 4$, $k = 4$, $\lambda = 2$. It is well known that a balanced incomplete block design is a non-orthogonal design and, therefore, the treatment sum of squares is adjusted for blocks and block sum of squares is adjusted for treatments. After deleting the two observations, the resulting design is again an incomplete block design with parameters $v = 7$, $b = 7$, $r_1 = 4$, $r_2 = 3$, $r_3 = 4$, $r_4 = 4$, $r_5 = 4$, $r_6 = 4$, $r_7 = 3$, $k_1 = 4$, $k_2 = 3$, $k_3 = 3$, $k_4 = 4$, $k_5 = 4$, $k_6 = 4$, $k_7 = 4$. But this design is not a balanced incomplete block design. The analysis of the data is given in the sequel. The data file for analysis using PROC GLM of SAS is given below:

DATA ancova_bibd;
INPUT blk trt cal;
/*In the input, blk denotes the block number; trt denotes the treatment number; cal denotes the total calorie obtained from the crop sequence*/
CARDS;

1	2	3325060
1	4	2606200
1	5	3279420
1	6	2330180
2	1	2992900
2	5	3348780
2	7	2982000
3	1	2639980
3	4	2467800
3	6	2196580
4	3	2602410

4	5	3696340
4	6	2388060
4	7	2921790
5	1	3055180
5	3	2653680
5	4	2501060
5	5	3594320
6	2	3380420
6	3	2760690
6	4	2522100
6	7	2961270
7	1	2921420
7	2	3380420
7	3	2677400
7	6	2594420

```
;
ODS RTF FILE = 'RESULT2.RTF';
PROC GLM;
CLASS trt blk;
MODEL cal = trt blk;
MEANS trt;
LSMEANS trt/PDIFF LINES;
RUN;
ODS RTF CLOSE;
```

4.4.5 Output of analysis

The results obtained by using PROC GLM for analysis of a general block design obtained after deleting the two missing observations are given in the Table 4.22.

Table 4.22: Result of analysis using PROC GLM

ANOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Treatments (adjusted)	6	3.2181177E12	536352954798	39.81	<0.0001
Blocks (adjusted)	6	152491189324	25415198221	1.89	0.1588
Error	13	175149529143	13473040703		
Corrected Total	25	4.1690508E12			

R-Square	CV	Root MSE	cal Mean
0.958	4.036	116073.400	2876149.000

It may be noted that the adjusted sum of squares due to treatments, adjusted sum of squares due to blocks and error sum of squares are exactly same in the two different approaches of analysis of data with missing observations. Consequently, the F-value and the Prob > F are also same in the two analysis. However, the corrected total sum of squares is different in the two cases. In one approach of analysis, the total number of observations is 28, because the two missing observations are taken as zero. In the second approach of analysis there are only 26 observations. The two missing observations are deleted from the data.

Table 4.23: Treatment wise unadjusted and least square means

Level of Treatments	N	cal	
		Unadjusted Mean	LS MEAN
1	4	2902370.00	2913120.00
2	3	3361966.67	3339875.71
3	4	2673545.00	2613851.43
4	4	2524290.00	2558294.29
5	4	3479715.00	3462608.57
6	4	2377310.00	2407921.43
7	3	2955020.00	2918418.57

It may be noticed that the unadjusted means of treatments 2 and 7 are different in the two approaches of analysis. The difference has occurred because in ANCOVA approach, the observations are treated as zero; so the total number of observations remains 4 for both these treatments. However, in the other approach, the two observations are deleted. Hence the number of observations for treatments 2 and 7 are 3 each and not 4.

Table 4.24: *t* comparison lines for least squares means of treatment

		LS-means with the same letter are not significantly different		
		cal LSMEAN	Treatment	LSMEAN Number
	A	3462608.57	5	5
	A	3339875.71	2	2
	B	2918418.57	7	7
	B	2913120.00	1	1
	C	2613851.43	3	3
D	C	2558294.29	4	4
D		2407921.43	6	6

Although the LS means obtained here are different from those obtained in the approach using ANCOVA, the inferences made are exactly the same.

4.4.6 Analysis using R

The R code is given for the benefit of readers who use R software for analysis. For saving the space and to avoid repetition, the results obtained from using the R code are not given.

R code

```
d7=read.table("ancova_bibd.txt",header=TRUE)
attach(d7)
names(d7)
#Treatment means and standard deviations
aggregate(cal~trt,data=d7,mean)
aggregate(cal~trt,data=d7,sd)
#Treatment wise box plot of cal
boxplot(cal~trt)
#ancova with class variable treatments and block and contiuous x1 and x2 variable
trt=factor(trt)
blk=factor(blk)
lm1=lm(cal~trt+blk+x1+x2)
#To get type III sum of squares, install car package
library(car)
Anova(lm1,type="III")
summary(lm1)
library(lsmeans)
lsm=lsmeans(lm1,"trt")
lsm
cld(lsm,Letters="ABCDEF")
pairs(lsm)
#Analysis after deleting missing observations
missing.obs=c(which(d7$x1==1),which(d7$x2==1))
d7=d7[-missing.obs,]
lm2=lm(cal~factor(trt)+factor(blk),data=d7)
Anova(lm2,type="III")
lsm2=lsmeans(lm2,"trt")
lsm2
cld(lsm2,Letters="ABCDEF")
detach(d7)
```

5

Incomplete Block Design without and with Interaction and Nested Classification

5.1 Introduction

In Chapter 2 it has been seen that if there is only one source of variability in the experimental material, then it can be taken care of by grouping (or blocking) the experimental units in such a way that within block variability is small and between blocks variability is large. Designs useful for such experimental situations are designs for one-way elimination of heterogeneity setting or block designs. The focus there was restricted to randomized complete block (RCB) designs, which have complete blocks in the sense that all the treatments appear precisely once in each block, *i.e.*, each block is a complete replication. However, there do occur experimental situations where it is not possible to accommodate all the treatments in every block. Incomplete block designs (block designs in which some or all blocks are incomplete in the sense that these blocks are not complete replication) are useful for these experimental situations. The purpose of this Chapter is to study such incomplete block designs.

In actual practice there do occur experimental situations where the number of treatments is large. We have come across actual experiments being conducted with number of treatments as large as 36 or even more than 80. When the number of treatments in an experiment is large, it is not possible to form blocks that contain as many homogeneous experimental units as the number treatments in the experiment. This is because of the fact that when the block size becomes large, since large number of treatments is to be accommodated within each block, it is difficult to maintain homogeneity within large blocks, leading thereby to large variability within blocks. Further, in hilly areas or wastelands, it may not be possible to have number of experimental units same as the number of treatments even when the experimenter is interested in comparing small number of treatments. If the experimenter forces the blocks to accommodate all the treatments so as to form a complete block or a replicate, then the purpose of blocking would be defeated. In that case the per plot variance would be very large and finally the CV of the design would be high. Many such experiments get rejected by the experimenters, particularly in the Initial Varietal Trials in the Crop Improvement Programme.

Intuitively, a remedy to this problem is to form blocks or groups of homogeneous experimental units by reducing the number of experimental units in each block. In other words, in order to overcome the problem of maintaining homogeneity among the experimental units within a block, form incomplete blocks instead of forming complete blocks, as in an RCB design. Such a design will be termed as an *incomplete block design*. An incomplete block design

is a block design in which there is at least one block that does not contain all the treatments. In other words, there is at least one block in the design which is not a complete block or a complete replication.

Alternatively, one can view a block design as a design involving two factors each having certain number of levels. The two factors are the treatments and the blocks and the levels are the number of treatments and number of blocks. In case of an RCB design, every level of treatment occurs with every level of blocks. In that sense the design is *orthogonal*. On the contrary, in case of an incomplete block design, every level of treatment does not appear with every level of block. In that sense, these designs are *non-orthogonal* designs. Therefore, when we use an incomplete block design, there would be confounding of the treatments with the blocks. Consequently, there would be a loss of information in estimating treatment contrasts from an incomplete block design, unlike the RCB design, which is an orthogonal design, where all the treatment contrasts are estimated without any loss of information. In this sense incomplete block designs may not appear to be useful. However, in these designs there is a substantial reduction in block size and consequently the per plot variance (intra block variance) also reduces considerably. This in itself is a very big advantage, which more than offsets the loss in precision because of incomplete blocks.

To make the exposition clear, suppose that an experiment is to be conducted as incomplete block design for $v = 7$ treatments in $b = 7$ blocks of size $k = 3$ each and replication of each treatment is $r = 3$. The blocks with treatment contents are (1, 2, 4); (2, 3, 5); (3, 4, 6); (4, 5, 7); (5, 6, 7, 1); (6, 7, 2); (7, 1, 3). The treatments are labeled as 1 through 7 and the arrangement is without any randomization. Suppose block 1 is compared with block 2, then treatments 1, 3, 4, 5 get mixed up with the block differences. Similarly, when blocks 5 and 7 are compared, treatments 3, 5, 6, 7 get mixed up with the block differences. Some other treatments would get mixed up with some other pair of block comparisons. Thus, treatment effects get mixed up or entangled with the block differences and the two cannot be separated. In other words, the treatment effects are said to be confounded with the block effects. Therefore, the use of an incomplete block design will always be accompanied with loss of information in estimating differences of treatment effects, unlike the RCB design, which is an orthogonal design, where all the treatment contrasts are estimated without any loss of information under the assumption that error variance is same for RCB design and incomplete block designs. Obviously, since RCB design is a complete block design, for any block comparisons, all the treatment effects get eliminated. Thus, there is no confounding of treatments with blocks and consequently, there is no loss of information in treatment comparisons. In this sense, incomplete block designs may appear not to be useful. However, following on the example above one can easily see that the block size has reduced from 7 in case of a complete block design to 3 in an incomplete block design. In other words, in an incomplete block design, there is a reduction of more than 50 per cent in the block size. This considerable reduction in block size leads to a sizeable reduction in the block variability or the per plot variance (intra block variance or error variance). This in itself is a very big advantage which more than offsets the loss in precision because of incomplete blocks.

A problem that immediately arises in the adoption of an incomplete block design is to get such a design. Because the blocks are incomplete, the block structure needs to be decided

on statistical considerations. An arbitrary arrangement of treatments in blocks may lead to a design that may not allow the estimation of all the pairwise differences among the treatments. The other problem associated with such designs is that the ease of analysis is lost. The analysis of data generated from incomplete block design would also become complicated. Since the design is non-orthogonal, for testing any hypothesis about treatment effects, the treatment sum of squares would need to be adjusted for the blocks. The unadjusted treatment sum of squares would not be the correct sum of squares for testing the null hypothesis about homogeneity of treatment effects. Similarly, for testing any hypothesis about block effects, the block sum of squares would need to be adjusted for treatments. The unadjusted block sum of squares would not be the correct sum of squares for testing the null hypothesis about the homogeneity of block effects. However, in case of orthogonal designs, the adjusted sum of squares is equal to the unadjusted sum of squares. Further, the unadjusted treatment means would not suffice to test the hypotheses about pairwise treatment comparisons, since the design is non-orthogonal. In this case one would have to work out the adjusted means called least square means (LS means). In case of orthogonal designs, the unadjusted means are same as the adjusted means.

From the point of view of the stakeholders or experimenters from other sciences, using an incomplete block design may not be an easy affair. A strong interaction with the statistician before the actual experiment is laid out is absolutely essential. But from practical considerations, the advantages are enormous.

The class of incomplete block designs is very wide. Among the class of incomplete block designs, a Balanced Incomplete Block (BIB) design is the simplest design. For making all the possible pairwise treatment comparisons, a BIB design is the most efficient design among all designs with given number of treatments, number of blocks and block sizes. Among the other incomplete block designs are partially balanced incomplete block designs, which form a very broad class of designs containing group divisible and extended group divisible designs, square and rectangular lattice, alpha designs, etc. It would be beyond the scope of this book to give details of such designs. We shall restrict ourselves to defining BIB design with some detail. For the analysis of data generated from any incomplete block design, the SAS and R commands would be given. Some examples would also be given to explain how the SAS commands are useful in analysis of such data. These SAS and R commands are similar for other classes of incomplete block designs.

5.2 Randomization of treatments in an incomplete block design

Since the blocks are incomplete in an incomplete block design, the randomization of treatments to the blocks is slightly different from that of a RCB design. The randomization of treatments to experimental units comprises of the following steps:

1. Randomize the treatment labels.
2. Randomize the blocks.
3. Randomize the treatments within each block.
4. A separate randomization is done in each block.

5.3 Analysis of incomplete block design

The model for analysis of data from an incomplete block design is

response = general mean + treatments effects + blocks effects +error

Alternatively, the mathematical model is

$$y_{iju} = \mu + \tau_i + \beta_j + e_{iju}, \; i = 1,2,\cdots,v; \; j = 1,2,\cdots,b; \; u = 0 \text{ or } 1$$

where y_{iju} is the observation recorded on the ith treatment in the jth block, $i = 1,2,\cdots,v$; $j = 1,2,\cdots,b$, μ is the general mean; τ_i is the effect of ith treatment, β_j is the effect of jth block, and e_{iju}'s are uncorrelated random error components assumed to be distributed normally with zero mean and constant variance σ^2. $u = 0$ or 1 indicates that the (i, j)th cell has no observation or one observation, *i.e.*, a treatment can appear at most once in a block.

In intra-block analysis we assume that the treatment effects and the block effects are fixed, though unknown and e_{ij}'s are uncorrelated random variables. Since an incomplete block design is a non-orthogonal design, there would be entanglement/confounding of effects of treatment differences with the block comparisons, and so the treatment sum of squares would have to be adjusted for blocks. For an incomplete block design with v treatments, b blocks, replication of treatments r and block sizes k, the ANOVA for testing the null hypothesis $H_o : \tau_1 = \tau_2 = \cdots = \tau_v = \tau$ (say) against an alternative hypothesis $H_1 : \tau_i \neq \tau_l$ for some $i \neq l = 1,2,\cdots,v$ takes the form given in Table 5.1.

Table 5.1: ANOVA table

Source	DF	SS	MS	F
Treatments (adjusted)	$v - 1$	SST	$s_t^2 = SST/(v-1)$	$\dfrac{s_t^2}{s_e^2}$
Blocks (unadjusted)	$b - 1$	SSB		
Error (Intra-block)	$bk - b - v + 1$	SSE	$s_e^2 = SSE/(bk - b - v + 1)$	
Total	$bk - 1$	$\sum\limits_{i=1}^{v}\sum\limits_{j=1}^{b} y_{ij}^2 - CF$		

If $F_{cal} > F_{\alpha;(v-1),(bk-b-v+1)}$, then the treatment effects are significant. Here $F_{\alpha;(v-1),(bk-b-v+1)}$ is the table value of Snedecor's F distribution with $v - 1$ and $bk - b - v + 1$ degrees of freedom and at α level of significance.

On the other hand for making block comparisons, the analysis of variance would be different. In this case the null hypothesis would be $H_o : \beta_1 = \beta_2 = \cdots = \beta_b = \beta$ (say) against an alternative hypothesis $H_1 : \beta_j \neq \beta_u$ for some $j \neq u = 1,2,\cdots,b$. The ANOVA in this case is given in Table 5.2.

Table 5.2: ANOVA table

Source	DF	SS	MS	F
Blocks (adjusted)	$b - 1$	SSB	$s_b^2 = SSB/(b-1)$	$\dfrac{s_b^2}{s_e^2}$
Treatments (unadjusted)	$v - 1$	SST		
Error (Intra-block)	$bk - b - v + 1$	SSE	$s_e^2 = SSE/(bk - b - v + 1)$	
Total	$bk - 1$	$\sum_{i=1}^{v}\sum_{j=1}^{b} y_{ij}^2 - CF$		

If $F_{cal} > F_{\alpha;(b-1),(bk-b-v+1)}$, then the block effects are significant. Here $F_{\alpha;(b-1),(bk-b-v+1)}$ is the table value of Snedecor's F distribution with $b-1$ and $bk - b - v + 1$ degrees of freedom and at α level of significance.

Remark 5.1. The following identity holds:

Treatment sum of squares (adjusted) + Block sum of squares (unadjusted) = Treatment sum of squares (unadjusted) + Block sum of squares (adjusted).

This is so because of the fact that the model sum of squares in both the cases would be same. As a consequence, the error sum of squares in both the cases would also be the same.

5.4 Balanced incomplete block design

A BIB design is an arrangement of v treatments in b blocks such that

i.   each treatment occurs at most once in a block, *i.e.*, a treatment either occurs or does not occur in a block and all treatments do not occur in a block,

ii.   each block contains $k\ (<v)$ treatments, *i.e.*, the block size is $k\ (<v)$,

iii. each treatment occurs in exactly $r\ (<b)$ blocks,

iv.   every pair of treatments occurs together in exactly $\lambda\ (<b)$ blocks.

The symbols v, b, r, k, λ are called the parameters of the design. These parameters are not independent. The following parametric relations hold:

(i) $vr = bk$, (ii) $\lambda(v-1) = r(k-1)$. Besides the two parametric relations, for a BIB design the inequality, $b \geq v$ holds. This is known as Fisher's inequality.

Let $d_{il} = \tau_i - \tau_l,\ i \neq l = 1,2,\cdots,v$, define the difference of two treatment effects. An incomplete block design is said to be balanced (variance balanced) if the variance of any estimated difference of treatment effects, $\hat{d}_{il}$ (or estimated elementary contrast of treatments), is constant, *i.e.* $V(\hat{d}_{il}) = $ constant for all $i \neq l = 1,2,\cdots,v$.

For a BIB design, $V(\hat{d}_{il}) = \dfrac{2k}{\lambda v}\sigma^2$, for all $i \neq j = 1,2,\cdots,v$, where σ^2 is the intra block variance or per plot variance or error variance.

Remark 5.2 If $\lambda = b$, i.e., every pair of treatments appears together in all the b blocks, then $v = k$ and $r = b$. In that case a BIB design reduces to RCB design.

Example 1 A BIB design for $v = 8, b = 14, r = 7, k = 4, \lambda = 3$ is the following:

Blocks													
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14
1	2	3	4	5	6	7	3	4	5	6	7	1	2
2	3	4	5	6	7	1	5	6	7	1	2	3	4
4	5	6	7	1	2	3	6	7	1	2	3	4	5
8	8	8	8	8	8	8	7	1	2	3	4	5	6

5.4.1 Symmetric BIB Design

A BIB design with $v = b$ (and consequently $r = k$) is called a symmetric BIB design. In a symmetric BIB design any two blocks have λ treatments common. The following is an example of a symmetric BIB design with $v = b = 11, r = k = 5, \lambda = 2$:

B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11
1	2	3	4	5	6	7	8	9	10	11
3	4	5	6	7	8	9	10	11	1	2
4	5	6	7	8	9	10	11	1	2	3
5	6	7	8	9	10	11	1	2	3	4
9	10	11	1	2	3	4	5	6	7	8

In this design there are always two treatments common in any two blocks.

5.4.2 Efficiency of BIB design

Let d_{il}, $i \neq l = 1,2,\cdots,v$, denote the difference of two treatment effects. The variance of the estimated difference between any two treatment effects from a BIB design is given by

$$V(\hat{d}_{il}) = \frac{2\sigma^2}{rE}, \text{ where } E = \lambda v / rk$$

The variance of the estimated difference between any two treatment effects from a RCB design with same number of treatments v and same replication r, as that in BIB design, is given by

$$V(\hat{d}_{il}) = \frac{2\sigma'^2}{r}$$

where σ'^2 is the per observation variance in case of RCB design. Thus the efficiency of BIB design as compared to RCB design is given by

$$\text{Efficiency} = E\frac{\sigma'^2}{\sigma^2}.$$

The quantity E is called the Efficiency Factor of BIB design. Here $E<1$, since $k < v$. However, σ^2 is expected to be very small compared to σ'^2 because the block size in incomplete block design is much small compared to the block size in a complete block design. As such the efficiency is expected to be greater than 1.

The estimated variance of the estimated difference between any two treatment effects, from a BIB design, is given by

$$\hat{V}(\hat{d}_{il}) = \frac{2}{rE}\hat{\sigma}^2 = \frac{2}{rE}s_e^2.$$

Therefore, the estimated standard error of estimated difference between any two treatment effects is given by

$$SE_d = \sqrt{\frac{2}{rE}s_e^2}.$$

5.4.3 Example 2

An experiment was conducted at Jorhat under the aegis of Project Directorate of Cropping Systems Research (now Indian Institute of Farming Systems Research), Modipuram using a BIB design with parameters $v = b = 7$, $r = k = 4$, $\lambda = 2$. The treatment details are as given in Table 5.3.

Table 5.3: Treatment details of an experiment

Season	Treatments (Crop Sequences)						
	T1	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**
Kharif	Rice	Rice	Rice	Rice	Rice	Rice	Rice
Rabi	-	Boro Rice	Mustard	Brinjal	Tomato	French Bean	Potato
Summer	Rice	-	Rice	Rice	Rice	Rice	Rice

117

The returns from each of the crop sequence were converted into calories / hectare. The block structure and the calories of the output per hectare obtained are given in the Table 5.4.

Table 5.4: Returns from different crop sequences in calories / ha

Block 1	T2 (3325060)	T4 (2606200)	T5 (3279420)	T6 (2330180)
Block 2	T1 (2992900)	T2 (3228180)	T5 (3348780)	T7 (2982000)
Block 3	T1 (3639920)	T4 (2467800)	T6 (2196580)	T7 (2730780)
Block 4	T3 (2602410)	T5 (3696340)	T6 (2388060)	T7 (2921790)
Block 5	T1 (3055180)	T3 (2653680)	T4 (2501060)	T5 (3594320)
Block 6	T2 (3380420)	T3 (2760690)	T4 (2522100)	T7 (2961270)
Block 7	T1 (2921420)	T2 (3380420)	T3 (2677400)	T6 (2594420)
T# denotes the treatment number and the figures in bracket are the calories / hectare				

The purpose of the experiment is to determine the crop sequence that produces maximum calories / ha.

5.4.4 Analysis of Data

The data has been generated using a BIB design with parameters $v = b = 7$, $r = k = 4$, $\lambda = 2$. The data has been analyzed using SAS software. The SAS commands and the data structure are given in the sequel.

```
DATA BIBD;
INPUT blk  trt    calorie;
CARDS;
1      2       3325060
1      4       2606200
1      5       3279420
1      6       2330180
2      1       2992900
2      2       3228180
2      5       3348780
2      7       2982000
3      1       2639980
3      4       2467800
3      6       2196580
3      7       2730780
4      3       2602410
4      5       3696340
4      6       2388060
4      7       2921790
5      1       3055180
5      3       2653680
```

```
5      4      2501060
5      5      3594320
6      2      3380420
6      3      2760690
6      4      2522100
6      7      2961270
7      1      2921420
7      2      3380420
7      3      2677400
7      6      2594420
;
PROC GLM;
CLASS blk trt;
MODEL calorie = trt  blk;
LSMEANS trt/PDIFF  LINES; /*pairwise comparisons using Fisher's protected LSD*/
LSMEANS trt / PDIFF ADJUST = BON LINES; /*pairwise comparisons using BONFERRONI
correction*/
LSMEANS trt / PDIFF ADJUST = TUKEY LINES; /*pairwise comparisons using Tukey's
Honest significant difference test*/
RUN;
```

5.4.5 Output of analysis

PROC GLM gives the analysis of variance for a block design.

Table 5.5: Results of PROC GLM for calorie data

ANOVA

Source	DF	SS	MS	F value	Prob > F
Model	12	4.1295465E12	344128876386	28.20	<0.0001
Error	15	183036843371	12202456225		
Corrected Total	27	4.3125834E12			

R-Square	CV	Root MSE	calorie Mean
0.958	3.831	110464.70	2883530

It is evident from these results that the model with treatments and blocks explains about 96 per cent of the total variability in the calorie values (data). This result is also supported by the fact that the model component is highly significant (p-value < 0.0001).

While dealing with non-orthogonal data, the way the terms are described in the model in the PROC GLM is very important. If the interest of the experimenter is in testing the homogeneity of treatment effects, then the statement is MODEL calorie = trt blk. For this description of model, the following ANOVA is produced:

Table 5.6: ANOVA table for testing homogeneity of treatment effects

ANOVA

Source	DF	Type I SS	Mean Square	F Value	Prob > F
Treatments	6	3.2867911E12 = 3.2867911×10^{12}	547798519371	44.89	<0.0001
Blocks	6	842755400400	140459233400	11.51	<0.0001
Error	15	183036843371	12202456225		
Corrected Total	27	4.3125834E12			

In ANOVA Table 5.6, the *F*-value for treatments component can be used for testing the equality of treatment effects, but the *F*-value for blocks component cannot be used for testing the equality of block effects. On the other hand, if the interest of the experimenter is in testing the homogeneity of block effects, then the statement is MODEL calorie = blk trt. The following ANOVA is then produced.

Table 5.7: ANOVA table for testing homogeneity of block effects

ANOVA

Source	DF	Type I SS	Mean Square	F Value	Prob > F
Blocks	6	195740712429	32623452071	2.67	0.0573
Treatments	6	3.9338058E12	655634300700	53.73	<0.0001
Error	15	183036843371	12202456225		
Corrected Total	27	4.3125834E12			

In the ANOVA Table 5.7, the *F*-value for blocks component can be used for testing the equality of block effects, but the *F*-value for treatments component cannot be used for testing the equality of treatment effects. As described earlier also, the error sum of squares in both the ANOVAs is same. This means the treatment sum of squares and the block sum of squares in both the ANOVAs, though different, have the same sum and that sum would be equal to the sum of squares due to model, as given in the first ANOVA above. In the first case where treatments appear before blocks in the model, the treatments sum of squares are adjusted for blocks, but the block sum of squares are unadjusted. In the second case where blocks appear before treatments in the model, the block sum of squares are adjusted for treatments, but the treatment sum of squares are unadjusted. So while using Type I sum of squares, one has to be very careful about the order in which the variables appear in the model.

A remedy to such a problem is that while analyzing the non-orthogonal data, one should use the Type III sum of squares, because the ANOVA produced has all the effects adjusted for all other remaining effects. The SAS code for obtaining Type III sum of squares is given below.

PROC GLM;

CLASS blk trt;

MODEL calorie = trt blk/SS3;

RUN;

This ANOVA is given in Table 5.8.

Table 5.8: ANOVA table with type III sum of squares

Source	DF	Type III SS	MS	F	Prob > F
Treatments	6	3.2867911E12	547798519371	44.89	< 0.0001
Blocks	6	195740712429	32623452071	2.67	0.0573
Error	15	183036843371	12202456225		
Total	27	4.3125834E12			

In ANOVA Table 5.8, both the treatments and blocks sums of squares are adjusted for blocks and treatments, respectively. One may also notice that the difference between the adjusted and the unadjusted sum of squares is quite large.

Such a thing, however, does not happen in case of an orthogonal design. For an orthogonal design, the Type I and Type III sum of squares are identical. So to conclude, it would always be better to use the Type III sum of squares.

For making the multiple comparisons in the form of pairwise treatment comparisons, LSMEANS trt / PDIFF LINES is used. This command makes by default the pairwise treatment comparisons using Student's *t* statistic and least square means. The unadjusted and the least squares means (LSMEANS) of the calorie value for each treatment are given in Table 5.9.

Table 5.9: Unadjusted and LS MEANS of the treatments

Treatment	Unadjusted Calorie Mean	Calorie LSMEAN
1	2902370.00	2917317.14
2	3328520.00	3309634.29
3	2673545.00	2609654.29
4	2524290.00	2553810.00
5	3479715.00	3467092.86
6	2377310.00	2403437.14
7	2898960.00	2923764.29

It is clearly visible from Table 5.9 that the unadjusted means are substantially different from the least square means. Once again, it is important to realize that the non-orthogonality has a major role to play and the adjustments need to be made for the non-orthogonality.

In order to make pairwise treatment comparisons through LSMEANS, Table 5.10 is produced.

Table 5.10: Least squares means for treatment effects

			Pr > |t| for H0: LSMean(i)=LSMean(j)				
			Dependent Variable: Calorie Values				
i/j	1	2	3	4	5	6	7
1		0.0003	0.0022	0.0006	<0.0001	<0.0001	0.9395
2	0.0003		<0.0001	<0.0001	0.0789	<0.0001	0.0003
3	0.0022	<0.0001		0.5138	<0.0001	0.0260	0.0019
4	0.0006	<0.0001	0.5138		<0.0001	0.0919	0.0005
5	<0.0001	0.0789	<0.0001	<0.0001		<0.0001	<0.0001
6	<0.0001	<0.0001	0.0260	0.0919	<0.0001		<0.0001
7	0.9395	0.0003	0.0019	0.0005	<0.0001	<0.0001	

It may be seen from Table 5.10 that mostly, the pairwise treatments are significantly different except the pairs (1, 7) and (3, 4) which are not significantly different from each other. The pairs of treatments (2, 5) and (4, 6) are also marginally not significantly different. Alternatively, these comparisons can also be made using the Table 5.11.

Table 5.11: *t* comparison lines for least squares means of treatments

		CALORIE LSMEAN	LSMEAN Number
	A	3467093	5
	A	3309634	2
	B	2923764	7
	B	2917317	1
	C	2609654	3
D	C	2553810	4
D		2403437	6

It may be seen that treatment 5 (Rice - Tomato - Rice sequence) is the cropping sequence that yields maximum calorific value and the treatment 6 (Rice - French Bean - Rice sequence) is the cropping sequence that yields lowest calorific value. It may also be seen from the table of *t*-comparison lines for least squares means of treatments that treatments 5 and 6 are significantly different from each other. In fact treatment 5 is significantly different from all other treatments except treatment 2 (Rice - Boro Rice Sequence). There is not much to choose between these two crop sequences. Similarly, treatment 6 is at par with treatment 4 but is significantly different from all other treatments.

LSMEANS trt / PDIFF ADJUST = BON LINES; LSMEANS trt / PDIFF ADJUST = TUKEY LINES; commands, respectively make the pairwise treatment comparisons using Bonferroni and Tukey's methods. The results obtained are given in the sequel.

Bonferroni method

Using Bonferroni method, the pairwise comparison of treatment effects are made using LS MEANS. The results are presented in Tables 5.12 and 5.13. It may be easily seen that using this method the significance levels are higher than those obtained using t statistic for pairwise treatment comparisons. This is so because the probability levels have been adjusted for the multiple comparisons. The pair of treatments that were not significantly different earlier are once again not significantly different, though the probability levels have increased. However, using Bonferroni method, the pair of treatments 3 and 6 is not significantly different, though these were significantly different using t statistic. The same results are also presented through the Table giving alphabets A, B, C, for representing the significance of the difference.

Table 5.12: Least squares means for treatment effects for calorie values

i/j	1	2	3	4	5	6	7
	Pr > |t| for H0: LSMean(i) = LSMean(j) Dependent Variable: Calorie values						
1		0.0060	0.0464	0.0119	0.0002	0.0004	1.0000
2	0.0060		<0.0001	<0.0001	1.0000	<0.0001	0.0070
3	0.0464	<0.0001		1.0000	<0.0001	0.5464	0.0396
4	0.0119	<0.0001	1.0000		<0.0001	1.0000	0.0102
5	0.0002	1.0000	<0.0001	<.0001		<0.0001	0.0002
6	0.0004	<0.0001	0.5464	1.0000	<0.0001		0.0003
7	1.0000	0.0070	0.0396	0.0102	0.0002	0.0003	

Table 5.13: Bonferroni comparison lines for least squares means of treatments

LS-means with the same letter are not significantly different		
	CALORIE LSMEAN	**LSMEAN Number**
A	3467093	5
A	3309634	2
B	2923764	7
B	2917317	1
C	2609654	3
C	2553810	4
C	2403437	6

Tukey's method

Using Tukey's method, the pairwise comparison of treatment effects is made using LS MEANS. The results are presented in Tables 5.14 and 5.15. It may be easily seen that using this method the significance levels are higher than those obtained using t statistic, but not as high as obtained using Bonferroni methods. The results obtained are similar to one obtained using Bonferroni method. The results are also depicted through the Table giving alphabets A, B, C, for representing the significance of the difference.

Table 5.14: Least squares means for treatment effects

i/j	1	2	3	4	5	6	7
Pr > |t| for H0: LSMean(i) = LSMean(j) **Dependent Variable: Calorie values**							
1		0.0042	0.0285	0.0080	0.0001	0.0003	1.0000
2	0.0042		<0.0001	<0.0001	0.5173	<0.0001	0.0048
3	0.0285	<0.0001		0.9926	<0.0001	0.2372	0.0246
4	0.0080	<0.0001	0.9926		<0.0001	0.5667	0.0069
5	0.0001	0.5173	<0.0001	<0.0001		<0.0001	0.0002
6	0.0003	<0.0001	0.2372	0.5667	<0.0001		0.0003
7	1.0000	0.0048	0.0246	0.0069	0.0002	0.0003	

Table 5.15: Tukey comparison lines for least squares means of treatments

LS-means with the same letter are not significantly different		
	CALORIE LSMEAN	**LSMEAN Number**
A	3467093	5
A	3309634	2
B	2923764	7
B	2917317	1
C	2609654	3
C	2553810	4
C	2403437	6

Alternatively, the pairwise treatment comparisons can be represented as given in Table 5.16. Means with at least one letter common are not statistically significant using Tukey's Honest Significant Difference. These results are same as obtained earlier using Bonferroni method or Tukey's method.

Table 5.16: Treatments with letter display

Treatment	Calorie value	
	Treatment LSMEAN	**Rank of Treatment**
1	2917317.14[B]	4
2	3309634.29[A]	2
3	2609654.29[C]	5
4	2553810.00[C]	6
5	3467092.86[A]	1
6	2403437.14[C]	7
7	2923764.29[B]	3
General Mean	2883530.00	.
p-Value	<0.0001	.
CV (%)	3.83	.

5.4.6 Analysis using R

The data has also been analyzed again using R software. The R code is given in the sequence. The results obtained are same as obtained using SAS and, therefore, have not been reported to avoid duplication.

R code
```
d8=read.table("bibd.txt",header=TRUE)
attach(d8)
names(d8)
#Treatment means and standard deviations
aggregate(calorie~trt,data=d8,mean)
aggregate(calorie~trt,data=d8,sd)
#Treatment wise box plot of calorie
boxplot(calorie~trt)
#anova with class variable treatments and block
trt=factor(trt)
blk=factor(blk)
lm1=lm(calorie~blk+trt)
anova(lm1)
lm2=lm(calorie~trt+blk)
anova(lm2)
library(car)
#produces both adj trtss and adj blkss
Anova(lm1,type="III")
Anova(lm2,type="III")
library(lsmeans)
lsm=lsmeans(lm1,"trt")
```

lsm
pairs(lsm)
#to provide letters for groups, need to install multcompView
library(multcompView)
cld(lsm,Letters="abcdef")
#Comparisons are based only on Tukey's HSD. It is also suggested that the number of letters
#here should be equal to expected number of groups or equal to the number of treatments,
#when number of treatments is more than 26, then one may add numerals.*/
detach(d8)

Here, the function cld() in R gives letter display of treatment comparisons based on Tukey's Honest Significant Different test.

Remark 5.3 As mentioned earlier in Section 5.1, as the experimenter moves away from a complete block design to an incomplete block design, obtaining the layout of the design becomes difficult for the experimenter. Unlike RCB design, the experimenter cannot write down the layout of an incomplete block design without the help of a statistician. However, for the benefit of experimenters, a catalogue of balanced incomplete block designs is available at Design Resources Server at www. iasri.res.in/ design/.

Remark 5.4 It may be worthwhile mentioning here that the procedure of analysis described above holds also for any incomplete block design, other than balanced incomplete block design. The other incomplete block designs could be partially balanced incomplete block design including group divisible design, rectangular design, etc. This analysis is also applicable to the incomplete block designs with unequal replications and unequal block sizes, as described in 5.4.

5.5 Other incomplete block designs

Incomplete block designs are used because of practical considerations. When the number of treatments is large, it may not be possible to form homogeneous complete blocks. But forming incomplete blocks of equal sizes may also lead to the problem of forming blocks which are not homogeneous. If soil salinity is a source of variability in the field, then different patches of salinity can form natural blocks. But the patches may be of unequal sizes, which may demand forming blocks of unequal sizes. Similarly litter mates of animals are the natural blocks. But the number of litter mates need not be equal. Once again, there would be a need to form blocks of unequal sizes. In the same spirit, it is indeed possible to have unequal replication of treatments because of not having enough replications possible for every treatment. In view of this, block designs with unequal block sizes or / and unequal replication of treatments are available in the literature. The following are the examples of a variance balanced incomplete block design with unequal replication of treatments or / and unequal block sizes:

Consider an incomplete block design D1 [$v = 8$, $b = 12$, $r = 5$, $k_1 = k_2 = k_3 = k_4 = k_5 = k_6 = k_7 = k_8 = 4$, $k_9 = k_{10} = k_{11} = k_{12} = 2$]:

1	5	2	6	3	7	4	8	1	2	3	4
2	6	7	3	8	4	1	5	5	6	7	8
3	7	8	4	1	5	6	2				
4	8	1	5	6	2	7	3				

This design has $n = 40$ experimental units. The design D1 is with unequal block sizes but has equal replications of treatments. This design is also variance balanced. The variance of the estimated contrast of the difference of any two treatment effects is given by $\mathrm{Var}(\hat{d}_{il}) = \sigma^2/2$ for all $i \neq l = 1,2,\cdots,8$.

Consider another example of a variance balanced design with unequal replication of treatments and unequal block sizes, D2 [$v = 6$, $b = 11$, $r_1 = r_2 = r_3 = r_4 = 4$, $r_5 = r_6 = 5$, $k_1 = k_2 = 4$, $k_3 = k_4 = k_5 = k_6 = k_7 = k_8 = k_9 = k_{10} = k_{11} = 2$]:

1	1	1	1	2	2	3	3	4	4	5
2	2	5	6	5	6	5	6	5	6	6
3	3									
4	4									

The parameters of this design are $v = 6$, $b = 11$, $r_1 = r_2 = r_3 = r_4 = 4$, $r_5 = r_6 = 5$, $k_1 = k_2 = 4$, $k_3 = k_4 = k_5 = k_6 = k_7 = k_8 = k_9 = k_{10} = k_{11} = 2$. The number of experimental units in this design is $n = 26$. This design is also variance balanced and the variance of the estimated contrast of the difference of any two treatment effects is given by $\mathrm{Var}(\hat{d}_{il}) = 2\sigma^2/3$ for all $i \neq l = 1,2,\cdots,6$.

These designs have strong potential for application. Although BIB designs have been used in agricultural experiments, these designs can also be used in experimentation without any problem. But as has been mentioned above, the researcher would have to interact with the statistician to get these designs. For the benefit of experimenters, however, the randomized layout of the design can be obtained from Design Resources Server at www.iasri.res.in/design. Designs with equal and unequal block sizes can be obtained from this web resource.

5.6 Crossed classification with interaction

This chapter has been devoted to incomplete block designs. These designs are in fact two way crossed classification in the sense that all the levels of one factor appear with all or some levels of the other factor. Since there can at most be one observation at the ith level of one factor and jth level of the other factor, it was not possible to fit/estimate the interactions. However, if there are more than one observation at the ith level of one factor and jth level of the other factor, then it is possible to fit the interactions also. But if like an incomplete block design, there are some cells in which there are no observations, then the interaction cannot be defined in that cell. Consider the following example to make the exposition clear.

Suppose that an agronomist conducts a series of experiments with four different fertilizer treatments on each of the five varieties of wheat. For each fertilizer-by-variety combination, the researcher plants several 5'×5' plots. At harvest time it is found that many of the plots have been lost due to some accident and the researcher is left with the data given in Table 5.17.

Table 5.17: Weight of wheat produced from 5'×5' plots

Fertilizer	Variety of Wheat				
	1	2	3	4	5
1	12	14	15	-	09
	09	13	12	-	12
	11	15	-	-	10
	-	-	-	-	08
2	11	17	-	14	12
	13	14	-	12	13
	10	16	-	10	-
	11	-	-	15	-
3	16	18	17	09	14
	13	20	17	11	-
	13	15	-	14	-
	-	13	-	08	-
4	20	19	18	15	18
	17	16	21	14	21
	-	18	-	16	16
	-	-	-	12	-

The suitable linear model for analyzing the data of the type just described is the following:

$$y_{iju} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{iju}$$

Here y_{iju} is the uth observation obtained from the jth level of factor B (variety) and the ith level of factor A (Fertilizer), μ is the general mean, α_i is the effect of the ith level of factor A, β_j is the effect of the jth level of factor B, γ_{ij} is the interaction effect due to the jth level of factor B (variety) and the ith level of factor A (Fertilizer), and ε_{iju} is the random error term associated with the observation y_{iju} and distributed independently and identically with mean zero and constant variance σ^2. Here $i = 1,2,\cdots,a; j = 1,2,\cdots,b$. Let n_{ij} denote the number of observations in the jth level of factor B and ith level of factor A. For the incomplete block designs described earlier, n_{ij} takes values 0 or 1 for all $i = 1,2,\cdots,a; j = 1,2,\cdots,b$. But in the present context, $n_{ij} \geq 0$, and

$\sum\limits_{j=1}^{b_i} n_{ij} = n_{i.}, \ \sum\limits_{i=1}^{a} n_{ij} = n_{.j}$ and $\sum\limits_{i=1}^{a} n_{i.} = \sum\limits_{j=1}^{b} n_{.j} = n.. = \sum\limits_{i=1}^{a}\sum\limits_{j=1}^{b} n_{ij}.$ Wherever $n_{ij} = 0$, the interaction is not

defined for that cell. The total number of interactions, therefore, would be less than or equal to *ab*. We shall let *s* denote the total number of cells in which $n_{ij} > 1$

For the example under consideration, $a = 4$, $b = 5$, $s = 18$, $n_{12} = 3, n_{13} = 2, n_{14} = 0, n_{15} = 4$,

$n_{21} = 4, n_{22} = 3, n_{23} = 0, n_{24} = 4, n_{25} = 2$, $n_{31} = 3, n_{32} = 4, n_{33} = 2, n_{34} = 4, n_{35} = 1$,

$n_{41} = 2, n_{42} = 3, n_{43} = 2, n_{44} = 4, n_{45} = 3$; $n_{1.} = 12, n_{2.} = 13, n_{3.} = 14, n_{4.} = 14$;

$n_{.1} = 12, n_{.2} = 13, n_{.3} = 6, n_{.4} = 12, n_{.5} = 10$; $n = 53$.

The SAS commands for the analysis of data generated are the following:


```
DATA twoway_interaction;
INPUT Factor A  FactorB   response;
CARDS;
.
INSRT DATA HERE
.
;
PROC GLM;
CLASS FactorA  FactorB;
MODEL response = FactorA  FactorB  FactorA*FactorB;
LSMEANS FactorA  FactorB  FactorA*FactorB / PDIFF ADJUST = TUKEY LINES
RUN;
```

The R code for similar analysis is given below.

```
twoway_interaction=read.table("twoway_interaction.txt",header=TRUE)
attach(twoway_interaction)
names(twoway_interaction)
lm1=lm(response~factor(A)+factor(B)+factor(A):factor(B),data= twoway_interaction)
anova(lm1)
#Tukey comparison
aov.fit=aov(response~factor(A)+factor(B)+factor(A):factor(B),data= twoway_interaction)
TukeyHSD(aov.fit,"factor(A)")
TukeyHSD(aov.fit,"factor(B)")
TukeyHSD(aov.fit,"factor(A):factor(B)")
detach(twoway_interaction)
```

The analysis of variance obtained from this analysis would be as shown in Table 5.18.

Table 5.18: ANOVA table for two way cross classification

ANOVA for the Example		ANOVA for two-way crossed classification	
Source	DF	Source	DF
Factor A	3	Factor A	$a - 1$
Factor B	4	Factor B	$b - 1$
FactorA*FactorB	10*	FactorA*FactorB	s
Error	35	Error	$n.. - a - b - s + 1$
Corrected Total	52	Corrected Total	$n.. - 1$

* indicates the number of interactions defined. In this case there are 12 cells but two cells are empty. So the number of interactions defined is 10.

For the example given above, the results obtained using the SAS commands are described in the sequel. It may be mentioned here that only the results are reported for the purpose of illustration only. There will be a reference to this type of analysis in subsequent Chapters as well where the analysis will be reported and discussed in detail.

Table 5.19: Result obtained using SAS commands

ANOVA

Source	DF	Type III SS	MS	F- Value	Prob > F
Fertilizer	3	243.655	81.218	20.95	<0.0001
Variety	4	152.547	38.137	9.84	<0.0001
Fertilizer*Variety	10	58.627	5.863	1.51	0.1763
Error	35	135.667	3.876		
Corrected Total	52	580.528			

R-Square	CV	Root MSE	yield Mean
0.766	13.969	1.969	14.094

From the analysis of variance, it is obvious that the fitted model has been able to explain about 77 per cent of the total variability in the data. The CV is slightly high, though (13.969). The effect of fertilizer and the variety are highly significant. But the interaction between fertilizer and variety is statistically not significant.

Table 5.20: LS MEANS for variety and fertilizer levels

Variety	yield LSMEAN		Fertilizer	yield LSMEAN
1	13.604		1	Non-est
2	15.958		2	Non-est
3	Non-est		3	14.400
4	Non-est		4	17.650
5	13.646			

The LS MEANS for various levels of fertilizers and varieties are given in Table 5.20. It may be seen that the levels 3 and 4 of the variety and levels 1 and 2 of the fertilizer are not estimable. The reason is that the interaction for these combinations are not defined because of no observation in two cells corresponding to level 1 of fertilizer, level 4 of variety and level 2 of fertilizer and level 3 of variety. It is for this reason that in Table 5.21 of LS MEANS of interaction, the combination (1, 4) and (2, 3) of variety and fertilizer is not included.

Table 5.21: LSMEANS for the combinations of variety and fertilizer levels

Fertilizer	Variety	yield LSMEAN		Fertilizer	Variety	yield LSMEAN
1	1	10.667		3	2	16.500
1	2	14.000		3	3	17.000
1	3	13.500		3	4	10.500
1	5	9.750		3	5	14.000
2	1	11.250		4	1	18.500
2	2	15.667		4	2	17.667
2	4	12.750		4	3	19.500
2	5	12.500		4	4	14.250
3	1	14.000		4	5	18.333

The pairwise comparison of variety×fertilizer interaction effect is made by using Tukey's adjustment. The highest yield is obtained from third level of variety used with fourth level of fertilizer, while the lowest yield is obtained from fifth level of variety and first level of fertilizer. The two responses are statistically significant.

Table 5.22: Tukey-Kramer comparison lines for least squares means of fertilizer*variety

				yield LSMEAN	Fertilizer	Variety	LSMEAN Number
		A		19.500	4	3	16
		A		18.500	4	1	14
		A		18.333	4	5	18
		A		17.667	4	2	15
B		A		17.000	3	3	11
B		A		16.500	3	2	10
B		A	C	15.667	2	2	6
B	D	A	C	14.250	4	4	17
B	D	A	C	14.000	1	2	2
B	D	A	C	14.000	3	1	9
B	D	A	C	14.000	3	5	13
B	D	A	C	13.500	1	3	3

131

LS-means with the same letter are not significantly different							
				yield LSMEAN	Fertilizer	Variety	LSMEAN Number
B	D	A	C	12.750	2	4	7
B	D	A	C	12.500	2	5	8
B	D		C	11.250	2	1	5
B	D		C	10.667	1	1	1
	D		C	10.500	3	4	12
	D			9.750	1	5	4
The LINES display does not reflect all significant comparisons. The following additional pairs are significantly different: (16,7); (10,5); (10,1).							

Remark 5.5: In the Example described just above, the interaction effect is found to be non-significant. Therefore, in this case the interpretation should have been based on a model without interaction effects, i.e., main effects alone. In presence of interactions, LSMEANS of main effects are non-estimable. However, just for the sake of completing the example so as to help the experimenters understand the way in which the results would be obtained, have the Tables 5.20, 5.21 and 5.22 given. This may not be done in practice.

5.7 Nested classification

In Chapter 2 while describing RCB design it has been seen that every level of one factor crosses with every level of the other factor. The two factors are the treatments and the blocks. In a Latin square design also there are three two-way classifications, *viz.* treatments *vs* rows, treatments *vs* columns and rows *vs* columns. In all the three classifications, like RCB design, every level of one factor crosses with every level of the other factor. In order to make the exposition clear, consider RCB design with $v = 6$ and $b = 5$. The 6×5 matrix **N** with cell entries giving the replication of each treatment in each block, known as incidence matrix, of this design is

$$\mathbf{N} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

In a BIB design described in this chapter, once again there are two classifications *viz.* treatments and blocks. But in an incomplete block design, be it a BIB design or any other incomplete block design, all the levels of one factor do not cross with every level of the other factor. It is for this reason these designs are non-orthogonal. To make the exposition clear, consider an incomplete block design with $v = 6$ and $b = 5$. The design is (1, 2, 11); (3, 4, 11); (5, 6, 11); (7, 8, 11); (9, 10, 11). The 11×5 incidence matrix **N** of this design is

$$\mathbf{N'} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

It may be seen from the above two examples that some levels of one factor do cross some levels of the other factor, if not all the levels. It does not happen that a given level of one factor crosses only one level of the other factor. However, there do occur experimental situations where it is not practically feasible to cross levels of one factor with each of the levels of the other factor. Levels of one factor can only be crossed with only one of the levels of the second factor. One such experimental situation is described in the sequel.

Consider an experiment in which 4 sires are mated to 15 dams. The first sire is mated to 4 dams, the second sire is mated to 3 dams, the third sire is mated to 3 dams and the fourth sire is mated to 5 dams. The birds produced lay eggs and the weight of eggs are recorded. The number of birds produced may vary from dam to dam between sires to which these dams are mated as well as within sires also. The details of the experiment are given in Table 5.23.

Table 5.23: Mating plan of sires and dams and the egg weight (in gms) of birds produced

Sire	Dams	Egg weight (in grams) of birds
1	1	52, 45, 43, 49
	2	50, 52, 49
	3	46, 53
	4	45, 49, 50, 47, 50
2	1	49, 50 , 54, 47
	2	55, 55, 50, 50, 54
	3	50, 47, 50
3	1	50, 47, 48
	2	43, 43, 48, 51, 47
	3	45, 46, 49
4	1	45, 44, 49
	2	50, 57, 45, 47, 50
	3	41, 37, 47, 48, 49
	4	45, 45, 48, 50
	5	58, 50, 41, 47, 37, 47

In this experiment the sire *vs* dams incidence matrix would be the following:

$$\mathbf{N} = \begin{bmatrix} 4 & 3 & 2 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 5 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 5 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 5 & 5 & 4 & 6 \end{bmatrix}$$

It may be seen that in this Example, the levels of the second factor appear with only one of the levels of the first factor. For example, levels 1, 2, 3, 4 of the second factor appear only with level 1 of the first factor and not with any other level of the first factor; similarly levels 5, 6, 7 of the second factor appear with only level 2 of the first factor and not with any other level of the first factor; levels 8, 9, 10 of the second factor appear with only level 3 of the first factor and not with any other level of the first factor; levels 11, 12, 13, 14, 15 of the second factor appear with only level 4 of the first factor and not with any other level of the first factor.

Sometimes, constraints prevent us from crossing every level of one factor with every or few levels of the other factor. In these cases one is to adopt what is known as a *nested* layout. We say we have a nested layout when fewer than all levels of one factor occur within each level of the other factor. The example given above is that of a nested layout wherein dams are nested within sires.

If Factor B is nested within Factor A, then a level of Factor B can only occur within one level of Factor A and there can be no interaction between factors A and B. The following model is used for a nested classification:

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + \varepsilon_{ijk}$$

Here y_{ijk} is the response on the *k*th unit of the *j*th level of factor B nested within the *i*th level of factor A, μ is the general mean, α_i is the effect of the *i*th level of factor A, β_{ij} is the effect of the *j*th level of factor B nested within the *i*th level of factor A and ε_{ijk} is the random error term associated with the observation y_{ijk} and distributed independently and identically with mean zero and constant variance σ^2. Here $i = 1, 2, \cdots, a; j = 1, 2, \cdots, b_i$. Let n_{ij} denote the number of observations in the *j*th level of factor B and *i*th level of factor A. Further

$$\sum_{i=1}^{a} b_i = s, \sum_{j=1}^{b_i} n_{ij} = n_{i.}, \sum_{i=1}^{a} n_{i.} = n..; \ e.g., \ a = 4, \ b_1 = 4, b_2 = 3, b_3 = 3, b_4 = 5 \text{ and } s = 15, \text{ in the example above.}$$

Moreover, $n_{11} = 4, n_{12} = 3, n_{13} = 3, n_{14} = 5$ and $n_{1.} = 15$. Similarly $n_{21} = 4, n_{22} = 5, n_{23} = 3$ and $n_{2.} = 12$; $n_{31} = 3, n_{32} = 5, n_{33} = 3$ and $n_{3.} = 11$; $n_{41} = 3, n_{42} = 5, n_{43} = 5, n_{44} = 4, n_{45} = 6$ and $n_{4.} = 23$. Further $n.. = 61$.

The SAS commands for the analysis of data generated are the following:

```
DATA Nested_Classification;
INPUT  Factor A   Factor B  response;
CARDS;
```

```
.
INSRT DATA HERE
.
;
PROC GLM;
CLASS FactorA  FactorB(FactorA);
MODEL response = FactorA  FactorB(FactorA);
LSMEANS FactorA/PDIFF  LINES; /*comparisons based on LSD*/
LSMEANS  FactorA / PDIFF  ADJUST = TUKEY  LINES; /*Comparisons based on Tukey's
HSD*/
RUN;
```

The R code for similar analysis is given below.

```
Nested_Classification =read.table("Nested_Classification.txt",header=TRUE)
attach(Nested_Classification)
names(Nested_Classification)
lm1=lm(response~factor(A)+factor(B)/factor(A),data= Nested_Classification)
anova(lm1)
#Tukey comparison
aov.fit=aov(response~factor(A)+factor(B)/factor(A),data= Nested_Classification)
TukeyHSD(aov.fit,"factor(A)")
detach(Nested_Classification)
```

The analysis of variance obtained from this analysis would be the following:

Table 5.24: ANOVA table for nested classification

ANOVA for the Example		ANOVA for nested classification	
Source	DF	Source	DF
Factor A	3	Factor A	$a - 1$
Factor B nested within Factor A	11	Factor B nested within Factor A	$s - a$
Error	45	Error	$n.. - b$
Corrected Total	59	Corrected Total	$n.. - 1$

For the example given above, the results obtained using the SAS commands are described in the sequel. It may be mentioned here that only the results are reported for the purpose of illustration only. There will be a reference to this type of analysis in subsequent Chapters as well where the analysis will be reported and discussed in detail.

Table 5.25: Output from SAS command for data in Table 5.23

ANOVA

Source	DF	Type III SS	MS	F-Value	Prob > F
Sire	3	106.11	35.37	0.40	0.75
Dam within Sire	11	961.12	87.37	0.99	0.47
Error	45	3975.38	88.34		
Corrected Total	59	5056.58			

R-Square	CV	Root MSE	eggwt Mean
0.214	20.24	9.39	46.42

It may be seen from the results above that the total variability explained by the model is very low (about 21 per cent). The sire effect is not significant. Similarly, the dam within sire effect is also no significant.

The LSMEANS for the sires are obtained and given in Table 5.26. Pairwise comparison of the sire effects reveals that sire 2 produces the highest egg weight while the lowest egg weight is produced by sire 4. The differences among sires are not statistically significant.

Table 5.26: Least square means for sires

sire	Eggweight LSMEAN	LSMEAN Number
1	46.32	1
2	48.60	2
3	47.80	3
4	45.10	4

Table 5.27: *t* comparison lines for least squares means of sires

LS-means with the same letter are not significantly different			
	eggwt LSMEAN	sire	LSMEAN Number
A	48.60	2	2
A	47.80	3	3
A	46.32	1	1
A	45.17	4	4

Table 5.28: Tukey-HSD comparison lines for least squares means of sires

	eggwt LSMEAN	sire	LSMEAN Number
LS-means with the same letter are not significantly different			
A	46.32	1	2
A	48.60	2	2
A	47.80	3	3
A	45.17	4	4

The pairwise comparison of sire effects was also made using the Tukey's adjustment. It is found that the difference between the egg weight of the highest producing sire and the lowest producing sire is not statistically significant. In fact, all the sire effects are at par so far as producing egg weight is concerned.

6

Designs with Nested Structure

6.1 Introduction

It has been emphasized over and again that the total variability in the data has two major explainable components of variability, *viz.* (a) due to treatments, and (b) due to experimental material. The remaining unexplainable component of variability is the experimental error. The variability in the experimental material is accounted for by forming groups of homogeneous experimental units or by forming rows and columns, etc. Designs with one blocking system *i.e.* with only one source of variability in the experimental material are called block designs. In these designs the experimental units are divided into groups called blocks, which comprise of homogeneous experimental units. The between blocks variability should be large and within blocks variability should be small. Designs with two cross classified blocking systems *i.e.* with two sources of variability in the experimental material are called row-column designs. In these designs, the experimental units are grouped in an array. The experimental units within the rows and within the columns are as homogeneous as possible. A block design is said to be a complete block design if each block is a complete replication. Further, a block design is said to be an incomplete block design if the design has at least one block that is not a complete replication in the sense that there is at least one treatment that does not appear in that block. In other words, each block has every treatment appearing at most once, if it appears in that block. We shall restrict our attention to such incomplete block designs only. There are, however, incomplete block designs where a treatment may appear more than once in a block. Similarly in a row - column set up, the rows or / and the columns may be complete in the sense that all the treatments appear in each row or / and column exactly once; in other words the rows or / and columns are complete replications. Similar to incomplete block designs, there may be incomplete row-column designs in the sense that there is at least one row or / and column that does not contain all the treatments. Like incomplete block designs, incomplete row-column designs may also have a treatment appearing more than once in a row or / and column. Most commonly used block designs are randomized complete block (RCB) design, balanced incomplete block (BIB) design, partially balanced incomplete block (PBIB) design, square lattice, rectangular lattice, alpha designs, etc. Similarly, commonly used row-column designs are Latin Square designs, Youden square designs, Generalized Youden designs, Pseudo Youden designs, etc. Henceforth in this Chapter, the focus would be restricted to block design set up.

Although incomplete block designs help in reducing the intra block variance to a considerable extent because of the reduction in block size, yet the agricultural experimenters are hesitant to adopt these designs because the blocks are not complete blocks (or complete replications). One major concern of the experimenters to adopt incomplete block designs for experimentation

could be the fear of analysis of data generated. However, with the advent of high speed computers and software and the presence of statisticians around, this fear is unfounded. The other major concern of the experimenters is to demonstrate the effect of all the treatments to the farmers/monitoring team at one place and for that reason the experimenters prefer a complete block design over an incomplete block design. But complete block designs are accompanied with large intra block variances. In the Agricultural Field Experiments Information System, which contains information about more than thirty three thousand experiments conducted by scientists of National Agricultural Research System in India, it is found that of the 70 per cent experiments conducted as randomized complete block design, the block (or replication) effects are not significantly different from each other (in other words block mean squares are not large as compared to mean square error) in about 70 per cent of the experiments. The field contour maps prepared for some experiments also indicate that it may not be possible to form long rectangular blocks because there are several fertility patches within the rectangular blocks. So it is always better to form long blocks and then have smaller sub blocks within each of the big blocks. The treatments may then be allocated randomly to the sub-blocks instead of allocating randomly to the plots within larger blocks. This type of blocking structure gives rise to a *nested* classification and the design is a nested block design. A particular class of nested designs, which could be of great interest to the experimenters, is the one in which the bigger block is a complete replication and the smaller blocks are incomplete blocks. In this way, the experimenters can demonstrate the effect of all the treatments at one place using the larger blocks which are complete replication and at the same time, the sub-blocks would address the problem of large blocks and would take care of the heterogeneity creeping in because of the large blocks.

Such designs in which the larger blocks are complete replication have been termed as resolvable designs in the literature. The concept of resolvability in incomplete block designs was introduced by Bose (1942). An incomplete block design is said to be resolvable if the blocks can be grouped in such a manner that each group of blocks is a complete replication. A block design with v treatments, b blocks, replication r and block size k ($<v$) is called resolvable if b blocks can be divided into r groups containing $\dfrac{b}{r} = x$ blocks of size k each such that each group is a complete replication; in other words the x blocks in each of the r groups contain all the v treatments exactly once. Obviously, $v = xk$ and $b = rx$. Square lattice, rectangular lattice designs are also resolvable block designs.

6.2 Example 1

An experimenter is interested in comparing $v = 20$ genotypes. A total of $n = 60$ experimental units are available that can be arranged in $b = 15$ blocks of size $k = 4$ experimental units each.

The 15 blocks can be so arranged that one gets $r = 3$ replications with five blocks of size four each in each replication. Here $\dfrac{b}{r} = x = 5$. A resolvable block design with the block contents is

Design 1														
Replication I						Replication II					Replication III			
B1	B2	B3	B4	B5		B6	B7	B8	B9	B10		B11	B12	B13	B14	B15
1	2	3	4	5		14	16	3	6	19		13	2	3	20	5
6	7	8	9	10		8	9	11	12	7		9	18	15	7	16
11	12	13	14	15		1	2	17		5		1	14	6	11	12
16	17	18	19	20		20	15	10	4	13		17	10	19	4	8

In another experiment, the experimenter is interested in comparing $v = 35$ genotypes. A total of $n = 105$ experimental units are available that can be arranged in $b = 21$ blocks of size $k = 5$ experimental units each. The 21 blocks can be so arranged that one gets $r = 3$ replications with seven blocks of size five each in each replication. Here $\dfrac{b}{r} = x = 7$. A resolvable block design with the block contents is

Design 2																	
Replication I						Replication II						Replication III					
B1	31	11	23	20	5	B8	24	3	10	17	31	B15	19	11	1	35	24
B2	13	7	25	33	15	B9	11	4	18	25	32	B16	2	29	25	12	20
B3	4	30	22	19	10	B10	6	13	34	27	20	B17	10	34	7	18	23
B4	18	28	9	29	3	B11	26	12	19	33	5	B18	13	3	21	26	30
B5	27	8	35	17	2	B12	9	30	16	2	23	B19	6	9	33	17	22
B6	21	12	6	24	32	B13	35	7	28	14	21	B20	31	4	27	15	14
B7	1	34	16	26	14	B14	1	29	15	8	22	B21	28	8	5	16	32

B# denotes the block number

Resolvable block designs have a lot of practical significance as these designs allow performing an experiment one replication at a time. For example, field trials with large number of crop varieties cannot always be laid out in a single location or a single season. It is desired that variation due to location or time periods may also be controlled along with controlling within location or within time period variation. This can be handled by using resolvable block designs. Locations or time periods can be taken as replications and the variation within locations or within time periods can be taken care of by blocks within replications. In some experimental situations, it may not be feasible to run the entire experiment in one session. The experiment may have to be run in different sessions. Resolvable designs become useful in these situations too. A single replication may be run in one session. If the experiment is discontinued after first session or second session, it would not matter much to the experimenter because all the treatments would have appeared till the time the experiment has been run. If the experiment is run up to second session and discontinued from the third session, the experimenter can still

analyze the data in the same way as it would have been analyzed had the experiment run for all the sessions. But if the experiment is discontinued after first session, then there would be a problem in the analysis of data because of single replication of treatments. The experimenter would not be able to get the experimental error and so testing of hypotheses would be a problem.

For resolvable block designs, the randomization has five steps. These are the following:

Randomize the treatments, *i.e.*, random allocation of treatment numbers (or labels) to treatments.

Randomize the replications.

Randomize the blocks within replications.

Randomize the treatments to experimental units within each block of all the replications; separate randomization should be done for each of the blocks.

An interesting feature of resolvable block designs is that these are run in complete replications. In that sense, these designs look like RCB designs. But in case of RCB designs there are no blocks within replications. The replication itself is a block. All the *v* treatments are randomized within the replication. However, in case of a resolvable design, the treatments are randomized within blocks in each replication. This is an advantage, which enables the experimenter to take out a part of the variability between blocks within replications from the error.

Following on the Example 1, the difference in randomization of treatments in an RCB design and a resolvable (nested) design is demonstrated below: The randomization in case of an RCB design is over the 20 experimental units within each replication. There would be no smaller blocks. The randomized layout for an RCB design could be

Replication I						Replication II						Replication III				
15	2	8	20	7		14	16	13	6	19		13	12	3	20	8
16	10	18	19	12		20	8	11	17	5		11	1	15	7	16
3	5	13	14	1		7	2	1	18	12		9	14	6	18	2
6	17	11	9	4		9	15	10	4	3		17	10	19	4	5

However, as a resolvable or nested design, the randomization is a restricted randomization. Following the steps of randomization given above, the randomized layout in this case would be

Replication III						Replication I						Replication II				
B3	B5	B1	B4	B2		B8	B6	B9	B10	B7		B14	B15	B12	B11	B13
15	8	9	4	14		3	1	4	5	2		6	19	16	14	3
6	12	17	7	2		8	6	9	10	7		12	7	9	8	11
19	5	1	20	10		13	11	14	15	12		18	5	2	1	17
3	16	13	11	18		18	16	19	20	17		4	13	15	20	10

It is obvious from this example that simply changing the randomization helps in controlling the experimental error.

6.2.1 Analysis

The statistical model for analysis of data generated from a resolvable design is the following:

response = general mean + treatments effects + blocks (or replications) effect + sub-blocks nested within blocks effect + error.

This model can also be written as

$$y_{ihjl} = \mu + \tau_i + \beta_h + \rho_{hj} + e_{ihjl}$$

where y_{ihjl} denotes the response from the *l*th experimental unit receiving *i*th treatment in sub-block *j* of block (or replicate) *h*. Thus, the triplet (*h, j, l*) identifies the experimental unit and the design is an allocation of a set of *v* treatments to these experimental units (*h, j, l*). The standard linear model for the responses incorporates a mean effect μ, block (or replicate) effect β_h, sub-block effect ρ_{hj}, treatment effect τ_i and mean zero, independent, equi-variable $(\sigma^2 > 0)$ random error terms e_{ihjl}, $i = 1,2,\cdots,v$; $h = 1,2,\cdots,r$; $j = 1,2,\cdots,x$; $l = 1,2,\cdots;v/x = k$.

For the Example 1 given above in Section 6.2, as well as for a general resolvable block design, the analysis of variance is given in Table 6.1.

Table 6.1: Skeleton ANOVA table for a resolvable design

ANOVA (*Resolvable Design*)		ANOVA (*Resolvable Design*) for *v, b, k, r*	
Source	**DF**	**Source**	**DF**
Treatments	19	Treatments	$(v-1)$
Blocks	14	Blocks	$(b-1)$
Replications	2	Replications	$(r-1)$
Sub-Blocks within Blocks (or replications)	12	Sub-Blocks within Blocks (or replications)	$r(x-1) = (b-r)$
Error	26	Error	$(vr-v-b+1)$
Total	59	Total	$(vr-1)$

However, in case of an RCB design with $v = 20$ treatments and $r = 3$ replications (or blocks), the analysis of variance is given in Table 6.2.

Table 6.2: Skeleton ANOVA table for a RCB design

ANOVA (*RCBD*)		ANOVA (*RCBD*) *v, r*	
Source	**DF**	**Source**	**DF**
Treatments	19	Treatments	$(v-1)$
Blocks (or Replications)	2	Blocks (or Replications)	$(r-1)$
Error	38	Error	$(v-1)(r-1)$
Total	59	Total	$(vr-1)$

It may be seen from these Tables that the use of a resolvable incomplete block design leads to a reduction in error on account of sub-blocks within blocks. In case of a RCB design this component of variability because of sub-blocks within blocks is merged with the error. A strong message that follows from here is that changing the randomization and allocation of treatments to smaller block structure in a RCB design by forming smaller blocks within large blocks helps in reducing the experimental error.

Remark 6.1 Resolvable block designs are in fact a special class of designs called nested block designs wherein there are two systems of blocking. There are bigger blocks, which are complete blocks (or complete replications) and each bigger block is subdivided into smaller blocks (or sub-blocks) to which the treatments are allotted randomly. The union of all the smaller blocks forms the bigger block which is a complete replication. Therefore, the resolvable designs are nested designs with the smaller (incomplete) blocks nested within the larger blocks (or complete replications). These designs have been shown to be very useful in on-farm agricultural research experiments (see for example, Nigam *et al.*, 2003).

Resolvable block designs have been studied extensively in the literature. For instance, there may be a resolvable BIB design. A BIB design with $v = 9$, $b = 12$, $r = 4$, $k = 3$, $\lambda = 1$ is resolvable. The block contents are {(1, 2, 3), (4, 5, 6), (7, 8, 9)}; {(1, 4, 7), (2, 5, 8), (3, 6, 9)}; {(1, 6, 8), (2, 4, 9), (3, 5, 7)}; {(1, 5, 9), (3, 4, 8), (2, 6, 7)}, here {.} denotes the set of blocks forming a complete replication. Even Partially Balanced Incomplete Block (PBIB) designs could be resolvable. Lattice designs introduced by Yates (1936) especially for varietal trials are resolvable incomplete block designs. Lattice designs are available for limited number of varieties and block sizes. The simple ($r = 2$) and triple ($r = 3$) Lattices require $v = s^2$, $k = s$ and number of blocks per replication is also s, where s is a prime number. Further conditions are imposed on v in quadruple and higher order Lattices. Harshbarger (1949) extended the principle of square lattice to simple and triple rectangular Lattice with $v = s(s-1)$, $k = s-1$ (see also Nair, 1951). Patterson and Williams (1976) generated special class of resolvable designs called alpha designs. These designs are less restrictive in the sense that the number of treatments and the number of blocks have a common integer multiple $s \geq 2$. The parameters of alpha design are $v = ks$, $b = rs$, r, k, s. Parsad *et al.* (2007a) obtained a catalogue of A-efficient and D-efficient alpha designs. For more details on alpha designs and also for getting a randomized layout of alpha design, please see Design Resources Server at www.iasri.res.in/design/. R software can be used to obtain alpha design as well. The following code in R generates a alpha design for 30 treatments with block size 3 in two replications.

library(agricolae)

design.alpha(trt=1:30,k=3,r=2,serie=2)

In order to give a better exposition of the subject of resolvable designs, another example is in order.

6.3 Example 2

An initial varietal trial was conducted to study the performance of 24 strains of Toria using

an alpha design at Pantnagar with three replications. The seed yield in kg/ha was recorded. The details of strains, design adopted and data obtained are given as under.

Table 6.3: Treatment details and seed yield data of initial varietal trial

RAU DT-01-03	1		TK–06-1	9		RH- 304	17
RAU DT-01-02	2		TK–06 - 2	10		TH– 302	18
BAUSM-92-24	3		TL-2027	11		JMT-05	19
RGN-186	4		TL-2013	12		**PT–303 (National Check)**	20
EJ-17	5		JMT–02-6	13		**Zonal Check**	21
NPJ–112	6		NDT–05-5	14		PTC–99-14	22
VLT-4	7		NDRE–2002-16	15		**JD–6 (Check)**	23
RRN-612	8		PT–2004-3	16		ORT-17-6-16	24

Replication 1						
Block 1	1 (1555.6)	5 (1160.5)	9 (1308.6)	13 (1382.7)	17 (987.7)	**21** (1135.8)
Block 2	2 (1284.0)	6 (1086.4)	10 (1284.0)	14 (1111.1)	18 (938.3)	22 (1308.6)
Block 3	3 (1234.6)	7 (419.8)	11 (1308.6)	15 (963.0)	19 (963.0)	**23** (987.7)
Block 4	4 (1234.6)	8 (987.7)	12 (1284.0)	16 (913.6)	**20** (1160.5)	24 (790.1)

Replication 2						
Block 5	1 (1481.5)	6 (1086.4)	11 (1308.6)	16 (1284.0)	19 (1111.1)	22 (1185.2)
Block 6	2 (987.7)	7 (308.6)	12 (1234.6)	13 (1308.6)	**20** (765.4)	**23** (938.3)
Block 7	3 (1012.3)	8 (864.2)	9 (1234.6)	14 (938.3)	17 (913.6)	24 (864.2)
Block 8	4 (1135.8)	5 (987.7)	10 (987.7)	15 (740.7)	18 (963.0)	**21** (1135.8)

Replication 3						
Block 9	1 (1284.0)	7 (333.3)	12 (1135.8)	15 (839.5)	18 (814.8)	24 (888.9)
Block 10	2 (1135.8)	8 (913.6)	9 (1456.8)	16 (1037.0)	19 (938.3)	**21** (1037.0)
Block 11	3 (963.0)	5 (1209.9)	10 (1259.3)	13 (1234.6)	**20** (963.0)	22 (1111.1)
Block 12	4 (1086.4)	6 (765.4)	11 (1111.1)	14 (1037.0)	17 (938.3)	**23** (938.3)

Figures in the parenthesis give the seed yield in kg/ha.

In the sequel we perform the analysis of the data generated using a resolvable design. The following analyses are performed: (a) analysis of variance to test the homogeneity of the treatment effects, (b) since the data are unbalanced or non-orthogonal, obtain the LSMEANS or adjusted means of the treatments and make pairwise treatment comparisons to identify the best performing treatment, and (c) to make comparisons of the 3 check strains (treatments 20, 21, 23) with the 21 new strains of Toria.

6.3.1 Analysis of data

This is an Example of a resolvable (nested) block design in which the bigger blocks are the replications and the four blocks of size 6 each within each replication are the sub-blocks. In this experiment, there are 21 new strains of Toria and 3 check strains. In order to compare new strains with check, appeal will be made to contrast analysis.

The design adopted is an alpha design, which is a resolvable incomplete block design. The parameters of the design are $v = 24$, $b = 12$, $r = 3$, Number of blocks per replication $= x = b/r = 4$. The data has been analyzed using SAS software. The analysis of data is in continuation below:

```
DATA resolvable_design;
INPUT blk sblk trt syield;
CARDS;
```

1	1	1	1555.6
1	1	5	1160.5
1	1	9	1308.6
1	1	13	1382.7
1	1	17	987.7
1	1	21	1135.8
1	2	2	1284.0
1	2	6	1086.4
1	2	10	1284.0
1	2	14	1111.1
1	2	18	938.3
1	2	22	1308.6
1	3	3	1234.6
1	3	7	419.8
1	3	11	1308.6
1	3	15	963.0
1	3	19	963.0
1	3	23	987.7
1	4	4	1234.6
1	4	8	987.7
1	4	12	1284.0
1	4	16	913.6
1	4	20	1160.5
1	4	24	790.1
2	1	1	1481.5
2	1	6	1086.4
2	1	11	1308.6

2	1	16	1284.0
2	1	19	1111.1
2	1	22	1185.2
2	2	2	987.7
2	2	7	308.6
2	2	12	1234.6
2	2	13	1308.6
2	2	20	765.4
2	2	23	938.3
2	3	3	1012.3
2	3	8	864.2
2	3	9	1234.6
2	3	14	938.3
2	3	17	913.6
2	3	24	864.2
2	4	4	1135.8
2	4	5	987.7
2	4	10	987.7
2	4	15	740.7
2	4	18	963.0
2	4	21	1135.8
3	1	1	1284.0
3	1	7	333.3
3	1	12	1135.8
3	1	15	839.5
3	1	18	814.8
3	1	24	888.9
3	2	2	1135.8
3	2	8	913.6
3	2	9	1456.8
3	2	16	1037.0
3	2	19	938.3
3	2	21	1037.0
3	3	3	963.0
3	3	5	1209.9
3	3	10	1259.3
3	3	13	1234.6
3	3	20	963.0
3	3	22	1111.1

3	4	4	1086.4
3	4	6	765.4
3	4	11	1111.1
3	4	14	1037.0
3	4	17	938.3
3	4	23	938.3

```
;
PROC GLM;
CLASS blk sblk trt;
MODEL syield = trt blk sblk(blk);
MEANS trt;
LSMEANS trt/PDIFF LINES;
CONTRAST 'Controls' trt 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 -1 0 0 0,
                      trt 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 -2 0;
CONTRAST 'Tests'      trt 1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 -2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 -3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 -5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 -6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 -7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 -8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 -9 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 -10 0 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 -11 0 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 -12 0 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 -13 0 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -14 0 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -15 0 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -16 0 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -17 0 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -18 0 0 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 -19 0 0 0,
                      trt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 1 0 -20;
CONTRAST 'Tests vs Controls' trt 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 -21 -21 3 -21 3;
RUN;
```

6.3.2 Output of analysis

The results obtained from the analysis are described in Table 6.4.

Table 6.4: Output of analysis using SAS command

ANOVA

Source	DF	SS	MS	F-value	Prob > F
Model	34	3535274.91	103978.67	12.79	<0.0001
Error	37	300877.64	8131.83		
Corrected Total	71	3836152.55			

R-Square	CV	Root MSE	Syield Mean
0.922	8.543	90.177	1055.564

ANOVA

Source	DF	Type-III SS	MS	F-value	Prob > F
Treatments	23	2555476.22	111107.66	13.66	<0.0001
Blocks	2	135161.75	67580.88	8.31	0.0010
Sub-Blocks within Blocks	9	194315.00	21590.56	2.66	0.0175
Error	37	300877.64	8131.83		
Corrected Total	71	3836152.55			

The model with nested classification has been able to explain about 92 per cent of the total variability in the data. It is evident from the ANOVA that the treatment effects are significantly different (p-value < 0.0001). But more prominently, it is relevant to note that the sub-blocks within blocks are highly significant (p-value = 0.0175) meaning thereby that it has been advantageous to form sub-blocks within blocks as it has resulted in considerable reduction in error mean square.

Remark 6.2. It may be noted from ANOVA in Table 6.4 above that the sum of squares due to the three components *viz.*, treatments, blocks and sub-blocks do not add up to the model sum of squares. The reason is simple. The blocks are incomplete and, therefore, the treatment sum of squares is adjusted for blocks. In this case the unadjusted sum of squares due to treatments is not the same as adjusted sum of squares due to treatments.

The summary of treatments and their LS Means are given in the Tables 6.5.

Table 6.5: Treatment wise mean and standard deviation of seed yield

Treatment	Syield		Treatment	Syield	
	Mean	**Standard Deviation**		**Mean**	**Standard Deviation**
1	1440.37	140.39	13	1308.63	74.05
2	1135.83	148.15	14	1028.80	86.69
3	1069.97	144.69	15	847.73	111.38
4	1152.27	75.46	16	1078.20	188.61
5	1119.37	116.67	17	946.53	37.73
6	979.40	185.33	18	905.37	79.40
7	353.90	58.39	19	1004.13	93.46
8	921.83	62.16	20	962.97	197.55
9	1333.33	113.15	21	1102.87	57.04
10	1177.00	164.40	22	1201.63	99.77
11	1242.77	114.03	23	954.77	28.52
12	1218.13	75.46	24	847.73	51.42

Treatment	Syield LS Mean		Treatment	Syield LS Mean		Treatment	Syield LS Mean
1	1403.80		9	1343.11		17	982.92
2	1135.83		10	1164.35		18	913.14
3	1074.43		11	1164.35		19	928.32
4	1198.44		12	1270.22		20	994.63
5	1136.00		13	1330.61		21	1112.75
6	912.96		14	1035.89		22	1115.35
7	384.24		15	872.74		23	984.53
8	941.29		16	1024.13		24	894.38

The pairwise treatment comparisons using the LS means are given in Table 6.6.

From Table 6.6, it is evident that the new strain RAU DT-01-03 (Treatment 1) is the most promising in terms of seed yield. It is significantly different from all the three check strains. It is, however, statistically at par with the new strains TK–06-1 (Treatment 9); TL-2013 (Treatment 12); JMT–02-6 (Treatment 13). The new strain VLT-4 (Treatment 7) is the lowest yielding in terms of seed yield. Its seed yield is significantly lower than even the check strains. The check strains (treatments 22 and 23) are statistically at par but are significantly different from the third check strain (Treatment 21).

Table 6.6: *t* comparison lines for least squares means of treatments

LS-means with the same letter are not significantly different

					Syield LS MEAN	Treatment	LSMEAN Number
			A		1403.80	1	1
	B		A		1343.11	9	9
	B		A		1330.61	13	13
	B		A	C	1270.22	12	12
	B		D	C	1198.44	4	4
	B	E	D	C	1193.56	11	11
	F	E	D	C	1164.35	10	10
G	F	E	D	C	1136.00	5	5
G	F	E	D	C	1121.76	2	2
G	F	E	D	C	1115.35	22	22
G	F	E	D	C	1112.75	**21**	21
G	F	E	D	H	1074.43	3	3
G	F	E	I	H	1035.89	14	14
G	F		I	H	1024.13	16	16
G			I	H	994.62	**20**	20
G			I	H	984.53	**23**	23
G			I	H	982.92	17	17
			I	H	941.29	8	8
			I	H	928.32	19	19
			I	H	913.14	18	18
			I	H	912.96	6	6
			I		894.38	24	24
			I		872.74	15	15
			J		384.24	7	7

In case the experimenter is interested in making further comparisons among tests, among controls and tests *vs* controls, then the results obtained are given in Table 6.7.

Table 6.7: Splitting of treatment sum of squares

Contrast	DF	Contrast SS	MS	F-value	Prob > F
Among Tests	20	2535129.68	126756.48	15.59	<0.0001
Among Controls	2	23375.33	11687.67	1.44	0.2505
Tests *vs* Controls	1	5660.36	5660.36	0.70	0.4095

It is once again very evident from Table 6.7 that the new strains are significantly different (p-value < 0.0001); but the check strains are not significantly different (p-value = 0.2505). The difference between new strains and the check strains is also not statistically significant (p-value = 0.4095). However, from Table 6.6, it can be seen that some of the pairwise comparisons of check with test varieties are significantly different. Using LS MEANS, one can see that based on the character yield, Check variety 21: Zonal Check is the best performing check and varieties numbered as 1,2,4,5,9,10,11,12,13 and 22 give higher yield than the best performing check. The varieties 2,4,5,10,11,12 and 22 are statistically at par with best performing check variety 21; whereas tests 1,9,13 are statistically significant from best performing check variety 21 at 5% level of significance. Therefore, one may conclude that the group of test varieties 1,9,13 is the best performing group.

6.3.3 Analysis using R

The data has also been analyzed using R software. The R code for the analysis of data is given in the sequel. The output obtained is similar to the one obtained using SAS and to avoid repetition, the same is not reported here.

```
d9=read.table("resolvable_design.txt",header=TRUE)
attach(d9)
names(d9)
#Treatment means and standard deviations
aggregate(syield~trt,data=d9,mean)
aggregate(syield~trt,data=d9,sd)
#Treatment wise box plot of calorie
boxplot(syield~trt)
#anova with class variable treatments and block
trt=factor(trt)
rep=factor(rep)
blk=factor(blk)
lm1=lm(syield~trt+rep+blk/rep-blk)
#anova(lm1)
library(car)
Anova(lm1,type="III")
library(lsmeans)
lsm=lsmeans(lm1,"trt")
lsm
```

```
pairs(lsm)
#to provide letters for groups, need to install multcompView
library(multcompView)
cld(lsm,Letters="abcdefgh")
controls=contrast
(lsm,list(con3  =c(1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con4=c(1,1,-2,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con5=c(1,1,1,-3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con6=c(1,1,1,1,-4,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con7=c(1,1,1,1,1,-5,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con8=c(1,1,1,1,1,1,-6,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con9=c(1,1,1,1,1,1,1,-7,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con10=c(1,1,1,1,1,1,1,1,-8,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con11=c(1,1,1,1,1,1,1,1,1,-9,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con12=c(1,1,1,1,1,1,1,1,1,1,-10,0,0,0,0,0,0,0,0,0,0,0,0,0),
          con13=c(1,1,1,1,1,1,1,1,1,1,1,-11,0,0,0,0,0,0,0,0,0,0,0,0),
          con14=c(1,1,1,1,1,1,1,1,1,1,1,1,-12,0,0,0,0,0,0,0,0,0,0,0),
          con15=c(1,1,1,1,1,1,1,1,1,1,1,1,1,-13,0,0,0,0,0,0,0,0,0,0),
          con16=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,-14,0,0,0,0,0,0,0,0,0),
          con17=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,-15,0,0,0,0,0,0,0,0),
          con18=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,-16,0,0,0,0,0,0,0),
          con19=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,-17,0,0,0,0,0,0),
          con20=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,-18,0,0,0,0,0),
          con21=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,-19,0,0,0),
          con22=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,0,1,0,-20)))
tests.vs.controls=contrast(lsm,list(con23=c(3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,-21,-21,3,-21,3)))
lht(lm1,controls@linfct)
lht(lm1,tests@linfct)
lht(lm1,tests.vs.controls@linfct)
detach(d9)
```

6.4 Nested designs

In Remark 6.1, there has been a mention of nested designs. It was mentioned there that the resolvable block designs are in fact nested block designs with blocks (bigger blocks) as

complete replications and the sub-blocks (smaller blocks) as incomplete blocks. Since the bigger blocks are complete replications, these nested designs find favour with the researchers and it is recommended that such designs should be used more often than not in field experimentation, particularly when it is not possible to have uniformity trials support to form blocks.

Since we have nested designs with bigger blocks as complete blocks, it is indeed possible to have nested designs with incomplete bigger blocks. The bigger blocks need not necessarily be complete blocks. These could very well be incomplete blocks. Kleczkouski (1960) devised a form of nested block design for $v = 8$ treatments for a series of experiments in which beans plants, in the two primary leaves stage, were inoculated with sap from tobacco plants infected with tobacco necrosis virus. The treatments were eight different virus concentrations. Each leaf had two inoculations, one for each half-leaf. Ignoring the leaf positions, plants and leaves were, respectively, the blocks of (size 4) and sub-blocks (of size 2) of a nested balanced incomplete block design. Preece (1967) for the case of two blocking systems, one nested within the other, introduced a Nested Balanced Incomplete Block (NBIB) design. An arrangement of v treatments each replicated r times in two systems of blocks is said to be a NBIB design with parameters (v, r, b_1, k_1, λ_1, b_2, k_2, λ_2, m) if the second system of blocks is nested within the first system of blocks, with each block from the first system containing exactly m blocks from the second system (sub-blocks); ignoring the second system of blocks leaves a balanced incomplete block (BIB) design with b_1 blocks each of size k_1 and λ_1 concurrences; ignoring the first system of blocks leaves a BIB design with $b_2 = b_1 m$ blocks each of size $k_2 = k_1/m$ units with λ_2 concurrences.

The parameters of a nested BIB design satisfy the following parametric relations:

$vr = b_1 k_1 = b_1 k_2 m = b_2 k_2;$

$\lambda_1(v - 1) = r(k_1 - 1); \lambda_2(v - 1) = r(k_2 - 1)$

$(\lambda_1 - m\lambda_2)(v - 1) = r(m - 1).$

6.5 Example 3

The following is a NBIB design with parameters $v = 8$, $r = 14$, $b_1 = 28$, $k_1 = 4$, $\lambda_1 = 6$, $b_2 = 56$, $k_2 = 2$, $\lambda_2 = 2$, $m = 2$:

(1, 5); (2, 3)	(1, 6); (4, 7)	(3, 5); (8, 6)	(2, 1); (8, 4)
(2, 6); (3, 4)	(2, 7); (5, 1)	(4, 6); (8, 7)	(3, 2); (8, 5)
(3, 7); (4, 5)	(3, 1); (6, 2)	(5, 7); (8, 1)	(4, 3); (8, 6)
(4, 1); (5, 6)	(4, 2); (7, 3)	(6, 1); (8, 2)	(5, 4); (8, 7)
(5, 2); (6, 7)	(5, 3); (1, 4)	(7, 2); (8, 3)	(6, 5); (8, 1)
(6, 3); (7, 1)	(6, 4); (2, 5)	(1, 3); (8, 4)	(7, 6); (8, 2)
(7, 4); (1, 2)	(7, 5); (3, 6)	(2, 4); (8, 5)	(1, 7); (8, 3)

The bigger blocks form a BIB design with parameters $v = 8$, $b_1 = 28$, $r_1 = 14$, $k_1 = 4$, $\lambda_1 = 6$ and the design is (1, 5, 2, 3); (2, 6, 3, 4); (3, 7, 4, 5); (4, 1. 5, 6); (5, 2, 6, 7); (6, 3, 7, 1); (7, 4, 1, 2); (1, 6, 4, 7); (2, 7, 5, 1); (3, 1, 6, 2); (4, 2, 7, 3); (5, 3, 1, 4); (6, 4, 2, 5); (7, 5, 3, 6); (3, 5, 8, 6); (4, 6, 8, 7); (5, 7, 8, 1); (6, 1, 8, 2); (7, 2, 8, 3); (1, 3, 8, 4); (2, 4, 8, 5); (2, 1, 8, 4); (3, 2, 8, 5); (4, 3, 8, 6); (5, 4, 8, 7); (6, 5, 8,1); (7, 6, 8, 2); (1, 7, 8, 3). Similarly, the sub-blocks form a BIB design with parameters $v = 8$, $b_2 = 56$, $r = 14$, $k_2 = 2$, $\lambda_2 = 2$ and the design is (1, 5); (2, 3); (2, 6); (3, 4); (3, 7); (4, 5); (4, 1); (5, 6); (5, 2); (6, 7); (6, 3); (7, 1); (7, 4); (1, 2); (1, 6); (4, 7); (2, 7); (5, 1); (3, 1); (6, 2); (4, 2); (7, 3); (5, 3); (1, 4); (6, 4); (2, 5); (7, 5); (3, 6); (3, 5); (8, 6); (4, 6); (8, 7); (5, 7); (8, 1); (6, 1); (8, 2); (7, 2); (8, 3); (1, 3); (8, 4); (2, 4); (8, 5); (2, 1); (8, 4); (3, 2); (8, 5); (4, 3); (8, 6); (5, 4); (8, 7); (6, 5); (8,1); (7, 6); (8, 2); (1, 7); (8, 3).

Thus far the focus has been on experimental settings in which one blocking system is nested within another blocking system. There may, however, be experimental settings where two cross classified factors cause variability in the experimental material and are nested within the blocking factor. Nested row-column designs have been developed for such situations. Consider the case of animal nutrition experiments where lactation number has been taken as a blocking factor. However, the age and stage of lactation within animals of same lactation number may contribute significantly to the variability and thus to the error variance and these two factors are cross classified with each other. Therefore, for such experimental situations, within blocking factor (lactation number), two cross classified factors, age (rows) and stage of lactation (columns) are nested.

For these experimental settings the experimental units are broadly classified into b bigger blocks such that within each bigger block the experimental units are arranged in p rows and q columns. The block sizes are $k = pq$ and the total number of experimental units are bpq. The number of treatments v may be equal to pq, in which case the bigger block is a complete replication. It is indeed possible that $v > pq$ and in that case the bigger block is an incomplete block.

To cope with these type of situations, Cochran and Cox (1957) suggested the use of repeated lattice square designs where each square can be considered a block (complete replication) within which are nested two other factors, denoted by rows and columns, so that one can eliminate two sources of variability within each block. These designs, however, demand that the number of treatments is $v = s^2$, where s is a prime number or power of a prime number. Similar to a NBIB design, for these situations, Singh and Dey (1979) introduced Balanced Incomplete Block Designs with Nested Rows and Columns (BIB-RC design). A block design with nested rows and columns with v treatments and b blocks, each containing p rows and q columns ($pq < v$) is said to be a BIB-RC design if the following conditions are satisfied:

1.   every treatment occurs at most once in a block;

2.   given a pair of treatments (i, j)

$$p\lambda_{r(i,j)} + q\lambda_{c(i,j)} - \lambda_{b(i,j)} = \lambda \ \text{ (a constant)}$$

where $\lambda_{r(i,j)}$, $\lambda_{c(i,j)}$ and $\lambda_{b(i,j)}$ denote the number of blocks in which treatments i and j occur together in the same row, same column and elsewhere respectively and λ is a constant independent of i and j.

It is easy to see that in a BIB-RC design every treatment occurs in exactly r blocks, where

$$r = \lambda(v-1)/(p-1)(q-1).$$

The following is a BIB-RC design with $v = 9$, $b = 4$, $r = 4$, $p = 3$, $q = 3$ and $\lambda = 2$ using $p\lambda_{r(i,j)} + q\lambda_{c(i,j)} - \lambda_{b(i,j)} = \lambda$ (a constant) and the fact that

$\lambda_{r(i,j)} = 1$, $\lambda_{c(i,j)} = 1$, $\lambda_{b(i,j)} = 4$ for all $i \neq j = 1,2,\cdots,9$.

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

It may be seen that the four blocks, ignoring rows and columns, are four complete blocks (or complete replications). Ignoring columns, and treating rows as blocks one gets a BIB design with parameters $v = 9$, $b = 12$, $r = 4$, $k = 3$, $\lambda^* = 1$. The blocks of this design are (1, 2, 3); (4, 5, 6); (7, 8, 9); (1, 4, 7); (2, 5, 8); (3, 6, 9); (1, 6, 8); (9, 2, 4); (5, 7, 3); (1, 9, 5); (6, 2, 7); (8, 4, 3). Similarly, ignoring rows and treating columns as blocks, one gets a BIB design with parameters $v = 9$, $b = 12$, $r = 4$, $k = 3$, $\lambda^{**} = 1$. The blocks of this design are (1, 4, 7); (2, 5, 8); (3, 6, 9); (1, 2, 3); (4, 5, 6); (7, 8, 9); (1, 9, 5); (6, 2, 7); (8, 4, 3); (1, 6, 8); (9, 2, 4); (5, 7, 3). It may be seen that both the BIBDs have the same blocks.

The following is another BIB-RC design with $v = 13$, $b = 26$, $r = 12$, $p = 2$, $q = 3$, $\lambda = 2$ using $p\lambda_{r(i,j)} + q\lambda_{c(i,j)} - \lambda_{b(i,j)} = \lambda$ (a constant) and the fact that $\lambda_{r(i,j)} = 2, \lambda_{c(i,j)} = 1, \lambda_{b(i,j)} = 5$ for all $i \neq j = 1,2,\cdots,9.$

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

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

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

7	9	2
5	3	10

13	1	4
12	11	8

8	10	3
6	4	11

1	2	5
13	12	9

9	11	4
7	5	12

2	3	6
1	13	10

10	12	5
8	6	13

3	4	7
2	1	11

11	13	6
9	7	1

4	5	8
3	2	12

12	1	7
10	8	2

5	6	9
4	3	13

13	2	8
11	9	3

6	7	10
5	4	1

One can see that each 2×3 array of 6 experimental units is a big block and these 26 blocks form a BIB design with parameters $v = 13$, $b = 26$, $r = 12$, $k = 6$, $\lambda^* = 5$. But each array has 2 rows of 3 experimental units each (or 3 columns of 2 experimental units each). In that respect, it is a BIB-RC design with parameters $v = 13$, $b = 26$, $r = 12$, $p = 2$, $q = 3$, $\lambda = 2$. In this Example, the bigger block is not a complete block because it has only 6 treatments whereas the total number of treatments in the design is 13.

6.6 Example 4

An agricultural field experiment was conducted in 9 treatments with 36 plots arranged in 4 complete blocks and a sample of harvested output from all the 36 plots had to be analyzed block wise by three technicians using three different operations. The data collected are given in Table 6.8.

Table 6.8: Data from an experiment conducted using a nested row column design

Block I			
	Technicians		
Operations	1	2	3
1	1 (1.1)	2 (2.1)	3 (3.1)
2	4 (4.2)	5 (5.3)	6 (6.3)
3	7 (7.4)	8 (8.7)	9 (9.6)

Block II			
	Technicians		
Operations	1	2	3
1	1 (2.1)	4 (5.2)	7 (8.3)
2	2 (3.2)	5 (6.7)	8 (9.9)
3	3 (4.5)	6 (7.6)	9 (10.3)

Block III			
	Technicians		
Operations	1	2	3
1	1 (1.2)	6 (6.3)	8 (8.7)
2	9 (9.4)	2 (2.7)	4 (4.8)
3	5 (5.9)	7 (7.8)	3 (3.3)

Block IV			
	Technicians		
Operations	1	2	3
1	1 (3.1)	9 (11.3)	5 (7.8)
2	6 (8.1)	2 (4.5)	7 (9.3)
3	8 (10.7)	4 (6.9)	3 (5.8)

The numbers in the cells are the treatment labels and figures in brackets are the responses. In the continuation of this example is presented the analysis of the data for identifying the best performing treatment.

6.6.1 Analysis

This experiment has been conducted as a block design with nested rows and columns. Of course the blocks are complete blocks. In that sense, if one ignores the rows and columns, the resulting design is a RCB design. The parameters of the design are $v = 9$, $b = 4$, $p = q = 3$. We give in the sequence the analysis of the data using SAS. The SAS commands and the data structure has also been given in detail.

```
DATA ncbrcd;
INPUT blk tech oper trt obs;
CARDS;
1 1  1 1    1.1
1 1  2 4    4.2
1 1  3 7    7.4
1 2  1 2    2.1
1 2  2 5    5.3
1 2  3 8    8.7
1 3  1 3    3.1
1 3  2 6    6.3
1 3  3 9    9.6
2 1  1 1    2.1
2 1  2 2    3.2
2 1  3 3    4.5
2 2  1 4    5.2
2 2  2 5    6.7
2 2  3 6    7.6
2 3  1 7    8.3
2 3  2 8    9.9
2 3  3 9   10.3
3 1  1 1    1.2
3 1  2 9    9.4
3 1  3 5    5.9
3 2  1 6    6.3
3 2  2 2    2.7
3 2  3 7    7.8
3 3  1 8    8.7
3 3  2 4    4.8
3 3  3 3    3.3
4 1  1 1    3.1
4 1  2 6    8.1
```

```
4 1  3 8 10.7
4 2  1 9 11.3
4 2  2 2  4.5
4 2  3 4  6.9
4 3 1 5  7.8
4 3 2 7  9.3
4 3 3 3  5.8
;
ODS RTF FILE='nestedcomplete.rtf';
PROC PRINT;
PROC GLM;
CLASS blk tech oper trt;
MODEL obs = blk tech(blk) oper(blk) trt/ss3;
MEANS trt;
LSMEANS trt oper(blk)/PDIFF LINES;
RUN;
ODS RTF CLOSE;
```

The results obtained from the analysis are given in Table 6.9.

Table 6.9: Results of analysis using SAS commands

ANOVA

Source	DF	SS	MS	F Value	Prob > F
Model	27	285.09	10.56	439.44	<0.0001
Error	8	0.19	0.02		
Corrected Total	35	285.28			

R-Square	CV	Root MSE	obs Mean
0.999	2.50	0.16	6.20

ANOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks	3	26.38	8.79	365.90	<0.0001
Technicians within blocks	8	0.40	0.05	2.06	0.1641
Operations within blocks	8	0.73	0.09	3.81	0.0380
Treatments	8	122.44	15.30	636.95	<0.0001
Error	8	0.19	0.02		
Corrected Total	35	285.28			

It may be observed that the model with nested terms explains almost 100 per cent variability in the data. The CV is also very low of the order of 2.50. The treatments and more importantly the blocks have been found to be highly significant (p-value < 0.0001). Further, the operations effect within blocks is significant but the technician effect within blocks is not significant (p-value = 0.1641).

The unadjusted and the adjusted (LS) treatment means are given in Table 6.10.

Table 6.10: Treatment wise unadjusted and least square means

Level of Treatment	N	Obs Mean		Level of Treatment	N	Obs Mean	
		Unadjusted Mean	LS Mean			Unadjusted Mean	LS Mean
1	4	1.875	2.150	6	4	7.075	7.167
2	4	3.125	3.133	7	4	8.200	8.150
3	4	4.175	4.017	8	4	9.500	9.533
4	4	5.275	5.250	9	4	10.150	9.917
5	4	6.425	6.483				

It may be seen that these means are quite different. The pairwise treatment comparisons are made and are presented in Table 6.11.

Table 6.11: *t* comparison lines for least squares means of treatments

LS-means with the same letter are not significantly different		
	obs LSMEAN	LSMEAN Number
A	9.917	9
B	9.533	8
C	8.150	7
D	7.167	6
E	6.483	5
F	5.250	4
G	4.017	3
H	3.133	2
I	2.150	1

It may be seen from Table 6.11 that all the treatments are pairwise significantly different from one another. However, treatment 9 is the best in terms of getting maximum response. Treatment 1 is the lowest yielding treatment.

Table 6.12: *t* comparison lines for least squares means of operation within block

Operation	Block	Observation LSMEAN	LSMEAN Number
1	1	5.200	1
2	1	5.167	2
3	1	5.567	3
1	2	6.217	4
2	2	6.417	5
3	2	6.633	6
1	3	5.317	7
2	3	5.733	8
3	3	5.650	9
1	4	7.417	10
2	4	7.350	11
3	4	7.733	12

Table 6.13: *t* comparison lines for least squares means of oper(blk)

		LS-means with the same letter are not significantly different			
		Observation LSMEAN	Operation	Block	LSMEAN Number
	A	7.733	3	4	12
B	A	7.417	1	4	10
B		7.350	2	4	11
	C	6.633	3	2	6
D	C	6.417	2	2	5
D		6.217	1	2	4
	E	5.733	2	3	8
F	E	5.650	3	3	9
F	E	5.567	3	1	3
F	G	5.317	1	3	7
	G	5.200	1	1	1
	G	5.167	2	1	2

The LS Means were also obtained for operations within blocks. The pairwise comparisons of the operation levels within each block were also made. These are summarized in the Tables 6.12 and 6.13. It may be seen that operation 3 within block 4 produces largest response, though statistically it is at par with operation 1 within block 4. It is significantly different from all other operations within blocks. The lowest observation is produced by operation 2 within block 1.

This is also statistically at par with operation 1 within block 1, and operation 1 within block 3. It is significantly lower than all other operations within blocks.

6.6.2 Analysis using R

In the sequel is given the R code for the analysis of data using R software. The results obtained are not reported to avoid duplication.

R code

```
d10=read.table("nrcd.txt",header=TRUE)
attach(d10)
names(d10)
#Treatment means and standard deviations
aggregate(obs~trt,data=d10,mean)
aggregate(obs~trt,data=d10,sd)
#Treatment wise box plot of calorie
boxplot(obs~trt)
#anova
trt=factor(trt)
tech=factor(tech)
oper=factor(oper)
blk=factor(blk)
lm1=lm(obs~trt+blk+tech/blk-tech+oper/blk-oper)
#anova(lm1)
library(car)
Anova(lm1,type="III")
library(lsmeans)
lsm=lsmeans(lm1,"trt")
lsm
pairs(lsm)
#to provide letters for groups, need to install multcompView
library(multcompView)
cld(lsm,Letters="abcdefghij")
#operator in block wise least square means and grouping
lsm2=lsmeans(lm1,~oper:blk)
lsm2
cld(lsm2,Letters="abcdefghij")
detach(d10)
```

6.7 Example 5

We continue once again with Example 4 in Section 6.6 and assume that there are 9 treatments to be allocated to 30 plots arranged in 4 blocks of sizes 6, 9, 9 and 6, respectively. Two blocks are complete blocks while remaining two blocks are incomplete blocks. A sample of harvested output from all the 30 plots has to be analyzed block wise using three different operations. The data collected is given in Table 6.14.

Table 6.14: Harvest data from thirty plots

Block I				Block II			
Operations	1	2		Operations	1	2	3
1	1 (1.1)	2 (2.1)		1	1 (2.1)	4 (5.2)	7 (8.3)
2	4 (4.2)	5 (5.3)		2	2 (3.2)	5 (6.7)	8 (9.9)
3	7 (7.4)	8 (8.7)		3	3 (4.5)	6 (7.6)	9 (10.3)

Block III				Block IV		
Operations	1	2	3	Operations	1	2
1	1 (1.2)	6 (6.3)	8 (8.7)	1	1 (3.1)	9 (11.3)
2	9 (9.4)	2 (2.7)	4 (4.8)	2	6 (8.1)	2 (4.5)
3	5 (5.9)	7 (7.8)	3 (3.3)	3	8 (10.7)	4 (6.9)

In the continuation of this example is presented the analysis of the data for identifying the best performing treatment.

6.7.1 Analysis

This design is a nested block design with unequal block sizes. Two bigger blocks are complete blocks with 9 treatments each while the other two bigger blocks are incomplete blocks with 6 treatments each. Since the bigger blocks are of unequal sizes, the sub-blocks within each block are also of unequal sizes. Three sub-blocks within each of the two incomplete bigger blocks have two treatments each while three sub-blocks within each of the two complete bigger blocks have three treatments each. The parameters of this design are $v = 9$, $b = 4$, block sizes as 6, 9, 9, 6 and sub-block sizes as 2, 2, 2, 3, 3, 3, 3, 3, 3, 2, 2, 2. We analyze the data using SAS. The commands are given below:

```
DATA nbib;
INPUT blk oper trt obs;
CARDS;
1    1    1    1.1
1    1    2    2.1
1    2    4    4.2
1    2    5    5.3
1    3    7    7.4
1    3    8    8.7
2    1    1    2.1
2    1    4    5.2
2    1    7    8.3
2    2    2    3.2
2    2    5    6.7
2    2    8    9.9
2    3    3    4.5
```

```
2       3       6       7.6
2       3       9       10.3
3       1       1       1.2
3       1       6       6.3
3       1       8       8.7
3       2       9       9.4
3       2       2       2.7
3       2       4       4.8
3       3       5       5.9
3       3       7       7.8
3       3       3       3.3
4       1       1       3.1
4       1       9       11.3
4       2       6       8.1
4       2       2       4.5
4       3       8       10.7
4       3       4       6.9
;
ODS RTF FILE='nestedincomplete.rtf';
PROC PRINT;
PROC GLM;
CLASS blk oper trt;
MODEL obs=blk oper(blk) trt/ss3;
MEANS trt;
LSMEANS trt oper(blk)/PDIFF LINES;
RUN;
ODS RTF CLOSE;
```

6.7.2 Output of analysis

The results obtained from the analysis are given in Table 6.15.

Table 6.15: Output of analysis from SAS commands

ANOVA

Source	DF	Type III SS	MS	F Value	Prob > F
Model	19	250.752	13.197	460.24	<0.0001
Error	10	0.287	0.0287		
Corrected Total	29	251.034			

R-Square	CV	Root MSE	Observation Mean
0.999	2.802	0.169	6.043

ANOVA

Source	DF	Type III SS	MS	F Value	Pr > F
Blocks	3	16.546	5.515	192.34	<0.0001
Operations within blocks	8	0.506	0.063	2.21	0.1201
Treatments	8	170.543	21.318	743.44	<0.0001
Error	10	0.287	0.0287		
Corrected Total	29	251.034			

It may be observed that the model with nested terms explains almost 100 per cent variability in the data. The CV is also very low of the order of 2.80. The treatments and more importantly the blocks have been found to be highly significant (p-value < 0.0001). However, the operations effect within blocks is not significant (p-value = 0.1201).

Table 6.16: Treatment wise unadjusted and least square means

Level of Treatment	N	Obs Mean	
		Unadjusted Mean	LS Mean
1	4	1.875	1.985
2	4	3.125	3.101
3	2	3.900	3.905
4	4	5.275	5.332
5	3	5.967	6.444

Level of Treatment	N	Obs Mean	
		Unadjusted Mean	LS Mean
6	3	7.333	6.985
7	3	7.833	8.299
8	4	9.500	9.487
9	3	10.333	9.891

It may be seen that the unadjusted means are quite different from the LS MEANS. The pairwise treatment comparisons are made and are presented in Table 6.17.

Table 6.17: *t* comparison lines for least squares means of treatments

LSMEANS with the same letter are not significantly different		
	observation LSMEAN	LSMEAN Number
A	9.891	9
B	9.487	8
C	8.299	7
D	6.985	6
E	6.444	5
F	5.332	4
G	3.905	3
H	3.101	2
I	1.985	1

It may be seen from comparisons in Table 6.17 that all the treatments are pairwise significantly different from one another. Treatment 9 is the highest yielding treatment while treatment 1 is the lowest yielding treatment.

The LSMEANS were also obtained for operations within blocks. The pairwise comparisons of the operation levels within each block were also made. These are summarized in the Tables 6.18 and 6.19.

Table 6.18: Least square mean for operation within block

Operation	Block	Observation LSMEAN	LSMEAN Number
1	1	5.216	1
2	1	5.021	2
3	1	5.315	3
1	2	6.154	4
2	2	6.415	5
3	2	6.698	6
1	3	5.406	7
2	3	5.684	8
3	3	5.610	9
1	4	7.421	10
2	4	7.415	11
3	4	7.549	12

It may be seen that operation 3 within block 4 produces largest observation, though statistically it is at par with operations 1 and 3 within block 4. It is significantly different from all other operations within blocks. The lowest observation is produced by operation 2 within block 1. This is also statistically at par with operation 1 within block 1, operation 3 within block 1 and operation 1 within block 3. It is significantly lower than all other operations within blocks.

Table 6.19: *t* comparison lines for least squares means of operation within block

LS-means with the same letter are not significantly different						
			Observation LSMEAN	Operation	Block	LSMEAN Number
	A		7.549	3	4	12
	A		7.421	1	4	10
	A		7.415	2	4	11
	B		6.698	3	2	6
C	B		6.415	2	2	5
C			6.154	1	2	4
	D		5.684	2	3	8

LS-means with the same letter are not significantly different						
			Observation LSMEAN	Operation	Block	LSMEAN Number
E	D		5.610	3	3	9
E	D	F	5.406	1	3	7
E	D	F	5.315	3	1	3
E		F	5.216	1	1	1
		F	5.021	2	1	2

6.7.3 Analysis using R

The R code for the analysis of data using R software is given in the sequence. To avoid duplication, the results obtained have not been included.

R code

```
d11=read.table("nbib.txt",header=TRUE)
attach(d11)
names(d11)
#Treatment means and standard deviations
aggregate(obs~trt,data=d11,mean)
aggregate(obs~trt,data=d11,sd)
#Treatment wise box plot of calorie
boxplot(obs~trt)
#anova
trt=factor(trt)
oper=factor(oper)
blk=factor(blk)
lm1=lm(obs~trt+blk+oper/blk-oper)
#anova(lm1)
library(car)
Anova(lm1,type="III")
library(lsmeans)
lsm=lsmeans(lm1,"trt")
lsm
pairs(lsm)
#to provide letters for groups, need to install multcompView
library(multcompView)
cld(lsm,Letters="abcdefghij")
#operator in block wise least square means and grouping
lsm2=lsmeans(lm1,~oper:blk)
lsm2
cld(lsm2,Letters="abcdefghij")
detach(d11)
```

7

Augmented Design

7.1 Introduction

So far the attention has been focused on designs in which the interest of the experimenter is to make all the possible pairwise treatment comparisons. However, this is not the case always. There exist situations where the treatments studied in an experiment comprise of two disjoint groups and each group of treatments has different importance to the researcher. The comparisons of treatments within groups may not be of interest or may be of less importance to the researcher while the comparisons of treatments between groups may be of interest or of more importance to the researcher. In other words, the interest of the experimenter is only in a subset of all the possible pairwise treatment comparisons. We illustrate this fact through examples studied in the literature.

Experimental Situation 1: (Federer, 1956). In a sugarcane breeding trial conducted at Hawaii, four sugarcane varieties *viz.*, *A, B, C,* or *D,* and eleven-seedling tests *e, f, g, h, i, j, k, l, m, n,* or *o,* were tried to evaluate the seedlings. Replicated plots on the individual seedlings were not possible because of the scarcity of seed cane and the large plots required for experimentation. One of the objectives of the trial was to make comparisons among the members of the two groups of sugarcane variety and seed cane. For obtaining the experimental error, sugarcane varieties were replicated. The trial was conducted in four blocks ($b = 4$) with seven plots in three blocks ($k_1 = k_2 = k_3 = 7$) and six plots in the fourth block ($k_4 = 6$). The layout of the design, without randomization, is the following:

B1	B2	B3	B4
A	*A*	*A*	*A*
B	*B*	*B*	*B*
C	C	C	C
D	D	D	D
e	h	k	n
f	i	l	o
g	j	m	

Experimental Situation 2: (Pearce, 1960). In a strawberry weed killer trial, it was intended to find out whether the application of any of the weed killers, *A, B, C,* or *D,* all of which were apparently suitable for controlling weeds in strawberry fields, would harm the growth of fruiting

strawberry plants. The trial was run in a design comprising of four blocks ($b = 4$) of seven plots each ($k = 7$) and there were four treatments comprising of four kinds of herbicides ($w = 4$) besides an untreated control (*O*). The layout, without randomization, is given as:

B1	B2	B3	B4
O	*O*	*O*	*O*
O	*O*	*O*	*O*
A	*A*	*A*	*A*
A	*B*	*B*	*B*
B	*C*	*B*	*C*
C	*D*	*C*	*C*
D	*D*	*D*	*D*

Interestingly, initially the design planned was randomized complete block (RCB) design with five test treatments (weed killers, *A, B, C, D,* or *E*) modified by adding two control replications in each block. But at last moment it was observed that supply of herbicide *E* still had not arrived and a decision had to be made quickly. It was thought of merging treatment *E* with any one of the treatments *A, B, C,* or *D,* in each block thus doubling the number of plots assigned to one of other substances. But sufficient supplies were not available for any of them (*i.e.* merging of treatments was not possible). Hence in desperation it was thought that in each of the four blocks treatment *E* will be exchanged with distinct treatments like *A, B, C,* and *D* in block I, II, III and IV, respectively. As a result of this a new class of designs, described above, was discovered.

Experimental Situation 3: In a trial, seven treatments (test treatments) were tried along with three controls. The purpose of the experiment was to make comparisons between the treatments in the two groups. The trial was laid out in ten blocks ($b = 10$), the first seven blocks having six plots each ($k_1 = k_2 = \cdots = k_7 = 6$) and the last three blocks having ten plots each ($k_8 = k_9 = k_{10} = 10$). The design adopted is the following:

B1	1	2	4	*A*	*B*	*C*				
B2	2	3	5	*A*	*B*	*C*				
B3	3	4	6	*A*	*B*	*C*				
B4	4	5	7	*A*	*B*	*C*				
B5	5	6	1	*A*	*B*	*C*				
B6	6	7	2	*A*	*B*	*C*				
B7	7	1	3	*A*	*B*	*C*				
B8	1	2	3	4	5	6	7	*A*	*B*	*C*
B9	1	2	3	4	5	6	7	*A*	*B*	*C*
B10	1	2	3	4	5	6	7	*A*	*B*	*C*

Experimental Situation 4: (Ture, 1994). A certain type of synthetic fiber is used in the production of various consumer goods. The research team of the manufacturer of this fiber has developed three new types of synthetic fibers that can be used for the same purpose. Each of these alternative fibers is more cost-efficient than the present one and can replace it if any one of them is proved to be stronger. An experiment was conducted to compare the breaking strengths of all these synthetic fibers. Suppose that 4 testing machines and 5 operators are available for the experiment. Because variability between the machines and the operators is suspected, the experiment must be designed to control such variability. Assuming that each operator can work on each testing machine once only, the following design may be used which is efficient for making tests *vs* controls comparisons.

Machine ↓	Operator →				
	A	**B**	**C**	**D**	**E**
I	0	3	0	1	2
II	1	2	0	0	3
III	3	1	2	0	0
IV	2	0	1	3	0

For the experimental settings considered here, all the possible pair-wise comparisons among treatments are not of equal interest to the researcher. In fact, the researcher is interested only in a subset of comparisons comprising of tests *vs* controls comparisons or pairwise comparisons among treatments belonging to the two groups. The comparisons among tests and among controls may be of little or no consequence to the researcher. For this experimental setting the variance-balanced designs for making all possible pair-wise treatment comparisons may not be useful.

Suppose that there are 6 treatments tried in an experiment *viz.*, A, B, C, D and 0, 1. If one is interested in making all the possible pairwise treatment comparisons, then there would be the following 15 comparisons: (A, B), (A, C), (A, D), (A, 0), (A, 1), (B, C), (B, D), (B, 0), (B, 1), (C, D), (C, 0), (C, 1), (D, 0), (D, 1), (0, 1). In these comparisons one may notice that every treatment appears 5 times in the 15 pairwise comparisons. So using any design described so far with equal or as far as possible equal replications will be an obvious choice. But suppose that the treatments A, B, C, D are tests (new treatments) and treatments 0, 1 are the controls (standard treatments or existing practices). The experimenter is not interested in making pairwise comparisons among treatments within groups. The interest is only in pairwise comparisons between the two groups. The following 8 pairwise comparisons of treatments are of interest now to the experimenter: (A, 0), (A, 1), (B, 0), (B, 1), (C, 0), (C, 1), (D, 0), (D, 1). Now it may be seen that treatments A, B, C, D appear twice in the comparisons but treatments 0, 1 appear 4 times each. Thus intuitively it is apparent that one needs to have designs with unequal replications of treatments with treatments 0, 1 replicated more times than the treatments A, B, C, D. Obviously then the designs with equal replications of treatments may not be good for these experimental situations.

The experimental situations described above can be classified in two broad categories *viz*. Category A: where it is not possible to replicate the test treatments (Experimental Situation 1) and Category B: where it is possible to have replication of test treatments as well (Experimental Situations 2,3, and 4). The designing and analysis of these experimental situations is described in the sequel.

7.2 Category A experiments with single replication of tests (Augmented designs)

Category-A designs are essentially augmented designs. In genetic resources environment, which is a field in the forefront of biological research, an essential activity is to test or evaluate the new germplasm / provenance / superior selections (test treatments), etc. with the existing provenance or released varieties (control treatments). A problem in these evaluation studies is that the quantity of the genetic material collected from the exploration trips is very limited or cannot be made available since a part of this is to be deposited in Gene Bank. The available quantity of seed is often not sufficient for replicated trials. Moreover, the number of new germplasm or provenance to be tested is very high (usually about 1000-2000 and sometimes up to 3000 accessions), and it is very difficult to maintain the within block homogeneity. These experimental situations may also occur in the fields of entomology, pathology, chemistry, physiology, microbiology, agronomy and perhaps others for screening experiments on new material and preliminary testing of experiments on promising material. In some other cases (*e.g.* physics), a single observation on new material may be desirable because of relatively low variability in the experimental material. These types of situations came to be known to Federer around 1955 in screening new strains of sugarcane and soil fumigants used in pineapples. Augmented (Hoonuiaku) designs were introduced by Federer (1956) to fill a need arising in screening new strains of sugarcane at Experimental Station of Hawaii Sugarcane Planters Association on the basis of agronomic characters other than yield.

Thus, we have seen that we have to design an experiment in which the experimental material for new (test) treatments is just enough for a single replication. However, the connectedness property of the design is ensured by augmenting any standard connected design in control treatments with new (test) treatments and replications of the control provide the estimate of error. More precisely, ***an augmented experimental design is any standard experimental design in standard (or control) treatments to which additional (new) treatments have been added***. The additional treatments require enlargement of the complete blocks or incomplete blocks in a block design set up or rows or / and columns in a row - column design set up, etc. The groupings (or blocks) in an augmented design may be of unequal sizes.

Augmented designs can be run in 0-way and 2-way elimination of heterogeneity settings also. Augmented designs eliminating heterogeneity in one direction are called augmented block designs and augmented designs eliminating heterogeneity in two directions are called augmented row-column designs. Federer (1956, 1961) gave the analysis, randomization procedure and construction of these designs by adding the new treatments to the blocks of RCB design and balanced lattice designs in control treatments. Federer (1963) gave procedures and designs useful for screening material inspection and allocation with a bibliography. Federer

and Raghavarao (1975), who obtained augmented designs using RCB design and linked block design for one-way heterogeneity setting, gave a general theory of augmented designs. They also gave a method of construction of augmented row-column design using a Youden Square design and also provided formulae for standard errors of estimable treatment contrasts. Federer *et al*. (1975) gave systematic methods of construction of augmented row column design. A procedure of analysis of data generated from these designs has also been given. The estimable contrasts in such designs may be (i) among new varieties (test treatments), (ii) among check varieties (control treatments), and (iii) among all check and new varieties simultaneously. Indeed it may be possible to estimate the contrasts between check and new varieties. We shall concentrate on augmented designs for 1-way elimination of heterogeneity settings. In general, the randomization procedure for an augmented block design is:

1. Follow the standard randomization procedure for the known design in control treatments or check varieties.

2. Test treatments or new varieties are randomly allotted to the remaining experimental units.

3. If a new treatment appears more than once, assign the different entries of the treatment to a block at random with the provision that no treatment appears more than once in a block until that treatment appears once in each of the blocks.

The analysis of variance of the data generated from an augmented block design with $v = u + w$ treatments comprising of w tests and u controls arranged in b blocks having k_1 plots (experimental units) in block 1, k_2 plots (experimental units) in block 2, and so on, and k_b plots (experimental units) in block b, such that $k_1 + k_2 + \cdots + k_b = n$, the total number of plots (experimental units) in the design, is sketched in Table 7.1.

Table 7.1: ANOVA table for augmented block design

Source	DF	SS	MS	F-value
Blocks (Eliminating treatments)	$b-1$	***ASSB***	*MSSB*	*MSSB/MSE*
Treatments (Eliminating blocks)	$v-1$	***ASST***		
Among Tests	$w-1$	*SST*	*MSST*	*MSST/MSE*
Among Controls	$u-1$	*SSC*	*MSSC*	*MSSC/MSE*
Tests vs Controls	1	*SSTC*	*MSSTC*	*MSSTC/MSE*
Error	$n-v-b+1$	***SSE***	*MSE*	
Corrected Total	$n-1$	***TSS***		

For making the exposition clear, we shall consider the Augmented Randomized Complete Block Design. Let us consider the experimental situation where w test treatments (*tth* test denoted by $N_t, t = 1,2,\cdots,w$) are to be compared with u control treatments (*sth* control denoted by $C_s, s = 1,2,\cdots,u$ via n experimental units arranged in b blocks such that *jth* block is of size $k_j (> u), j = 1,2,\cdots,b$. For an augmented randomized complete block design, each of the control treatments is replicated b times and occurs once in every block and test treatments occur only

once in any one of the b blocks. Therefore, it can be seen easily that in the jth block there will be $k_j - u = n_j$ test treatments, $j = 1, 2, \cdots, b$. The randomization procedure is given as follows:

1. Randomly allot u controls to u of the k_j experimental units in each block, $j=1,2,...,b$.
2. Randomly allot the w test treatments to the remaining experimental units.
3. If a new treatment appears more than once, assign the different entries of the treatment to a complete block at random with the provision that no treatment appears more than once in a complete block until that treatment occurs once in each of the complete blocks.

For augmented randomized complete block design standard errors for comparing mean differences are as follows

Estimated standard errors of the estimated difference

(i) Between two control treatments, $SE(1) = \sqrt{\dfrac{2MSE}{b}}$

(ii) Between two test treatments in the same block, $SE(2) = \sqrt{2MSE}$

(iii) Between two test treatments not in the same block, $SE(3) = \sqrt{2MSE(1+1/u)}$

(iv) Between a test treatment and a control treatment, $SE(4) = \sqrt{MSE(1+1/b+1/u-1/ub)}$

7.3 Example 1

An experiment was conducted with $w = 8$ new accessions (that were to be tested) denoted by N_t, $t = 1, \cdots, 8$ and $u = 4$ control treatments denoted by C_s, $s = 1, \cdots, 4$ of a genotype. There are 20 plots (experimental units) that could be arranged in three blocks ($b = 3$). There are 7 plots (4 for control treatments and 3 for new accessions) in the first and third block and 6 plots (4 for control treatments and 2 for new accessions) in the second block, *i.e.*, $k_1 = k_3 = 7$; $k_2 = 6$. For random allocation of these treatments in the experiment, we have to proceed as:

(i) Allot the 4 control treatments to each block randomly. In this process, say following is the arrangement:

Blocks   **Experimental units**

	1	2	3	4	5	6	7
1		C_3	C_4		C_1	C_2	
2	C_4	C_2	C_1	C_3			
3		C_3	C_1		C_4	C_2	

The 7th experimental unit is for blocks 1 and 3 and not for block 2. Of the 20 experimental units, 12 have been allotted to the control treatments. The remaining 8 will be allotted to the 8 new accessions.

8 new accessions are allotted randomly to the remaining experimental units of the 3 blocks. This way 4 controls and 8 new accessions randomly occupy 20 experimental units. The final arrangement looks as in Table 7.2.

Table 7.2: Data from an augmented block design

Blocks	Experimental units						
	1	2	3	4	5	6	7
1	N_8 (74)	C_3 (78)	C_4 (78)	N_3 (70)	C_1 (83)	C_2 (77)	N_7(75)
2	C_4 (91)	C_2 (81)	C_1 (79)	C_3 (81)	N_1 (79)	N_5 (78)	
3	N_4 (96)	C_3 (87)	C_1 (92)	N_2 (89)	C_4 (81)	C_2 (79)	N_6 (82)

The figures in the parenthesis represent the observed value of the character under study from an experiment conducted in the above layout. Source for this data is Federer (1956). The analysis of the data has been carried out and the ANOVA Table 7.3 is given.

Table 7.3: Analysis results of the data in Table 7.2

ANOVA

SOURCE	DF	SS	MS	F-value	Prob > F
Blocks(eliminating treatments)	2	69.500	34.750	1.29	0.3424
Treatments(eliminating blocks)	11	285.095	25.918	0.96	0.5499
Among Tests	7	215.169	30.738	1.14	0.4447
Among Controls	3	52.917	17.639	0.650	0.6092
Tests vs Controls	1	15.047	15.042	0.56	0.4834
Error	6	161.833	26.972		

R²	CV	Root MSE	Yield Mean
0.800	6.372	5.194	81.500

Estimated standard errors of the estimated difference

(i)   between two controls is 4.24.

(ii) between two tests in the same block is 7.34.

(iii) between two tests in different blocks is 8.21.

(iv) between a control and a test is 6.36.

We can use the adjusted values/means of the test treatments for comparison purpose. All those treatments for which yield levels are up to the satisfaction of breeder can be selected for further national level trials. In these kinds of experiments, generally, multiple characteristics

are observed. It may, therefore, be desirable to perform multivariate analysis of variance and use other related multivariate analytic techniques like cluster analysis, discriminant analysis, etc.

Keeping in view the importance of this design and for the ease of Biological Research Workers Agarwal and Sapra (1995) developed a user friendly program AUGMENT1 at the Documentation Unit of National Bureau of Plant Genetic Resources, New Delhi, to analyze the data of Augmented RCB design. It is interesting to note that the augmented RCB design is variance balanced with respect to tests *vs* controls comparisons.

7.4 Optimum replication of controls in a block

A survey of the literature reveals that generally the experiments described above are conducted using an augmented randomized complete block design. However, the experimenters would often like to know how many times the control treatments be replicated in each of the blocks so as to maximize the efficiency per observation for making test treatments *vs* control treatments(s) comparisons? An answer to this question was obtained by Parsad and Gupta (2000) and is described in the sequence.

Suppose that there are w test treatments which occur only once in the design and each of the u controls occurs in each of the b blocks, then to maximize the efficiency per observation the number of times each control appears in each of the blocks is

$$r = \frac{\sqrt{u+b-1}\sqrt{w}}{ub}$$

provided $b+u-1 \le w$. For example, consider the problem of obtaining the optimum number of

replications of the controls in an experiment with $w = 24$, $u = 3$, $b = 4$. We have $r = \frac{1}{3}\sqrt{\frac{6 \times 24}{4 \times 4}} = 1$.

Similarly, for $w = 98$, $u = 2$, $b = 7$, we have $r = \frac{1}{2}\sqrt{\frac{8 \times 98}{7 \times 7}} = 2$.

Remark 7.1 For a single control situation, *i.e.* $u = 1$, the above expression reduces to $r = \sqrt{\frac{w}{b}}$ and it can easily be seen that for $u = 1$, the condition $b+u-1 \le w$ becomes $b \le w$, which is always true.

There may, however, arise many combinations of w, u and b for which the above expression of r does not yield a positive integer value of r. In such situations, a question that arises is as to what integer value of r should be taken? To answer this question, the efficiency per observation has been calculated for $w \le 100$, $b \le 25$ and $u \le 10$ such that $b + u -1 \le w$ and r has been taken as r^* = $int(r)$ and $int(r) + 1$ besides taking $r = 1$. A close scrutiny reveals that if value of $r > $ #.42 then

take $r^* = int(r) + 1$ and for values of r smaller than or equal to #.42 take $r^* = int(r)$ for $u \geq 2$. For $u = 1$, the same rule applies but the value of r is taken as #.45 instead of #.42. Here # is the integral

part of $r = \dfrac{\sqrt{u + b - 1}\sqrt{w}}{ub}$.

7.5 Statistical package for augmented designs

A user friendly, menu driven, graphic user interface (GUI) based Statistical Package called "Statistical Package for Augmented Design" (SPAD) has been developed at IASRI, New Delhi. The package generates randomized layout of augmented designs and performs the analysis of data generated. For given number of test treatments, number of control treatments and number of blocks, it computes the optimum replication number of each control treatment in every block of the design such that the efficiency per observation of the test treatments *vs* control treatment(s) comparisons is maximum. The user may choose the optimum replication number. However, the package provides flexibility in choosing the replication number of each control treatment in every block. The user can choose the replication of each control treatment in every block according the resources available. It also asks the user to give the block sizes. One can have unequal block sizes as well. Once the user defines the number of test treatments, number of control treatments, and number of blocks in the design, the randomized layout of the design is generated. The package also provides the analysis of the data generated from augmented designs. A null hypothesis on any user-defined contrast can also be tested. This software is available at Design Resources Server. The URL is http://www.iasri.res.in/design/AugmentedDesigns/home.htm.

7.6 Example 2

An experiment was conducted at Directorate of Wheat Research during 2002-03 to compare 54 new accessions with 4 check varieties to see whether any of the check varieties can be replaced by any of the new accessions. The experiment was conducted using an augmented randomized complete block design with 6 blocks each of size 13 such that each of 4 check varieties are allocated in each of the six blocks and accessions are allocated only once in the design. The data on (i) days to 75% SE, (ii) FLL in centimeters and (iii) 1000 grain weight in grams is given in Table 7.4.

Table 7.4: Experiment data from an augmented RCB design

Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)		Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)
IC-028532	1	85	21.8	36.7		IC-079026	4	85	22.8	28.6
IC-028661	1	88	22.98	31.6		**C-3**	4	86	19.8	33.1
IC-028696	1	85	21	22.7		IC-079008	4	86	26.8	25.4

Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)		Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)
IC-028741	1	85	22.4	30.2		IC-079027	4	86	23.1	19.9
C-4	1	84	23.3	36		C-1	4	87	17.5	28.2
IC-028764	1	81	22.2	31.4		IC-079034	4	84	26.1	20.9
IC-028794	1	82	22.6	29.5		IC-079037	4	88	27.3	25.5
IC-028835	1	85	22.22	28.3		IC-079047	4	82	23.8	27.9
C-1	1	86	19.34	27.2		C-4	4	85	23.3	34.9
IC-028843	1	85	20.8	34.7		IC-079048	4	92	23.7	27
IC-028847	1	85	21	33.4		IC-079050	4	87	22.8	22.8
C-2	1	86	22.8	29.6		IC-079007	4	88	25.3	25.5
C-3	1	88	22.88	24.4		C-2	4	87	28.7	31.2
IC-036882	2	85	23.9	25.7		C-3	5	87	15.2	30.9
C-1	2	88	20.2	29.3		IC-082330	5	85	20.7	18.3
IC-036875	2	87	22.7	23.6		IC-082335	5	90	21.8	31.2
IC-042408	2	79	24.3	35.9		IC-082336	5	85	22.2	29
C-3	2	86	13.7	37.9		IC-082338	5	88	19.2	27.9
IC-036885	2	84	23.34	16.6		C-1	5	86	15.5	34.4
IC-041405	2	92	24.9	24.9		IC-082343	5	88	19.9	23.5
C-4	2	82	23.7	35.8		IC-082351	5	90	19.6	27.5
IC-036884	2	85	28.4	28.3		C-2	5	85	23.7	36.3
IC-042458	2	80	25.1	28.7		IC-082352	5	85	20.8	27.9
IC-036871	2	89	25.9	24.9		IC-082362	5	84	14.6	28.9
C-2	2	81	23.1	38.1		C-4	5	86	20.7	36.9
IC-042343	2	83	25.5	26.1		IC-082326	5	83	21.2	18.5
C-4	3	88	26.9	35.9		IC-104612	6	83	26.4	39.9
IC-060221	3	83	22.24	33.5		IC-104601	6	87	19.7	36.5
IC-073491	3	82	25.36	34.6		C-4	6	85	18.8	30.1
IC-063947	3	82	21.6	19.9		C-2	6	87	22	36.5
IC-066518	3	85	21.04	26.9		IC-104609	6	87	20.4	32.5
IC-073214	3	81	22.9	36.9		IC-104607	6	85	21.4	39.2
C-1	3	88	17.4	33.5		IC-104611	6	87	21.4	34.4
C-2	3	85	23.6	24.8		C-3	6	86	17.7	36.5
IC-073207	3	90	21.32	21.4		IC-104613	6	86	21.7	24.3
IC-073493	3	86	21.6	18.9		IC-104614	6	84	21.3	34.6
IC-060218	3	82	19.9	23		C-1	6	87	18.4	31.1

Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)		Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)
C-3	3	88	18.2	21.6		IC-104610	6	88	22.8	29.2
IC-060997	3	83	20.52	22.9		IC-104604	6	87	23.1	24.8

In what follows, the data are analyzed (a) to test the homogeneity of all the 58 treatment effects, (b) to test the equality of (i) all check varieties, (ii) all new accessions, and (iii) all new accessions with all check varieties, (c) to make all possible pairwise comparisons of check varieties with each of the new accessions to identify the best performing accession.

Remark 7.2 Using the online package, the following augmented design may be generated. It is indeed possible to generate another randomized layout of this design. The optimum replication of control in each block works out to one in this case. **C#** denotes the label of the control treatment and T# denotes the label of the new accession.

B1: (T4, T24, T46, T13, T17, **C4**, T34, T23, **C1**, T3, **C2**, **C3**, T16)

B2: (T12, T1, **C4**, **C3**, T54, **C1**, T10, T49, T37, **C2**, T43, T25, T30)

B3: (T21, T50, **C3**, T26, T47, **C2**, T48, **C4**, T20, **C1**, T40, T28, T29)

B4: (**C2**, T53, T22, T39, T5, T52, T11, **C3**, T18, **C4**, **C1**, T6, T42)

B5: (T14, T27, **C4**, T31, T9, **C1**, T45, **C3**, **C2**, T32, T36, T38, T2)

B6: (**C4**, **C2**, T8, T35, **C3**, T41, T15, **C1**, T7, T51, T44, T33, T19)

Remark 7.3 For preparing the data file for SAS, the treatment labels have to be given as numerals 1,2,3, ... For the Example 2 in Section 7.6, the numerals 1, 2, 3, 4 denote the controls (or check varieties) and the numerals 5, 6, 7, 8, . . . , 54, 55, 56, 57, 58 denote the 54 tests (or new accessions). While writing down the contrasts also, this has to be borne in mind.

7.6.1 Analysis of data

The parameters of the augmented design are given as:

Number of tests, $w = 54$; Number of controls, $u = 4$; Block sizes, $k_1 = k_2 = k_3 = k_4 = k_5 = k_6 = 13$

Replication of controls, $r_0 = 6$, Replication of tests, $r = 1$, Total number of observations, $n = 78$. The data has been analyzed using SAS software. The commands and the data preparation are given in the sequel.

```
DATA augmented;
INPUT block  trt  SE FLL GW;
CARDS;
```

1	1	86	19.34	27.2
1	2	86	22.8	29.6
1	3	88	22.88	24.4
1	4	84	23.3	36.0
1	5	85	21.8	36.7
1	6	88	22.98	31.6
1	7	85	21	22.7
1	8	85	22.4	30.2
1	9	81	22.2	31.4
1	10	82	22.6	29.5
1	11	85	22.22	28.3
1	12	85	20.8	34.7
1	13	85	21	33.4
2	1	88	20.2	29.3
2	2	81	23.1	38.1
2	3	86	13.7	37.9
2	4	82	23.7	35.8
2	14	85	23.9	25.7
2	15	87	22.7	23.6
2	16	79	24.3	35.9
2	17	84	23.34	16.6
2	18	92	24.9	24.9
2	19	85	28.4	28.3
2	20	80	25.1	28.7
2	21	89	25.9	24.9
2	22	83	25.5	26.1
3	1	88	17.4	33.5
3	2	85	23.6	24.8
3	3	88	18.2	21.6
3	4	88	26.9	35.9
3	23	83	22.24	33.5
3	24	82	25.36	34.6
3	25	82	21.6	19.9
3	26	85	21.04	26.9
3	27	81	22.9	36.9
3	28	90	21.32	21.4
3	29	86	21.6	18.9
3	30	82	19.9	23.0
3	31	83	20.52	22.9
4	1	87	17.5	28.2
4	2	87	28.7	31.2
4	3	86	19.8	33.1
4	4	85	23.3	34.9

4	32	85	22.8	28.6
4	33	86	26.8	25.4
4	34	86	23.1	19.9
4	35	84	26.1	20.9
4	36	88	27.3	25.5
4	37	82	23.8	27.9
4	38	92	23.7	27.0
4	39	87	22.8	22.8
4	40	88	25.3	25.5
5	1	86	15.5	34.4
5	2	85	23.7	36.3
5	3	87	15.2	30.9
5	4	86	20.7	36.9
5	41	85	20.7	18.3
5	42	90	21.8	31.2
5	43	85	22.2	29.0
5	44	88	19.2	27.9
5	45	88	19.9	23.5
5	46	90	19.6	27.5
5	47	85	20.8	27.9
5	48	84	14.6	28.9
5	49	83	21.2	18.5
6	1	87	18.4	31.1
6	2	87	22	36.5
6	3	86	17.7	36.5
6	4	85	18.8	30.1
6	50	83	26.4	39.9
6	51	87	19.7	36.5
6	52	87	20.4	32.5
6	53	85	21.4	39.2
6	54	87	21.4	34.4
6	55	86	21.7	24.3
6	56	84	21.3	34.6
6	57	88	22.8	29.2
6	58	87	23.1	24.8

```
;
ODS RTF FILE='Model1.rtf';
PROC GLM;
CLASS trt block;
MODEL SE FLL GW = trt block;
LSMEANS trt/PDIFF LINES;
```

CONTRAST 'Among New Accessions' trt 0 0 0 0 1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 -2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 -3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 -5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 -6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 -7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 -8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 -9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 -10 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 -11 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 -12 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 -13 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -14 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -16 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -17 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -18 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -19 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -20 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -21 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -22 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -23 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -24 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -25 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -26 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -27 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -28 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -29 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -30 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -31 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -33 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -34 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -35 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -36 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -37 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -38 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -39 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -40 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -41 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -42 0 0 0 0 0 0 0 0 0 0 0 0 0,

```
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -43 0 0
0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -44 0
0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -45
0 0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -46
0 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
-47 0 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 -48 0 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 -49 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 -50 0 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 -51 0 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 -52 0,
trt 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 -53;
CONTRAST 'Among Controls'
trt 1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0,
trt 1 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0,
trt 1 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0;
CONTRAST 'Control vs New Accessions'
trt  -27 -27 -27 -27 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2;
RUN;
ODS RTF CLOSE;
```

Note: It may appear difficult to generate the contrasts for each experimental situation. Therefore, a SAS Macro has been developed where user has only to enter the data file and variable names and with that information all other steps are generated automatically. This macro is available at http://www.iasri.res.in/sscnars/augblkdsgn.aspx.

7.6.2 Output of analysis

The results obtained from the analysis are given in Table 7.5.

Table 7.5: Results for the character SE

ANOVA for the character "Days to 75 % SE"

Source	DF	Type III SS	MS	F-Value	Prob > F
Treatments	57	432.564	7.5889	3.28	0.0069
Among New Accessions	53	405.251	7.646	3.31	0.0068
Among Controls	3	20.333	6.778	2.93	0.0676
Controls vs New Accessions	1	6.980	6.980	3.02	0.1027
Blocks	5	19.000	3.800	1.64	0.2087
Error	15	34.667	2.311		
Corrected Total	77	507.295			

R-Square	CV	Root MSE	SE Mean
0.932	1.777	1.520	85.551

It may be noted that the model explains about 93 percent of the total variability in the data pertaining to "Days to 75 % SE." The CV is also very low (CV = 1.78). The treatment effects are significantly different (p-value = 0.0069), but the block effects are not significant. The effect of new accessions is also significantly different (p-value = 0.0068), but the effects of controls and new accessions *vs* controls are not significantly different.

The pairwise treatment comparisons using LS MEANS produce Table 7.6.

Table 7.6: *t* comparison lines for least squares means of treatments

LS-means with the same letter are not significantly different								SE LSMEAN	LSMEAN of Treatment Number
				A				93.750	18
	B			A				91.750	38
	B			A		C		90.750	21
	B	D		A		C		90.000	42
	B	D		A		C		90.000	46

								SE LSMEAN	LSMEAN of Treatment Number
E	B	D		A		C		88.750	15
E	B	D		A		C		88.750	28
E	B	D				C		88.000	45
E	B	D				C		88.000	44
E	B	D		F		C		88.000	6
E	B	D		F		C		87.750	57
E	B	D		F		C	G	87.750	40
E	B	D		F		C	G	87.750	36
E	B	D		F		C	G	87.000	1
E	B	D		F		C	G	86.833	3
E	B	D		F	H	C	G	86.750	19
E	B	D	I	F	H	C	G	86.750	39
E	B	D	I	F	H	C	G	86.750	51
E	B	D	I	F	H	C	G	86.750	52
E	B	D	I	F	H	C	G	86.750	14
E	B	D	I	F	H	C	G	86.750	54
E	B	D	I	F	H	C	G	86.750	58
E	J	D	I	F	H	C	G	85.750	17
E	J	D	I	F	H	C	G	85.750	33
E	J	D	I	F	H	C	G	85.750	55
E	J	D	I	F	H	C	G	85.750	34
E	J	D	I	F	H		G	85.167	2
E	J	D	I	F	H		G	85.000	5
E	J	D	I	F	H		G	85.000	11
E	J	D	I	F	H		G	85.000	13
E	J	D	I	F	H		G	85.000	47
E	J	D	I	F	H		G	85.000	8
E	J	D	I	F	H		G	85.000	41
E	J	D	I	F	H		G	85.000	4
E	J	D	I	F	H		G	85.000	43
E	J	D	I	F	H		G	85.000	12
E	J	D	I	F	H		G	85.000	7
E	J		I	F	H	K	G	84.750	32
E	J		I	F	H	K	G	84.750	53

LS-means with the same letter are not significantly different

								SE LSMEAN	LSMEAN of Treatment Number
E	J		I	F	H	K	G	84.750	22
E	J		I	F	H	K	G	84.750	29
E	J		I	F	H	K	G	84.000	48
E	J		I	F	H	K	G	83.750	35
E	J		I	F	H	K	G	83.750	56
E	J		I	F	H	K	G	83.750	26
	J		I	F	H	K	G	83.000	49
	J		I		H	K	G	82.750	50
	J		I		H	K		82.000	10
	J		I		H	K		81.750	31
	J		I		H	K		81.750	37
	J		I		H	K		81.750	23
	J		I			K		81.750	20
	J					K		81.000	9
	J					K		80.750	16
	J					K		80.750	25
	J					K		80.750	24
	J					K		80.750	30
						K		79.750	27

The LINES display does not reflect all significant comparisons. The following additional pairs are significantly different: (18, 15) (38, 1) (38, 3) (38, 39) (21, 3) (21, 17) (42, 2) (42, 47) (42, 41) (42, 4) (42, 43) (46, 2) (46, 47) (46, 41) (46, 4) (46, 43) (28, 26) (1, 4) (1, 49) (1, 50) (3, 49) (3, 50) (39, 37) (14, 20) (17, 16) (2, 9) (2, 16) (2, 25) (2, 24) (2, 30) (4, 9) (4, 16) (4, 25) (4, 24) (4, 30) (29, 27)

Note: While interpreting the results care need to be taken to convert the treatment numbers back to new accessions and control varieties. The control varieties are labeled 1 – 4 and the new accessions are labeled 5 – 58. This means that treatment number 5 is actually new strain 1; treatment number 58 is actually new strain 54, and so on.

The estimated standard errors of various estimated comparisons can be obtained by using the online portal "Strengthening Statistical Computing for NARS" at www.iasri.res.in/sscnars/. The estimated standard errors are given below:

Estimated standard errors of the estimated difference

(i)   between two controls is 0.878 and Tukey's HSD at 5 % is 6.172.

(ii)   between two new accessions in the same block is 2.150 and Tukey's HSD at 5 % is 15.119.

(iii)   between two new accessions in different blocks is 2.404 and Tukey's HSD at 5 % is 16.904.

(iv)   between a control and a new accession is 1.783 and Tukey's HSD at 5 % is 12.536.

Table 7.7: Results for the character FLL

ANOVA for the character "FLL (cm)"

Source	DF	Type III SS	MS	F-Value	Prob > F
Treatments	57	425.265	7.461	1.26	0.3196
Among New Accessions	53	188.509	3.557	0.60	0.9116
Among Controls	3	179.234	59.745	10.10	0.0007
Control vs New Accessions	1	57.523	57.523	9.73	0.0070
Blocks	5	45.524	9.105	1.54	0.2366
Error	15	88.698	5.913		
Corrected Total	77	672.516			

R-Square	CV	Root MSE	SE Mean
0.868	11.067	2.432	21.972

It may be noted that the model explains about 87 percent of the total variability in the data pertaining to "FLL." The CV is little high compared to the one obtained for the character "days to 75 % SE." (CV = 11.067). The treatment effects are not significantly different (*p*-value = 0.3196); similarly the block effects are also not significant (*p*-value = 0.2366). The effect of new accessions is also not significantly different (*p*-value = 0.9116), but the effects of controls and controls vs new accessions are highly significant (*p*-values = 0.0007 and 0.0070, respectively).

Estimated standard errors of the estimated difference

i)   between two controls is 1.404.

(ii)   between two new accessions in the same block is 3.434.

(iii)   between two new accessions in different blocks is 3.845.

(iv)   between a control and a new accession is 2.851.

Note: While interpreting the results care need to be taken to convert the treatment numbers back to new accessions and control varieties. The control varieties are number 1 – 4 and the new accessions are numbered 5 – 58. This means that treatment number 5 is actually new strain 1; treatment number 58 is actually new strain 54, and so on.

Table 7.8: Results for the character 1000 grain weight

ANOVA for the character "1000 grain weight (gm)"

Source	DF	Type III SS	MS	F-Value	Prob > F
Treatments	57	1907.634	33.467	1.85	0.0946
Among New Accessions	53	1507.241	28.439	1.57	0.1694
Among Controls	3	74.508	24.836	1.37	0.2899
Controls vs New Accessions	1	325.884	325.884	17.98	0.0007
Blocks	5	144.933	28.987	1.60	0.2202
Error	15	271.817	18.121		
Corrected Total	77	2512.795			

R-Square	CV	Root MSE	GW Mean
0.892	14.582	4.257	29.192

It may be noted that the model explains about 89 percent of the total variability in the data pertaining to "1000 grain weight." The CV is little high compared to the one obtained for the character "days to 75 % SE." (CV = 14.582). The treatment effects are not significantly different (p-value = 0.0946); similarly the block effects are also not significant (p-value = 0.2202). The effect of new accessions is also not significantly different (p-value = 0.1694), similarly, the effect of controls is also not significantly different (p-value = 0.2899). However, the effect of controls vs new accessions is highly significant (p-value = 0.0007).

Estimated standard errors of the estimated difference

(i)   between two controls is 2.458.

(ii)   between two new accessions in the same block is 6.020.

(iii)   between two new accessions in different blocks is 6.731.

(iv)   between a control and a new accession is 4.992.

7.6.3 Analysis using R

In the sequence are give the R code for analysis of data generated from an augmented design. The results obtained are not given to avoid duplication.

R code

```
d12=read.table("augmented.txt",header=TRUE)
attach(d12)
names(d12)
#anova
trt=factor(trt)
block=factor(block)
```

```
lm1=lm(SE~trt+block)
#anova(lm1)
library(car)
Anova(lm1,type="III")
library(lsmeans)
lsm=lsmeans(lm1,"trt")
#to provide letters for groups, install and then load multcompView
library(multcompView)
cld(lsm,Letters="abcdefghij")
#generating the contrasts, trts 1-4 are the controls and 5-58 are new accessions
contrast.mat1=matrix(0,53,58)
for (i in 1:53)
{
  contrast.mat1[i,(5:(4+i))]=1
  contrast.mat1[i,(5+i)]=-i
}
contrast.mat2=matrix(0,3,58)
for (i in 1:3)
{
  contrast.mat2[i,1]=1
  contrast.mat2[i,(1+i)]=-1
}
controls.vs.newaccessions=contrast(lsm,list(con1=c(rep(-27,4),rep(2,54))))
Among.New.Accessions=contrast(lsm,list(apply(contrast.mat1,1,list)))
Among.controls=contrast(lsm,list(apply(contrast.mat2,1,list)))
lht(lm1,Among.New.Accessions@linfct)
lht(lm1,Among.controls@linfct)
lht(lm1,controls.vs.newaccessions@linfct)
lm2=lm(FLL~trt+block)
Anova(lm2,type="III")
lsm2=lsmeans(lm2,"trt")
controls.vs.newaccessions=contrast(lsm2,list(con1=c(rep(-27,4),rep(2,54))))
Among.New.Accessions=contrast(lsm2,list(apply(contrast.mat1,1,list)))
Among.controls=contrast(lsm2,list(apply(contrast.mat2,1,list)))
lht(lm2,Among.New.Accessions@linfct)
lht(lm2,Among.controls@linfct)
lht(lm2,controls.vs.newaccessions@linfct)
cld(lsm2,Letters="abcdefghij")
lm3=lm(GW~trt+block)
Anova(lm3,type="III")
lsm3=lsmeans(lm3,"trt")
controls.vs.newaccessions=contrast(lsm3,list(con1=c(rep(-27,4),rep(2,54))))
Among.New.Accessions=contrast(lsm3,list(apply(contrast.mat1,1,list)))
```

```
Among.controls=contrast(lsm3,list(apply(contrast.mat2,1,list)))
lht(lm3,Among.New.Accessions@linfct)
lht(lm3,Among.controls@linfct)
lht(lm3,controls.vs.newaccessions@linfct)
cld(lsm3,Letters="abcdefghij")
detach(d12 )
```

Remark 7.4 This Chapter has focused essentially on augmented design or Category A designs in which the test treatments have single replication. Category B designs are the ones in which both the test treatments and control treatments are replicated. In Category B a standard design (RCB design, BIB design, Latin Square, nested, etc) in test treatments is supplemented with the additional control treatments. Generally all the controls appear together. The analysis of such designs has been described at several places in the book. It is for this reason that the analysis of these designs has not been discussed separately in this Chapter. For instance in Section 2.3.2 in Chapter 2, the Example described is actually a Category B design. The analysis steps are almost same as category A experiments or augmented randomized complete block designs except that formulae for standard error of pairwise comparisons for tests, controls and test *vs* controls will change as per the design adopted.

8

Combined Analysis of Groups of Experiments

8.1   Introduction

We have so far studied design and analysis of experiments for single factor with many levels conducted at one place or in one season or in one lab. There could, however, be situations where a single factor experiment may be conducted over a number of locations (or places) and / or a number of seasons (or years) or in different labs. In large scale experimental programmes, it is necessary to repeat the trial of a set of treatments like varieties or manures at a number of places or in a number of seasons. For example, in a crop improvement programme, the trial may be run over different centres or locations. Similarly, a crop sequence programme may be run over seasons or years. The places where the trial is repeated are usually experimental stations located in the tract. The main objective of running the experiment over different time periods or over different locations or both is to study the performance of treatment effects over different locations or different time periods. Henceforth, we shall talk of experiments being repeated over different environments in place of writing that the experiments are being repeated over locations or seasons or both. More generally, the purpose of repetition is to find out treatments suitable for particular environment in which case the trials are carried out simultaneously on a representative selection of environments.

The purpose of the research carried out at experimental stations is also to formulate recommendations for the practitioners, which consist of a population quite extensive either over time or space or both. Therefore, it becomes imperative to ensure that the results obtained from researches are valid for at least several places in the future and over reasonably heterogeneous environments.

A single experiment will precisely furnish information about only one place where the experiment is conducted and about the season in which the experiment is conducted. It has, thus, become a common practice among the agricultural research trials to repeat an experiment over different environments to obtain valid recommendations taking into account the environment to environment variation. It also allows the experimenter to study the interaction of the treatments with environment. In such cases where the experiment is repeated over space or time, appropriate statistical methodologies would have to be followed for the combined analysis of data obtained from individual experiments. In combined analysis of data, the interest of the experimenter is to obtain answers to the following questions:

(a)   What is the estimate of the average response to given set of treatments?

(b)   Is there consistency of the responses from environment to environment, i.e., is the interaction of the treatments with environments present or absent?

The utility and the significance of the estimates of average response depend on whether the response is consistent from environment to environment or changes with it; in other words, it depends on the presence or absence of the interaction between treatments and environments.

The results of a set of trials may, therefore, be considered as belonging to one of the following four types:

a) the error variances pertaining to individual experiments are homogeneous and the interaction is absent;

b) the error variances pertaining to individual experiments are homogeneous and the interaction is present;

c) the error variances pertaining to individual experiments are heterogeneous and the interaction is absent;

d) the error variances pertaining to individual experiments are heterogeneous and the interaction is present.

Remark 8.1: The meaningfulness of average estimates of treatment responses would, therefore, depend largely upon the absence or presence of the interaction. If treatment × environment interaction is present then first identify whether the interaction is a cross-over (treatment ranks changes from one environment to another) or non-cross-over type where treatment differences change in magnitude but not in direction from one environment to another. In non-cross-over interaction, the treatments with superior means (or adjusted means) can be used in all the environments. If there is cross-over interaction, then the subsets of treatments are to be recommended only for certain environments. One way to identify the sub-sets of treatments for certain environments is to use the technique of biplots (generally Site Regression biplots).

In the sequel, we describe the analysis procedure for group of experiments conducted in different environments as a block design.

8.2 Analysis procedure

The model for the individual experiments would be the same as described in the earlier chapters. However, for the combined analysis of data from a block design, the model would be

response = general mean + environment effect + block effect(environment) + treatment effect + interaction between treatments and environments + error.

This model can alternatively be written as

$$y_{uij} = \mu + \pi_u + \tau_i + \beta_j + (\pi\tau)_{ui} + e_{uij}$$

where y_{uij} is the response to the ith treatment in the jth block on the uth environment, μ is the general mean, π_u is the effect of the uth environment, τ_i is the effect of the ith treatment, β_j is the effect of the jth block, $(\pi\tau)_{ui}$ is the interaction effect of the uth environment and the ith treatment and e_{uij} is the random error component associated with the observation y_{uij} and

follows a normal distribution with mean zero and variance σ_u^2. The e_{uij} 's are assumed to be independently distributed. Here $i = 1,2,\cdots, v$ $j = 1,2,\cdots, b$ $u = 1,2,\cdots, p$

The model has been described for combined analysis of data generated from block designs conducted across environments. Similar procedure can be used for other designs by modifying the model statement as per the design adopted. For example for a row-column design the model would be

response = general mean + environment effect + rows (environment) effect+ columns (environment) effect + treatment effect + interaction between treatments and environments + error.

For a resolvable block design, it can be written as

response = general mean + environment effect + replications (environments) effect+ blocks (replication, environment) effect + treatment effect + interaction between treatments and environments + error.

It may be noted here that treatment effect include treatment effect in case of single factor experiments and treatment combination effect in case of multi-factor experiments. In case of multi-factor experiments, one may include main effects and interaction effects in the model in place of treatment effect. In that case, interaction between treatments and environments may also be replaced with interaction of environments with main effects and interaction effects.

For the combined analysis of data, the following steps need to be followed:

Step 1. Construct an outline of combined analysis of variance over environments (over years or over locations or over artificially created environments), based on the basic design used.

Step 2. Perform the usual analysis of variance for the individual experiment (environment wise) depending upon the design adopted.

Step 3. Test the homogeneity of the error variances using the error mean squares obtained from the individual experiments. Suppose that there are v treatments in each experiment and the experiment is repeated over p environments. So there would be p error variances and it would be required to test the homogeneity of these p error variances using the p error mean squares. Let $\sigma_1^2, \cdots, \sigma_u^2, \cdots, \sigma_p^2$ be the error variances for the respective p environments. Further, $s_{e_1}^2, \cdots, s_{e_u}^2, \cdots, s_{e_p}^2$ let be the error mean squares obtained from the p experiments, with respective degrees of freedom as $n_1, \cdots, n_u, \cdots, n_p$. For $u = 1, 2, \ldots, p$, $s_{e_u}^2$ is an unbiased estimator of σ_u^2. We may have the following two cases:

Case 1: $p = 2$.

In this case, use Snedecor's F distribution to test the equality of two variances. In this case, the null and alternate hypotheses are H_0: $\sigma_1^2 = \sigma_2^2$ and H_1: $\sigma_1^2 \neq \sigma_2^2$, respectively. The test statistic is $F = \dfrac{s_{e_1}^2}{s_{e_2}^2}$. The computed value of F will be compared with the table value of Snedecor's F-distribution with n_1 and n_2 degrees of freedom and α level of significance. If the computed

value of F is greater than the table value of $F_{\alpha/2;n_1,n_2}$, then the null hypothesis of homogeneity of variances is rejected and it is concluded that the data are heterogeneous on the two environments; otherwise it is homogeneous.

Case 2. $p > 2$. In this case use Bartlett's chi-square to test the homogeneity of error variances. In this case, the null hypothesis and the alternate hypotheses are $H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_p^2$ and H_1: at least two of the σ_u^2 are not equal, respectively. The test statistic is

$$\chi_{p-1}^2 = \frac{\sum_{i=1}^p n_i \log \bar{s}_e^2 - \sum_{i=1}^p n_i \log s_{e_i}^2}{1 + \frac{1}{3(p-1)}\left(\sum_{i=1}^p \frac{1}{n_i} - \frac{1}{\sum_{i=1}^p n_i}\right)}, \text{ where } \bar{s}_e^2 = \frac{\sum_{i=1}^p n_i s_{e_i}^2}{\sum_{i=1}^p n_i}$$

Further, if $n_1 = \cdots = n_u = \cdots = n_p = n$, then

$$\chi_{p-1}^2 = \frac{n\left[p\log \bar{s}_e^2 - \sum_{u=1}^p \log s_{e_u}^2\right]}{1 + \frac{(p+1)}{3np}} \text{ and } \bar{s}_e^2 = \frac{\sum_{u=1}^p s_{e_u}^2}{p}.$$

Here χ_{p-1}^2 follows a χ^2-distribution with $p-1$ degrees of freedom. The computed value of χ^2 will be compared with the table value of χ^2-distribution with $p-1$ degrees of freedom and α level of significance. If the computed value of χ^2 is greater than the table value of $\chi_{\alpha/2;p-1}^2$, then the null hypothesis of homogeneity of error variances is rejected and it is concluded that the data are heterogeneous over different environments or different years; otherwise it is homogeneous.

Step 4. If error variances are heterogeneous, then for performing the combined analysis, one has to use weighted least squares. One choice of weights could be the reciprocals of the square root of error mean square of the individual experiments. The weighted analysis is carried

out by defining a new response variable as $z_{uij} = \frac{y_{uij}}{\sqrt{s_{e_u}^2}}\{i.e., (\text{newresvar})_u = \frac{\text{origresvar}}{\sqrt{s_{e_u}^2}}, u = 1,2,\cdots,p\}$. This

new variable is homogeneous and combined analysis can be performed on the transformed variable. If, however, the error variances are homogeneous, then no transformation is required.

Step 5: The group of experiments can now be viewed as a nested design with environments as the bigger blocks and the individual experiments nested within each bigger block. The replication wise data for each treatment on each environment provides useful information and one could work out the interaction of treatments with environments. There is a tendency among the experimenters to take the average of the replicated data for each treatment and environment. By doing so, one can't get the interaction.

Step 6: The next step in the analysis is to test for the significance of the treatments × environments interaction to see if the treatment effects differ or not from environment to environment. The significance of the interaction between treatments and environments is tested by comparing the mean square of the interaction with the error mean square of the combined analysis of variance table using *F*-statistic. If the interaction effect is not found to be significant, it means the interaction is absent. When the interactions are absent then the model gets reduced to a no-interaction model. In this case, the sum of squares due to interaction is pooled with the error sum of squares to get a more precise estimate of the error variance for testing the significance of treatment effects. If, however, the interaction effect is significant, meaning thereby that the treatments differ in effect over environments, then the treatment effects are tested for significance by using the mean square due to interactions. In this case, the valid error for testing the significance of treatment effects is the mean square due to interactions. The significance is tested using Snedecor's *F*- statistic.

Remark 8.2: It may be noted that to study the interaction of treatments and environments, the experimental unit wise data is required. If for each environment, only the average value of the observations pertaining to each treatment is given then it is not possible to study the interaction of treatment and environments.

Remark 8.3: The environments depending upon their nature may either be fixed or random effects. Generally the different environments or the years are natural environments. The natural environments are usually considered as random sample from the population. Therefore, the effect of environment may be considered as random. All other effects in the model that involve the environment either as nested or as crossed classification are also considered as random. For instance the interaction term would also be random. The assumption of these random effects helps in identifying the proper error terms for testing the significance of various effects. The combined analysis of data can easily be carried out using PROC GLM of SAS along with Random statement with TEST option or PROC MIXED of SAS.

Remark 8.4: Some other experimental situations that can be viewed as groups of experiments are those in which it is difficult to change the levels of one of the factors because of practical considerations. For example, consider an experimental situation where the experimenter is interested in studying the long-term effect of irrigation and fertilizer treatments on a given crop sequence. There are 12 different fertilizer treatments and three-irrigation treatments viz. continuous submergence, 1 day drainage and 3 days drainage. It is very difficult to change the irrigation levels and randomize them. Therefore, the three irrigation levels may be taken as 3 artificially created environments and the experiment may be conducted using RCB design with 12 fertilizer treatments with suitable number of replications in each of the 3 environments. The data from each of the three experiments may be analyzed individually and the mean square errors so obtained may be used for testing the homogeneity of error variances and combined analysis of data be performed. In case of artificially created environments, the environment effect also consists of the effect of soil conditions in field experiments, initial quality parameters in food processing experiments, etc. Therefore, it is suggested that the data on some auxiliary variables may also be collected. These auxiliary variables may be taken as covariate in the analysis.

8.3   Example

An initial varietal trial was conducted to study the performance of 9 new strains of quality mustard vis-a-vis 3 checks using an RCB design with three replications at each of the respective environments (centres) at Bathinda, Hisar, IARI New Delhi, Ludhiana, Navgaon and TERI, New Delhi. The seed yield in kg/ha was recorded. The details of strains, design adopted and data obtained are given in Table 8.1.

Table 8.1: Treatment details of a group of experiments

Treatment	Treatment No.		Treatment	Treatment No.
ELM-123	1		PRQ-2005-10	7
LET-5	2		RH(HO) 0501	8
VARUN(NC)	3		**ZONAL CHECK**	9
TERI HOJ-48	4		LET-14-1	10
ELM-108	5		ELM-134	11
KRANTI(NC)	**6**		RH(OH) 0502	12

Note: Strains of quality mustard in boldface are the three checks, i.e., treatment numbers 3, 6 and 9 are checks.

The data for each of the environments is given in Table 8.2.

Table 8.2: Data for the experiment

Environ	Treat	Replication			Environ	Treat	Replication		
		1	2	3			1	2	3
Bathinda	1	1794	2014	2581	Hisar	1	3286	2459	3286
	2	1134	1736	1898		2	2518	2364	2364
	3	718	764	880		3	757	993	875
	4	1852	1551	1887		4	2553	2388	2884
	5	2245	2361	2407		5	2908	2482	2884
	6	1111	1065	1111		6	1797	1560	2033
	7	1181	880	1528		7	1749	1537	1537
	8	1644	1991	2060		8	1501	2317	2577
	9	1551	1435	1991		9	1513	1608	2104
	10	1968	1551	2569		10	2447	2459	2813
	11	2662	2338	3056		11	2600	2884	2648
	12	1065	1227	1343		12	1631	1466	1844

Environ	Treat	Replication			Environ	Treat	Replication		
		1	2	3			1	2	3
IARI New Delhi	1	2600	2444	2711	Ludhiana	1	1370	1209	1320
	2	3289	2667	2889		2	904	729	1007
	3	2756	2511	2400		3	858	942	839
	4	2600	2444	2222		4	904	959	1155
	5	2689	2422	2444		5	1438	1456	1695
	6	2578	2400	2222		6	873	959	946
	7	3178	3044	2889		7	848	639	643
	8	3244	2911	3111		8	1668	1770	1607
	9	2444	2222	2667		9	910	907	1081
	10	3156	2978	2756		10	1558	1606	1705
	11	2667	2267	2111		11	1508	1389	1447
	12	2689	2444	2289		12	1280	1207	1256

Environ	Treat	Replication			Environ	Treat	Replication		
		1	2	3			1	2	3
Navgaon	1	2233	2222	2222	TERI, New Delhi	1	1666	1333	2222
	2	2222	2444	2722		2	1611	1389	1944
	3	2000	1778	1778		3	1389	1244	2056
	4	2667	3289	3333		4	1511	1778	1889
	5	2444	2000	2000		5	1644	1622	1711
	6	1778	1889	1556		6	1833	1822	2111
	7	1778	1722	1722		7	1788	2333	1711
	8	3000	2889	3222		8	1644	2220	2220
	9	1778	1611	1333		9	1889	1822	2444
	10	3778	3667	3556		10	2000	1556	1356
	11	3111	3111	3222		11	944	388	722
	12	2222	2000	2222		12	1488	1400	1356

The objectives of the experiment are (i) to study the significance of strains for each environment vis-a-vis checks, (ii) to identify which of the new strains among those tried in the experiment can be adopted in which of the environments (environments), and (iii) to identify, if possible, which of the new strains can be recommended for adoption considering their performance over all the environments (environments). In order to answer these questions, the combined analysis of groups of experiments needs to be done.

The analysis is described in the sequel.

8.3.1 Environment wise analysis of data

The combined analysis of data is done using SAS. This analysis is quite involved and, therefore, for the benefit of clarity and understanding, the SAS commands have been split into parts. The steps involved are described in the sequel. The SREG or GGE plot analysis will be taken up later.

First prepare a data file.
DATA combined_analysis_data;
INPUT env $ block varn syield;
/*env ~ indicates environment and $ indicates that the environment is expressed in alphabets; rep indicates replication; varn indicates the strains or the varieties; syeild indicates seed yield*/
CARDS;

Bathinda	1	1	1794
Bathinda	1	2	1134
Bathinda	1	3	718
Bathinda	1	4	1852
Bathinda	1	5	2245
Bathinda	1	6	1111
Bathinda	1	7	1181

.
.
.

IARINewDelhi	1	1	2600
IARINewDelhi	1	2	3289
IARINewDelhi	1	3	2756
IARINewDelhi	1	4	2600
IARINewDelhi	1	5	2689
IARINewDelhi	1	6	2578

.
.
.

Hisar	1	1	3286
Hisar	1	2	2518
Hisar	1	3	757
Hisar	1	4	2553
Hisar	1	5	2908

.
.
.

Ludhiana	1	1	1370
Ludhiana	1	2	904
Ludhiana	1	3	858

Ludhiana	1	4	904
.			
.			
.			
Navgaon	1	1	2233
Navgaon	1	2	2222
Navgaon	1	3	2000
Navgaon	1	4	2667
Navgaon	1	5	2444
.			
.			
.			
TERINewDelhi	1	1	1666
TERINewDelhi	1	2	1611
TERINewDelhi	1	3	1389
TERINewDelhi	1	4	1511
.			
.			
.			
TERINewDelhi	3	10	1356
TERINewDelhi	3	11	722
TERINewDelhi	3	12	1356

```
;
/* Sort the data with respect to the environments*/
PROC SORT;
BY env;
RUN;
 /*To perform the analysis of data for each of the environments separately use the following SAS
statements.*/
PROC GLM data = combined_analysis_data;
CLASS rep  varn;
MODEL syield = block  varn;
LSMEANS varn/pdiff adjust=tukey lines;
BY  env;
run;
```

If one wants to see the Tukey's HSD values in case of balanced data as generated by RCB design, then one may add the following statement /*Means varn/tukey*/.

8.3.2   Output of environment wise analysis

The results obtained through the analysis are described in the sequence. The first PROC GLM does the analysis of variance for each centre individually. The class variables are varieties and blocks (or replications), which means the analysis is done for a two way classified data (or

RCB design). The results are given in the sequence. Since the design adopted at all the centres is an RCB design, which is essentially an orthogonal design, Type I and Type III sum of squares are identical. The above SAS code will give results for each environment separately. To save on space, the ANOVA Tables for each environments and means/adjusted means along with multiple comparison procedures are presented in Table 8.3. The results obtained are reformatted as to be presented in research output (following changes in labelling has been done: block: Blocks; varn: Treatments)

Environment wise results

Table 8.3: ANOVA for Bathinda centre

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks	2	1071654.00	535827.00	11.99	0.0003
Treatments	11	10254463.42	932223.95	20.87	<0.0001
Error	22	982897.33	44677.15		
Corrected Total	35	12309014.75			

R-Square	CV	Root MSE	syield Mean
0.92	12.44	211.370	1698.58

It may be seen that the model has been able to explain 92 per cent of the variability in the data obtained from Bathinda centre. Both the treatment effects and the block effects are significantly different.

Table 8.4: ANOVA for Hisar centre

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks	2	509922.06	254961.03	3.98	0.0335
Treatments	11	13043363.89	1185760.35	18.50	<0.0001
Error	22	1410037.28	64092.60		
Corrected Total	35	14963323.22			

R-Square	CV	Root MSE	syield Mean
0.91	11.74	253.165	2156.28

The fitted model has been able to explain 91 per cent of the variability in the data obtained from this centre. The treatment effects are highly significant. The block effects are also significantly different.

Table 8.5: ANOVA for IARI New Delhi centre

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks	2	553955.17	276977.58	13.55	0.0001
Treatments	11	2515742.08	228703.83	11.19	<0.0001
Error	22	449781.50	20444.61		
Corrected Total	35	3519478.75			

R-Square	CV	Root MSE	syield Mean
0.87	5.40	142.98	2648.75

The fitted model has been able to explain 87 per cent of the variability in the data obtained from IARI, New Delhi centre. The treatment effects are highly significant. The block effects are also significantly different.

Table 8.6: ANOVA for Ludhiana centre

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks	2	36727.06	18363.53	2.18	0.1374
Treatments	11	3607440.22	327949.11	38.85	<0.0001
Error	22	185714.28	8441.56		
Corrected Total	35	3829881.56			

R-Square	CV	Root MSE	syield Mean
0.95	7.77	91.88	1183.11

The fitted model has been able to explain 95 per cent of the variability in the data obtained from Ludhiana centre. The treatment effects are highly significant. The block effects are, however, not significantly different.

Table 8.7: ANOVA for Navgaon centre

Source	DF	Type III SS	MS	F Value	Prob > F
Blocks	2	6589.06	3294.53	0.09	0.9181
Treatments	11	15204645.64	1382240.51	36.00	<.0001
Error	22	844617.61	38391.71		
Corrected Total	35	16055852.31			

R-Square	CV	Root MSE	syield Mean
0.95	8.15	195.94	2403.361

The fitted model has been able to explain 95 per cent of the variability in the data obtained from Navagaon centre. The treatment effects are highly significant. The block effects are not significantly different, though.

Table 8.8: ANOVA for TERI, New Delhi center

Source	DF	Type III SS	MS	F-Value	Prob > F
Blocks	2	381651.39	190825.69	2.39	0.1146
Treatments	11	4408968.89	400815.35	5.03	0.0006
Error	22	1753633.94	79710.63		
Corrected Total	35	6544254.22			

R-Square	CV	Root MSE	syield Mean
0.73	16.92	282.33	1668.22

The fitted model has been able to explain 73 per cent of the variability in the data obtained from TERI, New Delhi centre. The treatment effects are highly significant. The block effects are not significantly different, though.

Multiple comparisons of treatments are made using Tukey's Honest Significant Difference procedure of comparisons. The results obtained are presented in Table 8.9.

Table 8.9: Centre Wise Treatment Means and Multiple Comparison Procedure Results

Treatment Name	Bathinda SYIELD	Hisar SYIELD	IARI New Delhi SYIELD	Ludhiana SYIELD	Navgaon SYIELD	TERINewDelhi SYIELD
1	2129.67ABC	3010.33A	2585.00BC	1299.67BC	2225.67DE	1740.33A
2	1589.33CDE	2415.33ABC	2948.33AB	880.00EF	2462.67CD	1648.00A
3	787.33F	875.00E	2555.67BC	879.67EF	1852.00EF	1563.00A
4	1763.33BCD	2608.33AB	2422.00C	1006.00DE	3096.33AB	1726.00A
5	2337.67AB	2758.00AB	2518.33C	1529.67AB	2148.00DEF	1659.00A
6	1095.67EF	1796.67CD	2400.00C	926.00EF	1741.00EF	1922.00A
7	1196.33DEF	1607.67DE	3037.00A	710.00F	1740.67EF	1944.00A
8	1898.33BC	2131.67BCD	3088.67A	1681.67A	3037.00BC	2028.00A
9	1659.00CDE	1741.67CD	2444.33C	966.00EF	1574.00F	2051.67A
10	2029.33BC	2573.00AB	2963.33AB	1623.00A	3667.00A	1637.33A
11	2685.33A	2710.67AB	2348.33C	1448.00ABC	3148.00AB	684.67B
12	1211.67DEF	1647.00D	2474.00C	1247.67CD	2148.00DEF	1414.67AB
General Mean	1698.58	2156.28	2648.75	1183.11	2403.36	1668.22
p-Value	<.0001	<.0001	<.0001	<.0001	<.0001	0.0006
CV (%)	12.44	11.74	5.40	7.77	8.15	16.92
SE(d)	172.58	206.71	116.75	75.02	159.98	230.52
Tukey's HSD at 5%	627.78	751.92	424.67	272.88	581.95	838.54

It can be observed that at Bhatinda Centre Treatment 11 gives the highest seed yield though statistically it is at par with treatments 5 and 1. Treatment 3 gives the lowest seed yield. At Hisar Centre, Treatment 1 produces the highest seed yield, though statistically it is at par with treatments 5, 11, 4, 10, 2. Treatment 3 gives the lowest seed yield. At IARI, New Delhi, Treatment 8 gives the highest seed yield, though statistically it is at par with treatments 7, 10, 2. Treatment 11 gives the lowest seed yield. At Ludhiana Centre, Treatment 8 gives the highest seed yield, though statistically it is at par with treatments 10, 5, 11. Treatment 7 gives the lowest seed yield. For Navagaon Centre, Treatment 10 gives the highest seed yield, though statistically it is at par with treatments 11, 4. Treatment 9 gives the lowest seed yield. For TERI, New Delhi, Treatment 9 gives the highest seed yield, though statistically it is at par with all other treatments except treatment 11, which gives the lowest seed yield.

8.3.3 Combined analysis of data using mixed effects model

As discussed earlier location or seasons may be considered as random. Therefore, assuming environment effect as random, the blocks within environments and environment × treatment interaction are also random effects.

/*To perform the combined analysis of the above data set considering the environments as random effects one can perform the combined analysis as follows.;*/

/*Analysis using Proc Mixed*/

PROC MIXED ratio covtest DATA = combined_analysis_data;

class env block varn;

MODEL syield = varn;

RANDOM env block(env) env*varn;

LSMEANS varn/PDIFF adjust=tukey;

RUN;

However, if one wants estimates for all random effects, then one may use the modify the random statement by include option s as follows

RANDOM env rep(env) env*varn/s;

To save on space, the results obtained are reformatted as to be presented in research output. These are presented in the sequel(following changes in labelling has been done: block: Blocks; varn: Treatments; env:environment)

Table 8.10: Results from mixed model analysis

Covariance Parameter Estimates					
Covariance Parameter	**Ratio**	**Estimate**	**Standard Error**	**Z Value**	**Prob > Z**
Environment	6.413	273382	185538	1.47	0.0703
Environment*Treatment	3.939	167895	34770	4.83	<.0001
Block (Environment)	0.334	14229	7272.32	1.96	0.0252
Residual	1.000	42626	5246.94	8.12	<.0001

Fit Statistics	
-2 Res Log Likelihood	2975.5
AIC (smaller is better)	2983.5
AICC (smaller is better)	2983.7
BIC (smaller is better)	2982.6

Type 3 Tests of Fixed Effects				
Effect	**Numerator DF**	**Denominator DF**	**F Value**	**Prob > F**
Treatment	11	55	3.16	0.0023

It can be observed that variance components pertaining to Environment*Treatment and Block (Environment) are significant. Treatment effects are also significantly different at 5% level of significance. Tukey's HSD at 5% level of significance was used for performing multiple pair wise comparisons. The results based on LSD at 5% level of significance are also presented in Table 8.11. The results obtained after reformatting are summarised as follows:

Table 8.11: Overall treatment means and grouping of treatments using Tukey's HSD/LSD at 5% level of significance

Treatment Number	Adjusted Mean/ BLUP and Grouping based on Tukey's HSD	Adjusted Mean/ BLUP and Grouping based on Tukey's LSD
1	2165.11^{AB}	2165.11^{AB}
2	1990.61^{AB}	1990.61A^{BC}
3	1418.78^{B}	1418.78^{D}
4	2103.67^{AB}	2103.67^{ABC}
5	2158.44^{AB}	2158.44^{AB}
6	1646.89^{AB}	1646.89^{CD}
7	1705.94^{AB}	1705.94B^{CD}
8	2310.89^{A}	2310.89^{A}
9	1739.44^{AB}	1739.44^{BCD}
10	2415.50^{A}	2415.50^{A}
11	2170.83A^{B}	2170.83^{AB}
12	1690.50^{AB}	1690.50^{BCD}
Prob > F	0.0023	0.0023

It can be seen that using Tukey's HSD, treatments 10 and 8 are best performing and these two treatments are at par with treatments 1, 2 ,4, 5, 6, 7, 9, 11, 12 statistically. However, using Least Significant Difference, treatments 10, 8, 11, 1, 5, 4 and 2 are best performing treatments, which are not significantly different from each other. Since Treatment × Environment interaction is significant, it follows that the treatments do not perform equally well over all locations. As a consequence, the above described procedure should not be used for interpretation and Site Regression Biplots should be used to interpret the performance of the treatments over different locations. The procedure of generating SREG Biplots is described in the sequel.

Remark 8.5: In the above analysis, the structure of variance-covariance matrix for all random effects has been assumed as diagonal with constant diagonal elements. Depending upon the requirement of the experimental situation, one may assume different structure of variance-covariance matrix of random effects.

Remark 8.6: So far the focus of combined analysis of data has been on PROC MIXED of SAS. It may be worthwhile mentioning here that one could also use PROC GLM along with Random statement for random effects. The same is now described in the sequel. This would give the expected mean squares for each of the effects and perform the testing against appropriate error terms based on expected mean squares.

PROC GLM DATA = combined_analysis_data;

CLASS env block varn;

MODEL syield = env block(env) varn env*varn;

RANDOM env env*varn block(env)/TEST;

LSMEANS varn/PDIFF lines;

RUN;

8.4. SREG biplot

It has been seen above while doing the analysis of individual environments that the performance of treatments is different at different environments. The best performing treatments and the worst performing treatments are different at different environments. This fact has also been supported by the combined analysis of data. The environment*treatment interaction is significant both in the fixed effects model where the environment effect is assumed to be fixed and in the mixed effects model where the environment effect is assumed to be random. In such a situation, the same treatment cannot be recommended for all the environments and the environment specific recommendations need to be made. For this purpose appeal is made to SREG biplot approach described in the sequel.

8.4.1 SREG biplot

As varn × env (treatment × environment) interaction is significant, same treatment cannot be recommended for all environments. In such a situation, one can see the performance of treatment and treatment × environment interaction using SREG biplot or GGE biplot. For performing SREG Biplot, one needs to obtain means / adjusted means / best linear unbiased predictor (BLUP) for treatments across environment. For a balanced data (RCB design with same treatments and same number of replications in all the environments), it is means. If incomplete block design is used at some or all environments but the treatments are same in all the environments, then use adjusted mean of treatments for each environment. For an unbalanced data, when there are some empty cells in treatment × environment table, then one can use the best linear unbiased predictor (BLUP) of treatment × environment interaction as means. BLUPs can be obtained by using PROC MIXED of SAS by using the option 's' in Random statement. If some of the cells in treatment into environment table are missing, then one can obtain BLUP and use in place of lsmeans. The experimenters may exercise caution to ensure that no more than 20% of cells are empty in treatment × environment Table.

For performing SREG Biplot, create a data file named RAW where Environments are termed as ENV, treatment numbers as GEN and means for GEN as GYLD. The computer program used here for analysis is a slightly modified version of the program developed by Jose Crossa and his co-workers at CIMMYT, Mexico.

OPTIONS PS = 5000 LS = 78 NODATE;

/*after removing * one can get the output as a cgm file directly, which can be imported in PowerPoint or word documents for clarity. */

*FILENAME BIPLOT 'C:\Documents and Settings\owner\Desktop\comana.cgm'; *To have cgm files run it in BATCH;

*GOPTIONS DEVICE = CGMOF97L GSFNAME = BIPLOT GSFMODE = REPLACE;

/*one has to run the program twice, first time to see the portion of variation explained by two components in the output file, then one has to change the value of factor 1 and factor 2 in the file at appropriate place.*/

```
DATA RAW;
INPUT ENV $ GEN $ GYLD;
YLD=GYLD;
CARDS;
Bathinda        1        2129.66667
Bathinda        2        1589.33333
Bathinda        3        787.33333
Bathinda        4        1763.33333
Bathinda        5        2337.66667
Bathinda        6        1095.66667
Bathinda        7        1196.33333
Bathinda        8        1898.33333
```

Bathinda	9	1659
Bathinda	10	2029.33333
Bathinda	11	2685.33333
Bathinda	12	1211.66667
Hisar	1	3010.33333
Hisar	2	2415.33333
Hisar	3	875
Hisar	4	2608.33333
Hisar	5	2758
Hisar	6	1796.66667
Hisar	7	1607.66667
Hisar	8	2131.66667
Hisar	9	1741.66667
Hisar	10	2573
Hisar	11	2710.66667
Hisar	12	1647
IARINewD	1	2585
IARINewD	2	2948.33333
IARINewD	3	2555.66667
IARINewD	4	2422
IARINewD	5	2518.33333
IARINewD	6	2400
IARINewD	7	3037
IARINewD	8	3088.66667
IARINewD	9	2444.33333
IARINewD	10	2963.33333
IARINewD	11	2348.33333
IARINewD	12	2474
Ludhiana	1	1299.66667
Ludhiana	2	880
Ludhiana	3	879.66667
Ludhiana	4	1006
Ludhiana	5	1529.66667
Ludhiana	6	926
Ludhiana	7	710
Ludhiana	8	1681.66667
Ludhiana	9	966
Ludhiana	10	1623
Ludhiana	11	1448
Ludhiana	12	1247.66667
Navgaon	1	2225.66667
Navgaon	2	2462.66667
Navgaon	3	1852

Navgaon	4	3096.33333
Navgaon	5	2148
Navgaon	6	1741
Navgaon	7	1740.66667
Navgaon	8	3037
Navgaon	9	1574
Navgaon	10	3667
Navgaon	11	3148
Navgaon	12	2148
TERINewD	1	1740.33333
TERINewD	2	1648
TERINewD	3	1563
TERINewD	4	1726
TERINewD	5	1659
TERINewD	6	1922
TERINewD	7	1944
TERINewD	8	2028
TERINewD	9	2051.66667
TERINewD	10	1637.33333
TERINewD	11	684.66667
TERINewD	12	1414.66667

```
;
PROC GLM DATA = RAW OUTSTAT = STATS ;
CLASS  ENV GEN;
MODEL YLD =  ENV GEN ENV*GEN/SS4;
/*If this is required, then replace, MSE by the MSE in combined analysis, DFE with error degrees
of freedom in combined analysis, NREP number of replications at each environments*/
DATA STATS2;
SET STATS ;
DROP _NAME_ _TYPE_;
IF _SOURCE_ = 'ERROR' THEN DELETE;
MSE = 42626.4;          * MSE in combined analysis when environments are random;
DFE = 132;               * degrees of freedom in combined analysis;
NREP = 3;               * number of replications at each environments;
SS = SS*NREP;
MS = SS/DF;
F = MS/MSE;
PROB = 1- PROBF(F,DF,DFE);
PROC PRINT DATA = STATS2 NOOBS;
VAR _SOURCE_ DF SS MS F PROB;
PROC GLM DATA = RAW NOPRINT;
CLASS ENV GEN;
MODEL YLD = ENV / SS4 ;
```

```
OUTPUT OUT = OUTRES R = RESID;
PROC SORT DATA = OUTRES;
BY GEN ENV;
PROC TRANSPOSE DATA = OUTRES OUT = OUTRES2;
BY GEN;
ID ENV;
VAR RESID;
PROC IML;
USE OUTRES2;
READ ALL INTO RESID;
NGEN = NROW(RESID);
NENV = NCOL(RESID);
USE STATS2;
READ  VAR {MSE}  INTO MSEM;
READ  VAR {DFE}  INTO DFEM;
READ  VAR {NREP} INTO NREP;
CALL SVD (U,L,V,RESID);
MINIMO = MIN(NGEN,NENV);
L = L[1:MINIMO,];
SS=(L##2)*NREP;
SUMA = SUM(SS);
PERCENT = ((1/SUMA)#SS)*100;
MINIMO = MIN(NGEN,NENV);
PERCENTA = 0;
DO I = 1 TO MINIMO;
DF = (NGEN-1)+(NENV-1)-(2*I-1);
DFA = DFA//DF;
PORCEACU = PERCENT[I,];
PERCENTA = PERCENTA+PORCEACU;
PORCENAC = PORCENAC//PERCENTA;
END;
DFE = J(MINIMO,1,DFEM);
MSE = J(MINIMO,1,MSEM);
SSDF = SS||PERCENT||PORCENAC||DFA||DFE||MSE;
L12=L##0.5;
SCOREG1 = U[,1]#L12[1,];
SCOREG2 = U[,2]#L12[2,];
SCOREG3 = U[,3]#L12[3,];
SCOREE1 = V[,1]#L12[1,];
SCOREE2 = V[,2]#L12[2,];
SCOREE3 = V[,3]#L12[3,];
FACTOR1 = MAX(ABS(SCOREG1||SCOREG2));
FACTOR2 = MAX(ABS(SCOREE1||SCOREE2));
```

```
FACTOR = MAX(FACTOR1,FACTOR2);
SCOREG = (SCOREG1||SCOREG2||SCOREG3)*(1/FACTOR);
SCOREE = (SCOREE1||SCOREE2||SCOREE3)*(1/FACTOR);
SCORES = SCOREG//SCOREE;
CREATE SUMAS FROM SSDF;
APPEND FROM SSDF;
CLOSE SUMAS;
CREATE SCORES FROM SCORES;
APPEND FROM SCORES;
CLOSE SCORES;
DATA SS_SREG;
SET SUMAS;
SS_SREG = COL1;
PERCENT = COL2;
PORCENAC = COL3;
DF_SREG = COL4;
DFE = COL5;
MSE = COL6;
DROP COL1 - COL6;
MS_SREG = SS_SREG/DF_SREG;
F_SREG = MS_SREG/MSE;
PROBF = 1 - PROBF(F__SREG,DF_SREG,DFE);
PROC PRINT DATA = SS_SREG NOOBS;
VAR SS_SREG PERCENT PORCENAC;
PROC SORT DATA = RAW;
BY GEN;
PROC MEANS DATA = RAW NOPRINT;
BY GEN ;
VAR YLD;
OUTPUT OUT = MEDIAG MEAN=YLD;
DATA NAMEG;
SET MEDIAG;
TYPE = 'GEN';
NAME = GEN;
KEEP TYPE NAME YLD;
PROC SORT DATA = RAW;
BY ENV;
PROC MEANS DATA = RAW NOPRINT;
BY ENV ;
VAR YLD;
OUTPUT OUT = MEDIAE MEAN=YLD;
DATA NAMEE;
SET MEDIAE;
```

```
TYPE = 'ENV';
NAME1 = 'S'||ENV;
NAME = COMPRESS(NAME1);
KEEP TYPE NAME YLD;
DATA NAMETYPE;
SET NAMEG NAMEE;
DATA BIPLOT;
MERGE NAMETYPE SCORES;
DIM1=COL1;
DIM2=COL2;
DIM3=COL3;
DROP COL1-COL3;
TITLE1 'Biplot of Grain Yield';
PROC PRINT DATA = BIPLOT NOOBS;
VAR TYPE NAME YLD DIM1 DIM2 DIM3;
DATA labels;
SET BIPLOT;
retain xsys  '2' ysys  '2' ;
length function text $8 ;
text = name ;
IF type = 'GEN' THEN DO;
color = 'red ';
size = 1.0;
style = 'hwcgm001';
x = dim1;
y = dim2;
IF dim1 >=0
THEN position = '5';
ELSE position = '5';
function = 'LABEL';
OUTPUT;
END;
if type = 'ENV' then DO;
color = 'blue ';
size = 1.0;
style = 'hwcgm001';
x = 0.0;
y = 0.0;
function = 'MOVE';
output;
x = dim1;
y = dim2;
FUNCTION = 'DRAW';
```

```
OUTPUT;
IF dim1 >=0
THEN position = '6';
ELSE position = '4';
FUNCTION = 'LABEL';
OUTPUT;
END;
/*one has to run the program twice, first time to see the portion of variation explained by two
components, then change in the file at appropriate places for the factor 1 and factor 2 */
PROC GPLOT data = biplot;
PLOT  dim2*dim1 / Annotate=labels frame
Vref = 0.0 Href = 0.0
cvref = black chref = black
lvref = 3 lhref = 3
vaxis = axis2 haxis = axis1
vminor = 1  hminor = 1 nolegend;
symbol1 v=none  c=black  h=0.7;
symbol2 v=none  c=black  h=0.7;
axis2
length = 5.0 in
order = (-1 to 1.0 by 0.2)
/*one has to change the value for factor 2(.)*/
label = (f=hwcgm001 c=green h=1.2 a=90 r=0 'Factor 2 (15.25%)') /*please change the
percent  variation explained as per data*/
offset = (3)
value = (h = 1.0)
minor = none;
* length = 7.0 in FOR CGM files;
axis1
length = 7.0 in
order = (-0.8 to 1.0 by 0.2)
/*one has to change the value for factor 1(.)*/
label = (f = hwcgm001 c = green h=1.2 'Factor 1 (66.91%)') /*please change the percent variation
explained as per data*/
offset = (3)
value = (h = 1.0)
minor = none;
*length = 7.0 in FOR CGM files;
Title1 f=hwcgm001 c=Red h=2.0 'SREG biplot of the Grain Yield of Quality_zone2 at 6
Environments';
/*Give the title as is required in output*/
RUN;
```

The results obtained are given in the sequel. After performing Singular value Decomposition, the first two dimensions and variance explained by them are obtained. These dimensions are then used for plotting SREG biplot. First two dimensions /factors explain 66.91% and 15.25% variation which is more than 80%. THE SREG biplot obtained using these two dimensions is
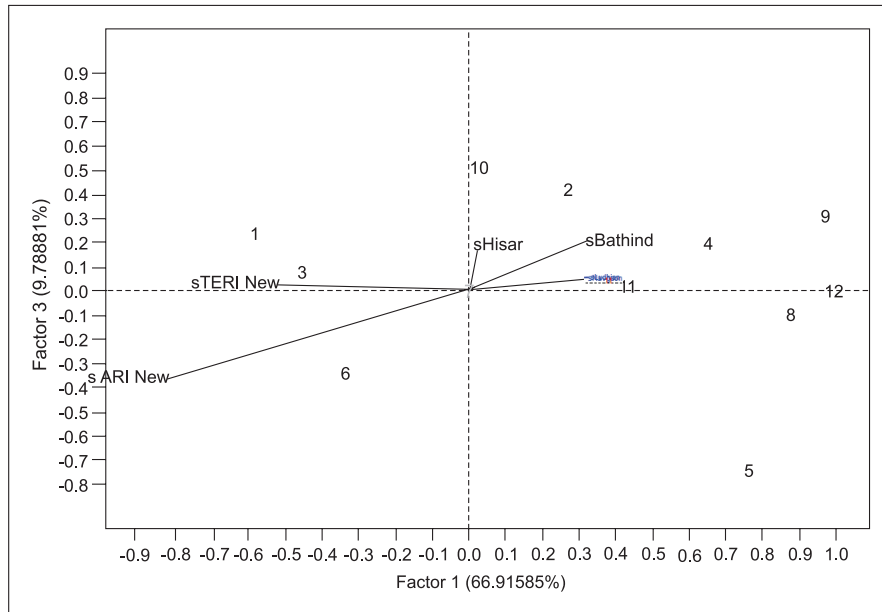
Figure 8.1: SREG biplot of grain yield at six environments

Remark 8.7: Since the error variances may be heterogeneous, there would be a need for transformation of data. Therefore, it would not be out of place here to describe the testing of homogeneity of error variances.

From environment wise analysis in Section 8.3.2, one can see that that there is a large variability among the error mean squares of different environments. It appears that the error variances over different environments may not be homogeneous. So it would be desirable to test for the homogeneity of error variances using Bartlett's chi square test. The SAS code for the same is given in the sequel.

/*To check the homogeneity of variances we apply Bartlett's Chi-square test*/

/* SAS Code for testing the homogeneity of variances, when variances and the degrees of freedom are given. It is useful for testing the homogeneity of error variances when the experiments are conducted over environments*/

SAS Code for applying Bartlett's test for homogeneity of variances

```
DATA mn;
INPUT df mse;
CARDS;
22 44677.15
22 64092.60
22 20444.61
22 8441.56
22 38391.71
22 79710.63
;
PROC IML;
USE mn;
READ ALL INTO a; /* use variances of residual variance putting it in m1 variable*/
*a =m1[2:nrow(m1),ncol(m1)-1:ncol(m1)];/*from m1 we extracting variances and number of
observations */
v =0; ct = 0; nchi = 0; St = 0;
DO i = 1 TO NROW(a);  /* computing pooled variance */
        St = St + (a[i,1]-1)*a[i,2];
        v = v + (a[i,1]-1);
        ct = ct + 1/(a[i,1]-1);
END;
S = St/v;
dchi = (1 + (1/(3*(nrow(a)-1)))*(ct-(1/v))); /*computing denominator of Bartlett's chi-square
statistic*/
DO i = 1 TO NROW(a);
nchi = nchi + (a[i,1]-1)*(log(S/a[i,2]));
END;
chi = nchi/dchi;
probability = 1 - probchi(chi,(nrow(a)-1)); /*computinf chi value and prob.*/
df = (NROW(a)-1);
PRINT probability chi df S; /* printing chi value, prob and degee of freedom*/
IF PROBABILITY >= 0.05 THEN Interpretation = "Data is Homogeneous at 5% level of
Significance";
ELSE Interpretation = "Data is Heterogeneous at 5% level of Significance";
PRINT Interpretation; /* testing and printing interpretation*/
pb = CHAR(probability);
```

The results obtained for testing the homogeneity of error variances using Bartlett's chi square are given in Table 8.12.

Table 8.12: Result of Bartlett's chi square test

probability	Chi square	DF	S
0.00003	28.41	5	42626.38

Since the probability of getting a chi square value of 28.41 at 5 degrees of freedom is so small (0.00003), it may be concluded that the error variances are significantly different and the data are heterogeneous. Once it is known that the errors are heterogeneous, it calls for transformation of data before doing the combined analysis of data from all the centres. In what follows is described the transformation of data using SAS package. The transformation involves dividing observations of each environment by the square root of MSE of that environment

/*In this example, error variances are heterogeneous; transform the data by dividing each observation by its corresponding square root of the error mean square and create a new variable new_var and use the following SAS statements for the combined analysis of data*/

DATA tranformed; /* This set of SAS statements transforms the data*/
SET combined_analysis_data;
IF env = "Bathinda" THEN
new_var = syield/sqrt(44677.15);
IF env = "Hisar" THEN
new_var = syield/sqrt(64092.6);
IF env = "IARINewDelhi" THEN
new_var = syield/sqrt(20444.61);
IF env = "Ludhiana" THEN
new_var = syield/sqrt(8441.56);
IF env = "Navgaon" THEN
new_var = syield/sqrt(38391.71);
IF env = "TERINewDelhi" THEN
new_var = syield/sqrt(79710.63);
RUN;

/*To perform the combined analysis of the above data set considering the environments as fixed effects one can use the following SAS statements*/

PROC GLM DATA = transformed;
CLASS env block varn;
MODEL new_var = env block(env) varn env*varn;
MEANS varn /TUKEY;
RUN;
RUN;
/*please note that for performing comparisons, transformed data (new_var) should only be used. However, for just having the original means, the analysis of syield should be seen*/

8.5 Analysis using R

In the sequence is given R code for the analysis of data using R software. The results obtained are similar to those obtained using SAS. To avoid duplication and to save space, the results obtained using R code are not reported.

```
d13=read.table("combined_analysis_data.txt",header=TRUE)
attach(d13)
names(d13)
#anova
env=factor(env)
levels(env)
out=by(d13,d13[,"env"],function(x) aov(syield~factor(rep)+factor(varn),data=x))
out=as.list(out)
sapply(out,summary)
lapply(out,TukeyHSD,"factor(varn)")
MSE=sapply(1:6,function(i,out) sum(out[[i]]$residual^2)/out[[i]]$df.residual,out)
residfs=sapply(1:6,function(i,out) out[[i]]$df.residual,out)
#Check homogeneity of error variances across environments
bartlett.test(out)
#Transform the observations as variances are heterogeneous
tsyield=unlist(tapply(syield,env,function(syield,MSE) syield/sqrt(MSE),MSE))
#for combined anova
varn=factor(varn)
rep=factor(rep)
lm2=lm(tsyield~env+varn+env/rep-rep+env:varn)
library(car)
Anova(lm2,type="III")
TukeyHSD(aov(tsyield~env+varn+env/rep-rep+env:varn),"varn")
#environments as random effects
#Need to install lme4 package
library(lme4)
lm3=lmer(syield~varn+(1|env:varn)+(1|env)+(1|env/rep))
anova(lm3)
library(car)
Anova(lm3,type="III")
summary(lm3)
#install pbkrtest package for testing random effects
library(lsmeans)
lsmeans(lm3,"varn")
detach(d13)
```

Diagnostics in Experimental Designs

9.1 Introduction

The interpretation of results based on analysis of variance (ANOVA) of data from a designed experiment is valid only when the assumptions on which ANOVA is based are satisfied. Some of these assumptions are as follows:

1. Additive Effects: Treatment effects and block (environmental) effects are additive.
2. Independence of errors: Experimental errors are independent.
3. Homogeneity of Variances: Errors have common variance.
4. Normal Distribution: Errors follow a normal distribution.

Further, the statistical tests like t, F, χ^2 and z are valid under the assumption of independence of errors and normality of error distribution. The departures from these assumptions make the interpretation based on these statistical techniques invalid. Therefore, before analysing the data, it is necessary to check whether or not these assumptions are met by the data. If one or more of these assumptions are violated, then there is a need to apply appropriate remedial measures. The assumption of independence of errors, *i.e.*, error of an observation is not related to or depends upon that of another, is usually assured with the use of proper randomization procedure. However, if there is any systematic pattern in the arrangement of treatments from one replication to another, errors may not be independent. This may be handled by using nearest neighbour methods in the analysis of experimental data. In the sequel are described the procedures for detecting violation of these assumptions and then the corresponding remedial measures to be used are also outlined.

9.2 Additivity of effects

The effects of two factors, say, treatment and replication, are said to be additive if the effect of one-factor remains constant over all the levels of the other factor. A hypothetical set of data from a randomized complete block (RCB) design, with 2 treatments and 2 replications, with additive effects is given in Table 9.1.

Table 9.1: A dataset with additivity of effects

Treatment	Replication		Replication Effect
	I	**II**	**I – II**
A	190	125	65
B	170	105	65
Treatment Effect (A – B)	20	20	

Here, the treatment effect is equal to 20 for both the replications and replication effect is 65 for both the treatments.

When the effect of one factor is not constant at all the levels of other factor, the effects are said to be non-additive. A common departure from the assumption of additivity in biological experiments is one where the effects are multiplicative. Two factors are said to have multiplicative effects if their effects are additive only when expressed in terms of percentages. Table 9.2 illustrates a hypothetical set of data with multiplicative effects.

Table 9.2: A dataset with multiplicative effects

Treatment		Replication		Replication Effect	
		I (log {I})	II (log {II})	I – II (log {I} – log{II})	100(I – II)/II
A (log {A})		200 (2.30103)	125 (2.09691)	75 (0.20412)	60
B (log {B})		160 (2.20412)	100 (2.0000)	60 (0.20412)	60
Treatment Effect	(A – B) (log {A} – log{B})	40 (0.09691)	25 (0.09691)		
100 (A – B)/B		**25**	**25**		

In this case, the treatment effect is not constant over replications and the replication effect is not constant over treatments. However, when both treatment effect and replication effect are expressed in terms of percentages, an entirely different pattern emerges. For such violations of assumptions, logarithmic transformation is quite suitable. For illustration, the logarithmic transformation of data in Table 9.2 is given in brackets.

This is, however a crude method for testing the additivity. This method cannot be applied when there are more than two factors and/or levels of factors. Tukey (1949) suggested a statistical test for testing the additivity in a RCB design. This test is known as one degree of freedom test for non-additivity. In this test, one degree of freedom is isolated from error and this degree of freedom is called as the degree of freedom for non-additivity. In the sequel, we describe the procedure in brief.

Suppose that an experiment has been conducted to compare v treatments using RCB design with b replications or blocks. Let y_{ij} denote the observed value of the response variable for ith treatment in jth replication; $i = 1,2, \ldots, v$; $j = 1,2, \ldots, b$. Arrange the data in a $v \times b$ table as given below.

Treatment	1	2	...	j	...	b	Treatment Total	Treatment Mean	Deviations from Grand Mean	Sum of Cross Product
1	y_{11}	y_{12}	...	y_{1j}	...	y_{1b}	$T_{1.}$	$\bar{y}_{1.}$	$d_{1.}$	C_1
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2b}	$T_{2.}$	$\bar{y}_{2.}$	$d_{2.}$	C_2
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ib}	$T_{i.}$	$\bar{y}_{i.}$	$d_{i.}$	C_i
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
v	y_{v1}	y_{v2}	...	y_{vj}	...	y_{vb}	$T_{v.}$	$\bar{y}_{v.}$	$d_{v.}$	C_v
Replication or Block Total	$R_{.1}$	$R_{.2}$	...	$R_{.j}$	...	$R_{.b}$	G (Grand total)			
Replication Mean	$\bar{y}_{.1}$	$\bar{y}_{.2}$	...	$\bar{y}_{.j}$	...	$\bar{y}_{.b}$	$GM = \dfrac{G}{vb}$			
Deviation from Grand Mean	$d_{.1}$	$d_{.2}$	...	$d_{.j}$	...	$d_{.b}$				

where $T_{i.} = \displaystyle\sum_{j=1}^{b} y_{ij}$; $\bar{y}_{i.} = T_{i.}/b$; $R_{.j} = \displaystyle\sum_{i=1}^{v} y_{ij}$; $\bar{y}_{.j} = R_{.j}/v$; $d_{i.} = \bar{y}_{i.} - GM$

$d_{.j} = \bar{y}_{.j} - GM$; $C_i = \displaystyle\sum_{j=1}^{b} y_{ij} \cdot d_{.j}$

Let $\quad L = \displaystyle\sum_{i=1}^{v} C_i d_{i.}$; $D_1 = \displaystyle\sum_{i=1}^{v} d_{i.}^2$; $D_2 = \displaystyle\sum_{j=1}^{b} d_{.j}^2$.

Now sum of squares of different sources can be calculated by the following formulae:

Sum of squares due to non-additivity (SSNA) $= \dfrac{L^2}{D_1 \times D_2}$

Sum of squares due to treatments (SST) $= \displaystyle\sum_{i=1}^{v} \dfrac{T_{i.}^2}{b} - \dfrac{G^2}{vb}$

Sum of squares due to replications (SSR) = $\sum_{j=1}^{b}\dfrac{R_{.j}^{2}}{v}-\dfrac{G^{2}}{vb}$

Total sum of squares (TSS) = $\sum_{i=1}^{v}\sum_{j}^{b}y_{ij}^{2}-\dfrac{G^{2}}{vb}$

Sum of squares due to Error (SSE) = TSS – SST – SSR – SSNA

An outline of ANOVA is given in Table 9.3.

<div align="center">

Table 9.3: ANOVA table for test of additivity

</div>

Source	DF	SS	MS
Treatments	$v-1$	SST	MST
Replications	$b-1$	SSR	MSR
Non-additivity	1	SSNA	MSNA
Error	$(v-1)(b-1)-1$	SSE	MSE
Total	$vb-1$	TSS	

Then the non-additivity is tested by *F*-statistic, given by $F=\dfrac{MSNA}{MSE}$, with 1 and $(v-1)(b-1)$ –1 degrees of freedom.

9.3 Normality of errors

The assumptions of homogeneity of variances and normality are generally violated together. To test the validity of normality of errors, both graphical and statistical procedures are now available. One of the most popular graphical tests is Normal probability plot. Among statistical tests, Kolmogorov-Smirnov test is most widely used test. Other statistical tests include Anderson-Darling test, D'Agostino's test, Shapiro-Wilk's test and Ryan-Joiner test. In general, moderate departures from normality are of little concern in the fixed effects ANOVA as *F*-test is slightly affected by moderate departures but in case of random effects, it is more severely impacted by non-normality. However, significant deviations of errors from normality make the inferences invalid. It is, therefore, important to detect non-normality of errors before analysing the data and take appropriate measures if non-normality is detected. For testing normality of errors, we need to estimate them. The estimates of errors are taken as the residuals. To obtain residuals, we have to first fit model corresponding to the design adopted, *i.e.*, to obtain the predicted values of the response after eliminating the effects of treatments, blocks, rows, etc. These residuals are then used for testing the normality of the errors. In other words, we want to test the null hypothesis H_{0}: errors are normally distributed against alternative hypothesis H_{1}: errors are not normally distributed. For details on these tests, one may refer to D'Agostino and Stephens (1986). Most of the standard statistical packages are capable of testing the normality of the data.

9.3.1 Correlation test for normality

In addition to visually assessing the appropriate linearity of the points plotted in a normal probability plot, a formal test for normality of the error terms can be conducted by calculating the coefficient of correlation between residuals e_i and their expected values under normality. A high value of the correlation coefficient is indicative of normality.

9.3.2 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test is used to decide if a sample comes from a population with a specific distribution. The Kolmogorov-Smirnov (K-S) test is based on the empirical cumulative distribution function (ECDF). Given N ordered data points $Y_1, Y_2, ..., Y_N$, the ECDF is defined as

$$E_N = n(i)/N \, ,$$

where $n(i)$ is the number of points less than Y_i and the Y_i are ordered from smallest to largest value. This is a step function that increases by $1/N$ at the value of each ordered data point. The K-S test is based on the maximum distance between these two curves. In designed experiments variable Y is taken as residuals obtained from the fitted model corresponding to a design adopted.

An attractive feature of this test is that the distribution of the K-S test statistic itself does not depend on the underlying cumulative distribution function being tested. Another advantage is that it is an exact test (the chi-square goodness-of-fit test depends on an adequate sample size for the approximations to be valid). Despite these advantages, the K-S test has several important drawbacks:

1. It applies to continuous distributions only.

2. It tends to be more sensitive near the center of the distribution than at the tails.

3. Perhaps the most serious limitation is that the distribution must be fully specified. That is, if location, scale, and shape parameters are estimated from the data, the critical region of the K-S test is no longer valid. It typically must be determined by simulation.

Due to limitations 2 and 3 above, many analysts prefer to use the Anderson-Darling goodness-of-fit test.

The Kolmogorov-Smirnov test is defined by:

H_0: The data follow a specified distribution

H_1: The data do not follow the specified distribution

The Kolmogorov-Smirnov test statistic is defined as

$$D = \max_{1 \le i \le N} \left(F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i) \right) \tag{9.1}$$

where F is the theoretical cumulative distribution of the distribution being tested which must be a continuous distribution (*i.e.*, no discrete distributions such as the binomial or Poisson), and it

must be fully specified (*i.e.*, the location, scale, and shape parameters cannot be estimated from the data).

The hypothesis regarding the distributional form is rejected if the test statistic, D, is greater than the critical value obtained from a table. There are several variations of these tables in the literature that use somewhat different scalings for the K-S test statistic and critical regions. These alternative formulations should be equivalent, but it is necessary to ensure that the test statistic is calculated in a way that is consistent with how the critical values were tabulated. These values are given in a Table in Appendix 1.

9.3.3 Anderson-Darling test

The Anderson-Darling test is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test. The K-S test is distribution free in the sense that the critical values do not depend on the specific distribution being tested. The Anderson-Darling test makes use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution. Currently, tables of critical values are available for the normal, lognormal, exponential, Weibull, extreme value type I, and logistic distributions.

The Anderson-Darling test is defined as:

H$_0$: The data follow a specified distribution.

H$_1$: The data do not follow the specified distribution

The Anderson-Darling test statistic is defined as $A^2 = -N - S$,

where $S = \sum_{i=1}^{N} \frac{(2i-1)}{N} [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))]$, F is the cumulative distribution function of the

specified distribution. Note that the Y_i are the *ordered* data. The critical values for the Anderson-Darling test are dependent on the specific distribution that is being tested. Tabulated values and formulas are available in literature for a few specific distributions (normal, lognormal, exponential, Weibull, logistic, extreme value type 1). The test is a one-sided test and the hypothesis that the distribution is of a specific form is rejected if the test statistic, A, is greater than the critical value.

9.4 Homogeneity of error variances

A crude method for detecting the heterogeneity of variances is based on scatter plots of means and variance or range of observations or errors, residual *vs* fitted values, etc. To be clearer, let Y_{ij} be the observation pertaining to *i*th treatment ($i = 1,2,...,v$) in the *j*th replication ($j = 1,2,...,r_i$). Compute the mean and variance for each treatment across the replications (the range can be used in place of variance) as

$$\text{Mean} = \bar{Y}_{i.} = \frac{1}{r_i} \sum_{j=1}^{r_i} Y_{ij} \; ; \; \text{Variance} = S_i^2 = \frac{1}{r_i - 1} \sum_{j=1}^{r_i} (Y_{ij} - \bar{Y}_{i.})^2$$

Draw the scatter plot of mean *vs* variance (or range). If $S_{i.}^{2}$'s ($i = 1,2,...,v$) are equal (constant) or nearly equal, then the variances are homogeneous. Based on these scatter plots, the heterogeneity of variances can be classified into two types:

1.   where the variance is functionally related to mean.

2.   where there is no functional relationship between the variance and the mean.

The first kind of variance heterogeneity is usually associated with the data whose distribution is non-normal *viz.,* negative binomial, Poisson, binomial, etc. The second kind of variance heterogeneity usually occurs in experiments, where, due to the nature of treatments tested, some treatments have errors that are substantially higher (lower) than others.

The scatter-diagram of means and variances of observations for each treatment across the replications gives only a preliminary idea about homogeneity of error variances. Statistical tests to confirm it are needed. Before, testing homogeneity of error variances, one should test normality of the errors. If the errors are normally distributed, Bartlett's test is used for testing homogeneity of error variances. On the other hand, if the error distribution is non-normal, one can use Levene or Modified Levene test. We now describe these tests.

9.4.1 Bartlett's test for homogeneity of variances

Suppose that there are *v* independent samples drawn from same population and *i*th sample is of size r_i and $(r_1 + r_2 + ... + r_v) = N$. In the present case, the independent samples are the residuals of the observations pertaining to *v* treatments and *i*th sample size is the number of replications of the treatment *i*. One wants to test the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_v^2$ against the alternative hypothesis H_1 : at least two of the σ_i^2's are not equal, where σ_i^2 is the error variance for treatment *i*.

Let e_{ij} denotes the residual pertaining to the observation of treatment *i* from replication *j*, then it can easily be shown that the sum of residuals pertaining to a given treatment is zero.

In this test $S_i^2 = \frac{1}{r_i - 1} \sum_{j=1}^{r_i} (e_{ij} - \bar{e}_{i.})^2 = \frac{1}{r_i - 1} \sum_{j=1}^{r_i} e_{ij}^2$ is taken as unbiased estimate of σ_i^2. The procedure

involves computing a statistic whose sampling distribution is closely approximated by the χ^2 distribution with $v - 1$ degrees of freedom. The test statistic is

$$\chi_0^2 = 2.3026 \frac{q}{c}$$

and null hypothesis is rejected when $\chi_0^2 > \chi_{\alpha, v-1}^2$, where $\chi_{\alpha, v-1}^2$ is the upper α percentage point of χ^2 distribution with $v - 1$ degrees of freedom.

The steps to compute χ_0^2 are given in the sequel.

Step 1: Compute mean and variance of all *v*-samples.

Step 2: Obtain pooled variance $S_p^2 = \dfrac{\sum_{i=1}^{v} (r_i - 1) S_i^2}{N - v}$

Step 3: Compute $q = (N - v)\log_{10} S_p^2 - \sum_{i=1}^{v} (r_i - 1)\log_{10} S_i^2$

Step 4: Compute $c = 1 + \dfrac{1}{3(v-1)} \left(\sum_{i=1}^{v} (r_i - 1)^{-1} - (N - v)^{-1} \right)$

Step 5: Compute χ_0^2.

Bartlett's χ^2 test for homogeneity of variances is a modification of the normal-theory likelihood ratio test. While Bartlett's test has accurate Type I error rates and optimal power when the underlying distribution of the data is normal, it can be very inaccurate if that distribution is even slightly non-normal (Box, 1953). Therefore, Bartlett's test is not recommended for routine use.

An approach that leads to tests that are more robust to the underlying distribution is to transform the original values of the dependent variable to derive a *dispersion variable* and then to perform analysis of variance on this variable. The significance level for the test of homogeneity of variance is the *p*-value for the ANOVA *F*-test on the dispersion variable. Commonly used test for testing the homogeneity of variance using a dispersion variable is Levene Test given by Levene (1960). The procedure is described in the sequel.

9.4.2 Levene test for homogeneity of variances

The test is based on the variability of the residuals. The larger the error variance, the larger the variability of the residuals will tend to be. To conduct the Levene test, we divide the data into different groups based on the number of treatments. If the error variance is either increasing or decreasing with the treatments, the residuals in one treatment will tend to be more variable than those in other treatments. The Levene test then consists of simply *F*-statistic based on one way ANOVA used to determine whether the mean of absolute / square root deviation from mean are significantly different or not. The residuals are obtained from the usual analysis of variance. The test statistic is given as

$$F = \frac{\dfrac{\left\{ \sum_{i=1}^{v} (r_i - 1) \right\} \left\{ \sum_{i=1}^{v} r_i (\bar{d}_{i.} - \bar{d}_{..})^2 \right\}}{v - 1}}{\sum_{i=1}^{v} \sum_{j=1}^{r_i} (d_{ij} - \bar{d}_{i.})^2} \sim F_{(v-1), \sum_{i=1}^{v}(r_i - 1)}$$

where $d_{ij} = \left| e_{ij} - \bar{e}_i \right|$; $\bar{d}_{i.} = \dfrac{\sum_{j=1}^{r_i} d_{ij}}{r_i}$; $\bar{d}_{..} = \dfrac{\sum_{i=1}^{v} \sum_{j=1}^{r_i} d_{ij}}{\sum_{j=1}^{r_i} r_i}$ and e_{ij} is the *j*th residual for the *i*th plot, and $\bar{e}_i$ is the mean of the residuals of the *i*th treatment.

It may be noted that this test was modified by Brown and Forsythe (1974). In the modified test, the absolute deviation is taken from the median instead of mean in order to make the test more robust.

In this Chapter, the Bartlett's χ^2-test has been used for testing the homogeneity of error variances when the distribution of errors is normal.

Remark 9.1: In a block design, it can easily be shown that the sum of residuals within a given block is zero. Therefore, the residuals in a block of size 2 will be same in magnitude but opposite in sign. As a consequence, q in Bartlett's test and numerator in Levene test statistic become zero for the data generated from experiments conducted to compare only two treatments in a RCB design. Hence, the tests for homogeneity of error variances cannot be used for the experiments conducted to compare only two treatments in a RCB design. Inferences from such experiments may be drawn using Fisher-Behren t-test. Further, Bartlett's test cannot be used for the experimental situations where some of the treatments are singly replicated.

Remark 9.2: In a RCB design, it can easily be shown that the sum of residuals from a particular treatment is zero. As a consequence, the denominator of Levene test statistic is zero for the data generated from RCB designs with two replications. Therefore, Levene test cannot be used for testing the homogeneity of error variances for the data generated from RCB designs with two replications.

9.5 Remedial measures

The purpose of this Section is to describe some remedial measures for non-normal and/or heterogeneous data in greater details.

Data transformation is the most appropriate remedial measure in the situation where the variances are heterogeneous and are some functions of means. With this technique, the original data are converted to a new scale resulting into a new data set that is expected to satisfy the homogeneity of variances. Because a common transformation scale is applied to all observations, the comparative values between treatments are not altered and comparison between them remains valid.

Error partitioning is the remedial measure of heterogeneity that usually occurs in experiments, where, due to the nature of treatments tested some treatments have errors that are substantially higher (lower) than others.

Here, we shall concentrate on those situations where character under study is non-normal and variances are heterogeneous. Depending upon the functional relationship between variances and means, suitable transformation is adopted. The transformed variate should satisfy the following:

1. The variances of the transformed variate should be unaffected by changes in the means. This is also called the variance stabilizing transformation.
2. It should be normally distributed.
3. It should be one for which effects are linear and additive.

4.   The transformed scale should be such for which an arithmetic average from the sample is an efficient estimate of true mean.

The following are the three transformations, which are being used most commonly, in agricultural research.

a)   Logarithmic transformation
b)   Square root transformation
c)   Arc Sine or angular transformation

a) Logarithmic transformation

This transformation is suitable for the data where (a) the variance is proportional to square of the mean, or (b) the coefficient of variation (S.D./mean) is constant, or (c) the effects are multiplicative. These conditions are generally found in the data that are whole numbers and cover a wide range of values. This is usually the case when analysing growth measurements such as trunk girth, length of extension growth, weight of tree or number of insects per plot, number of egg mass per plant or per unit area, etc.

For such situations, it is appropriate to analyse *log Y* instead of actual data, *Y*. When data set involves small values or zeros, log $(Y + 1)$, log$(2Y + 1)$ or log $\left(Y + \frac{3}{8}\right)$ should be used instead of *log Y*. This transformation would make errors normal, when observations follow negative binomial distribution like in the case of insect counts.

b) Square-root transformation

This transformation is appropriate for the data sets where the variance is proportional to the mean. Here, the data consists of small whole numbers. For example, data obtained in counting rare events, such as the number of infested plants in a plot, the number of insects caught in traps, number of weeds per plot, etc. This data set generally follows Poisson distribution and square root transformation approximates Poisson to normal distribution.

For these situations, it is better to analyse $\sqrt{Y}$ than analysing *Y*, the actual data. If *Y* is confirmed to small whole numbers then, $\sqrt{Y + \frac{1}{2}}$ or $\sqrt{Y + \frac{3}{8}}$ should be used instead of $\sqrt{Y}$.

c) Arc Sine transformation

This transformation is appropriate for the data on proportions, *i.e.,* data obtained from counts and expressed as decimal fractions and percentages. The distribution of percentages is binomial and this transformation makes the distribution normal. Since the role of this transformation is not properly understood, there is a tendency to transform any percentage

data using arc sine transformation. But only that percentage data that are derived from count data, such as % barren tillers (which is derived from the ratio of the number of non-bearing tillers to the total number of tillers), should be transformed and not the percentage data such as % protein or % carbohydrates, % nitrogen, etc. which are not derived from count data. For these situations, it is better to analyse $\sin^{-1}(\sqrt{Y})$ than Y, the actual data. A value of Y as 0% should be substituted by $\left(\dfrac{1}{4n}\right)$ and a value of Y as 100% should be substituted by $\left(100 - \dfrac{1}{4n}\right)$, where n is the number of units upon which the percentage data is based.

It may be noted here that not all percentage data need to be transformed, and even if they do, arc sine transformation is not the only transformation possible. The following rules, as given by Gomez and Gomez (1984), may be useful in choosing the proper transformation scale for percentage data derived from count data.

Rule 1: The percentage data lying within the range of 30 to 70% is homogeneous and no transformation is needed.

Rule 2: For percentage data lying within the range of either 0 to 30% or 70 to 100%, but not both, the square root transformation should be used.

Rule 3: For percentage data that do not follow the ranges specified in Rule 1 or Rule 2, the Arc Sine transformation should be used.

The rules just described may only be some indicative guidelines. However, it is always better to test the normality of errors and homogeneity of error variances. If the errors follow a normal distribution and error variances are homogeneous, then no transformation is required. If any one or both of these assumptions are not satisfied, then appropriate transformation may be used.

For performing different transformations of a variable X in SAS, one can use the following statements between INPUT and CARDS statement.

```
/*Arc sine transformation (variable X to be transformed is in % with values 0 to 100)*/
x1=ARSIN(SQRT(x/100))*180*7/22;
x2=100-1/400;
IF st=0 THEN x1=ARSIN(SQRT(1/400))*180*7/22;
IF st=100 THEN x1= ARSIN(SQRT(x2/100))*180*7/22;
However if variable X to be transformed is in proportions with values 0 to 1, then
X1=ARSIN(SQRT(x))*180*7/22;
/*Square root transformation for Variable X*/
X3=SQRT(X+0.5);
/*Logarithmic transformation for variable X;*/
X4=LOG(X+1);
```

The transformations discussed above are a particular case of the general family of transformations known as Box-Cox transformation.

d) Box-Cox transformation

The description thus far clearly implies that if the relation between the variance and the mean of the observations is known then this information can be utilized in selecting the form of the transformation. We now elaborate on this point and show how it is possible to estimate the form of the required transformation from the data. The transformation suggested by Box and Cox (1964) is a power transformation of the original data. Let Y_{ut} be the observation pertaining to the uth plot; then the power transformation implies that we use Y_{ut}'s as

$$Y_{ut}^* = Y_{ut}^\lambda .$$

The transformation parameter λ in $Y_{ut}^* = Y_{ut}^\lambda$ may be estimated simultaneously with the other model parameters (overall mean and treatment effects) using the method of maximum likelihood. The procedure consists of performing, for the various values of λ, a standard analysis of variance on

$$Y_{ut}^{(\lambda)} = \begin{cases} \dfrac{Y_{ut}^\lambda - 1}{\lambda \dot{Y}_{ut}^{\lambda-1}} & \lambda \neq 0 \\[2em] \dot{Y}_{ut} \log Y_{ut} & \lambda = 0 \end{cases} \tag{9.1}$$

where $\dot{Y}_{ut} = \log^{-1}\left[(1/n)\sum_{u=1}^{N}\sum_{t=1}^{n_u}\log Y_{ut}\right]$.

$\dot{Y}_{ut}$ is the geometric mean of the observations. The geometric mean of n numbers $x_1, x_2, ..., x_n$ is defined as the nth root of their product, i.e., $x = \left(x_1 x_2 x_n\right)^{1/n}$. The maximum likelihood estimate of λ is the value for which the error sum of squares, say SSE (λ), is minimum. Notice that we cannot select the value of λ by directly comparing the error sum of squares from analysis of variance on Y^λ because for each value of λ the error sum of squares is measured on a different scale. Equation (9.1) rescales the responses so that the error sums of squares are directly comparable. This is a very general transformation and the commonly used transformations follow as particular cases. The particular cases for different values of λ are given below.

λ	Transformation
1	No Transformation
½	Square Root
0	Log
−1/2	Reciprocal Square Root
−1	Reciprocal

Remark 9.3: If any one of the observations is zero then the geometric mean is undefined. In the expression (9.1), geometric mean is in denominator, so it is not possible to compute that expression. For solving this problem, one may add a small quantity to each of the observations.

It should be emphasized that transformation, if needed, must take place right at the beginning of the analysis, all fitting of missing plot values, all adjustments by covariance etc. being done with the transformed variate and not with the original data. At the end, when the conclusions have been reached, it is permissible to 're-transform' the results so as to present them in the original units of measurement. This may be done only to render the results more intelligible.

As a result of this transformation followed by back transformation, the means will rather be different from those that would have been obtained from the original data. A simple example is that without transformation, the mean of the numbers 1, 4, 9, 16 and 25 is 11. Suppose a square root transformation is used to give transformed values as 1, 2, 3, 4 and 5. The mean now is 3, which after back- transformation gives a value 9. Generally the difference will not be so big because the data do not usually vary as much as in this Example, but logarithmic and square root transformations always lead to a reduction of the mean, just as angles of equal formation usually lead to its moving away from the central value of 50%.

However, in practice, computing treatment means from original data is more frequently used because of its simplicity, but this may change the order of ranking of converted means for comparison. Although transformations make possible a valid analysis, they can be very awkward. For example, although a significant difference can be worked out in the usual way for means of the transformed data, none can be worked out for the treatment means after back- transformation.

9.6 Non-parametric tests in the analysis of experimental data

When the data remains non-normal and/or heterogeneous even after transformation, recourse is made to non-parametric test procedures. A lot of attention is being paid to develop non-parametric tests for analysis of experimental data. Most of these non-parametric test procedures are based on rank statistic. The rank statistic has been used in development of these tests as the statistic based on ranks is distribution free, easy to calculate, and simple to explain and understand.

Another reason for use of rank statistic is due to the well known result that the average rank approaches normality quickly as *n* (number of observations) increases, under the rather general conditions, while the same might not be true for the original data {see *e.g.* Conover and Iman (1976, 1981)}. The non-parametric test procedures available in the literature cover completely randomized design, randomized complete block design, balanced incomplete block design, design for bioassays, split plot design, cross-over design and so on. For an excellent and elaborate discussions on non-parametric tests in the analysis of experimental data, one may refer to Siegel and Castellan Jr. (1988), Deshpande, Gore and Shanubhogue (1995), Sen (1996), and Hollander and Wolfe (1999).

Kruskal-Wallis test can be used for the analysis of data from completely randomized design. Skillings and Mack test helps in analysing the data from a general block design. Friedman test and Durbin test are particular cases of this test. Friedman test is used for the analysis of data from RCB design and Durbin test for the analysis of data from BIB design. We now describe some of these tests for analysis of experimental data generated through block designs.

9.6.1 Kruskal-Wallis Test

This test is used for testing the equality of population mean of several groups or the treatment effects. Hence, this test is quite useful for the analysis of experimental data generated through a completely randomized design. Consider that there are v treatments and the ith treatment is replicated r_i times; $i = 1, 2, \ldots, v$, such that $N = \sum_{i=1}^{v} r_i$. The data generated through CRD can be represented by usual one-way classified linear model as

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \qquad\qquad i = 1, 2, \ldots, v;\ j = 1, 2, \ldots, r_i.$$

where y_{ij} is the yield or response of jth replication of ith treatment, μ is the general mean, τ_i is the ith treatment effect, ϵ_{ij} is the error due to ith treatment and jth observation.

The experimenter is interested to test the equality of treatments effects against the alternative that at least two of the treatments effects are not equal. In other words, we want to test the null hypothesis H_0: $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_v = \tau$ (say) against the alternative H_1: at least two of the τ_i's are different.

To test the above null hypothesis using the Kruskal-Wallis Statistic, we rank all N observations by giving rank 1 to smallest observation and N to largest observation. Once this is done, we obtain the sum and average of the ranks of the observations pertaining to each of the treatments. Now, if the treatment effects are equal, then the average ranks are expected to be same. The differences, if any, would be due to sampling fluctuations. The Kruskal-Wallis statistic is based on the assessment of the differences among the average ranks. This may be explained as below:

Let R_{ij} be the rank of y_{ij}; $i = 1, 2, 3, \ldots, v$; $j = 1, 2, \ldots, r_i$ and $R_i = \sum_{j=1}^{r_i} R_{ij}$ (the sum of the ranks of the observations pertaining to the ith treatment) and $\bar{R}_i = R_i / r_i$ (the average of the ranks of the observations pertaining to the ith treatment). Let $\bar{R}$ be the mean of the all $\bar{R}_i$. The Kruskal-Wallis statistic is, then, given by

$$T = \frac{12}{N(N+1)} \sum_{i=1}^{v} r_i (\bar{R}_i - \bar{R})^2$$

$$\Rightarrow T = \frac{12}{N(N+1)} \sum_{i=1}^{v} r_i \left(\frac{R_i}{r_i} - \frac{N+1}{2} \right)^2$$

$$= \frac{12}{N(N+1)} \sum_{i=1}^{v} \frac{R_i^2}{r_i} - 3(N+1)$$

The method of determining the significance of the observed values of T depends on the number of treatments (v) and their replications (r_i). The null distribution (distribution under null hypothesis) of T for $v = 3$ and $r_i \leq 5$ is extensively tabulated and available in several texts.

For a ready reference, the Table has been included as Appendix-I of this Chapter. In other cases, under null hypothesis, T may be approximated by the χ^2 with $(v-1)$ degree of freedom.

Remark 9.4: When ties occur between two or more observations, each observation is given the mean of ranks for which it is tied. The T-statistic after correcting the effect of ties is computed by using following formula

$$T = \frac{\left[\frac{12}{N(N+1)}\sum_{i=1}^{v}\frac{R_i^2}{r_i}\right]-3(N+1)}{1-\left[\sum_{s=1}^{g}(t_s^3-t_s)\right]/(N^3-N)}$$

where g is number of treatments of different tied ranks, t_s is the number of tied ranks in the sth treatment. The effect of correcting for ties is to increase the value of T and thus to make the result more significant than it would have been had no correction been made. Therefore, if one is able to reject null hypothesis without making the correction, one will be able to reject H_0 at an even more stringent level of significance if the correction is used.

Pair-wise comparisons

If the Kruskal-Wallis test rejects the null hypothesis of equality of v treatment effects, it indicates that at least two of the treatment effects are unequal. It does not tell the researcher which one are different from each other. Therefore, a test procedure for making pair wise comparisons is needed. For this, the null hypothesis $H_0 : \tau_i = \tau_{i'}$ against $H_1 : \tau_i \neq \tau_{i'}$ $\forall i \neq i' = 1, 2, ..., v$ can be tested at α level of significance by using the inequality

$$\left|\bar{R}_i - \bar{R}_{i'}\right| \geq z_p \sqrt{\frac{N(N+1)}{12}\left(\frac{1}{r_i}+\frac{1}{r_{i'}}\right)} \quad \forall i, i', i \neq i' = 1, 2, ..., v$$

where $p = \alpha/v(v-1)$ and z_p is the quantile of order $1-p$ under the standard normal distribution. From the above, we can say that the least significant difference between the treatments i and i' is

$$c_{ii'} = z_p \sqrt{\frac{N(N+1)}{12}\frac{1}{r_i}+\frac{1}{r_{i'}}}$$

Therefore, if $\left|\bar{R}_i - \bar{R}_{i'}\right| > c_{ii'}$ then the difference between i and i' treatment effects is considered significant at α level of significance. The above procedure is illustrated with the help of the following example.

Example 9.1: An experiment was conducted with 21 animals to determine if the four different feeds have the same distribution of weight gains on experimental animals. The feeds 1, 3 and 4 were given to 5 randomly selected animals and feed 2 was given to 6 randomly selected animals. The data obtained is presented in Table 9.4.

Table 9.4: Weight gain data of animals

Feeds	Weight gains (kg)					
1	3.35	3.80	3.55	3.36	3.81	
2	3.79	4.10	4.11	3.95	4.25	4.40
3	4.00	4.50	4.51	4.75	5.00	
4	3.57	3.82	4.09	3.96	3.82	

We use Kruskal-Wallis test to analyse the data in Table 9.4. We arrange the data in ascending order and give the ranks 1 to 21 to the observations. The ranks are then arranged feed wise as given in Table 9.5.

Table 9.5: Rank of the observations

Feeds	Ranks of Weight gains						Sum of ranks (R_i)	Average of Ranks ($\bar{R}$)
1	1	6	3	2	7		19	3.80
2	5	14	15	10	16	17	77	12.83
3	12	18	19	20	21		90	18.00
4	4	8.5	13	11	8.5		45	9.00

Then the Kruskal-Wallis test statistic is obtained as:

$T = \dfrac{NUM}{DEN}$, where

$$NUM = \left[\frac{12}{21*22} \left[\frac{(19)^2}{5} + \frac{(77)^2}{6} + \frac{(90)^2}{5} + \frac{(45)^2}{5} \right] - 3*22 \right] = 80.139 - 66.000 = 14.139,$$

$DEN = 1 - (2^3 - 2)/(21^3 - 21) = 1539/1540 = 0.999$. Here $t_s = 2$ for $s = 4$, because only two observations are tied for feed 4.

Thus $T = 14.139/0.999 = 14.153$

The tabulated value of χ^2 at 3 degree of freedom at 5% level of significance is 7.815 and the calculated value is 14.140. It, therefore, follows that the feed effects differ significantly.

Pair-wise comparisons for the feeds

Here $r_1 = r_3 = r_4 = 5$; $r_2 = 6$, $N = 21$, $v = 4$ and $v(v - 1) = 12$. Let $a = 0.05$. Then $p = 0.05/12 = 0.00417$. For this value of p, the value of z_p obtained from the tables is $z_p \approx 2.64$. We can compute $c_{ii'}$ as

$$c_{12} = c_{32} = c_{42} = 2.64 \sqrt{\frac{21*22}{12} \left(\frac{1}{5} + \frac{1}{6} \right)} = 9.919$$

$$c_{13} = c_{14} = c_{34} = 2.64 \sqrt{\frac{21*22}{12}\left(\frac{1}{5}+\frac{1}{5}\right)} = 10.360.$$

Thus $|\bar{R}_1 - \bar{R}_2| = 9.03$; $|\bar{R}_1 - \bar{R}_3| = 14.20$; $|\bar{R}_1 - \bar{R}_4| = 5.20$; $|\bar{R}_2 - \bar{R}_3| = 5.17$; $|\bar{R}_2 - \bar{R}_4| = 3.83$; $|\bar{R}_3 - \bar{R}_4| = 9.00$.

Here we see that $|\bar{R}_1 - \bar{R}_3| > c_{13}$ so the effect of feeds 1 and 3 are significantly different at 5% of level of significance while all other pairs of feeds effects do not differ significantly.

9.6.2 Friedman test

The Kruskal-Wallis test is useful for the data generated through completely randomized designs. A completely randomized design is used when experimental units are homogeneous in a block. However, there do occur experimental situations where one can find a factor (called nuisance factor), which, though not of interest to the experimenter, does contribute significantly to the variability in the experimental material. Various levels of this factor are used for blocking. For the experimental situations where there is only one nuisance factor, block designs are used. The simplest and most commonly used block design by the agricultural research workers is a randomized complete block (RCB) design. The problem of non-normality of data may also occur in RCB design as well. Friedman test is useful for such situations (see Friedman, 1937). Let there be v treatments and $N = vb$ experimental units arranged in b blocks of size v each. Each treatment appears exactly once in each block. The data generated through a RCB design can analysed using the following linear model

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \qquad i = 1, 2, \ldots, v; \; j = 1, 2, \ldots, b.$$

where y_{ij} is the yield (response) of the ith experimental unit receiving the treatment in jth block, τ_i is the effect due to ith treatment, β_j is the effect of jth block, ϵ_{ij} is random error in response. The interest of the experimenter is to test the equality of treatment effects. In other words, we want to test the null hypothesis H_0: $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_v = \tau$ (say) against the alternative H_1: at least two of the t_i's are different. For using Friedman test, we proceed as follows.

Arrange the observations in v rows (treatments) and b columns (blocks). The observations in the different rows are independent and those in different columns are dependent. Rank all the observations in a column (block), *i.e.* ranks are assigned separately for each block. Let R_{ij} be the rank of the observation pertaining to ith treatment in the jth block. Then $1 \leq R_{ij} \leq v$. As the ranking has been done within blocks from 1 to v, therefore sum of ranks in jth block is

$$R_j = \sum_{i=1}^{v} R_{ij} = \frac{v(v+1)}{2} \quad \text{and} \quad \bar{R}_j = \frac{v+1}{2} \quad \text{and the variance is} \quad \frac{v^2-1}{12}. \text{ The sum of ranks for } i\text{th}$$

treatment is $R_i = \sum_{j=1}^{b} R_{ij}$.

If all the treatment effects are same (under H_0), then we expect each R_i to be equal $b(v+1)/2$, that is, under H_0,

$$E(R_i) = \frac{1}{v} \frac{bv(v+1)}{2} = \frac{b(v+1)}{2}.$$

The sum of squared deviations of R_i's from $E(R_i)$ is, therefore, a measure of the differences in the treatment effects.

$$\text{Let} \quad S = \sum_{i=1}^{v} \left\{ R_i - \frac{b(v+1)}{2} \right\}^2$$

The Friedman test statistic is then defined as

$$T = \frac{12S}{bv(v+1)} = \frac{12}{bv(v+1)} \sum_{i=1}^{v} \left\{ R_i - \frac{b(v+1)}{2} \right\}^2$$

The method of determining the probability of occurrence of an observed value of T when H_0 is true depends upon the sizes of v and b. For small values of b and v, the null distribution of T has been tabulated. For a ready reference, the table has been included in Appendix-II. For large b and v, the associated probability may be approximated by the χ^2 distribution with $v - 1$ degrees of freedom.

Remark 9.5: When there are ties among the ranks for any given block, the statistics T must be corrected to account for changes in the sampling distribution. So if ties occur then we use following statistic

$$T = \frac{12 \sum_{i=1}^{v} R_i^2 - 3b^2 v(v+1)^2}{bv(v+1) + \dfrac{\left(bv - \sum_{j=1}^{b} \sum_{s=1}^{g_j} t_{js}^3 \right)}{(v-1)}} \sim \chi^2_{(v-1)}$$

where g_j is the number of sets of tied ranks in the jth block and t_{js} is the size of the jth set of tied ranks in the ith block.

Pair-wise comparisons

When the Friedman test rejects the null hypothesis that treatment effects are homogeneous, it is of interest to identify significant difference between the paired treatments. Therefore, a test procedure for making pair wise comparisons is needed. The null hypothesis $H_0 : \tau_i = \tau_{i'}$ against $H_1 : \tau_i \neq \tau_{i'} \ \forall i \neq i' = 1,2,...,v$ can be tested at α level of significance using the inequality.

$$|R_i - R_{i'}| \geq z_p \sqrt{\frac{bv(v+1)}{6}} \quad \text{for all } i,\, i' = 1,2,...,v,\, i \neq i'$$

where $p = \alpha/v(v-1)$ and z_p is the quantile of order $1 - p$ under the standard normal distribution. From the above, we can say that the least significant difference between the treatments i and i' is

$$c = z_p \sqrt{\frac{bv(v+1)}{6}}$$

If $|R_i - R_{i'}| > c$ then the difference between treatments i and i' is considered as significantly different at α level of significance. The above procedure is illustrated with the help of following example:

Example 9.2: An animal feeding experiment involving 8 different rations was laid out in a RCB design using 24 animals in 3 groups of size 8 each. The grouping was done on the basis of initial body weight.

Table 9.6: Data from animal feeding experiment

Rations	Block		
	1	2	3
1	138.46	400.00	461.50
2	387.60	369.20	384.60
3	436.90	430.80	467.70
4	424.60	461.50	421.50
5	430.70	455.30	406.10
6	424.61	384.60	384.50
7	424.59	350.80	307.60
8	360.00	400.10	400.00

The first step is to check the normality of observations. For testing the normality of observations we use the Shapiro-Wilk Test and Kolmogorov-Smirnov test. Here H_0: Observations come from normal population against H_1: Observations do not come from normal population.

Table 9.7: Result of normality test

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	DF	Significance level	Statistic	F	Significance level
Residual	0.200	24	0.014	0.873	24	0.010

We can see that the observations are non-normal at 5% level of significance. Now we use usual method of ANOVA and get the result as shown in Table 9.8.

Table 9.8: ANOVA results

Source	DF	SS	MS	F Value	Prob > F
Replication	2	3889.956	1944.978	0.405	0.6747
Treatment	7	32075.242	4582.177	0.954	0.4988
Error	14	67273.420	4805.244		
Total	23	103238.6190			

From this analysis, we can see that treatments are not significantly different. Now we use the Friedman Test for analysis of the same data. After ranking the observations within each block, we get Table 9.9.

Table 9.9: Ranked observations in each block

	Block			Sum of ranks (R_i)
Ration	1	2	3	
1	1	4	7	12
2	3	2	3	8
3	8	6	8	22
4	5	8	6	19
5	7	7	5	19
6	6	3	2	11
7	4	1	1	6
8	2	5	4	11

Here $b = 3$ and $v = 8$, so the Friedman statistic is

$$T = \frac{12}{3*8(8+1)}\left[(12)^2 + (8)^2 + (22)^2 + (19)^2 + (19)^2 + (11)^2 + (6)^2 + (11)^2\right] - 3*3*9$$

$= 94 - 81 = 13$.

The tabulated value of χ^2 at 7 degrees of Freedom and 10% level of significance is 12.017 while the calculated value of χ^2 is 13.000. So the rations are significantly different at 10% level of significance. Here the probability of getting a value of χ^2 greater than 13.000 is 0.0721.

Pair-wise comparisons

As mentioned above, here $b = 3$, $v = 8$ and $v(v - 1) = 56$. Let $\alpha = 0.1$ then $p = 0.1/56 = 0.00179$. For this value of p, the value of z_p obtained from the tables is $z_p = 2.1$. Then we calculate

$$c = 2.1\sqrt{\frac{3*8(8+1)}{6}} = 12.6$$

So the critical difference of rank sum is 12.6. If $|R_i - R_{i'}| > 12.6$, the treatments i and i' are significantly different. For example $|R_2 - R_3| = 14$ is more than 12.6, so we conclude that these

two treatments are significantly different at 10% level of significance. Similarly, $|R_3 - R_7|$, $|R_4 - R_7|$ and $|R_5 - R_7|$ are greater than 12.6. So these pair of treatments also differ significantly at 10% level of significance.

9.6.3 Durbin's Test

In this Section just above, we have discussed the non-parametric analysis of experimental data of a RCB design. In a RCB design, the number of experimental units required in each block is same as the number of treatments. However, when the number of treatments increases, the blocks become large and it is not possible to maintain homogeneity with blocks. If an experimenter persists with a RCB design it results into large intra block variances and hence reduced precision on treatment comparisons. To circumvent this problem, recourse is made to incomplete block designs. Many a time an experimenter may have to use an incomplete block design because of the nature of experimental units. The simplest design among the class of incomplete block designs is a balanced incomplete block (BIB) design described in Chapter 5. The standard notations used for describing the parameters of a BIB design are (i) v, the number of treatments, (ii) b, the number of blocks, (iii) r, the replication number of treatments, (iv) k, the number of experimental units per block ($k < v$), and (v) λ, the number of blocks in which a given treatment pair occurs together

The model of a BIB design is same as described in Chapter 5. For comparing the treatments from the non-normal data generated from a BIB design, Durbin (1951) proposed a test statistic. We want to test the equality of the treatment effects *i. e.* the null hypothesis H_0: $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_v = \tau$ (say) against the alternative H_1: at least two of the τ_i's are different.

To test the above hypothesis, rank the observations y_{ij} from 1 to k within a block. Let R_{ij} be the rank of y_{ij}. Then following the lines of Friedman test, Durbin test statistic is given by

$$T = \frac{12(v-1)}{rv(k-1)(k+1)} \sum_{i=1}^{v} \left[R_i - \frac{r(k+1)}{2} \right]^2$$

$$= \frac{12(v-1)}{rv(k-1)(k+1)} \sum_{i=1}^{v} R_i^2 - \frac{3r(v-1)(k+1)}{(k-1)} \sim \chi^2_{(v-1)}$$

The test rejects H_0 if T is more than the cut-off point. The cut off point is obtained by referring to the chi-square distribution with $(v-1)$ degree of freedom. The exact test can be obtained by rejecting H_0 when $T \geq m_\alpha$, where some values of m_α are given in Skillings and Mack (1981).

Pair-wise comparisons

When the Durbin test rejects the null hypothesis that the treatment effects are homogeneous, it is of interest to identify pairs of treatments that differ significantly. Therefore, a test procedure for making pair wise treatments comparisons is needed. The null hypothesis $H_0 : \tau_i = \tau_{i'}$ against $H_1 : \tau_i \neq \tau_{i'}$ $\forall i \neq i' = 1, 2, \ldots, v$ can be tested at α level of significance using the

$$\left| R_i - R_{i'} \right| \geq z_p \sqrt{\frac{vr(k+1)(k-1)}{6(v-1)}} \quad \forall \ i \neq i', i, i'=1,2,\dots,v$$

where $p = \alpha/v(v-1)$ and z_p is the quantile of order $1 - p$ under the standard normal distribution. From the above, we can say that the least significant difference between the treatments i and i' is

$$c = z_p \sqrt{\frac{vr(k+1)(k-1)}{6(v-1)}}$$

If $\left| R_i - R_{i'} \right| > c$ then the difference between treatments i and i' is considered significant at α level of significance. The above procedure is illustrated with the help of following example:

Example 9.3: In an experiment to compare palatability of four varieties of rice (cooked), four judges were asked to rank three varieties each. The results obtained are given in Table 9.10.

Table 9.10: Data from palatability study

Rice Variety	Judges				Sum of ranks (R_i)
	1	2	3	4	
1	1	2	1	-	4
2	3	1	-	2	6
3	2	-	2	1	5
4	-	3	3	3	9

Using the Durbin test statistic gives

$$T = \frac{12(4-1)}{3*4*(3-1)(3+1)}\left[(4-6)^2 + (6-6)^2 + (5-6)^2 + (9-6)^2\right] = 5.25$$

The tabulated value of χ^2 at 3 degree of freedom and at 5% level of significance is 7.815 while the calculated value of χ^2 is 5.250. So the treatments do not differ significantly at 5% level of significance. Here the probability of getting a value of χ^2 greater than 5.250 is 0.0724.

The R code for analysis of data is given below

```
d35=read.table("variety_ranking.txt",header=TRUE)
attach(d35)
names(d35)
library(agricolae)
out=durbin.test(judge,variety,rank)
out
detach(d35)
```

9.6.4. Skillings and Mack test

In some experimental situations, even the use of a BIB design may not be feasible and a recourse may have to be made to use of a partially balanced incomplete block (PBIB) design. In some other experimental situations, a non-proper (unequal number of experimental units in blocks) block design may be useful. Skillings and Mack (1981) proposed a Friedman-type test statistic, which is useful for the analysis of data generated from any binary block design. Let $v, b, r_1, r_2, \cdots, r_i, \cdots, r_v, k_1, k_2, \cdots, k_j, \cdots k_b)$ represent a binary block design in which v treatments are arranged in b blocks such that jth block contains $k_j \geq 2$ distinct treatments and ith treatment is replicated r_i times; $i = 1, 2, \ldots, v, j = 1, 2, \ldots, b$. For the analysis of experimental data generated through a binary block designs, we make use of test statistic given by Skillings and Mack (1981). To compute the test statistic, we find adjusted treatment sums for ranked data. For this we proceed as follows:

1. Within each block, rank the observations from 1 to k_j, where k_j is the number of experimental units (or treatments) in the jth block.

2. Let R_{ij} be the rank of y_{ij}, if the observation is present; otherwise, let $R_{ij} = (k_j+1)/2$

3. Compute an adjusted treatment sum of ranks for the ith treatment, namely

$$A_i = \sum_{j=1}^{b} \left[12/(k_j+1)\right]^{1/2} \left[R_{ij} - (k_j+1)/2\right]$$

Let Σ denote the covariance matrix (diagonal elements are variances and off-diagonal elements are covariances) of the random vector $\mathbf{A}' = (A_1, \ldots, A_v)$. The covariance structure of the R_{ij}'s is well known and in this case only minor modifications are required because of missing cells. In block j, under H_0: $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_v = \tau$ (say), we have

$$Var(R_{ij}) = \begin{cases} (k_j+1)(k_j-1)/12 & \text{if treatment } i \text{ is present in block } j \\ 0, & \text{otherwise} \end{cases},$$

$$Cov(R_{ij}, R_{i'j'}) = \begin{cases} -(k_j+1)/12 & \text{if } j = j', \ i \neq i' \ n_{ij} = 1 \ \text{and} \ n_{i'j} = 1 \\ 0, & \text{otherwise} \end{cases},$$

where n_{ij} is the number of times treatment i appears in block j. Thus

$$Var(A_i) = \sum_{j=1}^{b}(k_j-1)n_{ij}, \ i = 1, 2, \ldots, v$$

and $\quad Var(A_i) = \sum_{j=1}^{b}(k_j-1)n_{ij}, \ 1 \leq i \neq i' \leq v.$

Let $\lambda_{ii'}$ denote the number of blocks containing observations for both treatments i and i'. It can be seen by inspection that under H_0, the matrix $\Sigma = ((\sigma_{ii'}))$ can be rewritten as

$\quad \sigma_{ii'} = -\lambda_{ii'}, \ 1 \leq i \neq i' \leq v$

and

$$\sigma_{ii} = \sum_{\substack{i'=1 \\ i'\neq i}}^{v} \lambda_{ii'} = -\sum_{\substack{i'=1 \\ i'\neq i}}^{v} \sigma_{ii'} \, , \, i = 1, 2, \ldots, v.$$

Thus the elements of Σ are simple to obtain. The off-diagonal elements are $(-\lambda_{ii'})$, and the diagonal elements are the negative of the sum of off-diagonal elements in that row. We note that the covariance matrix Σ is singular, because the sum of the rows (columns) is always zero. In any connected block design, the rank of Σ will be $v - 1$. The proposed test statistic is of the form

$$T = \mathbf{A'}\Sigma^-\mathbf{A}$$

where Σ^- is a generalized inverse of Σ. T follows an approximate x^2-distribution with degrees of freedom as the rank of Σ. The above procedure is illustrated with the help of following example:

Example 9.4: An experimenter is interested in comparing the performance of four chemicals in a hilly area in terms of the effect of chemicals on the crop yield. There are 9 terraces in the hill, but these are of unequal sizes and it is not possible to accommodate all the chemicals on all the terraces. The data in kg/plot is given in Table 9.11.

Table 9.11: Chemical experiment data

Chemicals	Terraces (Blocks)								
	1	2	3	4	5	6	7	8	9
A	3.2	3.1	4.3	3.5	3.6	4.5		4.3	3.5
B	4.1	3.9	3.5	3.6	4.2	4.7	4.2	4.6	
C	3.8	3.4	4.6	3.9	3.7	3.7	3.4	4.4	3.7
D	4.2	4.0	4.8	4.0	3.9			4.9	3.9

Now we rank the data within each block. If an observation corresponding to some treatment is not present in a block, then we use $R_{ij} = (k_j +1)/2$. We get the Table 9.12 of ranks.

Table 9.12: Ranked data within each terrace (block)

Chemical	Terrace (Block)								
	1	2	3	4	5	6	7	8	9
A	1	1	2	1	1	2	1.5*	1	1
B	3	3	1	2	4	3	2	3	2*
C	2	2	3	3	2	1	1	2	2
D	4	4	4	4	3	2*	1.5*	4	3

* Represents rank for the missing observations.

Now we calculate adjusted treatment sums (A_i) as in Table 9.13.

Table 9.13: Adjusted treatment means

Block	Chemical			
	A	B	C	D
1	-2.324	0.775	-0.775	2.324
2	-2.324	0.775	-0.775	2.324
3	-0.775	-2.324	0.775	2.324
4	-2.324	-0.775	0.775	2.324
5	-2.324	2.324	-0.775	0.775
6	0.000	1.732	-1.732	0.000
7	0.000	1.000	-1.000	0.000
8	-2.324	0.775	-0.775	2.324
9	-1.732	0.000	0.000	1.732
Adjusted Treatments Sum of Ranks (A_i)	14.127	4.282	-4.282	14.127

So, $A_i = (-14.127 \quad 4.282 \quad -4.282 \quad 14.127)'$

The covariance matrix Σ is obtained by counting the number of times treatment pairs occur

together. Thus $\Sigma = \begin{bmatrix} 22 & -7 & -8 & -7 \\ -7 & 21 & -8 & -6 \\ -8 & -8 & 23 & -7 \\ -7 & -6 & -7 & 20 \end{bmatrix}$

A generalized inverse Σ^- of Σ is

$$\Sigma^- = \begin{bmatrix} 0.0256 & -0.0089 & -0.0078 & -0.0089 \\ -0.0089 & 0.0269 & -0.0077 & -0.0102 \\ -0.0078 & -0.0077 & 0.0245 & -0.0090 \\ -0.0089 & -0.1025 & -0.0090 & 0.0281 \end{bmatrix}$$

Now,

$$= T = \mathbf{A'\Sigma^- A}$$

$$[-14.127 \ 4.282 \ -4.282 \ 14.127] \begin{bmatrix} 0.0256 & -0.0089 & -0.0078 & -0.0089 \\ -0.0089 & 0.0269 & -0.0077 & -0.0102 \\ -0.0078 & -0.0077 & 0.0245 & -0.0090 \\ -0.0089 & -0.1025 & -0.0090 & 0.0281 \end{bmatrix} \begin{bmatrix} -14.127 \\ 4.282 \\ -4.282 \\ 14.127 \end{bmatrix} = 15.49$$

The tabulated value of χ^2 at 3 degrees of freedom and 5% level of significance is 7.815 while the calculated value of χ^2 is 15.49. So the treatments are significantly different at 5% level of significance. Here the probability of getting a value of χ^2 greater than 15.49 is 0.0014.

The statistic is quite general and the commonly used Friedman test statistic and Durbin test statistic discussed above are the particular cases of this test. For example, in a RCB designs all $n_{ij} = 1$, all $\lambda_{ii'} = b$ and $k_j = v$, therefore, for a RCB design

$$\Sigma = vb\left(\mathbf{I} - v^{-1}\mathbf{1}_v\mathbf{1}_v'\right) \text{ and } A_i = \left(\frac{12}{v+1}\right)^{1/2} \sum_{j=1}^{b}\left(R_{ij} - \frac{v+1}{2}\right).$$

Substituting, these values in T gives

$$T = (bv)^{-1}\sum_{i=1}^{v} A_i^2$$

$$T = \frac{12}{bv(v+1)}\sum_{i=1}^{v} R_i^2 - 3b(v+1)$$

Now in case of a BIB design, all blocks have $k < v$ observations and the number of blocks in which any pair of treatments occur together is λ. Therefore, for a BIB design $\Sigma = \left[\{r(k-1)+\lambda\}\mathbf{I} - \lambda\mathbf{1}_v\mathbf{1}_v'\right]$. The statistic T reduces to

$$T = (\lambda v)^{-1}\sum_{i=1}^{v} A_i^2$$

$$T = \frac{12(v-1)}{rv(k-1)(k+1)}\sum_{i=1}^{v} R_i^2 - \frac{3r(v-1)(k+1)}{k-1}, \text{ since } \lambda(v-1) = r(k-1)$$

9.6.5. Gore test for multiple observations per plot

The test discussed above can also be used for analysis of data generated from a general block design where only one observation per plot is available. However, sometime more than one observation is available from each cell. The appropriate model for such situation is

$$y_{ijs} = \mu + \tau_i + \beta_j + \epsilon_{ijs}, \ i = 1, 2, \ldots, v; \ j = 1, 2, \ldots, b; \ s = 1, 2, \ldots, n_{ij}.$$

where y_{ijs} is the sth observation in the (i, j)th cell, n_{ij} is the number of observations in the (i, j)th cell, that is the number of experimental units receiving ith treatment and ith block. ϵ_{ijs} are independent errors that follow a continuous distribution with a zero median. We are interested to test the equality of the treatment effects *i.e.* H_0: $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_v = \tau$ (say) against an alternative hypothesis that at least two of the τ_i's are different.

Now, suppose $N = \sum_i^v \sum_j^b n_{ij}$, denotes the total number of observations and

$$p_{ij} = \frac{n_{ij}}{N}, \ q_{ij} = \frac{1}{p_{ij}}, \ q_{i.} = \sum_{j}^{b} q_{ij}, \ q_{..}^* = \sum_{i}^{v} \frac{1}{q_{i.}}$$

Consider a pair of plots (i, j) and (i', j), located in the same column j. We can form $n_{ij} \times n_{i'j}$ pairs of observations by taking one observation from each of these cells. Suppose that $u_{ii'j}$ denotes the proportion of the $n_{ij} \times n_{i'j}$ pairs in which the observations from plot (i, j) are greater than the observations in the plot (i', j). Then define

$$u_i = \sum_{i' \neq i = 1}^{v} \sum_{j=1}^{b} u_{i'ij}$$

Using these, the test statistic is

$$T = \frac{12N}{v^2} \left[\sum_{i=1}^{v} \{u_i - (v-1)b/2\}^2 / q_{i.} - \left\{ \sum_{i=1}^{v} \{u_i - (v-1)b/2\} / q_{i.} \right\}^2 / q_{..}^* \right]$$

The test rejects H_0 if T is greater than the upper cut off point of the chi-square distribution with $(v - 1)$ degrees of freedom. The above procedure is illustrated with the help of the following example:

Example 9.5: Table 9.14 gives the number of days to maturity for three varieties of a cereal crop grown in two soil conditions.

Table 9.14: Number of days to maturity of three variety under two soil conditions

Variety	Soil type	
	Light	Heavy
A	130, 115, 123, 142	117, 125, 139
B	108, 114, 124, 106	91, 111, 110
C	155, 146, 151, 165	97, 108

In this Example $v = 3$, $b = 2$, $n_{11} = 4$, $n_{12} = 3$, $n_{21} = 4$, $n_{22} = 3$, $n_{31} = 4$, $n_{32} = 2$ and N = 20.

Now

$$p_{11} = p_{21} = p_{31} = 4/20 = 0.20; \; p_{12} = p_{22} = 3/20 = 0.15; \; p_{32} = 2/20 = 0.10.$$

Then

$$q_{11} = q_{21} = q_{31} = 1/0.20 = 5.00; \; q_{12} = q_{22} = 1/0.15 = 6.67; \; q_{32} = 1/0.1 = 10.00,$$

$$q_{1.} = 5.00 + 6.67 = 11.67; \; q_{2.} = 5.00 + 6.67 = 11.67; \; q_{3.} = 5.00 + 10.00 = 15.00,$$

$$q_{..}^* = \frac{1}{11.67} + \frac{1}{11.67} + \frac{1}{15.00} = 0.238$$

We now proceed to compute $u_{ii'j}$. Let us first illustrate computation of u_{121}. Here 16 (4×4) pairs of observations can be formed from row 1 and row 2 in column 1. Take each observation in cell (1,1) and compare it with four observations in cell (2, 1). Observation 130 is greater than all the four values in cell (2, 1). Hence, its contribution to u_{121} is 4. Similarly the observation 115 contributes 3, observation 123 contributes 3 and observation 142 contributes 4. The total is 14. So, u_{121} is 14/16.

Similarly $u_{122} = 9/9 = 1$; $u_{131} = 0$; $u_{132} = 1$; $u_{231} = 0$; $u_{232} = 4/6$; $u_{211} = 2/16$; $u_{212} = 0$; $u_{311} = 1$; $u_{312} = 0$; $u_{321} = 1$; $u_{322} = 2/6$.

Hence

$$u_1 = \frac{14}{16} + 1 + 0 + 1 = \frac{46}{16}\ ; \ u_2 = \frac{2}{16} + 0 + 0 + \frac{4}{6} = \frac{76}{96}\ ; \ u_3 = 1 + 0 + 1 + \frac{2}{6} = \frac{14}{6}.$$

The test statistic T is then computed as

$$T = \frac{12*20}{3^2}\left[(0.0656 + 0.1250 + 0.0074) - (0.0750 - 0.1035 + 0.0222)^2 / 0.238\right]$$

$$= 5.280.$$

The tabulated value of χ^2 at 2 degrees of freedom and at 5% level of significance is 5.991 and the calculated value is 5.280. It, therefore, follows that the varieties of cereal crop are not significantly different at 5% level of significance. Here the probability of getting a value of χ^2 greater than 5.280 is 0.071361.

A lot of efforts have also been made at IASRI to develop the non-parametric test procedures for the analysis of groups of experiments conducted in Randomized block designs and split plot designs. For details one may refer to Rai and Rao (1980, 1984).

9.7. Some more examples for testing normality and homogeneity of variances

Some examples of testing the assumptions of normality and homogeneity of errors and remedial measures are discussed in the sequence.

Example 9.6: Suppose an entomologist is interested in determining whether four different kinds of traps caught equivalent insects when applied to same field. Each of the traps is used six times on the field and resulting data (number of insects per hour) are as shown in Table 9.15 along with mean, variance and range.

Table 9.15: Data on insects per hour

Treatment (Trap)	Replication						Mean $\overline{Y_i}$	Variance S_i^2	Range
	I	II	III	IV	V	VI			
A	3	1	12	7	17	2	7	40.4	16
B	9	29	21	24	28	45	31	138.4	36
C	63	84	97	61	98	71	79	270.8	37
D	172	118	109	172	143	168	147	798.4	63

A scatter plot of mean and variance and mean versus range are given as follows:
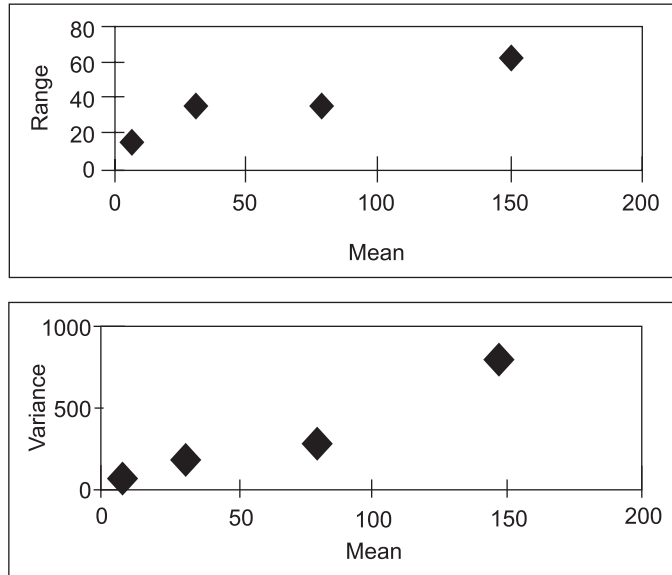
Figure 9.1: Scatter plot of mean versus range and variance

Both the plots indicate that variances are heterogeneous and variance is proportional to mean. The residuals for testing the normality and homogeneity of error terms are obtained and are given in Table 9.16.

Table 9.16: Residuals for the insect data

Treatment (Trap)	Replication						Mean	Variance
	I	II	III	IV	V	VI		S_i^2
A	−1.00	0.75	10.00	−1.25	3.25	−11.75	0	50.35
B	−14.00	9.75	0.00	−3.25	−4.75	12.25	0	94.85
C	−13.00	11.75	23.00	−19.25	12.25	−14.75	0	314.85
D	28.00	−22.25	−33.00	23.75	−10.75	14.25	0	650.20

The SAS code for testing normality of a variable y in a dataset mydata is given below.
PROC UNIVARIATE DATA=mydata NORMAL;
VAR y;
RUN;

The R code for testing normality of a variable y using Shapiro-Wilk test is shapiro.test(y). and using Kolmogorov-Smirnov test is ks.test(y, "pnorm", mu, sigma) where mu and sigma are the mean and standard deviation of the normal distribution.

Table 9.17: Result of normality tests

Shapiro-Wilk test		Kolmogorov-Smirnov test	
Statistic (SW)	p-value	Statistic (KS)	p-value
0.980	0.882	0.110	0.200

The errors were found to be normally distributed. Therefore, homogeneity of error variances was tested using Bartlett's test. It is described in the sequel.

Pooled Variance $\left(S_p^2\right) = \dfrac{5(50.35 + 94.85 + 314.85 + 650.20)}{20} = 277.5625$

$q = 20\log_{10} 277.5625 - 5[\log_{10} 50.35 + \log_{10} 64.85 + \log_{10} 314.85 + \log_{10} 650.20]$

$\quad = 3.916278$

$c = 1 + \dfrac{1}{9}\left(\dfrac{4}{5} - \dfrac{1}{20}\right) = 1.08333$

$\chi_0^2 = 8.324$

Since $\chi_{0.05,3}^2 = 7.81$, therefore, we reject the null hypothesis and conclude that the variances are unequal.

The SAS commands for Bartlett's test using SAS is given below.

```
PROC GLM data=insect;
CLASS treatment replication;
MODEL number = treatment replication;
MEANS treatment / HOVTEST=Bartlett;
RUN;
```

The R code for Bartlett's test using R software is

```
insect=read.table("insect.txt",header=TRUE)
bartlett.test(number ~ treatment, data = insect)
```

The values of $\dfrac{S_i^2}{\bar{Y}_{i.}}$ are 5.77, 5.32, 3.43 and 5.43, indicating that variance is proportional to mean. Therefore, square root transformation should be used. After application of square root transformation, the residuals are presented in Table 9.18.

Table 9.18: Residuals for the insect data after square root transformation

Treatment	Replication						Variance
	I	II	III	IV	V	VI	S_i^2
A	− 0.03614	− 0.92542	1.05800	0.20614	0.98287	− 1.28544	0.928
B	− 1.34939	0.87854	− 0.40473	− 0.12183	− 0.42993	1.42735	0.999
C	− 0.28226	0.78841	0.99143	− 1.08068	0.30794	− 0.72483	0.694
D	1.66779	− 0.74153	− 1.64469	0.99637	− 0.86087	0.58293	1.622

Results of test of normality of error terms on the transformed data are shown in Table 9.19.

Table 9.19: Result of normality tests after square root transformation

Shapiro-Wilk test		Kolmogorov-Smirnov test	
Statistic (SW)	p-value	Statistic (KS)	p-value
0.956	0.414	0.127	0.200

The errors remain normally distributed after transformation. The results of homogeneity of error variances using Bartlett's test are

Bartlett's Test (normal distribution): test statistic = 0.89, *p*-value = 0.828

Hence, we conclude that the errors are normally distributed and have a constant variance after transformation.

The results of analysis of variance with original and transformed data are given in Table 9.20.

Table 9.20: ANOVA with original data

Source	DF	Seq SS	Adj. SS	MS	F value	Prob > F
Replication	5	689.0	689.0	137.8	0.37	0.86
Treatment	3	70828.5	70828.5	23609.5	63.80	0.00
Error	15	5551.0	5551.0	370.1		
Total	23	77068.5				

R-Square	Root MSE
92.80%	19.237

Tukey Simultaneous tests for all pair wise treatment comparisons

	1	2	3	4
1	.			
2	0.3525	.		
3	0.0001	0.0013	.	
4	0.0000	0.0000	0.0001	.

ANOVA for transformed data

Source	DF	Seq SS	Adj. SS	MS	F-value	Prob > F
Replication	5	5.055	5.055	1.011	0.71	0.622
Treatment	3	326.603	326.603	108.868	76.98	0.000
Error	15	21.214	21.214	1.414		
Total	23	352.872				

R-Square	Root MSE
93.99%	1.18922

Tukey Simultaneous tests for all pair wise treatment comparisons

	1	2	3	4
1	.			
2	0.0091			
3	0.0000	0.0003		
4	0.0000	0.0000	0.0015	.

With transformed data treatments 1 and 2 are significantly different whereas with original data, they were not.

The Box-Cox transformation on the experiments was used, wherever, the data were found to be non-normal/heterogeneous. If the data becomes normal and homogeneous after transformation, then the analysis was performed on the transformed data. If the data remains non-normal after transformation, then the data were analysed using Skillings and Mack non-parametric test. For illustration the results of some of these experiments are discussed in the sequel.

Example 9.7: A varietal trial on Rapeseed-Mustard was conducted at Faizabad with 11 varieties using a randomized complete block design with 3 replications. The experimental data (Yield in kg/ha) obtained from the above experiment is given in Table 9.21.

Table 9.21: Yield data on Rapeseed-Mustard trial

Treatments	Replications		
	R1	**R2**	**R3**
MCN-157	952.380	1058.200	1079.364
MCN-158	846.560	634.920	687.830
MCN-159	529.100	687.830	687.830
MCN-160	1058.200	1005.290	952.380
MCN-161	1111.110	888.888	846.560
MCN-162	899.470	634.920	1005.290
MCN-163	1058.200	1164.020	952.380
MCN-164	687.830	740.740	529.100
MCN-165	952.380	952.380	867.724
MCN-166	1058.200	1058.200	529.100
MCN-167	1269.840	1164.020	1216.930

The analysis of variance of the original data is given in Table 9.22.

Table 9.22: ANOVA with original data

Sources	DF	SS	MS	F	Prob. >F
Replication	2	52534.9880	26267.4940	1.46	0.2563
Treatment	10	967055.0471	96705.5047	5.37	0.0007
Error	20	360218.589	18010.929		
Total	**32**	**1379808.624**			

R-Square	CV	RMSE	Yld Mean
0.738936	14.878	134.2048	902.035

The normality of error terms was tested, and the results obtained are given in Table 9.23.

Table 9.23: Result of tests of normality

Shapiro-Wilk test		Kolmogorov-Smirnov test	
Statistic (SW)	p-value	Statistic (KS)	p-value
0.9679	0.4249	0.1018	>0.1500

Since the data is normal, therefore, Bartlett's test is used for testing the homogeneity of error variances. The test gives Bartlett's test statistic 20.177 with *p*-value 0.0276. Hence, the errors were found to be heterogeneous. Therefore, it can be concluded that the data is heterogeneous and normal. Therefore, Box-Cox transformation was used as a remedial measure. In the sequel, we describe the results of the Box-Cox transformation.

For this, the data was transformed by varying λ from – 10 to + 10 with an increment of 0.01. The error sum of squares is computed for each value of λ. The value of λ with minimum error sum of squares is used for transformation given in (9.1). The minimum value of SSE is obtained for $\lambda = 2.38$. Therefore, reciprocal transformation was used.

The assumptions of normality and homogeneity of error variances are again tested using the transformed data. Normality of error terms was tested and the results obtained are given in Table 9.24.

Table 9.24: Result of tests of normality

Shapiro-Wilk test		Kolmogorov-Smirnov test	
Statistic (SW)	p-value	Statistic (KS)	p-value
0.984	0.8885	0.0867	>0.1500

Since the data is normal, therefore, Bartlett's test is used for testing the homogeneity of error variances. The test gives Bartlett's test statistic value of 15.725 with *p*-value 0.107757. Therefore, the transformed observations were found to be normal and homogeneous. So, ANOVA was performed on the transformed data. The results obtained are given in Table 9.25.

Table 9.25: ANOVA for transformed data

Sources	DF	SS	MS	F	Prob. >F
Replication	2	3.865471E13	1.93273335E13	1.62	0.2238
Treatment	10	7.8841391E14	7.8841391E13	6.59	0.0002
Error	20	2.3934391E14	1.1967195E13		
Total	32	1.0664125E15			

R-Square	CV	RMSE	Transformed Yld Mean
0.7756	29.563	3459363	11701777

It is seen that there is no change in the results of significance of treatment and replication effects. However, the transformed data satisfied the assumptions of ANOVA.

Example 9.8: [AFEIS Reference No. 1988(092)]. An experiment was conducted at Regional Research Station, Brahmavar (Karnataka) in 1988 to study on efficiency of large granular Urea on the growth and yield of paddy. The design adopted for this is Randomized Complete Block Design with 10 treatments in 3 replications. The net plot size is 3.60 x 2.40 m². The treatment details are given below.

Treatment	Treatment Details
1	Control (Untreated)
2	Untreated with N but treated with P @75kg/ha as Super Phosphate and K @50 kg/ha as Murate of Potash.
3	N @ 60 kg/ha as Urea Super Granule at planting.
4	N @ 120 kg/ha as Urea Super Granule at planting.
5	N @ 60 kg/ha as Urea Large Granule at planting.
6	N @ 120 kg/ha as Urea Large Granule at planting.
7	N @ 60 kg/ha as Prilled Urea in 3 splits as per recommendation.
8	N @ 120 kg/ha as Prilled Urea in 3 splits as per recommendation.
9	N @ 60 kg/ha as Large Granular Urea in 3 splits as per recommendation.
10	N @ 120 kg/ha as Large Granular Urea in 3 splits as per recommendation.

The experimental data (Yield in kg/plot) obtained from the above experiment is

Treatments ↓	Replications→		
	R1	R2	R3
T1	1.27	1.40	1.70
T2	1.71	1.65	1.87
T3	1.65	1.62	1.75
T4	1.65	1.62	1.65
T5	2.10	2.05	1.86
T6	1.57	2.00	1.77

T7	2.95	2.02	2.15
T8	1.97	2.45	2.05
T9	2.25	2.15	2.48
T10	1.87	1.85	1.90

The analysis of variance of the original data is given in Table 9.26

Table 9.26: ANOVA with original data

Source of variation	DF	SS	MS	F-value	Prob. >F
Replication	2	0.0068	0.0034	0.06	0.9373
Treatment	9	2.4315	0.2702	5.12	0.0016
Error	18	0.9489	0.0527		
Total	**29**	**3.3872**			

R-Square	CV	RMSE	Yld Mean
0.7199	12.088	0.2296	1.900

Shapiro-Wilk test was applied to test the assumption of normality of data. The data was found to be non-normal (Prob < W = 0.0482). Since the data is non-normal, therefore, Levene test is used for testing the homogeneity of error variances. The test resulted into heterogeneous errors (Prob > F as 0.0007). Therefore, we can conclude that the data is heterogeneous and non-normal. Therefore, Box-Cox transformation was used as a remedial measure. In the sequel we describe the results of the Box-Cox transformation.

For this, the data was transformed by varying λ from -10 to +10 with an increment of 0.01. The error sum of squares is computed for each value of λ. The value of λ with minimum error sum of squares is used for transformation given in (9.1). For illustration purposes, some values of λ and corresponding error sum of squares (SSE) are given below.

λ	SSE
-1.30	0.666901
-1.20	0.664440
-1.10	0.663444
-1.09	0.663409
-1.08	0.663386
-1.07	0.663375
-1.06	**0.663376**
-1.05	0.663389
-1.04	0.663414
-1.00	0.663629
-0.90	0.664997
-0.80	0.667553

From the above table, it may be seen that the minimum value SSE is 0.663375 for $\lambda = -1.07$. There data were transformed using $\lambda = -1.07$ in (9.1).

The assumptions of normality and homogeneity of errors are again tested using the transformed data. The transformed observations were found to be normal using Shapiro Wilk test (Prob. < W as 0.2624) and homogeneous using Bartlett's test (Prob. > Chi-square as 0.0679). Therefore, ANOVA was performed on the transformed data. The results obtained are given in Table 9.27.

9.27: ANOVA with transformed data

Source of variation	DF	SS	MS	F-value	Prob. >F
Replication	2	0.0033	0.0016	0.52	0.6031
Treatment	9	0.2005	0.0223	7.07	0.0002
Error	18	0.0567	0.0032		
Total	**29**	**0.2605**			

R-Square	CV	RMSE	Transformed Yld Mean
0.7822	10.7985	0.0561	0.520

We can see that there is no change in the results of significance of treatment and replication effects. However, the transformed data satisfied the assumptions of ANOVA.

Example 9.9: [AFEIS Reference No. 1985(015)]. An experiment was conducted at Agricultural Research station, Ponnempet (Karnataka) in 1985 to study the nitrogen management for low land pest and disease endemic area of paddy crop. The design adopted for this is Randomized Complete Block Design with 6 treatments in 5 replications. The net plot size used 5.00×3.00 m^2. The treatment details are given below.

Treatment	Treatment Details
1	Control
2	Split application of prilled (ordinary) Urea (50%Basal dose, 25% at 4 weeks after planting and 25% at panicle initiation stage)
3	Neem Cake Coated Urea
4	Coal Tar Coated Urea
5	Rock Phosphate Coated Urea
6	Urea Gypsum

The experimental data (yield in kg/plot) is

Treatments ↓	Replications →				
	R1	R2	R3	R4	R5
T1	1.00	1.37	0.73	0.53	2.85
T2	0.25	0.35	0.20	0.31	0.53
T3	0.73	0.62	0.76	0.24	0.38
T4	0.49	0.31	0.49	2.56	0.60
T5	0.80	0.8	0.52	0.95	0.37
T6	0.78	1.35	0.32	0.51	0.58

The results of the Analysis of Variance on the original data are given in Table 9.28.

Table 9.28: ANOVA with original data

Source of Variation	DF	SS	MS	F-value	Prob. >F
Replication	4	0.5815	0.1454	0.39	0.8142
Treatment	5	2.7135	0.5427	1.45	0.2496
Error	20	7.4799	0.3740		
Total	**29**	**10.7750**			

R-Square	CV	RMSE	Yld Mean
0.3058	82.345	0.6115	0.743

Shapiro Wilk test was applied to test the assumption of normality of data. The data was found to be non-normal (Prob. < W = 0.0009). Since the data is non-normal, therefore, Levene test is used for testing the homogeneity of error variances. The test resulted into homogeneous errors (Prob. > F as 0.1160). Therefore, we can conclude that the data is homogeneous and non-normal. As the data is homogeneous and non-normal, we can directly use the non-parametric test. But, here, we use the Box-Cox transformation to see its effect on normality. In the sequel, the results of the Box-Cox transformation are described.

The data was transformed by varying λ from -10 to $+10$ with an increment of 0.01. The error sum of squares are computed for each value of λ. The value of λ with minimum error sum of squares is used for transformation given in (9.1). The minimum value SSE is for $\lambda = -0.47$. The data were transformed using $\lambda = -0.47$ in (9.1).

The assumptions of normality and homogeneity of errors are again tested using the transformed data. The transformed observations were found to be normal using Shapiro Wilk test (Prob. < W as 0.7977) and homogeneous using Bartlett's test (Prob. > Chi-square as 0.6793). Therefore, ANOVA was performed on the transformed data. The results obtained are given in Table 9.29:

Table 9.29: ANOVA with transformed data

Source of Variation	DF	SS	MS	F-vlaue	Prob. >F
Replication	4	0.2107	0.0527	0.48	0.7489
Treatment	5	1.5571	0.3114	2.85	0.0423
Error	20	2.1868	0.1093		
Total	29	0.9545			

R-Square	CV	RMSE	Transformed Yld Mean
0.4470	24.8990	0.3307	1.328

We can see that treatment effects became significant at 5% level of significance where as the treatments were non-significant at 5% level of significance through the original data. CV% reduced about one third.

As the data in this experiment is non-normal and homogeneous. Therefore, one could have thought of applying the Skillings and Mack non-parametric test for testing equality of treatments effects. The result obtained from this test is

Degree of freedom   5

Statistic: 9.8000

Prob. >Chi-Square   0.0810

In can be observed that the treatments remain non-significant at 5% level of significance, where as these were significant with the transformed data.

Example 9.10: [AFEIS Reference No. 1991(040)]. An experiment was conducted at Agricultural Research station, Bagalkot (Karnataka) in 1991 to study the effect of crop residue on the growth and yield of Rabi Sorghum. The design is adopted for this is Randomized Complete Block Design with 6 treatments in 4 replications. The net plot size used is 4.20 × 10.20 m2. The treatment details are given below.

Treatment	Treatment Details
1	Control
2	50 kg/ha N as Urea + 25 kg/ha of P2O5 as Super Phosphate (Recommended dose)
3	Subabul Stalks @ 5 Tons/ha
4	Jowar Stubbles @ 5 Tons/ha
5	Jowar Stubbles @ 2.5 Tons/ha + Subabul Stalks @ 2.5 Tons/ha
6	Jowar Stubbles @ 1.25 Tons/ha + Subabul Stalks @ 3.25 Tons/ha

The experimental data (yield in kg/plot) is

Treatments ↓	Replications →			
	R1	R2	R3	R4
T1	8.90	9.30	9.20	9.10
T2	10.65	10.50	10.10	11.05
T3	10.60	11.25	10.25	11.50
T4	9.15	8.90	9.30	10.30
T5	13.00	9.91	10.14	8.85
T6	9.80	9.95	9.95	10.20

The results of the Analysis of Variance on the original data are given in Table 9.30.

Table 9.30: ANOVA with original data

Source of Variation	DF	SS	MS	F-value	Prob. >F
Replication	3	0.9523	0.3174	0.42	0.7392
Treatment	5	9.7680	1.9536	2.60	0.0690
Error	15	11.2547	0.7503		
Total	**23**	**21.9751**			

R-Square	CV	RMSE	Yld Mean
0.4878	8.5958	0.8662	10.077

Shapiro Wilk test was applied to test the assumption of normality of data. The data was found to be non-normal (Prob. < W = 0.0082). Since the data is non-normal, therefore, Levene test is used for testing the homogeneity of error variances. The test resulted into homogeneous errors (Prob. >F as 0.0842). Therefore, we can conclude that the data is homogeneous and non-normal. As the data is homogeneous and non-normal, we can directly use the non-parametric test. But, here, we use the Box-Cox transformation to see its effect on normality. In the sequel, the results of the Box-Cox transformation are described.

For this, the data was transformed by varying λ from -10 to +10 with an increment of 0.01. The error sum of squares are computed for each value of λ. The value of λ with minimum error sum of squares is used for transformation given in (9.1). The minimum value SSE is for $\lambda = -4.03$. Therefore data were transformed using $\lambda = -4.03$.

The assumptions of normality and homogeneity of errors are again tested using the transformed data. The transformed observations were found to be non-normal using Shapiro Wilk test (Prob. < W as 0.0253) and homogeneous using Bartlett's test (Prob. > Chi-square as 0.1376).

In this situation where data remains non-normal even after transformation, we cannot use usual analysis of variance for analysis of data. So, here, Skillings and Mack test (here Friedman) non-parametric test for testing equality of treatment effects was performed. The result obtained from this test is

Degree of freedom	5
Statistic:	12.8714
Prob. >Chi-Square	0.0248

When original observations were analysed by using usual analysis of variance, it was found that the treatment and replication effect is non-significant. The treatment became significant at 5% level of significance by using the Skillings and Mack non-parametric test.

The complete SAS code for testing the normality of errors, homogeneity of error variances, applying Box-Cox transformation, applying ANOVA on transformed data or using appropriate non-parametric test in analysis of data generated from RCB design is available at http://www.iasri.res.in/design/Analysis%20of%20data/Diagnostics%20_Remedial_measures_sas.html. This may be appropriately modified in case of other designed experiments.

APPENDIX – I

Kruskal-Wallis Test Statistic

Each table entry is the smallest of the Kruskal-Wallis T such that its right-tail probability is less than or equal to the value given on the top row for $v = 3$, each sample size less than or equal to five.

r_1	r_2	r_3	0.100	0.050	0.020	0.010	0.001
2	2	2	4.571	-	-	-	-
3	2	1	4.286	-	-	-	-
3	2	2	4.500	4.714	-	-	-
3	3	1	4.571	5.143	-	-	-
3	3	2	4.556	5.361	6.250	-	-
3	3	3	4.622	5.600	6.489	7.200	-
4	2	1	4.500	-	-	-	-
4	2	2	4.458	5.333	6.000	-	-
4	3	1	4.056	5.208	-	-	-
4	3	2	4.511	5.444	6.144	6.444	-
4	3	3	4.709	5.791	6.564	6.745	-
4	4	1	4.167	4.967	6.667	6.667	-
4	4	2	4.555	5.455	6.600	7.036	-
4	4	3	4.545	5.598	6.712	7.144	8.909
4	4	4	4.654	5.692	6.962	7.654	9.269
5	2	1	4.200	5.000	-	-	-
5	2	2	4.373	5.160	6.000	6.533	-
5	3	1	4.018	4.960	6.044	-	-
5	3	2	4.651	5.250	6.124	6.909	-
5	3	3	4.533	5.648	6.533	7.079	8.727
5	4	1	3.987	4.985	6.431	6.955	-
5	4	2	4.541	5.273	6.505	7.205	8.591
5	4	3	4.549	5.656	6.676	7.445	8.795
5	4	4	4.668	5.657	6.953	7.760	9.168
5	5	1	4.109	5.127	6.145	7.309	-
5	5	2	4.623	5.338	6.446	7.338	8.938
5	5	3	4.545	5.705	6.866	7.578	9.284
5	5	4	4.523	5.666	7.000	7.823	9.606
5	5	5	4.560	5.780	7.220	8.000	9.920

Friedman Test Statistic Critical value for the Friedman two-way analysis of variance by rank statistics, T

v	b	$a \leq 0.10$	$a \leq 0.05$	$a \leq 0.01$
3	3	6.00	6.00	---
	4	6.00	6.50	8.00
	5	5.20	6.40	8.40
	6	5.33	7.00	9.00
	7	5.43	7.14	8.86
	8	5.25	6.25	9.00
	9	5.56	6.22	8.67
	10	5.00	6.20	9.60
	11	4.91	6.54	8.91
	12	5.71	6.17	8.67
	13	4.77	6.00	9.39
	∞	4.61	5.99	9.21
4	2	6.00	6.00	---
	3	6.60	7.40	8.60
	4	6.30	7.80	9.60
	5	6.36	7.80	9.96
	6	6.40	7.60	10.00
	7	6.26	7.80	10.37
	8	6.30	7.50	10.35
	∞	6.25	7.82	11.34
5	3	7.47	8.53	10.13
	4	7.60	8.80	11.00
	5	7.68	8.96	11.52
	∞	7.78	9.49	13.28

Outliers in Designed Experiments

10.1 Introduction

Designing an experiment and the analysis of data generated form an integrated component of agricultural research. A properly designed experiment taking care of the variability in the experimental units, the subsequent analysis of data using appropriate methodologies satisfying all the assumptions and then drawing valid conclusions helps in improving the quality of agricultural research. Thus far the attention has been focused on the two components *viz.*, proper designing of experiment and analysis of data using appropriate methodologies to answer the questions of the experimenter. However, in practice there is a tendency to ignore the assumptions of the analysis and go ahead with the statistical processing of data as if the assumptions were satisfied. The assumptions, however, get violated during experimentation and as such the statistical analysis that is carried out no longer remains valid. At times, the inferences drawn are not the valid inferences. The different diagnostics procedures and remedial measures for validating the assumptions of ANOVA have been discussed in Chapter 9.

Another assumption of ANOVA is that the data do not have any abnormally high or low observation(s), or outliers. The purpose of this Chapter is to address this problem arising because of presence of outlier(s) in the data.

An outlier is an observation that lies at an abnormal distance from other values in a random sample from a population. An observation that "lies outside" (is much smaller or larger than) most of the other observations in a set of data is called an outlier. However, occurrence of outliers is very common wherever data collection is involved. For instance, during the experimentation, there might be an infestation of a disease or insect attack on some plots in the field, or there may be unintentional heavy irrigation on some particular block(s) or plot(s) of the experiment, or at times there may be mistakes creeping in during recording of data, etc. These result in occurrence of outliers in experimental data. It is important to detect these outlying observations and apply appropriate remedial measures.

Outliers in a set of data are defined to be sub-set of observations that appear to be inconsistent with the remainder set of data. Occurrence of outliers are very common in every field involving data collection and outliers arise from heavy tailed distributions or are simply bad data points due to error. When outliers are present in the data, the results obtained from the analysis of such data may lead to erroneous inferences. To be clearer, consider the following example.

Example 10.1: {Agricultural Field Experiments Information System (AFEIS) Reference No. 1985(015)}. An experiment in 6 treatments was conducted in a randomized complete block design in 5 replications at Agricultural Research station, Ponnempet (Karnataka) in 1985 to

study the nitrogen management for low land pest and disease endemic area of paddy crop [Net plot size used 5.00×3.00 m²]. The treatment details are

T1: Control

T2: Split application of prilled (ordinary) Urea (50% Basal dose, 25% at 4 weeks after planting and 25% at panicle initiation stage

T3: Neem Cake Coated Urea

T4: CT Coated Urea

T5: Rock Phosphate Coated Urea

T6: Urea Gypsum

Table 10.1 shows the data on yield in kilogram per plot for different treatments.

Table 10.1: Yield of paddy in kg/plot

Treatments ↓	Replications →				
	R1	R2	R3	R4	R5
T1	1.00	1.37	0.73	0.53	2.85
T2	0.25	0.35	0.20	0.31	0.53
T3	0.73	0.62	0.76	0.24	0.38
T4	0.49	0.31	0.49	2.56	0.60
T5	0.80	0.8	0.52	0.95	0.37
T6	0.78	1.35	0.32	0.51	0.58

The analysis of the data is presented in Table 10.2 below. It is observed that the treatment effects are not significant at 5% level of significance.

Table 10.2: ANOVA with original data

Source of Variation	DF	SS	MS	F-value	Prob. >F
Replication	4	0.5815	0.1454	0.39	0.8142
Treatment	5	2.7135	0.5427	1.45	0.2496
Error	20	7.4799	0.3740		
Total	**29**	**10.7750**			

It was then followed by residual analysis of this data. The studentized residuals may be obtained by using the following code:

```
DATA arsponem;
INPUT rep trt yield;
CARDS;
1       1       1.00
1       2       0.25
1       3       0.73
1       4       0.49
1       5       0.80
1       6       0.78
2       1       1.37
2       2       0.35
2       3       0.62
2       4       0.31
2       5       0.80
2       6       1.35
3       1       0.73
3       2       0.20
3       3       0.76
3       4       0.49
3       5       0.52
3       6       0.32
4       1       0.53
4       2       0.31
4       3       0.24
4       4       2.56
4       5       0.95
4       6       0.51
5       1       2.85
5       2       0.53
5       3       0.38
5       4       0.60
5       5       0.37
5       6       0.58
;
PROC glm;
CLASS rep trt;
MODEL yield = rep trt;
OUTPUT OUT=a STUDENT =s;
RUN;
PROC PRINT DATA=a;
RUN;
```

Studentized residuals are presented in Table 10.3.

Table 10.3: Studentized residuals

Treatments ↓	Replications →				
	R1	**R2**	**R3**	**R4**	**R5**
T1	-0.45728	0.03338	-0.65421	-1.74901	**2.82712**
T2	-0.02069	-0.07076	0.22297	-0.25100	0.11949
T3	0.50401	0.03338	0.90788	-0.82778	-0.61749
T4	-0.66556	-1.27638	-0.32176	**3.12952**	-0.86583
T5	0.35982	0.10948	0.14286	0.30975	-0.9219
T6	0.27971	1.1709	-0.29773	-0.61149	-0.54139

It is observed from the table that the observations pertaining to T1 in replication 5 and T4 in Replication 4 stand out because of their high value of studentized residuals. These two observations seem to be influential. The analysis is carried out again after removing these two observations. The results of this analysis are presented in the Table 10.4. The dramatic effect of removing these two observations is worth noticing. The treatment effects now become significant at 5% level of significance. Removal of any other observation or pair of observations does not affect the analysis. These two observations, therefore, definitely are influential.

Table 10.4: ANOVA after removing two observations

Source of Variation	DF	SS	MS	F-Value	Prob > F
Blocks (Adjusted)	4	0.4138	0.1035	1.63	0.2102
Treatments (Adjusted)	5	0.9082	0.1820	2.86	0.0451
Error	18	1.1429	0.0635		
Total	27	2.4813			

The present example clearly shows how the presence of outliers affects the analysis of the data and inferences drawn.

10.2 What is an outlier?

Daniel (1960) defines an outlier as "an observation whose value is not in the pattern of values produced by the rest of the data". A more comprehensive definition is due to Beckman and Cook (1983). They defined the following: Discordant observation is any observation that appears surprising or discrepant to the investigator. Contaminant observation is any observation that is not a realization from the target distribution. Outlier is a collective to refer to either a contaminant or discordant observation.

Influential cases: An outlier need not be influential in the sense that the result of an analysis may essentially remain unchanged when an outlying observation is removed. It is useful to regard an influential observation as a special type of outlier.

Approaches to outliers are divided into two broad categories, viz.

(i) To identify the outlier(s) for further study. This forms the detection part.

(ii) To accommodate the possibility of outlier(s) by suitable modifications of the models and / or method of analysis. The robust methods of estimation or analysis, which were created to modify least squares procedure so that the outliers do not have much influence, fall under this category.

10.3 Detection of outliers in designed experiments

Removal of the individual suspected case or group of suspected cases in turn may change the result of the analysis of the data. This is the idea behind the study of the influential cases in linear models. Using this idea many statistics have been developed for detecting outliers in linear regression model, viz., Cook-statistic given by Cook (1977, 1979), Q_k-statistic given by Gentleman and Wilk (1975) and AP-statistic given by Andrews and Pregibon (1978). However, these statistics cannot be applied to designed experiments as such due to rank deficiency of its design matrix. Moreover, in design of experiments we are generally interested in estimation of some contrasts of parameters rather than the whole set of parameters. Keeping these in mind Bhar and Gupta (2001) modified these statistics for application in designed experiments. However, here we concentrate on only Cook-statistic.

10.3.1 Cook statistic in designed experiments

Consider the mathematical model for a block design d (say)

$$y_{ij} = \mu + \tau_i + \beta_j + e_{ij}, \ i = 1, 2, \cdots, v; j = 1, 2, \cdots, b \tag{10.1}$$

where y_{ij} is the observation recorded on the ith treatment in the jth block, $i = 1, 2, \cdots, v$; $j = 1, 2, \cdots, b$; μ is the general mean; τ_i is the effect of ith treatment, β_j is the effect of jth block, and $e_{ij}'s$ are uncorrelated random error components assumed to be distributed normally with zero mean and constant variance σ^2.

Cook-statistic is based on the effect on the parameter estimates after removing the suspected outlier(s) from model (10.1). Since in designed experiments, we are interested in estimation of treatment effects or some suitable functions of treatments, Bhar and Gupta (2001) developed Cook statistic for detecting outliers in data obtained from a block design. We assume that the design d considered here is connected, i.e., all $(v-1)$ orthonomalized contrasts for the parameters of interest, i.e., τ are estimable. Let $\boldsymbol{\tau}$ be a vector of all v treatment effects and $\mathbf{P\tau}$ be the set of all $(v-1)$ orthonormalized contrasts for the parameters with best linear unbiased estimate as $\mathbf{P\hat{\tau}}$. Then Cook-statistic is defined as:

Definition 10.1: Cook-statistic for the set of contrasts $\mathbf{P\tau}$ is given by

$$D_k = \frac{(\mathbf{P\hat{\tau}} - \mathbf{P\hat{\tau}}_{(k)})'[D(\mathbf{P\hat{\tau}})]^{-1}(\mathbf{P\hat{\tau}} - \mathbf{P\hat{\tau}}_{(k)})}{\text{Rank}[D(\mathbf{P\hat{\tau}})]}$$

where $\mathbf{P\hat{\tau}}_{(k)}$ is the estimate of functions of treatment effects $\mathbf{P\tau}$ obtained after removing the k suspected outliers.

The statistic D_k provides a measure of the distance between $\mathbf{P}\hat{\boldsymbol{\tau}}_{(k)}$ and $\mathbf{P}\hat{\boldsymbol{\tau}}$ in terms of descriptive levels of significance, because D_k is actually $100(1-\alpha)$ % confidence ellipsoid for the vector $\mathbf{P}\hat{\boldsymbol{\tau}}$ under normal theory, which satisfies $D_k \leq F(k, n-k, 1-\alpha)$.

Suppose, for example, $D_k \approx F(k, n-k, 1-\alpha)$. Then the removal of the k data points moves the least squares estimate to the edge of the 50% confidence region for $\mathbf{P}\boldsymbol{\tau}$ based on $\mathbf{P}\hat{\boldsymbol{\tau}}$. Such a situation may be a cause for concern. For any analysis one would like each $\mathbf{P}\hat{\boldsymbol{\tau}}_{(k)}$ to stay well within 10%, say, confidence region (See Cook, 1977). Cook has also shown that this statistic can be used to assess the degree of influence for a subset of parameters as well as can be extended for more than one outlier. Thus, one can test the significance of an outlier by F-test because D_k is approximately distributed as F distribution with k and $n - k$ degrees of freedom.

Another approach for testing an outlier is due to Tukey (1977). The Cook-statistic here is obtained for a particular set of k outlying observations. In practice, we have to apply this statistic for all possible set of k outlying observations. Internal scaling defines extreme values of a diagnostic measure (Cook-statistic here) relative to the "weight of the evidence" provided by the given diagnostic series itself. The calculation of each Cook-statistic results in a series of values ${}^n C_k$ ($k =$1, 2, 3…). Following Tukey (1977), we compute inter-quartile range ($\overset{*}{S}$) for this series and indicate as extreme values that exceeds $7/2 (\overset{*}{S})$. This limit can be viewed as convenient point of departure in the absence of a more exact distribution theory.

Example 10.2: {AFEIS Reference No. 1989(451)}An experiment with twelve manurial treatments was conducted at Punjab Rao Deshmukh Krishi Vidyapeeth, Akola, Maharashtra in 1989 in a randomized complete block (RCB) design with three replications to study the effect of micnelf and other micro nutrients on the yield of groundnut crop [net plot size: 1.80m × 4.20m]. The treatment details are

$T_1 =$ Control

$T_2 =$ One spray of micnelf + magsulf at 20 days after sowing (DAS)

$T_3 =$ Two spray of micnelf + magsulf at 20 DAS and 40 DAS

$T_4 =$ Three spray of micnelf + magsulf at 20 DAS, 40 DAS and 55 DAS

$T_5 =$ One spray of micnelf + murate of potash at 55 DAS

$T_6 =$ Three spray of micnelf at 20 DAS, 40 DAS and 55 DAS

$T_7 =$ Two spray of borax at 40 DAS and 55 DAS

$T_8 =$ Three spray of urea+ dap at 20 DAS, 40 DAS and 55 DAS

$T_9 =$ Two spray of 0.5 ml/lit nutron at 20 DAS and 40 DAS

$T_{10} =$ Two spray of 1.0 ml/lit nutron at 20 DAS and 40 DAS

$T_{11} =$ Three spray of ferrous sulphate

$T_{12} =$ Water spray

Table 10.5 shows the data on yield per plot in kilogram for different treatments.

Table 10.5: Yield of groundnut in kg/plot

Replications	Treatments											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0.55	0.72	0.62	0.67	0.59	0.65	0.75	0.95	0.57	0.61	0.57	0.62
2	0.54	0.62	0.53	0.57	0.58	0.49	0.61	0.51	0.53	0.58	0.52	0.56
3	0.50	0.60	0.57	0.54	0.48	0.47	0.71	0.51	0.48	0.60	0.53	0.54

The ANOVA is obtained through SAS. The codes are given in Chapter 2 for analysis of a randomized complete block design.

Table 10.6: ANOVA with original data

Source	DF	SS	MS	F Value	Prob > F
Replication	2	0.09223889	0.04611944	9.53	0.0010
Treatment (adjusted)	11	0.09735556	0.00885051	1.83	0.1100
Error	22	0.10649444	0.00484066		
Total	35	0.29608889			

Analysis of this data is presented in Table 10.6. It is observed that the treatment effects are not significantly different at 5% level of significance. Cook-statistic was computed for each of the observations. For calculating the Cook-statistic, the following programme is written. This programme was run in SAS.

```
DATA test;
INPUT trt rep yield;
CARDS;
1    1        0.55
2    1        0.72
3    1        0.62
4    1        0.67
5    1        0.59
6    1        0.65
7    1        0.75
8    1        0.95
9    1        0.57
10   1        0.61
11   1        0.57
12   1        0.62
1    2        0.54
2    2        0.62
3    2        0.53
4    2        0.57
5    2        0.58
```

6	2	0.49
7	2	0.61
8	2	0.51
9	2	0.53
10	2	0.58
11	2	0.52
12	2	0.56
1	3	0.50
2	3	0.60
3	3	0.57
4	3	0.54
5	3	0.48
6	3	0.47
7	3	0.71
8	3	0.51
9	3	0.48
10	3	0.60
11	3	0.53
12	3	0.54

```
;
PROC iml  ;
USE outlier ;
READ ALL  var{rep trt yield} INTO x  ;
y =x[,3];
trt=x[,2];
d1=design(trt);
rep=x[,1];
d3=design(rep);
b1=max(rep);
v1=max(trt);
n=b1*v1;
o=j(n,1,1);
d2=o||d3;
b=i(n)-d2*ginv(t(d2)*d2)*t(d2);
c=t(d1)*b*d1;
p=i(n);
x1=o||d1||d3;
v=p-x1*ginv(t(x1)*x1)*t(x1);
 sigma=(t(y)*v*y)/(n-b1-v1+1);
ddd=b*d1*ginv(c)*t(d1)*b;
dd=diag(ddd);
ddv=diag(v);
w1=(t(y)*v);
```

```
w2=diag(w1);
w3=w2*inv(ddv)*dd*inv(ddv)*w2;
ww=(w3)*o;
cook=ww/((v1-1)*sigma);
PRINT y trt rep cook;
RUN;
```

This programme can be used for detecting a single outlier in any block design. Only data set has to be changed. Other things remain unchanged.

The values of Cook-statistic for each of the observations are presented in Table 10.7. It is observed from the Table that the observation at serial number 8 stands out. Comparing the statistic with the table value of $F_{(11, 22)}(0.90)$ is 0.472245, it was found that this observation is an outlier.

Table 10.7: Cook-statistics

Sl. No.	Replication	Treatment	Cook-statistics	Sl. No.	Replication	Treatment	Cook-statistics
1	1	1	0.0405781	19	2	7	0.036726
2	1	2	0.0000581	20	2	8	0.2051798
3	1	3	0.0093913	21	2	9	0.0182302
4	1	4	0.000428	22	2	10	0.0032059
5	1	5	0.0151393	23	2	11	0.001897
6	1	6	0.0270335	24	2	12	0.0048563
7	1	7	0.001993	25	3	1	0.0016231
8	1	8	**0.7569051**	26	3	2	0.0006272
9	1	9	0.0120946	27	3	3	0.0209725
10	1	10	0.0517893	28	3	4	0.002619
11	1	11	0.0263221	29	3	5	0.0135742
12	1	12	0.0093913	30	3	6	0.0107003
13	2	1	0.02597	31	3	7	0.05583
14	2	2	0.0003035	32	3	8	0.1739184
15	2	3	0.0022954	33	3	9	0.0006272
16	2	4	0.0009295	34	3	10	0.0292245
17	2	5	0.0573843	35	3	11	0.0140864
18	2	6	0.0037181	36	3	12	0.000741

Remark 10.1: For block designs with block size two, Cook-statistic cannot be applied to detect the outlying observations. Similarly, the AP- and Q_k- statistic also fail to detect an outlier in block designs with block size two. The reason is not far to see. The residuals in a block design with blocks of size 2 will have the same magnitude but would be opposite in sign (because the sum of residuals in a block is zero). As a consequence the Cook statistic for the two observations

in the same block would be the same. Therefore, some efforts are required for development of the procedure of detection of a single outlier in the experimental data generated from block designs with block size two.

Remark 10.2: Cook-statistic has also been applied to many real data sets taken from "Agricultural Field Experiments Information System (AFEIS)", IASRI, New Delhi. The experimental data from these experiments were investigated for the presence of any kind of problems like non-normality or heterogeneity of error variance under a project entitled 'A diagnostic survey of field experiments' conducted at IASRI. For more details, one may refer to Parsad et al. (2004). Based on the normality and homogeneity of errors, these data were grouped into the following groups:

(i)   Experiments having non-normal and heterogeneous error variance

(ii) Experiments having non-normal and homogeneous error variance

(iii) Experiments having normal and heterogeneous error variance

(iv) Experiments having normal and homogeneous error variance

Experiments from the first three groups have some problems. One of the reasons of such problems might be presence of outlier(s). We, therefore, took these experiments for applying the test statistic for detection of outlier(s). It was found that several of these experiments contain outliers.

10.4   Multiple outliers

In the case of multiple outliers, the well known problem of 'masking' hinders the identification problem. In masking, effect of one outlier is masked by the presence of another outlier; hence an observation cannot be detected as outlier if we apply single outlier detection procedure. Much attention was paid to this problem in linear regression model. A number of methodologies are now being emerged to handle such problems. Recently, Pena and Yohai (1995) proposed a new method to identify influential subsets in linear regression in presence of masking. As mentioned earlier, linear regression diagnostics cannot be applied directly to the case of designed experiments. Modification is required for the method to be applied in designed experiments. The modified method is based on influence matrix which is defined as the matrix of uncentred covariances of the effect on the whole data set deleting each observation. In the new statistic developed, influence matrix is defined on the basis of the vector of Cook statistic for all observations. In case of design of experiments, Cook statistic is developed using treatment contrast of interest. Therefore, defining influence matrix using Cook statistic is quite appropriate. The rest of the procedure is based on the "Eigen-structure" of the Influence matrix. The developed statistic was applied to experimental data from AFEIS. It was found in some experiments that individually some observations were not influential, but jointly with some other observations they become influential, i.e., some observations were masked by some other outlying observations and, therefore, were not detected when we apply single diagnostic procedure. The following example will clarify this point further.

Example 10.3: {AFEIS: Reference Number 1988(251)} An experiment with 10 treatments was carried out in the randomized block design with 4 replications at Sugarcane Research Institute, Shahjahanapur, Uttar Pradesh in 1988 with a view to find out the suitable herbicide to control weeds in Sugarcane (Net plot size: 8.00m × 5.40m.). The treatments (Weedicidal and Cultural) of the experiments are:

T_0 = Control weeded check

T_1 = Local conventional method

T_2 = Trash mulching

T_3 = 1.0 kg ai/ha of 2,4-D sodium salt and 0.50 kg a.i./ha of gramoxone at 3 weeks of

    planting followed by application of the same at 6-8 weeks of planting.

T_4 = 2.0 kg ai/ha of Atrazine as Pre-emergence spray

T_5 = 1.00 kg ai/ha of 2,4-D sodium salt at 8-10 weeks after planting

T_6 = 2.0 kg ai/ha of 2,4-D (Amine) as Pre-emergence spray followed by spray of the same at

    8-10 weeks after planting.

T_7 = 2.0 kg ai/ha of Atrazine as Pre-emergence spray followed by spray of glyphosate at 1.0

    kg ai/ha at 6-8 weeks after planting.

T_8 = 1.00 kg ai/ha of arochlor and 1.00 kg ai/ha of atrazine as pre-emergence spray

T_9 = 2.00 kg ai/ha of arochlor as pre-emergence spray

Table 10.8 shows the data on yield per plot in kilogram in different treatments.

Table 10.8: Yield of sugarcane in kg/plot

Replications	Treatments									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	2.52	2.82	2.42	2.67	2.50	3.01	2.65	2.62	2.18	2.57
2	2.77	2.77	2.52	3.69	3.21	3.05	2.64	2.53	2.47	2.82
3	2.32	2.38	2.44	2.30	1.90	2.46	2.35	2.47	2.15	2.26
4	2.31	2.14	2.38	2.13	2.51	2.79	2.21	2.52	2.66	2.35

Analysis of this data is presented in Table 10.9. It was observed that the treatment effects were not significant at 5% level of significance.

Table 10.9: ANOVA (with original data)

Source	DF	SS	M S	F Value	Prob > F
Replication	3	1.73105000	0.57701667	8.64	0.0003
Treatment	9	0.63781000	0.07086778	1.06	0.4206
Error	27	1.80225000	0.06675000		
Total	**39**	4.17111000			

Then the Cook statistic was computed for every observation. Cook statistics are presented in Table 10.10. It is observed from the Table that the observation number 14 stands out. It was tested with F value and found that this observation is statistically influential. No other observation is found to be influential.

Table 10.10: Cook-statistic values

Sl. No.	Replication	Treatment	Cook-statistic	Sl. No.	Replication	Treatment	Cook-statistic
1	1	1	0.0003126	21	3	1	0.0044407
2	1	2	0.0446265	22	3	2	0.0060796
3	1	3	0.0051954	23	3	3	0.0448183
4	1	4	0.0062219	24	3	4	0.022109
5	1	5	0.0065846	25	3	5	0.1292313
6	1	6	0.0124363	26	3	6	0.0147602
7	1	7	0.0134679	27	3	7	0.0120352
8	1	8	0.0005345	28	3	8	0.0233389
9	1	9	0.0491404	29	3	9	0.0002813
10	1	10	0.0000906	30	3	10	0.0000347
11	2	1	0.0003455	31	4	1	0.0009225
12	2	2	0.003801	32	4	2	0.0517879
13	2	3	0.043674	33	4	3	0.0048107
14	2	4	**0.3823402**	34	4	4	0.1526988
15	2	5	0.1122303	35	4	5	0.0111566
16	2	6	0.0063657	36	4	6	0.0080566
17	2	7	0.0145407	37	4	7	0.0110611
18	2	8	0.0818239	38	4	8	0.0121348
19	2	9	0.034714	39	4	9	0.1530533
20	2	10	0.0000742	40	4	10	0.0001498

The new statistic is then applied to identify the group of observations that are influential. It was found that observation numbers 14 and 39 are likely to be influential jointly. Cook statistic is applied for multiple outlier detection and it was found that these two observations are really

influential (Value of Cook-statistic is 0.4521055). The interesting point to note here is that though the observation number 14 was detected as outlier yet observation number 39 was not. It's effect was masked by the observation number 14. It is an interesting example of masking.

The data was reanalyzed after deleting these two observations. The result is presented in table 10.11. The dramatic effect to note here is that the treatment effects are now significant at 5% level of significance. Removal of any other pair of observations does not have any effect on the analysis.

Table 10.11: ANOVA (with 2 data points deleted)

Source	DF	SS	MS	F Value	Prob > F
Replication	3	1.20704269	0.40234756	11.58	<0.0001
Treatment	9	0.70698849	0.07855428	2.26	0.0519
Error	25	0.86835040	0.03473402		
Total	37	2.78238158			

10.5   Remedial measures

What to do with the outlier is the next question that needs to be answered once an outlier has been detected. The outlier can be deleted if it does not affect the connectedness property of the design, i.e., if it does not affect the estimation of all the elementary treatment contrasts. On the other hand, if there are too many outliers and removal of these observations affects severely the design properties, one may adopt a robust method of analysis. In robust method of analysis, analytical procedure is modified in such a way that the effect of outliers is minimized on the final results of the analysis. Here, we show both types of remedial measures.

10.5.1 Removal of observations

It is well known that a randomized block design remains connected after losing one observation. In example 10.2, we detected a single observation as outlier. We, therefore, carry out the analysis of this experiment again after removing this observation. The data is now treated as non-orthogonal and analysis is carried out using PROC GLM in SAS. The results of this analysis are presented in Table 10.12. The dramatic effects of removing this observation are worth noticing. The treatment effects now become significantly different at 5% level of significance. Removal of any other observation does not affect the analysis.

Table 10.12: ANOVA (after removing observation No. 8)

Source	DF	SS	MS	F Value	Prob > F
Replication	2	0.04919076	0.02459538	19.95	<.0001
Treatment	11	0.08356098	0.00759645	6.16	0.0002
Error	21	0.02588826	0.00123277		
Total	34	0.15864000			

10.5.2 Robust analysis

When the observations in the linear regression model are normally distributed, the method of least squares is a good parameter estimation procedure in the sense that it produces estimator of the parameters that has good statistical properties. However, there are many situations where we have evidence that the distribution of the response variable is considerable non-normal, and / or there are outliers that affect the regression model.

To deal with such type of situations, robust regression is an alternative. A robust regression procedure is one that dampens the effect of observations that would be highly influential if least squares were used. That is, a robust procedure tends to leave the residuals associated with outliers large, thereby making the identification of influential points much easy. In addition to insensitivity to outliers, a robust estimation procedure should produce essentially the same results as least squares when the underlying distribution is normal and there are no outliers. A widely used robust estimator is M-estimator. Another useful method is LMS estimator. These two methods in designed experiments are illustrated with some real experimental data sets.

10.5.3 M-estimator

The motivation for much of the work in robust estimation was due to Huber (1964). Subsequently, there are several type of robust estimators proposed. One of the most popular robust methods is M-estimation. In this method, the objective function to be minimized to get the parameter estimates is weighted according to the residual of each observation. Literature on robust regression particularly on M-estimation is now vast. A good number of objective functions to be minimized are proposed. Most of these functions are non-linear in nature and therefore, normal equations for solving the parameter estimates are also non-linear in parameters. Iteratively Re-weighted Least Squares methods are employed to solve these equations. One example where Huber's *t*-function has been applied to experimental data obtained from "Agricultural Field Experiments Information System (AFEIS)" is considered here. For illustration, consider the following example.

Example 10.4 {AFEIS Reference No. 1987(239)} An experiment with 6 treatments was carried out in the randomized complete block (RCB) design with 4 replications at Mahatma Phule Agricultural University, Rahuri, Maharashtra in 1987 with a view to test the validity of fertilizer adjustment equation in Groundnut (Net plot size: 3.00m × 3.75m). The treatments of the experiments are as follows.

T_0 = Control (No Fertilizer)

T_1 = 25Kg/Ha N + 50Kg/Ha P_2O_5

T_2 = As Per Soil Test (38Kg/Ha N + 50Kg/Ha P_2O_5)

T_3 = 15 Qt/Ha Target (11Kg/Ha N + 16Kg/Ha K_2O)

T_4 = 20 Qt/Ha Target (32Kg/Ha N + 51Kg/Ha P_2O_5 +31Kg/Ha K_2O)

T_5 = 25 Qt/Ha Target (52Kg/Ha N + 10Kg/Ha P_2O_5 +56Kg/Ha K_2O)

The data on yield per plot in quintals for different treatments is given in Table 10.13.

Table 10.13 Yield of sugarcane in kg/plot

Replication	Treatments					
	1	2	3	4	5	6
1	3.70	4.50	4.63	4.08	4.15	4.06
2	3.43	3.90	3.77	3.80	3.79	4.00
3	2.60	3.67	2.53	4.35	3.27	3.13
4	3.37	3.62	3.37	3.43	3.47	3.38

Analysis of this data is presented in Table 10.14. It is observed that the treatment effects are not significantly different at 5% level of significance. To begin with the Huber's t-function is applied. M-estimation procedure is then applied to the data. The result is presented in Table 10.15. The dramatic effect to note here is that the treatment effects are now significant at 5% level of significance.

Table 10.14: ANOVA with original data

Source	DF	SS	MS	F Value	Prob > F
Replication	3	3.01043333	1.00347778	7.62	0.0025
Treatment	5	1.15808333	0.23161667	1.76	0.1823
Error	15	1.97661667	0.13177444		
Total	23	6.14513333			

Table 10.15: ANOVA (M-estimation)

Source	DF	SS	MS	F Value	Prob > F
Replication	3	3.0003420	1.0001473	10.51	.00056
Treatment	5	1.585680	0.31713	3.3339427	0.03187
Error	15	1.4268512	0.0951234		
Total	**23**	4.37683636			

The Cook-statistic is also applied to identify outlying observations, if any. It is found that observation numbers 15 and 16 are influential (Value of Cook-statistic is 0.5757).

The data is reanalyzed after deleting these two observations. The results are presented in Table 10.16. Once again the dramatic effect to note is that the treatment effects are now significant at 5% level of significance. The result is similar to that obtained through M-estimation.

Table 10.16: ANOVA (with 2 data points deleted)

Source	DF	SS	MS	F Value	Prob > F
Replication	3	2.99734470	0.99911490	27.42	<.0001
Treatment	5	0.90585278	0.18117056	4.97	0.0092
Error	13	0.47363889	0.03643376		
Total	21	4.37683636			

10.5.4 LMS-estimator

It is well known that least squares (LS) model can be distorted even by a single outlying observation. The fitted line or surface might be tipped so that it no longer passes through the bulk of the data. A complete name for the LS method would perhaps be least sum of squares, but apparently few people have objected to the deletion of the word "sum" – as if the only sensible thing to do with *n* positive numbers would be to add them. Perhaps as a consequence of its historical name, several people have tried to make this estimator robust by replacing the square by something else, not touching the summation sign (M-estimator belongs to this group). Why not, however, replace the sum by a median, which is very robust? This yields the least median of squares (LMS) estimator, given by

$$\underset{\theta}{Minimize} \quad \underset{i}{med} \; r_i^2 \tag{10.2}$$

where θ is the parameter vector and r_i is the *i*th residual. This estimator was proposed by Rousseeuw (1984). It turns out that this estimator is very robust with respect to outliers.

Though LMS estimator has some good properties, yet this method did not get much popularity in designed experiments. LMS method gives parameter estimates based on clean observations only and thus, outliers or distributional extreme observations cannot create any problem in parameter estimation or rather they do not have any impact on parameter estimation. One of the possible reasons why LMS method is not being used in designed experiments might be its computational difficulties. There is no exact formula for computing this estimator explicitly in linear regression models. Rousseeuw (1984) provided an algorithm for computing this estimator in linear regression models. As mentioned earlier, by this algorithm all possible subsets of size *p*, where *p* is the number of parameters in the model, are fitted separately. Residuals from each of these fitted models are calculated. The median of the squared residuals for each set is calculated. The subset that gives minimum median is chosen as the final set and analysis is carried out on this subset. Application of this algorithm to designed experiments possesses some problems. The main problem is the problem of connectedness of the design. If we choose the size of subset as *p*, the design may become disconnected for some subsets or all subsets. Connectedness property is a very important property for designed experiments. Secondly, in case of design of experiments, we are mainly interested in estimation of some functions of treatment effects. This will also be severely affected if we choose a very small subset of data for estimating the treatment effects. Combating all these problems, we propose an appropriate LMS procedure for application to designed experiments.

As mentioned earlier, since the connectedness is the main problem in designed experiments, LMS method as such cannot be applied. Therefore, this is appropriately modified and then applied to experimental data taken from AFEIS. The LMS method is primarily designed to tackle the problem of outliers. In case of designed experiments, generally one or two outlying observations are present in a particular data set. We, therefore, proposed LMS method in the following manner:

i) Consider the size of the subset as $n - 1$ or $n - 2$. Here, we assume that the design remains connected after losing one or two observations.

ii) Obtain least squares residuals for each subset. There will be in total $^nC_{n-1}$ or $^nC_{n-2}$ subsets of data.

iii) Square the residuals and obtain the median for each subset.

iv) Retain that subset which yields minimum median among all subsets.

v) Carry out usual analysis on the chosen subset

It is well known that all randomized complete block designs are robust against the loss of any two observations, *i.e.*, these designs remain connected even after losing two observations. Therefore, there is no problem to apply LMS technique to RCB design, by taking the size of the subset as $n - 2$. There are also many block designs that are robust against the loss of one or two observations. However, this size of subset can be increased for those designs that are robust against the loss of more than two observations. We illustrate this method with the following example.

Example 10.5: {AFEIS Reference No. 1993(049)} An experiment with 10 weeding treatments was carried out in randomized complete block design with 3 replications at G.K.V.K., Bangalore, Karnataka with a view to study the integrated weed management in cowpea (Net plot size: 3.60m × 2.80m). The treatments of the experiment are as follows:

T_0 = Weedy check

T_1 = Weed free

T_2 = Sowing at 30 cm row spacing

T_3 = 0.75

T_4 = 1.00

T_5 = 1.25 Kg a.i/ha of pendimethalin

T_6 = Hand weeding at 3 weeks after sowing (w.a.s.)

T_7 = Interculturing at 3 w.a.s

T_8 = T3+ Hand weeding at 3 w.a.s

T_9 = T3+ Interculturing at 3 w.a.s

The data on grain yield per plot in quintals for different treatments is given in Table 10.17.

Table 10.17: Yield of cowpea in quintal/plot

Treatments	Replication		
	1	2	3
1	0.36	0.68	1.52
2	1.35	1.50	1.35
3	1.15	1.31	0.48
4	0.97	1.10	0.59
5	1.15	1.40	1.05
6	0.75	1.25	0.80
7	0.88	1.30	0.67
8	0.80	1.15	0.6
9	1.10	1.45	1.41
10	0.95	1.72	0.98

The usual analysis of the data is first carried out. The analysis of variance table is given in Table 10.18. From the table it is observed that the treatment effects are not significant at 5% level of significance, whereas block effects are significant at 5% level of significance.

Table 10.18: ANOVA with the original data

Source of variation	DF	SS	MS	F	Prob > F
Treatment	9	1.136	0.126	1.50	0.223
Block	2	0.772	0.386	4.58	0.024
Error	18	1.5195	0.084		
Total	29	3.428			

The LMS technique is then applied to this data set. Size of the subset is chosen as $n - 1$. The analysis of variance is done on the chosen subset. The result is presented in Table 10.19.

Table 10.19: ANOVA through LMS

Source	DF	SS	MS	F	Pr.>F
Treatment	9	1.784	0.198	6.690	0.0004
Block	2	0.920	0.460	15.529	0.0001
Error	17	0.503	0.029		
Total	28	3.208			

It is observed from the table that the treatment effects are now significantly different and block effects have now become highly significant. The Cook statistic is then applied to check whether any outlier is present or not. It is found that this data set contains one outlier. Incidentally in the chosen subset, this outlying observation was deleted.

As the analysis is quite involved and requires long SAS / R Codes, these codes are not presented here. However, the interested reader may get in touch with authors to know more about these codes. For more details on outliers in designed experiments, one may refer to Bhar *et al.* (2008).

11

Multivariate Analysis of Variance

11.1 Introduction

The focus of discussion so far has been on experiments where only one response variable from one experimental unit is recorded for observations. But there do occur many experimental situations where more than one response variable is recorded for observations from each experimental unit. Industrial experiments are generally designed to record observations on more than one response variable. Similarly, agricultural research more often than not involves design of experiments where observations on more than one response variable are recorded. Even when more than one response variable is recorded for observations, every response variable is analyzed individually assuming that all the response variables are independent. Another very common approach adopted for analysis of data from several response variables is that many response variables are converted into a single response variable by taking some suitable function of all the response variables, thus converting multi-response (multivariate) experiment to a single response (univariate) experiment. If all the response variables are measured in terms of yield, then one possible function could be the monetary value of the total produce from the entire response variables. Other function could be the total calorie value of the produce. Yet another function could be the energy value of the total produce, the energy value being computed through calorie value. Once some index or a function of the entire response variables is developed, the data are analyzed in the usual way for a single response variable. All the inferences made are then applicable to the new response variable generated as a function of the original response variables.

Consider an agricultural experiment where the two response variables recorded for observations are the grain yield and the straw yield. The other response variables on which the data are generally recorded are the plant height, number of green leaves, germination count, etc. Generally, the analysis is carried out only on the grain yield in conformity with the design adopted. The best treatment is identified on the basis of grain yield alone. The straw yield is generally not taken into consideration during the analysis. However, although the grain yield is important for feeding the human population, the straw yield is also important either for the cattle feed or for mulching or manuring, etc. Therefore, while analyzing the data, the straw yield should also be taken into consideration. Similarly, in varietal trials, the data are collected on several plant characteristics and quality parameters. In these experimental situations the data is generally analyzed separately for each of the characters. The best treatment or genotype is identified separately for each of the characters. The problems starts when different treatments are identified for different response variables.

When several response variables are recorded and the data are analyzed separately for each response variable as if they are independent or if the data are analyzed by converting several response variables into a single response variable by a suitable function, then the advantage of correlations present among several response variables is lost. In order to derive advantage from the correlations among several response variables, it would be appropriate to analyze all the response variables simultaneously. Such experiments in which several response variables are recorded for observations and analyzed simultaneously are known as multi-response experiments. In these situations, multivariate analysis of variance (MANOVA) can be helpful and advantageous. A general procedure of performing MANOVA on the data generated from RCB design is described and then illustrated with the help of an example. The procedure of SAS used for MANOVA is exactly the same for any general block design and similarly can be extended to any experimental design setting.

11.2   Multivariate analysis of variance

Consider an experiment conducted to compare v treatments using a randomized complete block (RCB) design with b blocks (or replications) and the observations (or data) are collected on p response variables. We shall assume throughout that observations on all the p variables are recorded on each experimental unit. In other words, there is neither any missing response on any of the p variables on any experimental unit nor all the p responses are missing on one or more experimental units. Let y_{ijk} denote the observed value of the kth response variable for the ith treatment in the jth block, $i = 1, 2, \cdots, v;\ j = 1, 2, \cdots, b; k = 1, 2, \cdots, p$. The total number of observations is actually vb, with each observation being a p-component vector corresponding to p response variables. But if one unrolls the p-variate observations, then the total number of observations would be vbp. Alternatively, therefore, we can write the vb observation vectors as $\mathbf{y}_{ij}$, where $\mathbf{y}_{ij}$ is a p-component vector of responses on all the p response variables pertaining to ith treatment in jth block. In fact $\mathbf{y}_{ij} = (y_{ij1}, y_{ij2}, \cdots, y_{ijk}, \cdots, y_{ijp})'$. For performing the MANOVA, the data from the experimental set up can be rearranged as in Table 11.1.

Table 11.1: Data layout for MANOVA

Treatments	Blocks						
	1	2	...	j	...	b	Treatment Mean
1	$\mathbf{y}_{11}$	$\mathbf{y}_{12}$	...	$\mathbf{y}_{1j}$	...	$\mathbf{y}_{1b}$	$\overline{\mathbf{y}}_{1.}$
2	$\mathbf{y}_{21}$	$\mathbf{y}_{22}$	...	$\mathbf{y}_{2j}$	...	$\mathbf{y}_{2b}$	$\overline{\mathbf{y}}_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
i	$\mathbf{y}_{i1}$	$\mathbf{y}_{i2}$	...	$\mathbf{y}_{ij}$	...	$\mathbf{y}_{ib}$	$\overline{\mathbf{y}}_{i.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
v	$\mathbf{y}_{v1}$	$\mathbf{y}_{v2}$	...	$\mathbf{y}_{vj}$	...	$\mathbf{y}_{vb}$	$\overline{\mathbf{y}}_{v.}$
Block Mean	$\overline{\mathbf{y}}_{.1}$	$\overline{\mathbf{y}}_{.2}$	...	$\overline{\mathbf{y}}_{.j}$	...	$\overline{\mathbf{y}}_{.b}$	$\overline{\mathbf{y}}_{..}$

We shall let $\bar{y}_{i.k} = \frac{1}{b}\sum_{j=1}^{b} y_{ijk}$ denotes the mean of the ith treatment pertaining to the kth response variable and $\bar{y}_{.jk} = \frac{1}{v}\sum_{i=1}^{v} y_{ijk}$ denotes the mean of the jth block pertaining to the kth response variable. We shall also let $\bar{\mathbf{y}}_{i.} = (\bar{y}_{i.1}, \bar{y}_{i.2}, \cdots, \bar{y}_{i.k}, \cdots, \bar{y}_{i.p})'$ denote a p-component vector of ith treatment means, $i = 1, 2, \cdots, v$; and $\bar{\mathbf{y}}_{.j} = (\bar{y}_{.j1}, \bar{y}_{.j2}, \cdots, \bar{y}_{.jk}, \cdots, \bar{y}_{.jp})'$ denote a p-component vector of jth block means, $j = 1, 2, \cdots, b$. The treatment mean and block mean vectors can alternatively be written as $\bar{\mathbf{y}}_{i.} = \frac{1}{b}\sum_{j=1}^{b} \mathbf{y}_{ij}$; $\bar{\mathbf{y}}_{.j} = \frac{1}{v}\sum_{i=1}^{v} \mathbf{y}_{ij}$, Further, let $\bar{y}_{..k} = \frac{1}{vb}\sum_{i=1}^{v}\sum_{j=1}^{b} y_{ijk}$ denote the grand mean of the kth response variable, $k = 1, 2, \cdots, p$ Let $\bar{\mathbf{y}}_{..} = (\bar{y}_{..1}, \bar{y}_{..2}, \cdots, \bar{y}_{..k}, \cdots, \bar{y}_{..p})'$ denote the vector of overall means of the p response variables. The vector of overall means can alternatively be written as $\bar{\mathbf{y}}_{..} = \frac{1}{vb}\sum_{i=1}^{v}\sum_{j=1}^{b} \mathbf{y}_{ij}$.

For the sake of completeness, the procedure of analysis is described in the sequel. The experimenters may like to skip this portion and proceed straight to Section 11.4. The observations can be represented by a two-way classified multivariate model $\boldsymbol{\Omega}$

$$\boldsymbol{\Omega}: \mathbf{y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \boldsymbol{\beta}_j + \mathbf{e}_{ij} \quad i = 1, 2, \ldots, v; j = 1, 2, \ldots, b$$

where y_{ij} are the p-variate vector of observed values from the experimental unit receiving the treatment i in block j. This indicates that unlike in the univariate case where the experimenter observes only one response from the treatment i in block j, in multi-response experiment, the experimenter observes a vector of p components as a response from the treatment i in block j. Let μ_k denote the general mean corresponding to the kth response variable, and $\boldsymbol{\mu} = (\mu_1 \quad \mu_2 \quad \cdots \quad \mu_k \quad \cdots \quad \mu_p)'$ be the $p \times 1$ vector of general means. Similarly, let τ_{ik} denote the effect of the ith treatment corresponding to the kth response variable, and $\boldsymbol{\tau}_i = (\tau_{i1} \quad \tau_{i2} \quad \cdots \quad \tau_{ik} \quad \cdots \quad \tau_{ip})'$ be the p-component vector of the ith treatment effect. Further, let β_{jk} denote the effect of the jth block corresponding to the kth response variable, and $\boldsymbol{\beta}_j = (\beta_{j1} \quad \beta_{j2} \quad \cdots \quad \beta_{jk} \quad \cdots \quad \beta_{jp})'$ be the p-component vector of jth block effect. $\mathbf{e}_{ij} = (e_{ij1} \quad e_{ij2} \quad \cdots \quad e_{ijk} \quad \cdots \quad e_{ijp})'$ is a p-component random vector associated with y_{ij} and assumed to be having a p-variate normal distribution $N_p(\mathbf{0}, \boldsymbol{\Sigma})$. Here $\mathbf{0}$ is a vector of all zeros indicating that the mean of all the components of e_{ij} is zero. It may be worthwhile mentioning here that the variance covariance matrix $\boldsymbol{\Sigma}$ has the elements as $V(e_{ijk}) = \sigma_{ii}$, $Cov(e_{ijk}, e_{i'j'k}) = 0$ and $Cov(e_{ijk}, e_{ijk'}) = \sigma_{ij}$. We want to test the equality of treatment effects i.e., $H_0 : (\tau_{i1} \quad \tau_{i2} \quad \cdots \quad \tau_{ik} \quad \cdots \quad \tau_{ip})' = (\tau_1 \quad \tau_2 \quad \cdots \quad \tau_k \quad \cdots \quad \tau_p)'$ (say) $\forall i = 1, 2, \cdots, v$, i.e., $\boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_i = \cdots = \boldsymbol{\tau}_v = \boldsymbol{\tau}$ against the alternative H_1 : at least two of the treatment effects are unequal. Under the null hypothesis, the model reduces to $\boldsymbol{\Omega}_0 : \mathbf{y}_{ij} = \boldsymbol{\alpha} + \boldsymbol{\beta}_j + \mathbf{e}_{ij}$ where $\alpha_k = \mu_k + \tau_k$ and $\boldsymbol{\alpha} = (\mu_1 + \tau_1, \mu_2 + \tau_2, \cdots, \mu_p + \tau_p)'$.

An outline of MANOVA Table for testing the equality of treatment effects and block effects is given in Table 11.2.

Table 11.2: MANOVA table

Source	DF	SSCPM (Sum of Squares and Cross Product Matrix)
Treatment	$v - 1 = h$	$\mathbf{H} = b\sum_{i=1}^{v} (\bar{y}_{i.} - \bar{y}_{..})(\bar{y}_{i.} - \bar{y}_{..})'$
Block	$b - 1 = t$	$\mathbf{B} = v\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})(\bar{y}_{.j} - \bar{y}_{..})'$
Residual	$(v-1)(b-1) = s$	$\mathbf{R} = \sum_{i=1}^{v}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})'$
Total	$vb - 1$	$\mathbf{T} = \sum_{i=1}^{v}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})' = \mathbf{H} + \mathbf{B} + \mathbf{R}$

Here $\mathbf{H}, \mathbf{B}, \mathbf{R}$ and $\mathbf{T}$ are the sum of squares and sum of cross products matrices of treatments, blocks, errors (residuals) and totals respectively. The residual sum of squares and cross products matrix for the reduced model $\boldsymbol{\Omega}_0$ is denoted by $\mathbf{R}_0$ and is given by

$$\mathbf{R}_0 = \mathbf{R} + \mathbf{H}$$

We reject the null hypothesis of equality of treatment mean vectors if the ratio of generalized variance (*Wilk's lambda* statistic) $\Lambda = \dfrac{|\mathbf{R}|}{|\mathbf{H} + \mathbf{R}|}$ is too small. Here $|\mathbf{U}|$ denotes the determinant of the matrix U. Assuming multivariate normal distribution, Rao (1973) showed that under null hypothesis Λ is distributed as the product of independent beta variables. A better but more complicated approximation of the distribution of is Λ

$$\frac{1 - \Lambda^{1/b}}{\Lambda^{1/b}} \frac{(ab - c)}{ph} \sim F(ph, ab - c)$$

where $a = \left(s - \dfrac{p - h + 1}{2}\right)$, $b = \sqrt{\left\{(p^2 h^2 - 4)/(p^2 + h^2 - 5)\right\}}$, $c = \dfrac{ph - 2}{2}$

For some particular values of h and p, it reduces to exact Snedecor's F-distribution. The special cases are given below:

For $h = 1$ and any p, the distribution of reduces to

$$\frac{(1 - \Lambda)}{\Lambda} \frac{(s - p + 1)}{p} \sim F(p, s - p + 1).$$

For $h = 2$ and any p, the distribution of reduces to

$$\frac{(1 - \sqrt{\Lambda})}{\sqrt{\Lambda}} \frac{(s - p + 1)}{p} \sim F\{2p, 2(s - p + 1)\}.$$

For $p = 2$ and any h, the distribution of reduces to

$$\frac{(1 - \sqrt{\Lambda})}{\sqrt{\Lambda}} \frac{(s - 1)}{h} \sim F(2h, 2(s - 1)).$$

For $p = 1$, the statistic reduces to the usual variance ratio statistics or usual Snedecor's *F*.

The hypothesis regarding the equality of block effects can be tested by replacing Λ by $\dfrac{|\mathbf{R}|}{|\mathbf{B} + \mathbf{R}|}$ and h by t in what has been described above.

Several other criteria viz. Pillai's Trace, Hotelling-Lawley Trace or Roy's Greatest Root are available in literature for testing the null hypothesis in MANOVA. Wilks' Lamda is, however, the commonly used criterion. In this Chapter, however, we shall restrict to the use of Wilks' Lamda criterion. For further details on MANOVA, a reference may be made to Seber (1983) and Johnson and Wichern (1988). Some more examples of MANOVA may be seen from Parsad *et al.* (2004).

Remark 11.1 One complication of multivariate analysis that does not arise in the univariate case is due to the ranks of the matrices. The rank of **R** should not be smaller than p or in other words error degrees of freedom s should be greater than or equal to p ($s \geq p$).

11.3 Multivariate treatment contrast analysis

If the treatments are found to be significantly different through MANOVA, then the next question is "which treatments are significantly different?" This question can be answered through multivariate treatment contrast analysis. In the literature, the multivariate treatment contrast analysis is generally carried out using the χ^2-statistic. The χ^2-statistic is based on the assumption that the error variance-covariance matrix is known. The error variance-covariance matrix is, however, generally unknown. Therefore, the estimated error variance-covariance matrix is used in place of variance-covariance matrix. The error variance-covariance matrix is estimated by sum of squares and cross products (SSCP) matrix for error divided by the error degrees of freedom. As a consequence, the test based on χ^2-statistic is an approximate solution. A multivariate treatment contrast analysis using the Wilk's Lambda criterion is described in the sequel. For the sake of completeness, the procedure based on χ^2-statistic is also described. Suppose we want to test $H_0: \tau_i = \tau_{i'}$ against $H_1: \tau_i \neq \tau_{i'}$. The above hypothesis can be rewritten as

$$H_0: = (\tau_i - \tau_{i'}) = \mathbf{0} \text{ against } H_1: = (\tau_i - \tau_{i'}) \neq \mathbf{0},$$

where $(\tau_i - \tau_{i'})' = (\tau_{i1} - \tau_{i'1} \quad \tau_{i2} - \tau_{i'2} \quad \cdots \quad \tau_{ik} - \tau_{i'k} \quad \cdots \quad \tau_{ip} - \tau_{i'p})$. Here τ_{ik} denotes the effect of treatment i for the dependent variable k. The best linear unbiased estimate of $(\tau_i - \tau_{i'})$ is

$$(\bar{y}_{i.} - \bar{y}_{i'.}) = (\bar{y}_{i1} - \bar{y}_{i'1} \quad \bar{y}_{i2} - \bar{y}_{i'2} \quad \cdots \quad \bar{y}_{ik} - \bar{y}_{i'k} \quad \cdots \quad \bar{y}_{ip} - \bar{y}_{i'p})$$

where $\bar{y}_{ik}$ is the mean of treatment i for variable k.

11.3.1 χ^2 test

The test statistic based on χ^2-statistic, requires covariance matrix of the contrast of interest. The covariance matrix, in case of a RCB design for elementary treatment contrasts is obtained by dividing the SSCP matrix for errors obtained in MANOVA by half of the product of error degrees of freedom and the number of replications. Let this variance-covariance matrix be denoted by Σ_c. Under null hypothesis, $x_i = \bar{y}_{i.} - \bar{y}_{i'.}$ follows p-variate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix Σ_c. Applying the Aitken's transformation, it can be shown that $\mathbf{z}_i = \Sigma_c^{-1/2} \mathbf{x}_i$ follows a p-variate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix I_g, where I_g, denotes the identity matrix of order g. Then using the results of quadratic forms, it can easily be seen that $\mathbf{z}_i' \mathbf{z}_i = \mathbf{x}_i' \Sigma_c^{-1} \mathbf{x}_i$ follows a χ^2 distribution with p degrees of freedom.

11.3.2 Wilk's Lambda criterion

For testing the null hypothesis $H_0: = (\tau_i - \tau_{i'}) = \mathbf{0}$ against $H_1: = (\tau_i - \tau_{i'}) \neq \mathbf{0}$, we obtain a sum of squares and products matrix for the above elementary treatment contrast. Let the SSCP matrix for above elementary treatment contrast be $\mathbf{G}_{p \times p}$. The diagonal elements of $\mathbf{G}$ are then obtained by

$$g_{kk} = \left(\frac{b}{2}\right)\left(\bar{y}_{ik} - \bar{y}_{i'k}\right)^2 \quad \forall \quad k = 1,2,...,p; \ i \neq i' = 1,2,...,v$$

and the off-diagonal elements are obtained by

$$g_{kk'} = \frac{b}{2}\left(\bar{y}_{ik} - \bar{y}_{i'k}\right)\left(\bar{y}_{ik'} - \bar{y}_{i'k'}\right).$$

We reject the null hypothesis if the value of Wilk's Lambda $\Lambda^* = \frac{|\mathbf{R}|}{|\mathbf{G}+\mathbf{R}|}$ is small, where R is the SSCP matrix due to residuals as obtained through MANOVA. The hypothesis is then tested using the following Snedecor's *F*-test statistics based on Wilk's Lambda for $h = 1$

$$\frac{1-\Lambda^*}{\Lambda^*}\frac{edf - p + 1}{p} \sim F(p, \ s - p + 1)$$

where *edf* denote error degrees of freedom.

Remark 11.2 The above procedure is for the situations when the experiment is conducted using a RCB design. There, however, exist situations, where the use of RCB design may not be possible and one has to use incomplete block design. It seems that a stepwise procedure of analysis of multi-response data from incomplete block designs is not available. Keeping this in view, Nandi (2007) developed a stepwise procedure of the analysis of data generated from complete multi-response experiments conducted in block designs. The general procedure is beyond the scope of this book and the readers interested in detail may refer to Nandi (2007).

11.4 Example

The purpose of this section is to give the results obtained from bivariate analysis of variance of the data generated from the experiments conducted under the aegis of Project Directorate of Cropping Systems Research (PDCSR) now known as of Indian Institute of Farming Systems Research (IIFSR), Modipuram, where the data on grain yield and straw yield were observed in the multi-response experiment.

An experiment *entitled Studies on the experimentation on conservation of organic carbon in the soil to improve soil condition* was conducted at Bhubaneshwar on rice crop. The experiment was initiated in the year 1997. The data on grain and straw yield used for the illustration pertains to the Rabi season of 2000. Ten treatments were tried in the experiment. The details of the treatments are given below:

T1 - Recommended N 100%

T2 - Recommended N 100% out of which 10 Kg at first ploughing

T3 - Recommended N 100% out of which 20 Kg at first ploughing

T4 - Recommended N 100% and add 10 Kg N/ha at first ploughing

T5 - Recommended N 100% and add 20 Kg N/ha at first ploughing

T6 - Recommended N + 10 Kg N/ha

 T7 - Recommended N + 20 Kg N/ha

T8 - Recommended N + cellulose decomposing enzyme (FYM)

T9 - Recommended N + FYM 5 t/ha during Kharif

T10 - Recommended N + FYM 5 t/ha during Rabi.

The data in q/ha is given in Table 11.3.

Table 11.3: Grain and straw yield data

Treatment	Grain Yield in q/ha				Straw Yield in q/ha			
	Block 1	Block 2	Block 3	Block 4	Block 1	Block 2	Block 3	Block 4
1	36.10	26.60	32.20	23.85	40.40	30.00	36.25	26.75
2	32.00	44.20	33.20	37.70	37.10	51.25	38.50	43.70
3	32.00	40.00	47.15	41.90	45.50	47.25	56.00	49.00
4	46.80	38.25	44.65	52.60	55.25	45.00	52.50	62.10
5	49.50	41.75	53.55	51.85	59.25	50.25	64.25	62.20
6	49.75	53.50	49.40	58.50	60.30	64.70	59.75	70.75
7	57.25	57.00	44.00	53.75	69.75	70.00	53.50	65.50
8	59.50	52.00	47.00	57.50	69.00	60.75	54.50	61.00
9	51.60	57.25	50.25	51.75	61.40	68.10	59.75	59.50
10	62.50	59.00	43.00	45.75	74.50	70.10	51.25	51.25

Let us perform separate analysis for each of the two characters and multivariate analysis of variance taking both the characters into consideration simultaneously.

11.4.1 Analysis of Data

In what follows are given the SAS commands and the data structure for the analysis of bivariate data.

```
DATA multi_response_experiment;
INPUT rep trt  gyld syld;
CARDS;
1       1       36.10   40.4
1       2       32.00   37.1
1       3       32.00   45.5
1       4       46.80   55.25
1       5       49.50   59.25
```

1	6	49.75	60.30
1	7	57.25	69.75
1	8	59.50	69.00
1	9	51.60	61.40
1	10	62.50	74.50
2	1	26.60	30.00
2	2	44.20	51.25
2	3	40.00	47.25
2	4	38.25	45.00
2	5	41.75	50.25
2	6	53.50	64.70
2	7	57.00	70.00
2	8	52.00	60.75
2	9	57.25	68.10
2	10	59.00	70.10
3	1	32.20	36.25
3	2	33.20	38.50
3	3	47.15	56.00
3	4	44.65	52.50
3	5	53.55	64.25
3	6	49.40	59.75
3	7	44.00	53.50
3	8	47.00	54.50
3	9	50.25	59.75
3	10	43.00	51.25
4	1	23.85	26.75
4	2	37.70	43.70
4	3	41.90	49.00
4	4	52.60	62.10
4	5	51.85	62.20
4	6	58.50	70.75
4	7	53.75	65.50
4	8	57.50	61.00
4	9	51.75	59.50
4	10	45.75	51.25

```
;
PROC GLM data = multi_response_experiment;
CLASS rep trt;
MODEL gyld syld = rep trt / ss3;
MEANS trt/LSD;
CONTRAST   '1 vs 2'   trt  1 -1  0  0  0  0  0  0  0  0;
CONTRAST   '1 vs 3'   trt  1  0 -1  0  0  0  0  0  0  0;
CONTRAST   '1 vs 4'   trt  1  0  0 -1  0  0  0  0  0  0;
```

```
CONTRAST    '1 vs 5'    trt  1 0 0 0 -1 0 0 0 0 0;
CONTRAST    '1 vs 6'    trt  1 0 0 0 0 -1 0 0 0 0;
CONTRAST    '1 vs 7'    trt  1 0 0 0 0 0 -1 0 0 0;
CONTRAST    '1 vs 8'    trt  1 0 0 0 0 0 0 -1 0 0;
CONTRAST    '1 vs 9'    trt  1 0 0 0 0 0 0 0 -1 0;
CONTRAST    '1 vs 10'   trt  1 0 0 0 0 0 0 0 0 -1 ;
CONTRAST    '2 vs 3'    trt  0 1 -1 0 0 0 0 0 0 0;
CONTRAST    '2 vs 4'    trt  0  1 0 -1 0 0 0 0 0 0;
CONTRAST    '2 vs 5'    trt  0 1 0 0 -1 0 0 0 0 0;
CONTRAST    '2 vs 6'    trt  0 1 0 0 0 -1 0 0 0 0;
CONTRAST    '2 vs 7'    trt  0 1 0 0 0 0 -1 0 0 0;
CONTRAST    '2 vs 8'    trt  0 1 0 0 0 0 0 -1 0 0;
CONTRAST    '2 vs 9'    trt  0 1 0 0 0 0 0 0 -1 0;
CONTRAST    '2 vs 10'   trt  0 1 0 0 0 0 0 0 -1 0;
CONTRAST    '3 vs 4'    trt  0 0 1 -1 0 0 0 0 0 0;
CONTRAST    '3 vs 5'    trt  0 0 1 0 -1 0 0 0 0 0;
CONTRAST    '3 vs 6'    trt  0 0 1 0 0 -1 0 0 0 0;
CONTRAST    '3 vs 7'    trt   0 0 1 0 0 0 -1 0 0 0;
CONTRAST    '3 vs 8'    trt   0 0 1 0 0 0 0 -1 0 0;
CONTRAST    '3 vs 9'    trt   0 0 1 0 0 0 0 0 -1 0;
CONTRAST    '3 vs 10'   trt   0 0 1 0 0 0 0 0 0 -1;
CONTRAST    '4 VS 5'    trt   0 0 0 1 -1 0 0 0 0 0;
CONTRAST    '4 vs 6'    trt   0 0 0 1 0 -1 0 0 0 0;
CONTRAST    '4 vs 7'    trt   0 0 0 1 0 0 -1 0 0 0;
CONTRAST    '4 vs 8'    trt   0 0 0 1 0 0 0 -1 0 0;
CONTRAST    '4 vs 9'    trt   0 0 0 1 0 0 0 0 -1 0;
CONTRAST    '4 vs 10'   trt   0 0 0 1 0 0 0 0 0 -1;
CONTRAST    '5 vs 6'    trt   0 0 0 0 1 -1 0 0 0 0;
CONTRAST    '5 vs 7'    trt   0 0 0 0 1 0 -1 0 0 0;
CONTRAST    '5 vs 8'    trt   0 0 0 0 1 0 0 -1 0 0;
CONTRAST    '5 vs 9'    trt   0 0 0 0 1 0 0 0 -1 0;
CONTRAST    '5 vs 10'   trt   0 0 0 0 1 0 0 0 0 -1;
CONTRAST    '6 vs 7'    trt   0 0 0 0 0 1 -1 0 0 0;
CONTRAST    '6 vs 8'    trt   0 0 0 0 0 1 0 -1 0 0;
CONTRAST    '6 vs 9'    trt   0 0 0 0 0 1 0 0 -1 0;
CONTRAST    '6 vs 10'   trt   0 0 0 0 0 1 0 0 0 -1;
CONTRAST    '7 vs 8'    trt   0 0 0 0 0 0 1 -1 0 0;
CONTRAST    '7 vs 9'    trt   0 0 0 0 0 0 1 0 -1 0;
CONTRAST    '7 vs 10'   trt   0 0 0 0 0 0 1 0 0 -1;
CONTRAST    '8 vs 9'    trt   0 0 0 0 0 0 0 1 -1 0;
CONTRAST    '8 vs 10'   trt   0 0 0 0 0 0 0 1 0 -1;
CONTRAST    '9 vs 10'   trt   0 0 0 0 0 0 0 0 1 -1;
MANOVA h = rep trt;
RUN;
```

Remark 11.3 It may be worthwhile mentioning here that the MODEL Statement along with MANOVA statement will provide univariate as well as multivariate analysis. If univariate analysis is not required, then one may use option NOUNI in MODEL statement. For instance, in this example, the statements would be

MODEL gyld syld = rep trt / ss3;
MANOVA h = rep trt;
/*The two SAS commands provide univariate analysis and multivariate analysis, i.e., the two statements will provide the analysis for every individual response variable separately and also for all the response variables taken together*/
or
MODEL gyld syld = rep trt / ss3 NOUNI;
MANOVA h = rep trt;
/*The two SAS commands provide multivariate analysis only, i.e., the two statements will provide the analysis for all the response variables taken together. It will not provide the analysis for every individual response variable separately*/

The results obtained are given in Table 11.4. First the results for each of the two characters are presented separately.

Table 11.4: Univariate analysis output using SAS commands

ANOVA: Grain Yield (GYLD)

Source	DF	SS	MS	F-Value	Prob > F
Model	12	2617.2218	218.1018	5.94	<.0001
Error	27	991.5130	36.7227		
Corrected Total	39	3608.7348			

R-Square	CV	Root MSE	gyld Mean
0.725	12.990	6.060	46.653

Source	DF	SS	MS	F-Value	Prob > F
Block	3	68.2782	22.7594	0.62	0.6083
Treatment	9	2548.9435	283.2159	7.71	<0.0001
Error	27	991.5130	36.7227		
Corrected Total	39	3608.7348			

ANOVA: Straw Yield (SYLD)

Source	DF	SS	MS	F-Value	Prob > F
Model	12	4029.807	335.817	6.88	<0.0001
Error	27	1317.860	48.810		
Corrected Total	39	5347.667			

R-Square	Coeff Var	Root MSE	syld Mean
0.754	12.657	6.986	55.196

Source	DF	SS	MS	F- Value	Prob > F
Block	3	111.048	37.016	0.76	0.5273
Treatment	9	3918.759	435.418	8.92	<0.0001
Error	27	1317.860	48.810		
Corrected Total	39	5347.667			

It can be seen that for both the attributes (grain yield and straw yield), the model with treatment and block effects has been able to explain about 73 and 75 per cent variability in the data, respectively. For both the attributes, the block (or replication) effects are not significantly different whereas the treatment effects are significantly different (*p*-value < 0.0001). Therefore, for making all the possible pair wise treatment comparisons, the least significant difference procedure of multiple comparisons was used. The results are given in Table 11.5.

Table 11.5: Multiple comparison of treatments using LSD

t Tests (LSD) for Grain yield

Alpha	0.05
Error Degrees of Freedom	27
Error Mean Square	36.7227
Critical Value of t	2.0518
Least Significant Difference	8.7921

Treatments with same alphabet are not significantly different				
t Grouping		Mean	N	Treatment
	A	54.000	4	8
	A	53.000	4	7
	A	52.788	4	6
	A	52.713	4	9
	A	52.563	4	10
	A	49.163	4	5
B	A	45.575	4	4
B	C	40.263	4	3
D	C	36.775	4	2
D		29.688	4	1

t Tests (LSD) for Straw yield

Alpha	0.05
Error Degrees of Freedom	27
Error Mean Square	48.8096
Critical Value of t	2.05183
Least Significant Difference	10.136

Treatments with same alphabet are not significantly different				
t Grouping		Mean	N	Treatment
	A	54.000	4	8
	A	53.000	4	7
	A	52.788	4	6
	A	52.713	4	9
	A	52.563	4	10
	A	49.163	4	5
B	A	45.575	4	4
B	C	40.263	4	3
D	C	36.775	4	2
D		29.688	4	1

t Tests (LSD) for Straw yield

Alpha	0.05
Error Degrees of Freedom	27
Error Mean Square	48.8096
Critical Value of t	2.05183
Least Significant Difference	10.136

Treatments with same alphabet are not significantly different					
t Grouping			Mean	N	Treatment
	A		64.688	4	7
	A		63.875	4	6
B	A		62.188	4	9
B	A		61.775	4	10
B	A		61.313	4	8
B	A	C	58.988	4	5
B		C	53.713	4	4
	D	C	49.438	4	3
E	D		42.638	4	2

It can be concluded that the treatment T8 (recommended N + cellulose decomposing enzyme) is highest yielding for grain yield but the treatment T7 (recommended N + 20 kg / ha) is the highest yielding for straw yield, although the two treatments are statistically at par with each other both for the grain yield and the straw yield. Treatment T1 (Recommended N) is the lowest yielding for grain yield but treatment T2 (Recommended N 100% out of which 10 Kg at first ploughing) is the lowest yielding for straw yield. Similarly, treatments T5 and T3 are significantly different for grain yield but are not significantly different for straw yield. Similarly variations can be seen for other pairs of treatments. In this example, it is easier to infer as for both the response variables, best performing treatments are statistically at par. However, as mentioned earlier, this may not hold true in general and the best treatment for different response variables may be different. Therefore, to rank the treatments collectively for both the characters, the multivariate analysis of variance is to be carried out. The results obtained are given in Table 11.6.

Table 11.6: Multivariate analysis of variance of grain yield and straw yield data

H = Type III SSCP Matrix for treatment		
	Grain Yield	Straw yield
Grain yield	2548.9435	3129.0410
Straw yield	3129.0410	3918.7588

E = Error SSCP Matrix		
	gyld	**syld**
gyld	991.513	1115.0758
syld	1115.078	1317.8599

MANOVA test criteria and Snedecor's F approximations for the Hypothesis of no overall treatment effect

Statistic	Value	F-Value	Numerator DF	Denominator DF	Prob > F
Wilks' Lambda	0.1200	5.45	18	52	<0.0001
Pillai's Trace	1.2550	5.05	18	54	<0.0001
Hotelling-Lawley Trace	4.2100	5.91	18	39.714	<0.0001
Roy's Greatest Root	3.2475	9.74	9	27	<0.0001
NOTE: F Statistic for Roy's Greatest Root is an upper bound					
NOTE: F Statistic for Wilks' Lambda is exact					

From Table 11.6, it can be concluded that the treatment effects are significantly different (p-value <0.0001).

The null hypothesis about the equality of treatment effects has been tested. The null hypothesis has been rejected. The next question is to make pair wise treatment comparisons for ranking the treatments and selecting the best treatment in terms of both the grain yield and straw yield. It has been seen while analyzing each attribute separately that treatment T8 is best

291

for grain yield and treatment T7 is best for straw yield. But the next question is to make the pair wise treatment comparisons for ranking the treatments and picking up the best treatment by considering both the grain yield and the straw yield. This question can be answered through multivariate contrast analysis. The results of multivariate treatment contrast analysis for making all possible pair wise comparisons of the treatments are given in Table 11.7.

Table 11.7: Probabilities of significance of all possible paired treatment comparisons using Wilks' Lambda criterion

Treatment	1	2	3	4	5	6
1	.	0.1525	**0.0006**	**0.0010**	**<0.0001**	**<0.0001**
2	0.1525	.	0.0388	0.0938	**0.0055**	**0.0004**
3	**0.0006**	**0.0388**	.	0.1673	0.1352	**0.0270**
4	**0.0010**	0.0938	0.1673	.	0.3945	0.0631
5	**<0.0001**	**0.0055**	0.1352	0.3945	.	
6	**<0.0001**	**0.0004**	**0.0270**	0.0631	0.5497	.
7	**<0.0001**	**0.0001**	**0.0194**	0.3271	0.3271	0.8742
8	**<0.0001**	**0.0020**	**0.0006**	0.0531	**0.0200**	**0.0058**
9	**<0.0001**	**0.0023**	**0.0181**	0.2604	0.5636	0.3653
10	**<0.0001**	**0.0030**	**0.0159**	0.2904	0.4866	0.2667

Treatment	7	8	9	10
1	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
2	**0.0001**	**0.0020**	**0.0023**	**0.0030**
3	**0.0194**	**0.0006**	**0.0181**	**0.0159**
4	0.3271	0.0531	0.2604	0.2904
5	0.3271	**0.0200**	0.5636	0.4866
6	0.8742	**0.0058**	0.3653	0.2667
7	.	**0.0017**	0.1651	0.1113
8	**0.0017**	.	0.1264	0.1828
9	0.1651	0.1264	.	0.9755
10	0.1113	0.1828	0.9755	.

*bold face type shows the treatment pairs that are significantly different

From the above results we can see that treatments T7 and T8 are significantly different where as they were not found to be significantly different when analyzed for individual characters. The ranking of the treatments is not possible using the multi-characters. If the ranks of the treatments change with different attributes, then possibly biplot can help in identifying the subset of treatments which are good for a given subset of attributes. For this the matrix of means/adjusted means of treatment versus attributes may be used. This type of analysis has been done in Chapter 13.

The experimenters seldom use MANOVA and carry out the univariate contrast analysis on the combined average values of all the dependent variables. For this purpose, the data is converted into univariate by defining an index. The index may be net returns, total calories, total energy, etc. or some weighted average of all the response variables, the weights being the relative importance of the response variables, decided in consultation with the subject matter specialist. Sometimes, to avoid the bias due to subject matter experts, the first principal component score is taken as an index. The analysis obtained using principal component analysis is given below:

To account for the correlation structure between the two variables, the principal component analysis was carried out using the following code:

```
PROC PRINCOMP data=mult cov;
VAR gyld syld;
RUN;
```

The results obtained are summarized in Table 11.8.

Table 11.8: Eigen values and eigen vectors of the covariance matrixEigen values

	Eigenvalue	Difference	Proportion	Cumulative
1	227.914864	226.178402	0.9924	0.9924
2	1.736462		0.0076	1.0000

Eigenvectors

	Principal Component 1	Principal Component 2
Grain yield	0.633586	0.773672
Straw yield	0.773672	-.633586

It can be seen that the first principal component explains 99.24% of the variance. Therefore, the principal component scores of the observations for the first principal component are obtained and the univariate analysis of variance is carried out. The results obtained are shown in Table 11.9.

Table 11.9: ANOVA of first principal component scores

Source	DF	SS	MS	F-Value	Prob > F
Model	12	6608.630	550.719	6.52	<0.0001
Error	27	2280.047	84.446		
Corrected Total	39	8888.676			

R-Square	CV	Root MSE	Principal Component 1 Score Mean
0.744	12.717	9.1895	72.2622

Source	DF	SS	MS	F-Value	Prob > F
Block	3	172.124	57.375	0.68	0.5723
Treatment	9	6436.506	715.167	8.47	<0.0001
Error	27	2280.047	84.446		
Corrected Total	39	8888.676			

It can be seen that the treatment effects are highly significant (p-value <0.0001). Therefore, multiple comparisons using the least significant difference procedure was used.

Table 11.10: Multiple comparison of treatments using LSD
t Tests (LSD) for prinscore1

Alpha	0.05
Error Degrees of Freedom	27
Error Mean Square	84.4462
Critical Value of t	2.0518
Least Significant Difference	13.333

Treatments with same alphabet are not significantly different				
t Grouping		Mean	N	Treatment
	A	83.627	4	7
	A	82.864	4	6
	A	81.649	4	8
	A	81.511	4	9
	A	81.096	4	10
B	A	76.786	4	5
B	A	70.432	4	4
B	C	63.758	4	3
D	C	56.288	4	2
D		44.612	4	1

The treatment T7 (recommended N + 20 kg /ha) gets the first rank and is not significantly different from T8 (recommended N + cellulose decomposing enzyme). The treatments T4 and T2 are significantly different among themselves. This procedure answers the questions to some extent. But a multivariate contrast analysis is the best answer for this situation.

11.4.2 Analysis using R

In the sequel is given the R code for multivariate analysis of variance.

```
d28=read.table("mult.txt",header=TRUE)
attach(d28)
names(d28)
```

```
lm1=lm(gyld~factor(rep)+factor(trt),data=d28)
anova(lm1)
library(agricolae)
LSD.test(lm1,"factor(trt)",console=TRUE)
lm2=lm(syld~factor(rep)+factor(trt),data=d28)
anova(lm2)
LSD.test(lm2,"factor(trt)",console=TRUE)
#manova
lm3=manova(cbind(gyld,syld)~factor(rep)+factor(trt),data=d28)
#For pairwise comparison among treatments after manova, install RVAideMemoire
library(RVAideMemoire)
pairwise.manova(cbind(gyld,syld), factor(trt), p.method = "none")
#principal component analysis
pc.out=princomp(~gyld+syld,data=d28)
summary(pc.out)
pc.out$loadings
pc1=(d28$gyld)*pc.out$loadings[1]+(d28$syld)*pc.out$loadings[2]
lm4=lm(pc1~factor(rep)+factor(trt),data=d28)
anova(lm4)
LSD.test(lm4,"factor(trt)",console=TRUE)
detach(d28)
```

Remark 11.4 The MANOVA described in Sections 11.2 and 11.3 can be employed usefully in the experimental situations where the experiment is continued for several years / seasons with same treatments and same randomized layout. For a detailed discussion on this one may refer to Parsad et al. (2004).

A Typical Experimental Situation

12.1   Introduction

It is a well-established fact that designing an experiment is an inevitable part of every research endeavor, particularly the agricultural research. In the National Agricultural Research and Education System, the scientists conduct a large number of experiments by using some design or the other. More often than not, these experiments are designed without the involvement of statisticians at the planning stage. The involvement of statistician(s) is required, if at all, during the analysis of data. This may often lead to difficulties in analyzing the data generated to answer the questions for which the experiment was designed. At that stage, a statistician at best would be able to do a post mortem study by doing some analysis that can closely answer the questions for the purpose of which the experiment was designed. The most common problems encountered in the designed experiments conducted by the experimenters without the involvement of statistician are (a) improper choice of treatments in the experiment, and (b) not properly accounting for the variability in the experimental material by way of local control (forming blocks or nested structures or more than one system of blocks, etc.) and / or if accounted for, not doing it properly and running the experiment using some convenient design. At times, there are problems of designing the experiment properly because of the practical difficulties. In this case, even the statistician would find it hard to suggest a proper design for experimentation and one may have to use a naïve design keeping in mind the practical considerations.

The purpose of this Chapter is to highlight and illustrate through a typical example that running an experiment without the involvement of a statistician at the time of planning an experiment may lead to problems difficult to handle.

In the sequel we describe a typical experimental setting.

12.2   An interesting experimental situation

We now illustrate through an Example the importance of the choice of treatments in an experiment. Although this part of the book is restricted to single factor experiments, yet it would not be out of place to describe the importance of choice of treatments particularly with respect to factorial experiments run as a block design. In cropping systems research, a sequence of two crops is grown: one in kharif season followed by another in rabi season. More than two crops may also be grown in a crop sequence experiment, but we focus our attention to only two crops grown in the sequence. In these experiments, two different sets of treatments are applied in succession: one set applied in kharif crop and the other set applied in rabi crop. The observations are recorded in both the crop seasons. In two-crop sequence experiments, the

interest of the experimenter is in direct effects of treatments applied in kharif and rabi season and residual effects of kharif treatments. The interaction between the residual effect of kharif crop treatments and the direct effect of rabi crop treatments may or may not be of interest to the experimenter. These experiments may be run in a block design to account for the variability in the experimental material.

Crop sequence experiments run in block designs can be viewed as block designs with factorial structure of treatments. For example if *p* treatments are applied in kharif crop and *q* treatments are applied in rabi crop, then a maximum of *pq* treatment combinations may appear and the treatment structure is factorial in nature. These experiments are classified into two broad categories. *Category* I experiments are those in which all the possible *pq* treatment combinations applied in both the crops are taken for experimentation; in other words, one has a complete factorial experiment run in a block design. In these experiments, the interaction between the residual effect of kharif crop treatments and the direct effect of rabi crop treatments are also of interest to the experimenter. *Category* II experiments are similar to the *category* I experiments with the difference that after the application of *q* treatments in the second crop, all the *pq* treatment combinations do not appear. In other words, it is a fractional factorial plan rather than a complete factorial experiment run in a block design. Further, the interaction between residual effect of kharif treatments and direct effect of rabi treatments is not of interest to the experimenter. We restrict here to *category* II experiments.

12.2.1 Category II experiments

Some experiments are conducted to develop suitable integrated nutrient supply system of a crop sequence or crop rotation. One set of treatments is applied to the kharif crop and another set of treatments is applied to the rabi crop. In these experiments, the treatment combinations are smaller than *pq* (fractional factorial) and the estimation of interactions between the residual effects of kharif treatments and the direct effects of rabi treatments is not of interest to the experimenter. Another experimental situation, in which the treatment combinations are smaller than *pq* is dryland farming where the kharif is generally left as fallow and experiments on crop rotations are conducted: one crop for the 1st year (rabi season) and another rabi crop for the second year. Due to fallow kharif season, it is expected that direct effect of treatments applied in second crop year do not interact with the residual effect of treatments applied in first year.

12.2.2 An example

An experiment was conducted on oilseeds (safflower) with an objective to find out the better method of phosphorus (P) management in safflower based cropping system to increase P-use efficiency. This is a crop rotation experiment with one crop for each year in rabi season. The rotation is chickpea - safflower (Series-I) and safflower – chickpea (Series-II). Since it takes two years to complete one cycle, the experiment was conducted in two series so that at the end of two years, we have 2 cycles of data. RCB design was used with 12 treatments and 2 replications (See Table 12.1).

Table 12.1: Treatment details for safflower-chickpea experiment

Treatment	Chickpea (Safflower)	Safflower (Chickpea)
1	No Phosphorus	No Phosphorus
2	100% Recommended P	100% Recommended P
3	50% Recommended P	100% Recommended P
4	50% Recommended P	50% Recommended P
5	50% Recommended P+PSB	50% Recommended P+PSB
6	No Phosphorus	100% Recommended P
7	5 ton FYM/ha	100% Recommended P
8	PSB + 5 ton FYM/ha	100% Recommended P
9	100% Recommended P	50% Recommended P
10	100% Recommended P	No Phosphorus
11	100% Recommended P	5 ton FYM/ha
12	100% Recommended P	PSB + 5 ton FYM/ha

Note: *PSB ~ Phosphate solubilizing bacteria; FYM ~ Farm Yard Manure*

The data generated on the crop sequences is bivariate. The bivariate data is transformed into a univariate data by converting the returns into an economic index (gross returns, net returns), energy equivalent or protein equivalent, etc. Converted data is analyzed as per RCB design procedure. This will give the cumulative effect of the treatments applied in two seasons.

The experimenter, however, is interested to compare the direct effects of treatments applied to first and second crops, respectively and residual effect of the treatments applied to first crop using the data from second crop. This is not answered by converting the bivariate data into an economic index and analyzing it as RCB design.

From the structure of the treatment combinations in the sequence, it is seen that there are six distinct treatments applied in each of the two crops. But instead of 36 possible treatment combinations, only 12 distinct treatment combinations have been used in the crop sequence. It may be noted that a treatment combination is formed of the treatments applied to both the crops in the crop sequence. The 6 distinct treatments for individual crops are given in Table 12.2.

Table 12.2: Treatment details

Phosphorus	Distinct Treatments	Replication in kharif crop	Replication in rabi crop
No Phosphorus	T_1	2	2
50% Recommended P	T_2	2	2
100% Recommended P	T_3	5	5
50 % Recommended P+ PSB	T_4	1	1
5 ton FYM/ha	T_5	1	1
PSB + 5 ton FYM/ha	T_6	1	1

For comparing the direct effects of treatments applied to chickpea, the data obtained for this crop may be analyzed by general block design ANOVA for 6 treatments run in 2 replications (blocks) of size 12 each such that the r_{ij} treatment is replicated r_{ij} times in the jth block, with r_{ij} as given in the Table 12.2.

To test the residual effects of phosphorus treatments of chickpea on safflower in Series-I and phosphorus treatments of safflower on chickpea in Series-II, and the direct effects of phosphorus treatments applied to safflower in Series-I and phosphorus treatments applied to chickpea in Series-II, a connection is developed between structurally complete/incomplete row-column designs (with single or multiple observations per cell) and designs with treatments applied in sequence. The treatments can be divided into two sets *viz.*,

	First Set	**Second Set**
Series-I	Chickpea	Safflower
Series-II	Safflower	Chickpea

Consider a row-column design with the replications of the original design as rows and the treatments applied to the chickpea crop (first set of treatments) as the columns. The cells of the array have the treatments applied to the safflower (second set of treatments). According to this structure, we get Table 12.3 from the Table 12.1 describing treatment combinations in the sequence.

Table 12.3: Treatment combinations

Series-I	First Set →					
	1	**2**	**3**	**4**	**5**	**6**
Replication - I	1, 3	3, 2	3, 2, 1, 5, 6	4	3	3
Replication - II	1, 3	3, 2	3, 2, 1, 5, 6	4	3	3

Series-II	Second Set →					
	1	**2**	**3**	**4**	**5**	**6**
Replication - I	1, 3	2, 3	3, 2, 1, 5, 6	4	3	3
Replication - II	1, 3	2, 3	3, 2, 1, 5, 6	4	3	3

The columns are the six distinct treatments in the first set (or first crop in the sequence). The numbers listed in the cells are the second set of treatments. The replication of treatment 'No Phosphorous' in the first set is two. Treatment 1 of the first set appears with treatments 'No Phosphorous' and 100% 'Recommended Phosphorous' of the second set. These treatments are Treatments 1 and 3. So the first entry in first replication is 1, 3. The replication of treatment 3 '100% Recommended Phosphorous' is five in the first set. Treatment 3 of the first set appears with treatments 'No Phosphorous,' '50% Recommended Phosphorous,' '100% Recommended Phosphorous,' '5 ton FYM/ha' and 'PSB + 5 ton FYM/ha' of the second set. These treatments are Treatments 1, 2, 3, 5 and 6. So the third entry in first replication is 1, 2, 3, 5 and 6. Similarly all other entries are obtained. Since the replications 1 and 2 have the same 12 treatments, the entries for replication 1 and replication 2 are same. The direct effects of the phosphorous

treatments applied to the first crop are the effects of first set of treatments (6 distinct treatments with unequal replications of the first set of treatments) and are analyzed using the data from the first crop. Similarly, the direct effects of phosphorus treatments applied to second crop are the effects of second set of treatments (6 distinct treatments with unequal replications of the second set of treatments) and are analyzed using the data from the second crop. The residual effects of first set of treatments are on the second crop and are analyzed using the data from the second crop and using all the 12 treatment combinations.

The above hypothesis can be tested using the model:

Response = General Mean + Replication effect + effect of First Set of Treatments + effect of Second Set of Treatments + Error.

The coefficient matrix of reduced normal equations for first set of treatments (C_f) and the coefficient matrix of reduced normal equations for second set of treatments (C_s) are obtained as:

$$
\mathbf{C}_f = \mathbf{C}_s = \begin{bmatrix}
2.6 & -0.4 & -1.4 & 0.0 & -0.4 & -0.4 \\
-0.4 & 2.6 & -1.4 & 0.0 & -0.4 & -0.4 \\
-1.4 & -1.4 & 3.6 & 0.0 & -0.4 & -0.4 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
-0.4 & -0.4 & -0.4 & 0.0 & 1.6 & -0.4 \\
-0.4 & -0.4 & -0.4 & 0.0 & -0.4 & 1.6
\end{bmatrix}
$$

It may be noted that the fourth row (or the fourth column) of the matrix C_f (C_s) has all elements zero. This implies that the design adopted is disconnected in first set of treatments and also in second set of treatments. Treatment 4 (50% Recommended P + PSB) is disconnected with rest of the treatments. A similar phenomenon is observed for Series-II. Therefore, it is not possible to estimate all the possible pair wise comparisons of first set and second set of treatments. This implies that the treatment structure (the fractional factorial) chosen for experimentation is not proper. So for answering the questions for which the experiment was conducted, the choice of fraction is not proper.

In the above set up of 12 treatment combinations, however, if the combination of 50% P + PSB and 50% P + PSB is removed from both the replications and instead 50% P + PSB and 5 ton FYM/ha is given in replication I and 5 ton FYM/ha and 50% P + PSB is given in replication II, then the design becomes treatment connected and it would be possible to estimate all the effects. The new treatment combinations for the two replications are given in Table 12.4.

Table 12.4: Treatment combinations

	Replication I	
Treatment	**Chickpea (Safflower)**	**Safflower (Chickpea)**
1	No Phosphorus	No Phosphorus
2	100% Recommended P	100% Recommended P
3	50% Recommended P	100% Recommended P
4	50% Recommended P	50% Recommended P
5	50% Recommended P+PSB	5 ton FYM/ha
6	No Phosphorus	100% Recommended P
7	5 ton FYM/ha	100% Recommended P
8	PSB + 5 ton FYM/ha	100% Recommended P
9	100% Recommended P	50% Recommended P
10	100% Recommended P	No Phosphorus
11	100% Recommended P	5 ton FYM/ha
12	100% Recommended P	PSB + 5 ton FYM/ha

	Replication II	
Treatment	**Chickpea (Safflower)**	**Safflower (Chickpea)**
1	No Phosphorus	No Phosphorus
2	100% Recommended P	100% Recommended P
3	50% Recommended P	100% Recommended P
4	50% Recommended P	50% Recommended P
5	5 ton FYM/ha	50% Recommended P+PSB
6	No Phosphorus	100% Recommended P
7	5 ton FYM/ha	100% Recommended P
8	PSB + 5 ton FYM/ha	100% Recommended P
9	100% Recommended P	50% Recommended P
10	100% Recommended P	No Phosphorus
11	100% Recommended P	5 ton FYM/ha
12	100% Recommended P	PSB + 5 ton FYM/ha

From the structure of the treatment combinations in the sequence, it is seen that in replication I, five distinct treatments are applied to the first crop and six distinct treatments to the second crop, while in replication II, six distinct treatments are applied to both the crops. But instead of 36 possible treatment combinations, only 13 distinct treatment combinations have been used in the crop sequence. The distinct treatments for individual crops are given in Table 12.5.

Table 12.5: Treatment combinations for individual crops

Replication I			
Phosphorus	Distinct Treatments	Replication in Kharif crop	Replication in Rabi crop
No Phosphorus	T_1	2	2
50% Recommended P	T_2	2	2
100% Recommended P	T_3	5	5
50 % Recommended P+ PSB	T_4	1	1
5 ton FYM/ha	T_5	1	1
PSB + 5 ton FYM/ha	T_6	1	1
Replication II			
Phosphorus	Distinct Treatments	Replication in Kharif crop	Replication in Rabi crop
No Phosphorus	T_1	2	2
50% Recommended P	T_2	2	2
100% Recommended P	T_3	5	5
50 % Recommended P+ PSB	T_4	0	1
5 ton FYM/ha	T_5	2	1
PSB + 5 ton FYM/ha	T_6	1	1

It may be noted that now the new design in two replications is not in 12 but in 13 treatment combinations and is, therefore, not a RCBD but an incomplete block design. The total number of observations generated from the design, and, therefore, the total number of experimental units remains the same. For the resulting design, we get Table 12.6 from Table 12.1 of treatment combinations in the sequence.

Table 12.6: Treatment combinations

Series-I	First Set					
	1	**2**	**3**	**4**	**5**	**6**
Replication - I	1, 3	3, 2	3, 2, 1, 5, 6	5	3	3
Replication - II	1, 3	3, 2	3, 2, 1, 5, 6	-	4, 3	3
Series II	Second Set					
	1	**2**	**3**	**4**	**5**	**6**
Replication - I	1, 3	2, 3	3, 2, 1, 5, 6	-	4, 3	3
Replication - II	1, 3	2, 3	3, 2, 1, 5, 6	5	3	3

It can easily be seen that the design is connected for first set of treatments as well as for second set of treatments.

The alternative suggested ensures that the total number of observations (or the block sizes) of the original design does not change although the number of treatments (treatment combinations) have increased from 12 to 13. So recourse was made to an incomplete block design with block size 12 only.

In practice, however, if one wishes to go for a complete block design in two replications and the experimenter has resources to get additional observations, then yet an alternative way of running the experiment so that design is connected, is the following:

Add in both the blocks (or replications) the treatment combination 50% of Recommended phosphorus + PSB and 5 ton FYM/ha along with the already existing 12 treatment combinations making it a total of 13 treatment combinations. This is once again a RCBD in 13 treatments and 2 blocks. The treatments combinations now are shown in Table 12.7.

Table 12.7: Treatment combinations for two crops

Treatment	Chickpea (Safflower)	Safflower (Chickpea)
1	No Phosphorus	No Phosphorus
2	100% Recommended P	100% Recommended P
3	50% Recommended P	100% Recommended P
4	50% Recommended P	50% Recommended P
5	50% Recommended P+PSB	50% Recommended P+PSB
6	No Phosphorus	100% Recommended P
7	5 ton FYM/ha	100% Recommended P
8	PSB + 5 ton FYM/ha	100% Recommended P
9	100% Recommended P	50% Recommended P
10	100% Recommended P	No Phosphorus
11	100% Recommended P	5 ton FYM/ha
12	100% Recommended P	PSB + 5 ton FYM/ha
13	50% Recommended P+PSB	5 ton FYM/ha

From the structure of the treatment combinations in the sequence, it is seen that there are six distinct treatments applied in each of the two crops. But instead of 36 possible treatment combinations, only 13 distinct treatment combinations have been used in the crop sequence.

The 6 distinct treatments for individual crops are as in Table 12.8.

Table 12.8: Treatment combinations for individual crops

Phosphorus	Distinct Treatments	Replication in kharif crop	Replication in rabi crop
No Phosphorus	T_1	2	2
50% Recommended P	T_2	2	2
100% Recommended P	T_3	5	5
50 % Recommended P+ PSB	T_4	2	1
5 ton FYM/ha	T_5	1	2
PSB + 5 ton FYM/ha	T_6	1	1

According to this structure of treatment combinations, we get Table 12.9.

Table 12.9: Structure of treatment combinations

Series-I	First Set →					
	1	**2**	**3**	**4**	**5**	**6**
Replication - I	1, 3	3, 2	3, 2, 1, 5, 6	4, 5	3	3
Replication - II	1, 3	3, 2	3, 2, 1, 5, 6	4, 5	3	3

Series-II	Second Set →					
	1	**2**	**3**	**4**	**5**	**6**
Replication - I	1, 3	2, 3	3, 2, 1, 5, 6	4	3, 4	3
Replication - II	1, 3	2, 3	3, 2, 1, 5, 6	4	3, 4	3

The resulting design is a treatment connected design.

For category II experiments, the choice of a proper fraction is very important. As seen an improper choice of fraction would lead to a disconnected design and all the questions of the experimenter cannot be answered. There is a connection between row-column designs and block designs with two sets of treatments applied in succession. This connection in fact unifies the research efforts made in the literature in two different directions viz. row-column designs and block designs for two sets of treatments applied in succession. Now with the help of a simple mapping one can easily obtain a block design for two sets of treatments applied in succession from a row column design and vice versa (see *e.g.* Parsad *et al.*, 2003).

Introduction to SAS

I.1 Introduction

SAS (once stood for Statistical Analysis System) software is comprehensive software which deals with many problems related to Statistical analysis, Spreadsheet, Data Creation, Graphics, etc. SAS was initially developed in the early 1970 at North Carolina State University with the objective of management and analysis of agricultural field experiments. Nowadays SAS is widely used in many disciplines such as medical sciences, biological sciences and social sciences because of its versatility for various kinds of data management, analysis and data analytics. Latest version of SAS is 9.4. If not SAS 9.4, SAS 9.3 is available in most of the Institutes of National Agricultural Research and Education System in India. SAS comes with a number of products namely Base SAS, SAS/STAT and SAS/GRAPH to name a few. Out of these Base SAS is the base product of SAS which supports data management and a number of basic data analysis procedures. The SAS/STAT product has provision for most of the standard and advanced statistical analysis and the SAS/GRAPH is designed to produce quality graphics. There are specialized products for operation research (SAS/OR), econometrics and time series analysis (SAS/ETS) and so on.

I.2 The SAS windows

One can start SAS, by clicking Start → Programs → SAS. After opening the SAS, it can be seen that there are four windows, *viz.*, (a) the explorer window on the left hand side, (b) the output window, (c) the log window, and (d) the program editor on the right hand side. The Windows of SAS are shown in Figure I.1.

The program editor window is used to type the SAS commands. The editor supports text editing features like select, copy, cut, paste, moving the cursor, etc. The enhanced program editor gives color-coded procedures, statements, and options that help to find errors in the code before even running it. Once the code is written in the editor window or the enhanced editor window, the code is run clicking run symbol on the menu bar. The log window is very important and the errors, if any, in the SAS commands can be found in the log window after the code has been run. Therefore, it is always a good practice to check the log window to see if the SAS program contains any errors or not. As the name suggests, the output window is the place where the output after running SAS commands appears. Unlike previous versions of SAS, version 9.3 output by default uses HTML. One can change this option to text output. The explorer window can be used to open/view data that is read into SAS. The explorer window contains libraries which includes the work folder containing any datasets that has been read or created in SAS in that session. Generally all the datasets created in the work folder are temporary, which means once the current SAS session is closed, all temporary datasets created are deleted automatically.
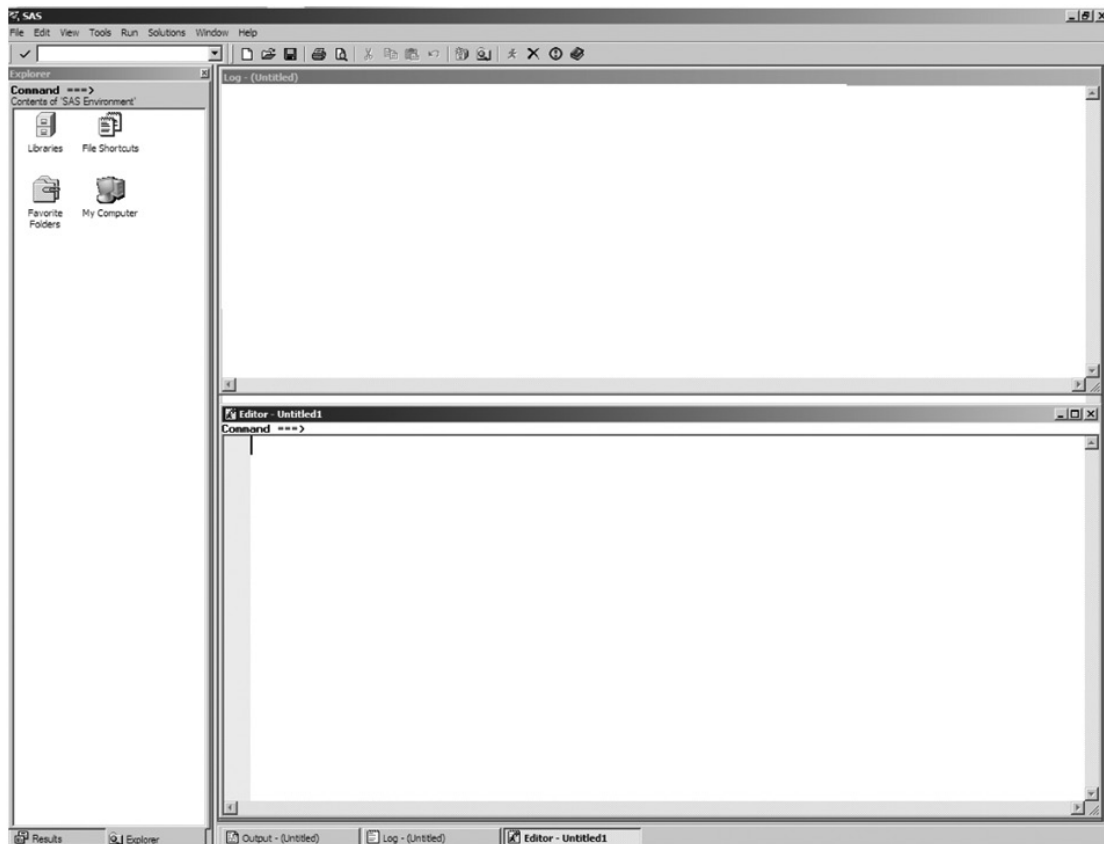
Figure I.1: Windows of SAS software

From the above it is clear that three windows are very important in SAS: the program editor window, the log window and the output window. After writing a code in SAS and getting the output, it is a good practice to save the codes and the outputs from the analysis. To save the program written in the program editor window, the cursor should be placed in the program editor window so that this window becomes active and then program can be saved with File ® save command from menu. Similarly, to save the outputs, go to the output window and select File → Save. One can also save the log window contents if desired, in the same way.

I.3 Rules for SAS programs

SAS programs like any other programming language has some rules. These rules are called syntax. Some basic syntax that must be followed are given in the sequence.

i   SAS is not case sensitive.

ii   Every line must end with a semi-colon (;).

iii   Variable names must be 32 characters or less and may be combination of letters, numbers and underscore character but should begin with an alphabet or an underscore.

iv   Words should be separated by at least one space.

v   One block of code should end with a run statement.

vi   Lines starting with /* are treated as comments and are not processed. Comments are closed by */.

vii   Data steps start with the word DATA and procedures start with the word PROC.

viii   Cards sections can be invoked with word CARDS or DATALINES

Note that the words DATA, PROC, CARDS and DATALINES have been written in Capital letters. Though SAS is not case sensitive but capital letters will be used to differentiate between usual texts and SAS syntax throughout this Annexure. But this distinction has also been made in the main Chapters of the book.

A semi-colon at the end of line tells SAS that the current statement is finished, and invokes SAS to read the next line. Data statement is used to specify a name of a data set. This name of the data set can be used later in the program to refer to it. The names of SAS datasets and variables are not case sensitive and hence a data set called mydata, MyData, MYDATA are all equivalent. Similarly, a variable name AGE, age, Age all refer to a same variable. After a block of code is written, then that block of code should end with a run statement; otherwise SAS may not process the code. Once the block of code with a run at end is written, the user needs to click on 🏃 icon in the toolbar at the top of the screen to process the codes (or press F8 to submit the code) and get the outputs in the output window.

I.4 Basic structure of SAS programs

A SAS program is mainly composed of two parts: a data step and a procedure step. The data step deals with reading, manipulating, formatting and cleaning data. For example, data step can be used to extract some of the variables or some subset of the observations of an existing data set for further analysis. It may be used to convert a dataset from a different format to SAS compatible format. Data step can also be used to manipulate data in a number of ways so that resultant data can be used for further analysis. For example, data may need to be transformed into appropriate format for a subsequent analysis. To summarize, the data step is used to prepare data for use by one of the procedures.

An example of data step used to create a data set is given below:

DATA mydata;

INPUT var1 var2 var3;

CARDS;

10 12 15

15 25 14

35 25 16

12 24 24

;

RUN;

The details of this program will be clearer in the next Section.

The procedure step performs statistical analyses and/or produce graphical results. Generally, a SAS program is composed of one or more (statistical) procedures. Each procedure is a unit, although some are needed to run others. When a procedure is submitted to run in SAS, the results will go to the output window unless the output is not suppressed in the code itself. An example of a procedure step which obtain basic summary statistics is given below

PROC MEANS DATA = mydata;

RUN;

I.5 Reading data into SAS

First we need to know how data set is defined in SAS. A data set in SAS consists of two components: variables and observations. These terms are defined below.

Variable: A variable refers to an entity or characteristic that can take a set of values, *e.g.* leaf area. Here, leaf area denotes a variable and if there are 100 leaves available then we have 100 values of leaf area. When defining a variable name in a SAS data set, some rules need to be followed. A variable name can have up to a maximum of 32 characters and must begin with a letter or underscore. Note that no special characters except underscore can appear in a SAS variable name. Blank spaces are also not allowed in SAS names. So a valid variable name for leaf area variable in SAS can be leafarea or leaf_area or la or var1 or x or y. Now, in SAS, variables can mainly be of two types:

i) Character Variable: A character variable can take values which are combination of alphabets, numbers and special characters or symbols. For example, a variable country can take values "India", "Australia", "US", "Germany", "Netherland", "UK", so the variable country is a character variable.

ii) Numeric Variable: A numeric variable can take values only as numbers which may be with or without decimal points and with + or – signs. For example, a variable weight can take values 40, 60, 55, 45, 76, 60, etc. So this variable is a numeric variable.

Observation: The set of values taken by different variables on an individual object or subject or unit is the observation on that object or subject or unit. For example, if there are five variables namely name, address, age, weight and height, then an observation on an individual may be "Samir", "Mumbai", 35, 68, 145; another observation on another individual may be "Dipak", "Bangalore", 40, 64, 142.

Datasets can be read in SAS by at least three basic ways.

i) Using INFILE statement

The INFILE statement can be used to read data from external files located in a drive (e.g., D:, C:, F:) and make them available for the entire SAS session.

Reading data from an external text file

To read data from a text file located in drive D, the following command can be used:

DATA d1;

INFILE 'D:\mydata.txt';

INPUT x1 x2 x3;

RUN;

Here d1 denotes the name of the dataset that is created in SAS. The input statement is used to specify names of the variables in the SAS dataset. In the current example, the names of variables are x1, x2 and x3. Here, it must be mentioned that the data in the external file mydata.txt should be in the following format:

10 20 30

24 25 .

65 64 35

25 .   25

30 25 36

Here note that since we are reading three variables, the external file should have three columns with first column referring to the values of x1 variable, second column to values of x2 variable and third column to values of x3 variable. Two values in a row should be separated by at least one space. Any missing values for numeric variables should be denoted as and missing character values should be a blank space; otherwise there may be error in reading the data set. Each row of the data represents one observation for the data set.

Reading data from an external ASCII file

Data can be read from an external (ASCII) file as shown below:

DATA ex2;

INFILE 'd:\mydata';

INPUT group $ x y z;

RUN;

or

DATA ex2a;

FILENAME abc 'd:\mydata';

INFILE abc;

INPUT group $ x y z;

RUN;

*Reading data from an external *.csv file*

One can easily read data from an external comma separated value (.csv) file using infile statement as shown in the following example:

DATA ex4;

INFILE 'C:\mydata.csv' DLM=',' ;

INPUT snloc $ year season $ crop $ rep trt gyield syield return kcal;

RUN;

ii) Using the import wizard:

To import the data from a drive, one can go to "File", then "Import data". An Import wizard window will pop up with the option to select the format of the file to import. After the file type (*i.e.* Excel, Access, text, etc.) has been selected, it asks for the location of the file. Then the file to be imported is to be selected. Thereafter a proper SAS name is to be specified to the data set to be created in SAS. It may be mentioned here that SAS 9.2 can only read Excel files saved under Excel 2003 or earlier version, (*i.e.* with .xls extension). To read an .xlsx file with 9.2, the file should be saved as .xls version before reading it into SAS. However, SAS 9.3 supports importing .xlsx files as well.

One can also use proc import statement to read data from excel files using following commands:

PROC IMPORT DATAFILE = 'C:\descriptive_stats.xls'

OUT = descriptive_stats REPLACE;

RUN;

In the PROC IMPORT statement above, one needs to specify the name of the excel file with its location and a name of the SAS data set to be created.

iii) Using the CARDS or DATALINES statement:

Another easy way of reading data into SAS is using CARDS or DATALINES statement. This is useful when the dataset is small. An example to read a data set using CARDS statement is:
DATA d2;
INPUT x1 x2;
CARDS;
3 5.1
5 10
4 14.7
2 3.3
;
RUN;

Similarly DATALINES statement can be used.

DATA d3;

INPUT x1 x2;

DATALINES;

3 5.11

5 10

4 14.7

2 3.3

;

RUN;

Reading data in various formats

We have seen that data can be read in SAS using INFILE statement, CARDS or DATALINES statement and through import wizard. While using the INFILE and CARDS or DATALINES statements, note that both use INPUT statement. The INPUT statements are part of data section. This statement provides the SAS system the name of the variables with the format, if it is formatted. The examples given above show what is called list directed input.

List directed input:

In list directed input

- Data are read in the order of variables given in INPUT statement.

- Data values are separated by one or more spaces.

- Missing values of numeric variables are represented by period (.).

- Missing values of character variables are represented by blank and

- Character variables are followed by $ (dollar sign).

Often the data may be read where the data values are not separated by space. In this case INPUT statement should specify the starting and ending column numbers that is occupied by values of a variable. For example, the following example tells SAS that variable ID is available in columns 1 to 3, sex in column 4, height in columns 5-6 and weight in columns 7-11.

DATA mydata;
INPUT ID 1-3 sex $ 4 height 5-6 weight 7-11;
CARDS;
001M68155.5
2F6199

3M5333.5

;

RUN;

Alternatively, starting column of the variable can be indicated along with its length as

DATA mydata;

INPUT @1 ID 3.

@4 sex $ 1.@5height 2. @7weight5. ;

CARDS;

001M68155.5

2F6199

3M5333.5

;

RUN;

Reading more than one line per observation for one record of input variables

It is possible to read more than one line per observation. The following example shows that the variables ID, age and height are to be read from line one and the variables sbp and dbp are to be read from line two from each block of two lines in the data.

DATA mydata;

INPUT # 1 ID 1-3 age 5-6 height 10-11

2 sbp5-7 dbp8-10;

CARDS;

001 56   72

14080

;

RUN;

Reading the variable more than once

It is possible to read one variable more than once. For example, if a variable Id which is of length 6 contains information about state code, which is the last two columns of Id variable, then both Id and state can be read as shown below:

DATA mydata;

INPUT @ 1 Id 6. @ 5 state 2.;

CARDS;

123401

124102

254103

;

RUN;

Or

DATA mydata;

INPUT ID 1-6 state 5-6;

CARDS;

123401

124102

254103

;

RUN;

Formatted lists

The data can also be read by groups of variables with the starting column for the first variable in the group and the number of columns occupied by each variable in the groups. As an example, the following commands read two groups of variables namely x1, x2 and y1, y2 where the first variable in the first group starts at column one and each of the variables in the first group occupies one column, the first group starts from column four and each of the variables in second group occupies three columns.

DATA B;

INPUT ID @1(x1-x2)(1.)

@4(y1-y2)(3.);

CARDS;

11 563789

22 567987

;

RUN;

Data can also be read using @ in the INPUT statement. With this, SAS can read values of variables only once from each observation in the data part after the cards section and leaves the other values unread. The following example reads four observations for three variables in one line. Note that observations are read alternatively variable wise, *i.e.*, first value goes to first variable, second value to second variable, third value to third variable, then rest of 9 observations remains unread.

DATA C;

INPUT x y z @;

CARDS;

1 1 1 2 2 2 5 5 5 6 6 6

1 2 3 4 5 6 3 3 3 4 4 4

;

RUN;

Data can similarly be read using @@ in the INPUT statement. With this, SAS can read more than one observations per line in the data part after the CARDS section. The following example reads four observations for three variables in one line. Note that observations are read alternatively variable wise, *i.e.*, first value goes to first variable, second value to second variable, third value to third variable, then fourth value to first variable, and so on.

DATA D;

INPUT x y z @@;

CARDS;

1 1 1 2 2 2 5 5 5 6 6 6

1 2 3 4 5 6 3 3 3 4 4 4

;

RUN;

Creating a permanent SAS data set

All the methods discussed above create data sets which are temporary in the sense that the data sets are removed once the user closes the current session of SAS. Often one may be interested to create a dataset which is permanent and is not removed after the current session of SAS is closed. This is done using LIBNAME statement as shown below:

LIBNAME xyz 'c:\SASDATA'; /* these statements create a file named example in library named xyz /*

DATA xyz.example;

INPUT group x y z;

CARDS;

1 10 20 30

1 15 25 32

2 24 25 26

2 21 25 26

;

RUN;

The above program reads data following the CARDS statement and creates a permanent SAS data set in a subdirectory named \SASDATA on the C: drive.

The following example illustrates how to use a permanent SAS data set:

LIBNAME xyz 'c:\SASDATA';

PROC MEANS DATA=xyz.EXAMPLE;

RUN;

I.6 Basic data manipulation using the DATA step

As we have seen that the data step deals with reading, manipulating, formatting and cleaning data. Some examples of data manipulation in SAS using data step are provided in the sequel. For this purpose, consider the following sample data set:

DATA d1;

INPUT name $ age gender $ height weight;

CARDS;

sam 15 M 140 45

john 18 M 135 42

madhu 21 F 128 48

kanika 24 F 135 46

peter 30 M 140 49

soumi 29 F 128 51

;

RUN;

Note that we have used $ symbol after the variables name and gender. The $ symbol is used to specify that those two variables are character variables. Suppose we want to get a subset of

data with males only. The following SAS code can be used to create a new data set called d1m with gender as males only.

DATA d1m;

SET d1;

IF gender = 'M';

RUN;

The SET d1 statement after the DATA d1m statement tells SAS to make a copy of the dataset d1 and save it as d1m. The IF statement tells to take only those observations where gender is equal to M. The same subset of data can be alternatively created as

DATA d1m;

SET d1;

IF gender = 'F' then delete;

RUN;

Further subsetting of data can be done based on numerical values. For example, to create a subset of data with persons at least 18 years old, the following code can be used:

DATA d18;

SET d1;

IF age > 18;

RUN;

Variables to be included in a subset of data can be selected using DROP or KEEP statement. For example, to drop the name variable the following code can be used:

DATA d1wn;

SET d1;

DROP name;

RUN;

The same subset of data can be created with KEEP statement as follows:

DATA d1wn;

SET d1;

KEEP age gender height weight;

RUN;

I.7 The procedure step

The procedure step performs statistical analyses and/or produce graphical results. When a procedure is submitted to run in SAS, the results will go to the output window unless the output is not suppressed in the code itself. Some frequently used procedures for statistical analysis are explained in detail below:

PROC PRINT procedure can be used to print a dataset in the output window after a dataset is created in a DATA step. For example, to print the dataset d1, the SAS commands are

PROC PRINT data=d1;

RUN;

To print only some of the variables, a VAR statement can be included. For example, to print only age, gender, height and weight variables of the dataset d1, the following code can be used:

PROC PRINT data=d1; /*here d1 denotes the data name*/

VAR age gender height weight;

RUN;

PROC UNIVARIATE is a procedure which is used for elementary statistical analysis. This procedure computes the basic statistics of one or more variables in a dataset and has optional statements to generate some plots like histograms and qqplots. For example, the following code can be used to get univariate statistics of variables height and weight and to produce their histograms:

PROC UNIVARIATE data=d1;

VAR height weight;

HISTOGRAM;

RUN;

If VAR statement is not used, the procedure will compute univariate statistics for all numeric variables in the dataset.

A number of plots can be produced in the PROC univariate procedure. Following statements can be used in the PROC univariate procedure to get the plots mentioned below:

Plot type	Statement	Example
histogram, normal density plot	histogram/kernel normal	PROC UNIVARIATE DATA=d1; VAR x; HISTOGRAM/KERNEL NORMAL; RUN;
probabilty plot	probplot	PROC UNIVARIATE DATA=d1; VAR x; PROBLPLOT; RUN;

Plot type	Statement	Example
quantile quantile plot	qqplot	PROC UNIVARIATE DATA=d1; VAR x; QQPLOT x/NORMAL SQUARE; RUN;
cumulative density plot	cdfplot	PROC UNIVARIATE DATA=d1; VAR x; CDFPLOT x/NORMAL; RUN;

PROC SORT can be used to sort the observations in a dataset by some variables in either ascending or descending order. For example to sort the data set d1 by age, the following code can be used:

PROC SORT DATA=d1 OUT=d2;

BY age;

RUN;

The observations of dataset d1 are sorted in ascending order, by default, of the variable age, and the sorted data is saved in a dataset named d2. Without the OUT=d2 option, the unsorted dataset named d1 will be replaced by the sorted dataset. The observations can be sorted in descending order by specifying the descending option in the BY statement, *e.g.*, BY descending age. To sort the data by more than one variable, the variable names should be listed in the BY statement. For example, to sort the data set d1 by age and height, the following code can be used:

PROC SORT DATA=d1 OUT=d2;

BY age height;

RUN;

PROC MEANS procedure is used to produce simple univariate descriptive statistics for numeric variables. It also calculates confidence limits for the mean and identifies extreme values and quartiles. For example, the following commands can be used to produce mean, median, minimum, maximum and number of observations of a data set.

PROC MEANS data=d1 mean median min max n;

RUN;

The mean, median, minimal value, maximal value and sample size will be computed for all the numerical variables in the data set d1. To compute these statistics for some of the variables in the dataset, VAR statement can be used. For example, the following code computes mean and median for two variables height and weight:

PROC MEANS DATA=d1 MEAN MEDIAN;

VAR height weight;

RUN;

Often data sets can have grouping variables. One may be interested in getting these statistics for each group of the data set created by these grouping variables. For example, in the data set d1, gender is a grouping variable and one may be interested to know mean and median height and weight for both males and females. This can be done by using following code:

PROC SORT DATA=d1 OUT=d2;

BY gender;

RUN;

PROC MEANS data=d2 MEAN MEDIAN;

VAR height weight;

BY gender;

RUN;

One word of caution is in order here. Before using BY statement in any procedure, the data needs to be sorted by the variables which appear in the BY statement. Otherwise the statement will not run. Alternatively, one can use the statement CLASS gender; between PROC and VAR statement in place of BY statement.

PROC SUMMARY computes descriptive statistics on numeric variables in a SAS dataset and outputs the results to a new SAS dataset. For example, the descriptive statistics of height and weight variables for males and females in the data set d2 can be saved as a dataset d3 using the following code:

PROC SUMMARY data=d2 PRINT;

VAR height weight;

BY gender;

OUTPUT OUT=d3;

RUN;

Alternatively one may use

PROC SUMMARY data=d2 PRINT;

CLASS gender;

VAR height weight;

OUTPUT OUT=d3;

RUN;

A PRINT option or the output statement must be specified in the PROC SUMMARY, otherwise it may not run properly.

PROC FREQ procedure produces frequency tables and contingency tables. To illustrate the use consider the following data:

DATA mydata;

INPUT gender $ income class $;

CARDS;

M	5000	small
F	3900	small
M	6000	medium
M	5400	medium
F	4000	small
M	12000	big
M	15000	big
F	10000	big

;

RUN;

To create a two-way frequency table by variables gender and class, you can use the following SAS commands:

PROC FREQ DATA=mydata;

TABLES gender*class/MISSING CHISQ;

RUN;

PROC CORR calculates the correlation coefficients between quantitative variables along with their simple summary statistics. For example, to compute the correlation coefficient between height and weight variables, one can use the following commands:

PROC CORR DATA=d1;

VAR height weight;

RUN;

This will produce a correlation coefficient matrix along with probability level of significance.

The PROC TTEST has many applications in testing of hypothesis. Some of them are described in the sequel. PROC TTEST procedure is used for testing that the population mean is equal to some specified value. For example, the following code tests the null hypothesis that average weight of population of leaves is 2 gm.

PROC TTEST h0=2;

VAR leafwt;

RUN;

This PROC can also be used for comparing the means of two populations based on independent samples drawn from the two populations. An example of comparing the average weights of populations of male and female fish is given below.

PROC TTEST;

CLASS sex;

VAR fishwt;

RUN;

The CLASS statement specifies that the variable sex is used to classify the values of fishwt variable into two samples, *i.e.*, PROC TTEST divides the observations into the two groups for the t test using the levels of this variable. One can use either a numeric or a character variable in the CLASS statement.

PROC TTEST computes the group comparison t statistic based on the assumption that the variances of the two groups (or populations) are equal. It also computes an approximate t based on the assumption that the variances are unequal (the Behrens-Fisher problem). One can use COCHRAN option for this situation, *e.g.*,

PROC TTEST COCHRAN;

CLASS sex;

VAR fishwt;

RUN;

The COCHRAN option uses the Cochran and Cox approximation of the probability level of the approximate t statistic for the unequal variances situation.

To perform a paired two-sample t-test, PROC TTEST can be used with the PAIRED statement. For example, the following commands test whether the average weight of fish population before and after an experiment are same or not.

PROC TTEST;

PAIRED wtbefore*wtafter;

RUN;

The PAIRED statement identifies the variables to be compared in paired comparisons. In the above example the variables are wtbefore and wtafter. One or more paired variables can be used in the PAIRED statement. Variables or lists of variables are separated by an asterisk (*) or a colon (:). Examples of the use of the asterisk and the colon are shown in the following table:

The PAIRED *Statements*	*Comparisons made*
PAIRED a*b;	a-b
PAIRED a*b c*d;	a-b and c-d
PAIRED (a b)*(c b);	a-c, a-b and b-c
PAIRED (a1-a2)*(b1-b2);	a1-b1, a1-b2, a2-b1 and a2-b2
PAIRED (a1-a2):(b1-b2);	a1-b1 and a2-b2

PROC ANOVA is used to perform analysis of variance (ANOVA), multivariate analysis of variance and repeated measures analysis of variance in case of balanced data. In the case of ANOVA, a CLASS statement needs to be used for categorical variables before the model statement. The PROC ANOVA procedure should be used when the data is balanced with respect to the variables listed in the CLASS statement. For example, to study the effect of gender on height of persons, it may be noted that the dataset d1 is balanced and hence the following code can be used to perform ANOVA.

PROC ANOVA data=d1;

CLASS gender;

MODEL height=gender;

RUN;

It tests whether the height is the same for females and males.

Since PROC ANOVA is one of the procedures that is used in analyzing data from designed experiments, a detailed syntax of PROC ANOVA is given in Annexure-1 of this Chapter.

PROC GLM procedure performs general linear model fitting which includes simple and multiple regression, analysis of variance (ANOVA), analysis of covariance (ANCOVA), multivariate analysis of variance, and repeated measures analysis of variance. As an example, the following code fits a linear model for weight on age and height.

PROC GLM DATA=d1;

MODEL weight=age height;

OUTPUT OUT=d4 p=pred r=resid;

RUN;

Here the MODEL statement is used to specify the dependent and independent variables and the dependent and independent variables are separated by ' =' sign. In this example weight is dependent variable and age and height are independent variables. The OUTPUT statement is used to save the analyzed results in a new data set. In the current example the name of the output data set is d4 and the predicted values are stored as variable named pred and residuals are stored as variable named as resid. Since PROC GLM is one of the procedures that is used in analyzing data from designed experiments, a detailed syntax of PROC GLM is given in Appendix-2 of this Chapter.

As in case of PROC ANOVA, a CLASS statement needs to be used for categorical variables before the MODEL statement. For example, to study the effect of gender on height of persons, the following code can be used. Other options of PROC ANOVA or PROC GLM are given in the Appendix-1 and Appendix-2.

PROC GLM DATA=d1;

CLASS gender;

MODEL height=gender;

RUN;

PROC REG is used to fit linear regression models. It allows multiple MODEL statements in one procedure, can do model selection, and even plot summary statistics and normal qq-plots.

PROC OPTEX is used to search for optimal design in mixture experiments with linear models. To generate a design, PROC OPTEX and MODEL statements are used. Other statements are used as needed. A CLASS statement naming classification variables must precede the MODEL statement that uses those variables. As the OPTEX procedure is interactive, all other statements (except the PROC OPTEX statement) can be used after the first RUN statement.

There is option for several PLOT statements for each MODEL statement. For example, the following commands fit a regression model for weight on age and height and plot the predicted values and residuals against age and height variables.

PROC REG DATA=d1;

MODEL weight=age height;

PLOT weight*age

PLOT weight*height;

PLOT PREDICTED.*age;

PLOT RESIDUAL.*age;

PLOT PREDICTED.*height;

PLOT RESIDUAL.*height;

RUN;

As in PROC GLM, PROC REG also has a MODEL statement and usage of MODEL statement is similar to that of PROC GLM. Note that in the above code predicted. and residual. refer to predicted values and residuals respectively.

The PROC RSREG procedure uses the method of least squares to fit the full quadratic or second order response surface regression models. Response surface models are a kind of general linear model in which attention focuses on characteristics of the fitted response function and in particular, where optimum estimated response values occur. The following example fits a three factor response surface model with the independent variables x1, x2 and x3.

PROC RSREG;

MODEL y=x1 x2 x3;

RUN;

In addition to fitting a quadratic function, one can use the RSREG procedure to test for lack of fit, to test for the significance of individual factors, to analyze the canonical structure of the estimated response surface, to compute the ridge of optimum response and predict new values of the response. The detailed syntax of PROC RSREG is given in Appendix-3 of this Chapter.

I.8   Options in the procedures and statements

We have seen that various procedures perform various kinds of analysis. All these procedures by default produce some standard output. However, a number of additional analyses can be performed by specifying some options in these procedures. For example, by default, SAS produces 95% confidence intervals. If someone is interested to get 99% confidence interval, one can use the following code:

PROC GLM DATA = d1 ALPHA=.01;

CLASS gender;

MODEL height=gender;

RUN;

Here, the ALPHA =.01 option is used to get 99% confidence interval. It also performs hypothesis tests at 1% significance level.

Another example of options in procedure statement is the PLOT option in PROC UNIVARIATE procedure. A number of plots can be produced with PROC UNIVARIATE procedure. For example, you can get a stem and leaf plot, box plot and normal probability plots using PLOT option.

PROC UNIVARIATE DATA=test NORMAL PLOT;

RUN;

All the SAS Procedures have a number of different statements to perform various kinds of specific analysis. These statements are to be specified when using a procedure to perform a required analysis. All these statements have provision for some further analysis by way of specifying Options. Unless options are specified, these optional analysis results are not produced by SAS. Some commonly used statements are described below.

i)   The VAR statement: The VAR statement is used to specify which variables out of all the available variables in a dataset should be used in a procedure. For example, to get univariate statistics for only height and weight variables in data set d1, the following command uses VAR statement.

PROC UNIVARIATE DATA=d1;

VAR height weight;

RUN;

ii) The BY statement: The PROC SORT procedure uses the BY statement to sort a data set using the variables specified in the BY statement. Once a dataset has been sorted by some variables, then the BY statement can further be used in another procedure to perform analysis for each combination of values of the variables specified in the BY statement. It is important to note here that the variables specified in the BY statement in a procedure must be from those variables by which the data was sorted. For example, data d1 is first sorted by gender and name.

PROC SORT DATA=d1;

BY gender name;

RUN;

Now suppose we want to get the summary statistics for height and weight for each gender separately, then the following commands can be used:

PROC UNIVARIATE DATA=d1;

VAR weight height;

BY gender;

RUN;

iii) The CLASS statement: The CLASS statement in a procedure is used to denote that some variables are categorical. For example, if in an experiment 5 treatments are applied to 15 experimental units in 3 blocks and each block has 5 experimental units, where each of the unit receives one distinct treatment, then here treatment and block are categorical variables. The CLASS statement is particularly used for performing analysis of variance using PROC ANOVA or PROC GLM. The code for analysis of such data is given below assuming that the variable names in the data are treatment, block and observation and the SAS data set name is experiment.

PROC ANOVA data=experiment;

CLASS treatment block;

MODEL observation = treatment block;

RUN;

iv) The MODEL statement: The MODEL statement is used whenever some model fitting is done in SAS. For example, PROC REG, PROC GLM, PROC ANOVA etc. use MODEL statement. The MODEL statement is used to specify which variables are to be used in the MODEL and which are dependent and independent variables. Dependent variable appears on the left of '=' sign and independent variables appear on the right of '=' sign in a MODEL statement. The example described just above has observation as dependent variable and treatment and block as independent variables. When there are multiple dependent variables specified in a MODEL statement then one model is fitted for each of the dependent variable separately unless a MANOVA statement is used in the procedure.

Some illustrations are given for various types of model statements that PROC ANOVA and PROC GLM can handle. Here it is assumed that A, B, C are CLASS variables and X1, X2, X3 are quantitative, regression variables.

Simple linear regression	MODEL Y = X1;
Multiple regression	MODEL Y = X1 X2 X3;
Polynomial regression	MODEL Y = X1 X1*X1 X1*X1*X1;
One way anova	MODEL Y = A;
Two-way, main effects only	MODEL Y = A B;
Two-factor factorial with interaction	MODEL Y = A B A*B;
Two-factor factorial with interaction using "|" notation	MODEL Y = A | B;
Three-factor complete factorial	MODEL Y = A B C A*B A*C B*C A*B*C;
Three-factor complete factorial using "|" notation	MODEL Y = A | B | C;

A number of options are available in the MODEL statement in PROC GLM. For example, in the PROC GLM, MODEL statement, one can get different types of sums of squares using SS1, SS2, SS3, SS4 options and confidence intervals using CLI and CLM options. Detail of different sum of squares in PROC GLM are provided at the end of Annexure-2. One can also specify desired confidence interval level using ALPHA option. As an example, following commands produces type III sum of squares in the analysis of variance table.

PROC ANOVA DATA=experiment;

CLASS treatment block;

MODEL observation=treatment block /SS3;

RUN;

v) The MEANS and LSMEANS statements: The MEANS and LSMEANS statements are very important and they are commonly used in PROC ANOVA and PROC GLM to get means or least square means of the dependent variables for each class of categorical variables which are listed in the CLASS statement in these procedures. The MEANS statement is generally used when the data are balanced and there are no missing values and no covariates. If the data are not balanced or the data contain missing values or the data have covariates, then the LSMEANS statement should be used. Moreover, the MEANS statement has options for multiple comparisons among the main effects whereas the LSMEANS statement has options for multiple comparisons among the main effects and the interactions. As an example, the following code performs Tukey's test and Duncan's multiple range test for the treatment means:

PROC GLM DATA= experiment;

CLASS treatment block;

MODEL observation =treatment block;

MEANS treatment / TUKEY DUNCAN;

RUN;

The MEANS statement will perform means comparisons for all five treatment groups. The options TUKEY and DUNCAN will produce multiple comparisons among the treatment means. There are a number of multiple comparison tests available in the MEANS statement *e.g.*, Bonferroni (BON), Dunnet (DUNNET), least significant difference (LSD), Scheffe (SCHEFFE), Student-Newman-Kuels (SNK) to name a few.

To specify options for multiple comparisons in the LSMEANS statement, the name of the test should be specified after ADJUST= . For example, one can use the following code for performing Bonferroni test. The default option is Fisher' LSD.

PROC GLM data = experiment;

CLASS treatment block;

MODEL observation = treatment block;

LSMEANS treatment / ADJUST=BON STDERR;

RUN;

The STDERR option above will produce the standard errors of least square means.

Other options available with this statement are:

PDIFF: Prints the p - values for the tests of equality of all pairs of class means.

singular: tunes the estimability checking.

LINES: gives the letters for the treatments in same or significantly different groups

vi) The CONTRAST statement: The CONTRAST statement can be used to test pre-planned hypothesis. The basic form of the CONTRAST statement is given below.

CONTRAST 'label' effect name< ... effect coefficients ></options>;

Here label is a character string used for labeling output, effect name is CLASS variable (which is independent) and effect coefficients is a list of numbers that specifies the linear combination of parameters in the null hypothesis. The contrast is a linear function such that sum of the elements of the coefficient vector sum is equal to zero for each effect. While using the CONTRAST statements, one should keep the following points in mind. If there are more levels of that effect in the data than the number of coefficients specified in the CONTRAST statement, the PROC GLM adds trailing zeros. Suppose there are 5 treatments in a completely randomized design denoted as τ_1, τ_2, τ_3, τ_4, τ_5 and null hypothesis to be tested is

H_o: $\tau_2 + \tau_3 = 2\tau_1$ or $-2\tau_1 + \tau_2 + \tau_3 = 0$

Suppose in the data, treatments are classified using trt as CLASS variable, then a valid CONTRAST statement is

CONTRAST 'τ_1 vs τ_2 and τ_3' trt -2 1 1 0 0;.

Suppose last 2 zeros are not given, the trailing zeros can be added automatically. The use of this statement gives sum of squares with 1 degree of freedom and F-value against error as residual mean squares until specified. The name or label of the contrast must be 20 characters or less.

The available contrast statement options are

E   prints the entire vector of coefficients in the linear function, *i.e.*, contrast.

E = effect specifies an effect in the model that can be used as an error term

ETYPE = *n* specifies the types (1, 2, 3 or 4) of the E effect.

Multiple degrees of freedom contrasts can be specified by repeating the effect name and coefficients as needed separated by commas. Thus the statement for the above example

CONTRAST 'all' trt–2 1 1 0 0, trt 0 1 -1 0 0;

This statement produces sum of squares due to both the contrasts with two degrees of freedom. One can use this feature to obtain partial sums of squares for effects through the reduction principle, using sums of squares from multiple degrees of freedom contrasts that include and exclude the desired contrasts. Although only *t* - 1 linearly independent contrasts exist for *t* classes, any number of contrasts can be specified.

vi) The ESTIMATE statement: This statement is specific to PROC GLM. It can be used to estimate linear functions of parameters that may or may not be obtained by using CONTRAST or LSMEANS statement. For the specification of the statement only word CONTRAST is to be replaced by ESTIMATE in CONTRAST statement.

vii) The TEST statement: This statement is used in PROC GLM. In general F-tests of hypotheses in ANOVA use the residual mean squares as the error term. The TEST statement is used to test the significance of effects where the residual mean square is not the appropriate term, for example, testing the significance of main-plot effects in split-plot experiment. PROC GLM provides the TEST statement which is identical to the TEST statement available in PROC ANOVA. The PROC GLM also allows specification of appropriate error terms in MEANS, LSMEANS and CONTRAST statements. For illustration, consider a split plot experiment involving the yield of different irrigation (irrigat) treatments applied to main plots and cultivars (cultvar) applied to subplots. The data so obtained can be analyzed using the statements given below.

DATA splitplot;

INPUT rep irrigat cultvar yield;

CARDS;

. . .

. . .

. . .

 ;

RUN;

PROC GLM;

CLASS rep irrigat cultvar;

MODEL YIELD = rep irrigat rep*irrigat cultvar irrigat* cult;

TEST h = irrigat e = rep*irrigat;

CONTRAST 'IRRIGATI vs IRRIGAT2' irrigat 1 -1 / e = rep* irrigat;

RUN;

Here, the irrigation effects are tested using error (A) which is sum of squares due to rep*irrigat, as taken in TEST statement and CONTRAST statement respectively.

In TEST statement h = numerator source of variation and

e = denominator source of variation

Note here that the PROC GLM can also be used to perform analysis of covariance. For analysis of covariance, the covariate should be defined in the model without specifying under CLASS statement.

viii) The OUTPUT statement: The OUTPUT statement is available in a number of procedures, for example, in PROC GLM, PROC REG. This statement is used to create a new dataset which contains outputs of a procedure. The output can contain some variables, by default, depending on the procedure being used and some optional variables which can be specified using options in the OUTPUT statement. For example, the following statements create an output data set called new which contains all the original variables in the data set experiment and a variable resid containing residuals from the fitted models.

PROC GLM DATA = experiment;

CLASS treatment block;

MODEL observation = treatment block;

MEANS treatment;

OUTPUT OUT = new r = RESID;

RUN;

I.9 Exporting data from SAS to other formats

Exporting a data set to another program is the reverse of the import process. For exporting a data, go to "File" and then select "Export Data". An export wizard window will pop up. Then just follow the wizard through the following steps.

Step 1: Choose a data set from the WORK library (where the SAS datasets are stored automatically by SAS) and click 'Next' button.

Step 2: Choose the file type you want to export to. Available types include Excel, Access, dBase, delimited file, and many others. Then click next.

Step 3: Type in the directory path where you want to save your data file. If you are not sure of the path, click on the browse button and find the location. Then click OK. If exporting to Excel, the wizard will ask you to assign a name to the exported table. This name will appear as the Sheet name tab at the bottom of the Excel workbook. At this time, you may click on the FINISH button.

I.10 Titles, footnotes and labels

Titles

One can enter up to 10 titles at the top of output using TITLE statement in your procedure.

PROC PRINT;

TITLE 'height-dia study';

TITLE3 '1999 statistics';

RUN;

Footnotes

One can enter up to 10 footnotes at the bottom of your output.

PROC PRINT DATA=diaht;

FOOTNOTE '1999';

FOOTNOTE5 'study results';

RUN;

For obtaining output as RTF file, use the following statements:

ODS RTF FILE='xyz.rtf' STYLE =JOURNAL;

ODS RTF CLOSE;

For obtaining output as CSV/PDF/HTML file, replace rtf with csv or pdf or html in the above statements.

If we want to get the output in continuous format, then we may use

ODS RTF FILE='xyz.RTF' STYLE =JOURNAL bodytitle startpage=no;

Labelling the variables

DATA dose;

TITLE 'yield with factors N P K';

INPUT N P K Yield;

LABEL N = "Nitrogen";

LABEL P = "Phosphorus";

LABEL K = "Potassium";

CARDS;

...

...

...

;

PROC PRINT;

RUN;

We can define the line size in the output using statement options. For example, if we wish that the output should have the line size (number of columns in a line) as 72, use Options LINESIZE =72; in the beginning.

I.11 Getting help

SAS comes with an excellent documentation for each and every procedure and statement available in SAS. Under *SAS Help and Documentation*, one can go to the *Index* tab or *Search* tab and can search for the topic. One can get help on a topic in SAS website and there are many other sources for help online.

Appendix-1: Syntax of PROC ANOVA

PROC ANOVA <options> ;

    CLASS variables </ option> ;

    MODEL dependents=effects </ options> ;

    ABSORB variables ;

    BY variables ;

    FREQ variable ;

    MANOVA <test-options></ detail-options> ;

    MEANS EFFECTS </ options> ;

    REPEATED factor-specification </ options> ;

    TEST <H=effects>E=effect ;

To perform analysis of variance using PROC ANOVA, one muse use CLASS and MODEL statements, and they must appear before the first RUN statement. The CLASS statement must appear before the MODEL statement. If the ABSORB, FREQ, or BY statements are used, they must appear before the first RUN statement. The MANOVA, MEANS, REPEATED, and TEST statements must be placed after the MODEL statement, and they can be specified in any order. These four statements can also appear after the first RUN statement.

Table A I.1 summarizes the function of each statement (other than the PROC statement) in the ANOVA procedure.

Table AI.1: Statements in the ANOVA procedure	
Statement	**Description**
ABSORB	To absorb classification effects in a model
BY	Separate analysis is done by the levels of the variables specified
CLASS	Used to declare classification or grouping variables
FREQ	To declare frequency variables
MANOVA	To perform multivariate analysis of variance, useful when MANOVA needs to be performed
MEANS	For computing and comparing means based on different multiple comparison procedures
MODEL	To define the dependent and independent variables/factors in the model to be fitted
REPEATED	For performing multivariate and univariate repeated measures analysis of variance
TEST	For performing tests that use the sums of squares for effects and the error terms specified

Appendix-2: Syntax of PROC GLM

PROC GLM <options> ;

CLASS variables </ option> ;

MODEL dependents=independents </ options> ;

ABSORB variables ;

BY variables ;

FREQ variable ;

ID variables ;

WEIGHT variable ;

CONTRAST 'label' effect values <...effect values></ options> ;

ESTIMATE 'label' effect values <...effect values></ options> ;

LSMEANS effects </ options> ;

MANOVA <test-options></ detail-options> ;

MEANS EFFECTS </ options> ;

OUTPUT <OUT=SAS-data-set> keyword=names <...keyword=names></ option> ;

RANDOM effects </ options> ;

REPEATED FACTOR-SPECIFICATION </ options> ;

TEST <H=effects>E=effect </ options> ;

There are a number of statements and options available in PROC GLM. However, only few of them are mostly required. To use PROC GLM, the PROC GLM and MODEL statements are required. At least one MODEL statement must be specified. If the model contains classification effects, CLASS statement must have the classification variables, and the CLASS statement must be placed before the MODEL statement. In addition, if a CONTRAST statement is used in combination with a MANOVA, RANDOM, REPEATED, or TEST statement, the CONTRAST statement must be entered first in order for the contrast to be included in the MANOVA, RANDOM, REPEATED, or TEST analysis. A number of statements in PROC GLM statement must appear before or after some specific statements. For example, the statements ABSORB, BY, FREQ, ID, WEIGHT must appear before first RUN statement, the CLASS statement must appear before the MODEL statement, the CONTRAST statement must appear before MANOVA, REPEATED or RANDOM statement, the MODEL statement must appear before the CONTREAST, LSMEANS, MEANS, ESTIMATE statements, the TEST statement must appear before MANOVA or REPEATED statement. Similarly, the statements CONTRAST, ESTIMATE, LSMEANS, MEANS, MANOVA, OUTPUT, RANDOM, TEST must appear after the MODEL statement.

Table A I.2 summarizes the function of each statement in the GLM procedure.

Table A I.2: Statements in the GLM procedure	
Statement	**Description**
ABSORB	To absorb classification effects in a model
BY	For specifying variables by whose levels analysis is to be done
CLASS	Useful in declaring classification or grouping variables
CONTRAST	Helpful in performing tests of linear functions of the parameters
ESTIMATE	For estimating linear functions of the parameters
FREQ	To specify frequency variables
ID	To identify observations on output
LSMEANS	For computing least squares (marginal) means, these are adjusted means
MANOVA	For performing multivariate analysis of variance
MEANS	For computing and optionally comparing arithmetic means using different multiple comparison procedures
MODEL	For defining the dependent and independent variables/factors in the model to be fitted
OUTPUT	To create an output data set containing diagnostics for each observation
RANDOM	To declare effects which are to be treated as random and computes expected mean squares
REPEATED	For performing multivariate and univariate repeated measures analysis of variance
TEST	To perform tests that use the sums of squares for effects and the error term specified by user
WEIGHT	To specify a variable to give weight to the observations

Different sum of squares in PROC GLM

PROC GLM can provide four types of sums of squares namely Type I, Type II, Type III and Type IV sum of squares while performing ANOVA. For illustration, consider a model with two factors A and B and hence, there are two main effects A, B and an interaction AB. The sum of squares due to full model is represented as SS(A,B,AB). Therefore, the notation SS(A,B) represents sum of squares due to A and B and no interaction, SS(A,AB) represents sum of squares due to A and AB.

The Type I sums of squares are the sequential sums of squares obtained by adding the terms to the model one by one in some sequence. The sum of squares for an effect is adjusted for only those effects which appear before it in the model. Thus, the sums of squares and their expectations depends on the order of terms in the model. For example, SS(A) compute sum of square for the main effect of factor A, SS(B | A) compute sum of square for the main effect of factor B after the main effect of A and SS(AB | B, A) compute the sum of square for the interaction effect AB after both the main effects of A and B.

The Type II, III and IV are 'partial sums of squares' and each of the sum of squares is adjusted for all other classes of the effects in the model. However the adjustment is done in each sum of squares according to different rules. All the three types of sum of squares follow one general rule: the estimable functions that generate the sums of squares for one class of effects will not

involve any other classes of effects except those that "contain" the class of effects in question. For example, the estimable functions that generate SS (AC) in a three-factor factorial will have zero coefficients on main effects A and B and the $(A \times B)$ and $(B \times C)$ interaction effects. They will contain non-zero coefficient for the $(A \times B \times C)$ interaction effects, because the interaction $A \times C$ is contained in the $A \times B \times C$ interaction.

Type II sum of squares for a given effect is adjusted for all effects that do not contain the given effect. For example, sum of squares due to main effects A and B will both be adjusted for each other because none of them contains the other, but will not be adjusted for AB since it contains both A and B. Sum of squares due to effect AB will be adjusted for both the main effects A and B. Type II sum of squares is most powerful when there is no interaction and hence, should not be used in factorial designs.

Type III sum of squares due to a term is the sum of squares that would be obtained for each variable if the term were entered last into the model. In other words, Type III sum of squares due to a variable is evaluated after all other factors have been accounted for. Type III sum of squares should not be used when there are missing cells in the design.

Type IV sum of squares, a variation of Type III sum of squares is specifically developed for designed to deal with missing cells.

If there are no empty cells (no $n_{ij} = 0$), Type III and Type IV sums of squares are equal. The hypothesis being tested is the same as when the data are balanced. When there are empty cells, the hypotheses being tested by the Type III and Type IV sums of squares may be different. The Type III criterion of orthogonality reproduces the same hypotheses one obtains if sum of the add to zero. When there are empty cells this is modified to "the effects that are present are assumed to be zero". The Type IV hypotheses make use of balanced subsets of non-empty cells and may not be unique. Consider a 2×3 factorial where the terms to the model are added in the order A, B, AB. Then various types of sums of squares can be explained as follows:

Effect	Type I	Type II	Type III	Type IV
General Mean	R(m)	R(m)		
A	R(A/ m)	R(A/ m,B)	R(A/m,B,AB)	
B	R(B/m,A)	R(B/m,A)	R(B/m,A,AB)	
A*B	R(A*B/ m,A,B)	R(A*B/m,A,B)	R(AB/m,A,B)	

R (A/m) is sum of squares adjusted for m, and so on.

The four types of sums of squares and four types of data structures (balanced and orthogonal, unbalanced and orthogonal, unbalanced and non-orthogonal (all cells filled), unbalanced and non-orthogonal (empty cells)) are related to each other. To see this, let n_{IJ} denote the number of observations in level *I* of factor A and level *J* of factor B. Then, the following table gives the relationship between different types of sums of squares in a two-way classified data.

Data Structure Type

Effect	1	2	3	4
	Equal n_{IJ}	Proportionate n_{IJ}	Disproportionate non-zero n_{IJ}	Empty Cell
A	I=II=III=IV	I=II, III=IV	III=IV	
B	I=II=III=IV	I=II, III=IV	I=II, III=IV	I=II
A*B	I=II=III=IV	I=II=III=IV	I=II=III=IV	I=II=III=IV

In general,

I=II=III=IV	(balanced data);	II=III=IV	(no interaction models)
I=II, III=IV	(orthogonal data);	III=IV	(all cells filled data).

Appendix-3: Syntax of PROC RSREG

PROC RSREG <options>;

MODEL RESPONSES = independents </ options> ;

RIDGE <options> ;

WEIGHT variable ;

ID variables ;

BY variables ;

The PROC RSREG and MODEL statements are required. The BY, ID, MODEL, RIDGE and WEIGHT statements are optional and they can appear in any order.

The important options available with the MODEL statement are:

Option	Purpose
NOCODE	To analyse the original data without any coding.
ACTUAL	To specify the actual values from the input data set.
COVAR= n	To declare that the first n variables on the independent side of the model are simple linear regression (covariates) rather than factors in the quadratic response surface.
LACKFIT	To perform lack of fit test. For this the repeated observations must appear together. Therefore, if repeated observations do not appear together, then SORTING is must before using LACKFIT
NOANOVA	To suppress the printing of the analysis of variance and parameter estimates from the model fit.
NOOPTIMAL (noopt)	To suppress the printing of canonical analysis for quadratic response surface.
NOPRINT	To suppress both anova and the canonical analysis.
PREDICT	To specify the values predicted by the model.
RESIDUAL	To specify the residuals.
RIDGE	To compute the ridge of the optimum response. Following important options available with ridge statement : max: To compute the ridge of maximum response. min: To compute the ridge of the minimum response.

Introduction to R

II.1 Introduction

R is a powerful programming language for statistical analysis. This software is an implementation of S programming language which was designed by John Chambers at Bell Labs. R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand. It is currently developed by R Development Core Team. The name 'R' is from the first names of first two authors and partly due to its inheritance from 'S'. Recently R has become one of world's popular statistical analysis software, because of the following reasons:

i) R is free, open source and capable of almost any statistical analysis including the most recently developed statistical methodologies.

ii) R provides very good graphical facilities.

iii) R is easily extensible through new contributions from statisticians and researchers around the globe, which also makes R quite different from other popular statistical analysis software. In fact, R community is highly active in terms of new contributions in the form of packages to R.

II.2 Getting and installing R

To be able to use R, it needs to be installed in computer. R is available for free download from any one of the mirror sites of Comprehensive R Archive Network (CRAN) in http://cran.r-project.org/. For downloading, it is better to select a mirror located nearer to you. R is available for installation in Windows/Macintosh/Unix platforms. To install R in a given machine, first double-click the downloaded file R.exe, then select language as 'English'. R setup wizard window will appear. Select on 'Next' and accept most of the default settings during the installation. Latest version as on 10.12.2015 available is 3.2.3.

II.3 Using R

II.3.1 Starting R

To start R, click on start menu → all programs → R → R 3.2.3 and a screen as shown in Figure 1 appears. The white blank screen is called R Console and this is the place where all R codes are written and outputs appear, unless outputs are directed to some external files.
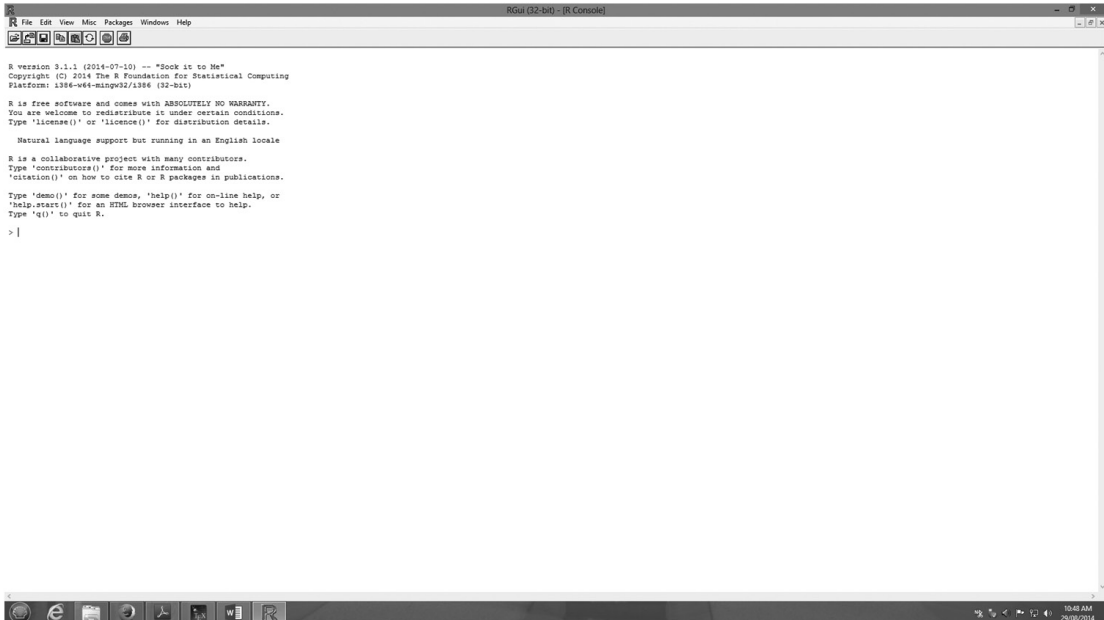
Figure II.1: R window

There will be a toolbar at the top of the Console and a few menus in the R window. To know what the buttons on the toolbar does, hold your mouse on the button for some time, a description of the button will appear. There is an R editor which can be used to write and edit R codes. R editor window is just like a text editor with facilities for select, cut, copy, paste, typing text, deleting text etc. This window opens by clicking on File → New Script in the menu bar. The codes written in R editor window needs to be passed to the R console for execution by clicking on 'Run line or selection button' 🔲 on the toolbar in the R editor window

II.3.2 R commands

There is a '>' symbol in the Console. The commands are typed after this symbol and then the Enter button needs to be pressed. When a command is written in the Console and the Enter button is pressed, R reads the commands and returns some results or some error message on the Console. For example, if you type

> 2+8

[1] 10

In this case R added 2 and 8 and returned the result 10. Now, type

> 2+*5

Error: unexpected '*' in "2+*"

This time R has returned an error message, because '+*' is not a defined operator in R.

Therefore, one should know what to type in the Console, otherwise, it will always give an error message. So R works interactively by returning results to the commands one by one, like in the following example:

> x=2

> x*5

[1] 10

R is mainly an *expression language* and comes with a syntax. Some common rules of R commands are

i)   R commands are *case sensitive*, so AB, Ab, aB and ab are different objects.

ii)   All alphanumeric characters, '.' and '_' are allowed as symbols. In some countries accented letters are also allowed.

iii)   An R name must start with '.' or a letter, and if it starts with '.' the second character must not be a digit. Names are unlimited in length.

iv)   Elementary commands consist of either *expressions* or *assignments*. An expression is evaluated, printed (unless specifically made invisible), and the value is lost. An assignment also evaluates an expression and passes the value to a variable but the result is not automatically printed.

v)   Commands are separated either by a semi-colon (';'), or by a newline.

vi)   Braces ('{' and '}') are used to create a block of codes.

vii)   *Any line starting with a # till the end of the line is a comment and are not evaluated. Comments can be placed anywhere.*

viii)   If a command is not complete at the end of a line, R will give a different prompt, by default '+' on second and subsequent lines and continue to read input until the command is syntactically complete. This prompt may be changed by the user.

ix)   Length of a command at the Console is limited to about 4095 bytes (not characters).

II.3.3 Working directory

The working directory refers to the directory or folder where R is currently working. By default the working directory is "My documents" or 'Documents". You can get the working directory by using code

> getwd()

[1] "C:/Users/User/Documents"

R can read and open files from working directory directly without specifying any path. Similarly, it can save files and write to files in the working directory directly. One can reset the working directory to a different folder using the code below.

> setwd("C:/Users/User/Documents/DOE handbook/")

In the beginning of an R session, it is better to set the working directory to a folder where most of the data files and codes are located.

II.3.4 Data types in R

R is an object oriented language and therefore, all data types in R are some kind of object. Objects may be variables, vectors, matrices, arrays, character strings, functions, or more general structures built from such components.

During an R session, objects are created and stored by name. One can use the command

> objects()

to display the names of the objects which are currently stored within R. The collection of objects currently stored is called the *workspace*. One can remove objects using the function rm(). For example, the following code removes objects x and y form workspace.

> rm(x, y)

An object created during an R session can be saved in a file for use in future R sessions. The entire workspace of an R session and the history of all the commands used during the session can also be saved. Some commonly encountered objects are discussed below.

a) Vectors: Simplest object in R is a vector. A vector is a collection of elements. For example,

> x = c(10, 15, 20, 25, 26)

creates a vector of 5 numbers. Here the object x contains those numbers and the function c() is used to assign those numbers to the object x. Vectors can be of three types i) numeric ii) character and iii) logical. A numeric vector contains numbers, a character vector contains characters and a logical vector can contain values TRUE, FALSE or NA.

b) Matrices: A matrix object also is a collection of elements but it has two dimensions. They can also be numeric, character or logical in nature. Following is an example of creating a matrix.

> x=matrix(c("a", "b", "c", "d"),nrow=2)

> x

[,1] [,2]

[1,] "a" "c"

[2,] "b" "d"

c) Arrays: Arrays are multi-dimensional generalization of vectors and matrices. A two-dimensional array is a matrix. Arrays can have more than two dimensions.

d) Factors: Factor objects are used to specify categorical or classificatory or grouping variables. For example, males and females are two levels of a variable gender. Then gender can be thought of a factor object.

> gender=c("M", "F", "M")

> gender=as.factor(gender)

> levels(gender)

[1] "F" "M"

Factor variables are particularly useful in analysis of variance and in linear model with grouping variables.

e) Lists: A list is a collection of objects where each object can be of different type. For example, a list can have first object as a vector, second object as a matrix and third object as a data frame.

> mylist=list(x=c(10,20,30),y=matrix(1:6,nrow=3))

> mylist

$x

[1] 10 20 30

$y

   [,1] [,2]

[1,] 1   4

[2,] 2   5

[3,] 3   6

> mylist[[1]]

[1] 10 20 30

f) Data frames: A data frame is a two dimensional object. But unlike matrices, different columns of data frame can be different types, for example some columns can be numeric, some columns can be character, some columns can be factors. Here a column generally refers to a variable.

> age=c(20,25,28,30,26)

> weight=c(50,53,54,55,51)

> mydata=data.frame(age,weight)

> mydata

 age weight

1 20   50

2 25   53

3 28   54

4 30   55

5 26   51

The data.frame() function is used to create a data frame.

g) Functions: Functions in R are a kind of objects which takes one or more inputs and produces some result(s) as output. R has a number of in-built functions. R also provides facility to create new functions by users. R has huge number of in-built functions. As a simple example, to obtain the mean and variance of a set of numbers 10, 13, 21, 34, 51, 32, 45, 32, 17, 29, 41, 52, the following code can be used.

> x=c(10,13,21,34,51,32,45,32,17,29,41,52)

> mean(x)

[1] 31.41667

> var(x)

[1] 200.9924

Here, c(), mean() and var() are in-built functions of R. The function c() assigns those numbers to the object x. The commands mean(x) and var(x) computes the mean and variance of an object x. Here, x is the input, also called argument, to the function mean() and var().

A complete list of in-built functions is available in the document R reference manual. The R reference manual opens by clicking on Help → Manuals (in PDF) → R reference. It opens the full reference manual. It contains a complete list of all the functions and objects in base R. Apart from in-built functions, a large number external functions are available in contributed packages. Contributed packages are nothing but a collection of functions written by the authors of the packages to perform specific analysis. The manual of a package contains the details of the functions provided in that package.

To know what are the argument(s) of a function and how to use it, you can always type help(functionname) where functionname is the name of the function. This opens an html page in browser containing the details of the function. For example, help(lm) gives the details of the usage of the function lm().

II.4 R packages

Though most of the standard statistical analysis are available in base R, but sometimes some contributed R packages are needed to do some specific analysis. An R package is a bundle of functions and codes for performing some statistical or mathematical analysis which are generally not covered in base R. This facility of R packages extends the usefulness of R greatly. A

large number of packages, as of 09 February 2016 around 7,800 packages are available on CRAN for a variety of analysis and the number is increasing day by day.

II.4.1 Downloading and installing a package

To use an R package, download the package from CRAN and then install and load it in an R session. A package can be downloaded from within R or from outside R. On a Windows machine which is connected to internet, a package can be installed by clicking on Packages → Install packages(s) from the menu bar. This will open a list of mirrors. Select a mirror and then, from the available list of packages in the website, select the desired packages. They will be installed into R.

To keep a copy of the downloaded package, you can visit any CRAN mirror web page and download the package. You can then install it by clicking on 'Packages' and then clicking on 'Install package(s) from local zip files...' and then select the zip file containing the package.

After you have installed a package, you need to load it to R. For this click on 'Packages' and then click on 'Load package...'. Alternatively, you can type library (packagename) in the console to load a package where packagename is the name of the package. For example, to load a package agricolae, you need to type

> library(agricolae)

A package is to be installed just once, but to use it for analysis, it needs to be loaded every time R is started. To use a package, you should download the manual of the package. The package manual contains documentation on functions which are available in that package. Sometimes more than one package may be needed for analysis. It is always better to load only the required packages. There is no need to load packages which are not required in a session, because loading a number of packages slows down R.

When some analysis are not available directly in base R software, then only we need to use some R package. Sometimes more than one R package may be available for similar kind of analysis and the user should see which of the packages does exactly what and then use the suitable package for the analysis. Another important point to consider while installing an R package is to check the compatibility of the package with the R version being used. It can be seen from the package manual what version of R is required for using the package. Generally, most of the packages installed on a given date will run on most recent version of R software on that date.

II.5 Reading data in R

i) Loading data in R directly

Data with few variables and few observations can be read in R by typing in the Console R as shown in the following example.

> month<-c('Jan','Feb','Mar','Apr','May','Jun','Jul', 'Aug','Sep','Oct','Nov','Dec')

> rainfall<-c(5,4,8,7,9,20,30,35,24,15,10,8)

> mydata=data.frame(month,rainfall)

> mydata

  month rainfall

1 Jan   5

2 Feb   4

3 Mar   8

4 Apr   7

5 May   9

6 Jun   20

7 Jul   30

8 Aug   35

9 Sep   24

10 Oct   15

11 Nov   10

12 Dec   8

Note that data.frame() function combines the vectors month and rainfall into a data frame called mydata. Note that a dataset in R is always in the form of a two-dimensional array with columns representing variables and rows representing individual observations. Sometimes one may be interested to know the names of variables in a data set loaded in R. For example, to know the names of the variables in data set mydata one can use following command:

> names(mydata)

[1] "month"   "rainfall"

The scan() function can also be used to read data directly typed in R console. For example,

> y<-scan()

1: 393 55 32 40

5: 2 1 3 5

9:

Read 8 items

> y

[1] 393 55 32 40 2 1 3 5

When entering data from keyboard using scan() function, one has to hit enter button when one does not want to type any more data. Then R stops scanning and loads the data into the object. The function scan() is also able to read data from external file.

ii) Loading data in R from an external file

Most often the data may not be just a few observations. There may be quite many variables and observations. In that case, the data may be in a spreadsheet or some other external file, or from some other statistical software or from some web page. R provides facilities for loading data from each of them.

Reading data from text file

Data in text file should be kept such that the individual observations are separated with a delimiter. Some commonly used delimiters are ',',':','t',' ' i.e., blank space, '\~', '@', '\&', '*' etc. But be sure that none of the observations or variables in the data set have any of those characters, otherwise data will be loaded improperly and there may be error in loading of data. Consider a text file with following observations with comma(',') as a delimiter.

Jan,5

Feb,4

Mar,8

Apr,7

May,9

Jun,20

Jul,30

Aug,35

Sep,24

Oct,15

Nov,10

Dec,8

Let the file name is "rainfall.txt" and is kept in the working directory. This data can be loaded in R by using the function read.table() as follows:

> mydata2=read.table("rainfall.txt",header=TRUE,sep=",")

> mydata2

  month rainfall

1 Jan   5

2 Feb   4

3	Mar	8
4	Apr	7
5	May	9
6	Jun	20
7	Jul	30
8	Aug	35
9	Sep	24
10	Oct	15
11	Nov	10
12	Dec	8

First argument of read.table() refers to the external file. The second argument header=TRUE tells R that there is header in the rainfall.txt file, and those are used as variable names for the data. If there is no header in a text file, then header=FALSE should be used. Third argument sep=" , " tells R that observations are separated by a ' , '. There are other arguments to read. table() function, but these three are essential. The details of the usage of the function read. table() is available with help(read.table) in the Console.

There are some other functions to read files with specific delimiters. The function read.csv() function loads comma separated value (csv) files, i.e., files with comma delimited observations, read.csv2() function loads data from semicolon (';') delimited files, read.delim() and read. delim2() functions load data from tab delimited files.

Reading data from a webpage

Suppose some data is available on a webpage. To read a dataset from a web page the function read.table() can be used with the complete address of the page. For example,

> webdata=read.table("http://data.princeton.edu/wws509/datasets/effort.dat")

> webdata

	setting	effort	change
Bolivia	46	0	1
Brazil	74	0	10
Chile	89	16	29
Colombia	77	16	25
CostaRica	84	21	29
Cuba	89	15	40
DominicanRep	68	14	21

Ecuador	70	6	0
ElSalvador	60	13	13
Guatemala	55	9	4
Haiti	35	3	0
Honduras	51	7	7
Jamaica	87	23	21
Mexico	83	4	9
Nicaragua	68	0	7
Panama	84	19	22
Paraguay	74	3	6
Peru	73	0	2
TrinidadTobago	84	15	29
Venezuela	91	7	11

Loading data from a spreadsheet

To load data from an excel file to R, the relevant worksheet may be saved into a tab delimited text file or into a csv file and then the text file or .csv may be loaded using read.table() or read.csv() function. However, if to read the data from excel directly into R, a package called RODBC is needed. An example of loading data from excel is shown below.

> library(RODBC)

> connection = odbcConnectExcel("myfile.xlsx")

> sqlTables(connection)$TABLE_NAME # show the worksheets

> df = sqlFetch(connection, "Sheet1") #Read worksheet 1

Alternative way

> df = sqlQuery(connection, "select * from [Sheet1 $]")

> close(connection) #close the connection to the file

There is also an xlsx package which offers reading and writing excel files. For example,

> library(xlsx)

> read.xlsx("myfile.xlsx", sheetName = "Sheet1")

II.6 Some statistical analysis examples

In this Section, some examples of statistical analysis using R is given. For this purpose, we use the PlantGrowth data set available in R.

```
> data(PlantGrowth)
> PlantGrowth
   weight group
1    4.17  ctrl
2    5.58  ctrl
3    5.18  ctrl
4    6.11  ctrl
5    4.50  ctrl
6    4.61  ctrl
7    5.17  ctrl
8    4.53  ctrl
9    5.33  ctrl
10   5.14  ctrl
11   4.81  trt1
12   4.17  trt1
13   4.41  trt1
14   3.59  trt1
15   5.87  trt1
16   3.83  trt1
17   6.03  trt1
18   4.89  trt1
19   4.32  trt1
20   4.69  trt1
21   6.31  trt2
22   5.12  trt2
23   5.54  trt2
24   5.50  trt2
25   5.37  trt2
26   5.29  trt2
27   4.92  trt2
```

28 6.15 trt2

29 5.80 trt2

30 5.26 trt2

Basic summary of the dataset can be found using summary() function.

> summary(PlantGrowth)

	weight	group
Min. :	3.590	ctrl:10
1st Qu. :	4.550	trt1:10
Median :	5.155	trt2:10
Mean :	5.073	
3rd Qu.:	5.530	
Max. :	6.310	

Note the summary() function has returned mean, median, 1st and 3rd quartiles and minimum and maximum of the weight variable. Since the group variable is factor object, so it has returned number of observations for each level of the group variable.

To get the variance of weight variable, the var() function can be used.

> var(PlantGrowth$weight)

[1] 0.49167

To have an idea of distribution of weight for each group in the data, boxplots can be obtained using boxplot() function as shown below.

> boxplot(weight~group,data=PlantGrowth)

This produces the plot as shown in Figure II.2 in R Graphics Device window.
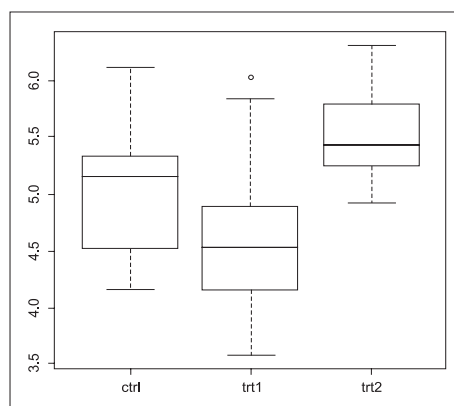
Figure A II.2: Box plot of weight for each group

351

To get the mean and standard deviation of weight for each group, the aggregate() function can be used as shown below.

> aggregate(weight~group,data=PlantGrowth,mean)

  group weight

1 ctrl 5.032

2 trt1 4.661

3 trt2 5.526

> aggregate(weight~group,data=PlantGrowth,sd)

  group   weight

1 ctrl 0.5830914

2 trt1 0.7936757

3 trt2 0.4425733

R has function for performing t-tests. The function t.test() is available for this purpose. For example, to test the hypothesis that the population mean of weight variable is 5, the following command can be used.

> t.test(PlantGrowth$weight,mu=5)

One Sample t-test

data: PlantGrowth$weight

t = 0.5702, df = 29, p-value = 0.5729

alternative hypothesis: true mean is not equal to 5

95 percent confidence interval:

 4.811171 5.334829

sample estimates:

mean of x

5.073

Two-independent sample *t*-test can also be performed using t.test() function. For example, the following example tests the hypothesis whether the population mean of weights of trt1 group is equal to the population mean of weights of trt2 group.

> t.test(weight~group,data=PlantGrowth,subset=group!="ctrl",

var.equal=TRUE)

Two Sample t-test

data: weight by group

t = -3.0101, df = 18, p-value = 0.007518

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -1.4687336 -0.2612664

sample estimates:

mean in group trt1 mean in group trt2

4.661   5.526

A paired *t*-test can be performed with an additional argument paired=TRUE in the above function. However that should be done only when the sample observations for the two groups are matched or paired.

Since the above data is from an experiment and the objective was to see whether the average weight of three groups differ significantly or not, an analysis of variance (ANOVA) can be performed. R provides aov() function for performing ANOVA for balanced data. If the data is not balanced with respect to the grouping variables, then it is better to use the lm() function which fits linear model. The following commands show the uses of both the functions.

```
> aov1=aov(weight~group,data=PlantGrowth)
> summary(aov1)
          Df Sum Sq Mean Sq F value Pr(>F)
group      2  3.766  1.8832   4.846 0.0159 *
Residuals 27 10.492  0.3886

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> lm1=lm(weight~group,data=PlantGrowth)
> anova(lm1)
Analysis of Variance Table
Response: weight
          Df  Sum Sq Mean Sq F value  Pr(>F)
group      2  3.7663  1.8832  4.8461 0.01591 *
Residuals 27 10.4921  0.3886

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that the function anova() has been used on the fitted lm1 object to get ANOVA table. The ANOVA tables are similar from both aov1 and lm1 objects. Both these objects contain a number of other terms such as fitted values, residuals, coefficients, degrees of freedoms etc.

Since the ANOVA suggests that the groups differ significantly at 5% level with respect to weight, a post hoc test (Tukey's Honest significant difference test) can be performed using TukeyHSD() function to compare the groups pairwise.

> TukeyHSD(aov1)

  Tukey multiple comparisons of means

   95% family-wise confidence level

Fit: aov(formula = weight ~ group, data = PlantGrowth)

$group

        diff   lwr   upr   p adj

trt1-ctrl -0.371 -1.0622161 0.3202161 0.3908711

trt2-ctrl 0.494 -0.1972161 1.1852161 0.1979960

trt2-trt1 0.865 0.1737839 1.5562161 0.0120064

The result suggest treatment 1 and treatment 2 groups are significantly different from each other. Note that the TukeyHSD() function takes an object of class "aov" as argument. It will not work on an object of class "lm".

To get least significant difference (LSD) or to perform other post hoc tests such as Duncan's multiple range test, additional packages need to be used. Some important packages with respect to this book are agricolae, car and lsmeans.

LSD can be computed after installing and loading the package in R.

> library(agricolae)

> LSD.test(aov1,"group",console=TRUE)

> # Or alternatively

> LSD.test(lm1,"group",console=TRUE)

Study: aov1 ~ "group"

LSD t Test for weight

Mean Square Error: 0.3885959

group, means and individual (95 %) CI

	weight	std	r	LCL	UCL	Min	Max
ctrl	5.032	0.5830914	10	4.627526	5.436474	4.17	6.11
trt1	4.661	0.7936757	10	4.256526	5.065474	3.59	6.03
trt2	5.526	0.4425733	10	5.121526	5.930474	4.92	6.31

alpha: 0.05 ; Df Error: 27

Critical Value of t: 2.051831

Least Significant Difference 0.5720126

Means with the same letter are not significantly different.

Groups, Treatments and means

a   trt2   5.526

ab   ctrl   5.032

b   trt1   4.661

The package lsmeans is useful for computing least square means as is done in SAS. For example

> library(lsmeans)

> lsmeans(aov1,"group")

> #or alternatively

> lsmeans(lm1,"group")

group	lsmean	SE	df	lower.CL	upper.CL
ctrl	5.032	0.1971284	27	4.627526	5.436474
trt1	4.661	0.1971284	27	4.256526	5.065474
trt2	5.526	0.1971284	27	5.121526	5.930474

Confidence level used: 0.95

Pairwise comparisons of treatments is possible using pairs() function in lsmeans pacakge.

> lsm=lsmeans(aov1,"group")

> pairs(lsm)

contrast	estimate	SE	df	t.ratio	p.value
ctrl - trt1	0.371	0.2787816	27	1.331	0.3909
ctrl - trt2	-0.494	0.2787816	27	-1.772	0.1980
trt1 - trt2	-0.865	0.2787816	27	-3.103	0.0120

P value adjustment: tukey method for a family of 3 means

Often we are interested in compact letter display of the treatments. Under compact letter display, treatments with same letters are not significantly different. This display is particularly useful if number of treatments is more. For the above data, one can have compact letter display using following code.

> cld(lsm, Letters="ABCDE")

group	lsmean	SE	df	lower.CL	upper.CL	group
trt1	4.661	0.1971284	27	4.256526	5.065474	A
ctrl	5.032	0.1971284	27	4.627526	5.436474	AB
trt2	5.526	0.1971284	27	5.121526	5.930474	B

Confidence level used: 0.95

P value adjustment: tukey method for a family of 3 means

significance level used: alpha = 0.05

Often type III sum of squares are desired. The package car is useful for such situations.

> library(car)

> Anova(aov1,type="III")

> # or alternatively

> Anova(lm1,type="III")

Anova Table (Type III tests)

Response: weight

	Sum Sq	Df	F value	Pr(>F)	
(Intercept)	253.210	1	651.6029	< 2e-16	***
group	3.766	2	4.8461	0.01591	*
Residuals	10.492	27			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Testing the significance of contrasts is often required. Both the packages lsmeans and car are useful for this. For example, to test the significance of the contrast 2*ctrl – trt1 – trt2, one can use the following commands.

> lsm=lsmeans(lm1,"group")

```
> con1=contrast(lsm, list(con1 = c(2,-1,-1)))
> con1
 contrast estimate      SE df t.ratio p.value
 con1     -0.123 0.4828639 27  -0.255  0.8009
> lht(lm1,con1@linfct)
Linear hypothesis test
```

Hypothesis:
- grouptrt1 - grouptrt2 = 0

Model 1: restricted model
Model 2: weight ~ group

```
Res.Df   RSS Df Sum of Sq     F Pr(>F)
1    28 10.517
2    27 10.492  1  0.025215 0.0649 0.8009
```

The function lht() is for linear hypothesis test and is available in car package. All these packages have many other functions. To get details about the available functions in these packages, please refer to the manuals of these packages. The reader is referred to the E-manual on R-Scripts For Statistical Analyses using R-Studio by Parsad et al (2015) available at http://www.iasri.res.in/sscnars/content_rmanual.htm

II.7 RStudio

RStudio is an integrated development environment (IDE) for doing statistical analysis and other tasks using computing power of R software. RStudio is available in two editions: RStudio Desktop, where the program is run locally as a regular desktop application; and RStudio Server, which allows accessing RStudio using a web browser while it is running on a remote Linux server. RStudio Desktop is available for Microsoft Windows, Mac OS X, and Linux.

RStudio is available in open source and commercial editions and runs on the desktop (Windows, Mac, and Linux) or in a browser connected to RStudio Server or RStudio Server Pro (Debian/Ubuntu, RedHat/CentOS, and SUSE Linux). The free edition of RStudio desktop can be downloaded from *https://www.rstudio.com/*.

RStudio provides more user friendly interface to use R. For using RStudio in machine, base

R must be installed in that machine. The interface of RStudio is given in Figure II.3.

The top left window in Figure II.3 has the editor for writing R codes. Bottom left window
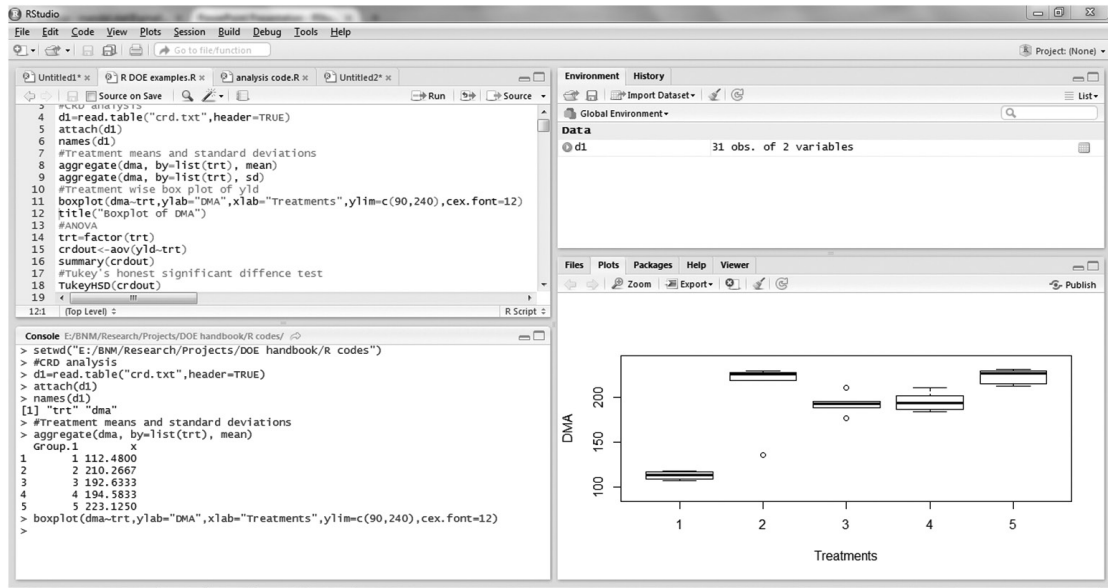
Figure II.3: RStudio interface

is similar to R console where R codes get executed. Top right window of RStudio has two tabs: Environments and History, respectively. In the Environment tab, we can see which objects are currently loaded in R. The History tab gives the history of commands already executed. Bottom right window of RStudio has several tabs namely Files, Plots, Packages, Help and Viewer. The Files tab can be used to explore different files, Plots tab is used to view plots produced from R codes, Packages tab shows the packages already installed and also allows easy installation and loading of packages in a session. Help tab can be used to see the help on R functions and Viewer tab can be used to see local web content.

Multiple Comparison Procedures

III.1 Introduction

In the usual analysis of variance the key test statistic used is Snedecor's *F*. This statistic is used for testing the null hypothesis that the population means of all the treatment effects are same against the alternative that there is at least one inequality between two treatment effects. In other words, the null hypothesis is H_0: $\tau_1 = \cdots = \tau_i = \cdots = \tau_v$ and the alternate hypothesis is H_1: $\tau_i \neq \tau_l$ for at least one pair ($i, l; i \neq l = 1, 2, \ldots, v$), where τ's denote the population means of treatment effects and v is the total number of treatments. The *F* test actually determines whether there exists a significant difference among treatments or not. When the null hypothesis that there is no difference among treatment effects is not rejected, it implies that the factor levels do not influence the response. In that case there is nothing which the researcher can learn and the analysis ends there. When the null hypothesis of equality of treatment effects is rejected, the researcher knows that there is no equality amongst the treatment effects. The implication of this is that there is a relation between the factor levels and the response. But the researcher does not get any knowledge about the form of the inequality of the alternate hypothesis. In that case the researcher needs to thoroughly investigate the nature of the factor levels relation with the response.

The objective of an experiment is never so narrow as to determine whether or not all the treatments have similar effects in terms of similar response. The purpose of an experiment has to be much broad and may demand getting answers to some more probing questions given below:

(a)   Does the effect of treatment 1 differ from that of treatment 2 (H_0: $\tau_1 - \tau_2 = 0$), or does the effect of treatment 5 differ from that of treatment 8 (H_0: $\tau_5 - \tau_8 = 0$), or does the effect of treatment *i* differ from that of treatment *l* (H_0: $\tau_i - \tau_l = 0$); $i \neq l = 1, 2, \ldots, v$?

(b)   Does the average effect of treatments 1 and 2 differ from the average effect of treatments 3 and 4 (H_0: $\tau_1 + \tau_2 - \tau_3 - \tau_4 = 0$), or does the average effect of treatments 1, 2, 3, and 4 differ from the effect of treatment 9 (H_0: $\tau_1 + \tau_2 + \tau_3 + \tau_4 - 4\tau_9 = 0$)? Or does the average effect of a subset of treatments differ from the average effect of another subset of treatments (H_0: $2\tau_1 + 2\tau_2 + 2\tau_3 - 3\tau_4 - 3\tau_5 = 0$)?

(c)   Which of the treatments differ from a standard or control treatment or from each other (comparisons of interest $\tau_2 - \tau_1, \tau_3 - \tau_1, \cdots, \tau_v - \tau_1$ given that τ_1 is the control treatment)?

(d)   Which is the best treatment? (Best may be with highest (adjusted) mean or lowest adjusted (mean) depending on the attribute under study). For yield, etc. the treatment with highest mean is best, whereas for disease intensity in a plant protection trial, the

treatment with lowest mean is the best. Further, the treatments which are statistically at par with the best treatment need to be identified. The recommendation may be given based on the feasibility/availability/returns over the variable cost among the treatments which are statistically at par with the best performing treatment.

The implication of what has been just described is that when the *F* test rejects the null hypothesis, we usually want to undertake a thorough analysis of the nature of the factor-levels effects. Contrast analysis is a way to probe further the type of factor levels response relationship and has been dealt with in Chapter 3. This is a very powerful technique and has the capacity to answer almost all the questions for which the researcher needs an answer. Another useful and popular way to explore the cause of rejection of the null hypothesis or the type of relation between the factor levels and response is the multiple comparison procedure. In contrast to analysis of variance, which simply tests the null hypothesis that the population means of all treatment effects are equal, multiple comparisons procedures help one determine where the differences among the population means occur. These methods examine or compare more than one pair of treatment effects simultaneously. It may be noted that making pairwise comparisons of treatments effects over and over again does not work in general, because the significance level is not as specified for a single pair comparison. If *c* hypotheses are to be tested simultaneously, each at significance level α, then the probability that at least one hypothesis is incorrectly rejected is at most *cα*. Similarly, while obtaining *c* simultaneous confidence intervals, one for each of comparison of treatment effects with a $100(1 - \alpha)$ % confidence coefficient, then the probability that the *c* confidence intervals will be simultaneously correct is at least $1 - c\alpha$. The probability *cα* is the overall significance level or experiment wise error rate. Several multiple comparison procedures are such that these can be used after examining the results. The comparisons need not be defined prior to conducting the experiment.

Many methods exist to detect differences between population means of treatment effects. Some commonly used procedures for comparison of population means are (a) Fisher's Protected Least Significant Difference, (b) Duncan's Multiple Range Test (DMRT), (c) Bonferroni's adjustment method, (d) Scheffe's method, (e) Tukey's Honest Significant Difference (HSD) method, (f) Dunnett's method.

Multiple comparison tests have the same assumptions of ANOVA *viz.*, normality and independence of errors and homogeneity of error variances. Though these tests are somewhat robust, nonparametric multiple comparison tests exist if the assumptions are seriously violated. All multiple comparison tests work best if sample sizes are equal.

It may be worthwhile mentioning here that in a broader sense, multiple comparison procedures can be applied for testing of differences in population parameters other than the mean. The logic remains much the same. Multiple comparison procedures rest on application of the same test to each pair of populations. Of course, the more tests (comparisons) a researcher does, the greater the chance that an apparently extreme difference appears just by chance. When several hypotheses are to be tested, the probability that at least one hypothesis is incorrectly rejected can be very high.

Hence, main problems here are

i.   Establishing and maintaining an appropriate probability of family-wise type I error and

ii.   Based on the family-wise type I error, calculating the probability of comparison-wise type I errors.

III.2   Methods of multiple comparisons

The purpose of this section is to describe various multiple comparisons methods for testing hypotheses after the analysis of variance has been done. We begin with the most popular Fisher's Least Significant Difference.

III.2.1   Fisher's Least Significant Difference (LSD)

Fisher's LSD (to be written as LSD, henceforth) is a technique used for making pairwise comparison between the population mean of one treatment effect with the population mean of the other treatment effect. This test is needed to explore and compare the effects of the treatments pairwise after the analysis of variance has rejected the null hypothesis that all the treatment effects are same or equal. The LSD test is essentially a set of individual Student's *t* tests, and unlike other multiple comparison methods, it does not make any correction for multiple comparisons. The only difference between a set of *t* tests and LSD is that *t* tests compute the pooled standard error from only the two treatments being compared, while the LSD test computes the pooled standard error from all the treatments (or groups).

The LSD test considers the square root of the error (residual) mean square from the analysis of variance as the pooled standard error. It then takes account of the sample sizes (or replications) of the two treatments (or groups) being compared and computes the estimated standard error of the estimated difference of the two treatment effects. Thereafter, it computes *t* statistic as ratio of estimated difference of two population means of two treatment effects to the estimated standard error of this difference. The LSD, however, does not account for the multiple comparisons. LSD, however, controls only individual error rate. Therefore, it should be used only when null hypothesis about equality of treatment effects through ANOVA is rejected.

Consider a design with v treatments and replications (or sample sizes) $r_1, r_2, \ldots r_i, \ldots r_v$, r_i being the replication of the ith treatment. For the pair of treatments i and l, let the difference be denoted as $d_{il} = \tau_i - \tau_l$. The estimated difference is $\hat{d}_{il}$ with estimated standard error as $se_d = \sqrt{\left(\dfrac{1}{r_i} + \dfrac{1}{r_l}\right) s_e^2}$, where s_e^2 is the error (or residual) mean square obtained from the analysis of variance. The *t* statistic is defined as $= \dfrac{\hat{d}_{il}}{se_d}$. The statistic *t* follows Student's *t* distribution with error degrees of freedom ($= \zeta$, say). Let $t_{\varsigma, \alpha}$ denote the table value of Student's *t* at ζ degrees of freedom and α level of significance. Then

$$\text{LSD}\,(\tau_i,\tau_l) = \ t_{\varsigma,\alpha} \times se_{d}.$$

If $\hat{d}_{il} > \text{LSD}(\tau_i,\tau_l)$, then the differences are significant; otherwise the differences are not significant. LSD is also known as CD (Critical Difference).

If $r_1 = r_2 = \cdots = r_i = \cdots = r_v = r$, then $se_d = \sqrt{2s_e^2 / r}.$

For designs with equal replications *i.e.* $r_i = r\ \forall\,\text{i} = 1,2,\cdots,\text{v}$, the LSD will be same for each pair of treatments and will be given by

$$\text{LSD}(\tau_i,\tau_l) = t_{\varsigma,\alpha} \times \sqrt{2s_e^2 / r}.$$

The SAS code for pairwise comparison of treatments with LSD method is given below.

DATA mult_comp;

/*mult_comp is the name of the dataset; it could be any name*/

INPUT treatment yield;

CARDS;

. . . . .Enter Data Here . . . .

;

PROC GLM;

CLASS treatment;

MODEL yield = treatment;

MEANS treatment / LSD LINES;

/*When the data are unbalanced or non-orthogonal, then use LSMEANS instead of MEANS. For balanced data both MEANS and LSMEANS are identical. So it is advisable that the researcher always use LSMEANS*/

LS MEANS treatment / PDIFF LINES;

/*MEANS statement performs multiple comparisons only for main-effects; however, in contrast LSMEANS statement performs multiple comparisons for main effects as well as interaction effects. In other words LSMEANS is powerful than MEANS statement*/

/*However, if data is balanced, one is required to give the value of minimum significant difference, and in that case one may opt for means statement*/

RUN;

The R code for comparing treatments using LSD is as follows.

library(agricolae)

treatment=as.factor(treatment)

lm1=lm(yield~treatment)

anova(lm1)

LSD.test(lm1,"treatment",group=FALSE,console=TRUE)

LSD.test(lm1,"treatment",console=TRUE)

III.2.2 Duncan's Multiple Range Test (DMRT)

This is another multiple comparison test based on Individual Error Rate. It was given by Duncan (1955). Unlike a single value for all pairwise comparisons in case of LSD, DMRT involves computation of a series of values each corresponding to a specific set of pair comparisons depending on the difference in ranks of the treatment effects being compared.

The least significant Range $R_p = \dfrac{r_\alpha(p,edf) * SE_d}{\sqrt{2}}$, where α is the desired significance level, *edf* is the error degrees of freedom and *p = 2, …, v* is one more than the distance in rank between the pairs of the treatment means/effects to be compared.

If the two treatment means have consecutive rankings, then *p = 2* and for the highest and lowest means, it is *v*. The values of $r_\alpha(p,edf)$ can be obtained from Duncan's table of significant ranges.

III.2.3 Bonferroni method

The Bonferroni method is the simplest method that makes corrections for the multiple comparisons. It is important to note that this method does not have any pre-requisite that null hypothesis in the analysis of variance is rejected, *i.e.*, the treatment effects are not homogeneous. This test is used to correct any number of *p*-values for multiple comparisons. To begin with a *p*-value is obtained for every pairwise comparison. These *p*-values are not adjusted for making multiple comparison while these computations are being made. The family wise significance threshold is defined and is set at generally 0.05 or 0.01 (5% or 1% level of significance). The significance threshold defined is divided by the total number of comparisons. If the significance threshold is α (0.05 or 0.01) and the total number of comparisons is *c* then the new threshold value is α /*c* (or 0.05/*c* or 0.01/*c*). For example, if significance threshold is 0.05 and the total number of comparisons is say *c = 21*, then the new threshold value is 0.05/21 = 0.0024. We shall then say that a comparison is significant if its *p*-value is smaller than or equal to α /*c* (or ≤ 0.0024, say). Otherwise, the comparison is not significant. This method has a limitation,

though, that *c* should be small. This method can also be extended straightway to obtaining multiple confidence intervals.

In the sequel is given SAS code for multiple comparisons with Bonferroni method. BON is the SAS code that performs multiple comparison tests between treatments using Bonferroni method and is used with LSMEANS statements as follows.

```
DATA mult_comp;
INPUT treatment yield;
CARDS;
. . . Enter Data Here. . .
;
PROC GLM;
CLASS treatment;
MODEL yield = treatment;
LSMEANS treatment/ PDIFF = ALL ADJUST = BON LINES;
RUN;
```

The R code for comparing treatments using Bonferroni method is given below.

```
library(agricolae)
treatment=as.factor(treatment)
lm1=lm(yield~treatment)
anova(lm1)
LSD.test(lm1,"treatment",group=FALSE, p.adj="bonferroni",console=TRUE)
LSD.test(lm1,"treatment", p.adj="bonferroni",console=TRUE)
```

Although this test avoids any false discovery, yet it is very conservative. Sometimes the difference between treatments is significantly different but may not be detected through the use of this test.

III.2.4 Scheffe's method

Scheffe's method is applicable to the set of estimates of all possible contrasts among treatment effects and is not restricted to just pairwise comparisons of treatment effects. Therefore, unlike Bonferroni method, this method is not restricted to small values of *c*. This method uses the fact that every possible treatment contrast can be written as linear combination of the $v - 1$ "treatments vs control" contrasts like $\tau_2 - \tau_1, \tau_3 - \tau_1, \ldots\ldots, \tau_v - \tau_1$.

Chapter 3 of the book has been devoted to contrast analysis, but to explain Scheffe's method, it may not be out of place to define a contrast again here. An arbitrary contrast of treatment effects

is defined as $B = \sum_{i=1}^{v} p_i \tau_i$, where $\sum_{i=1}^{v} p_i = 0$. A contrast is not unique and there can be infinite

number of choices of set of numbers $P_1, P_2, \cdots, P_i, \cdots, P_v$ such that $\sum_{i=1}^{v} p_i = 0$. Thus, there can be infinite number of contrasts. The contrast B can be estimated as $\hat{B} = \sum_{i=1}^{v} p_i \hat{\tau}_i = \sum_{i=1}^{v} p_i \overline{T}_i$, where $\overline{T}_i$ is the sample mean of the ith treatment, $i = 1, 2 \ldots, v$. The estimated variance of the estimated contrast is $s_{\hat{B}}^2 = s_e^2 \sum_{i=1}^{v} p_i^2 / r_i$, where s_e^2 is the error mean square in the analysis of variance and r_i is the replication number of the ith treatment.

After the data has been generated and the analysis of variance is done, confidence bounds can be obtained for every possible treatment contrast with simultaneous confidence coefficient exactly $100(1-\alpha)\%$. The overall confidence level remains fixed. Thus, a set of overall simultaneous correct $100(1-\alpha)\%$ confidence intervals for all possible treatment contrasts is given by

$$\left(\hat{B} \pm \sqrt{(v-1)F_{\alpha, v-1, n-v}} \; s_{\hat{B}}\right).$$

Here $F_{\alpha,(v-1),(n-v)}$ is the value of Snedecor's F distribution at $(v - 1)$, $(n - v)$ degrees of freedom and α level of significance.

The SAS code for multiple comparisons with Scheffe's method is given in the sequel. SCHEFFE is the SAS code that is used for multiple comparisons test by Scheffe's method and is used with LSMEANS statements as follows:

```
DATA mult_comp;
INPUT treatment yield;
CARDS;
. . . Enter Data Here. . .
;
PROC GLM;
CLASS treatment;
MODEL yield = treatment;
LSMEANS treatment / PDIFF = ALL ADJUST = SCHEFFE LINES;
RUN;
```

The R code for comparing treatments using Scheffe's method is given below.
```
library(agricolae)
treatment=as.factor(treatment)
lm1=lm(yield~treatment)
anova(lm1)
scheffe.test(lm1,"treatment",group=FALSE, console=TRUE)
scheffe.test(lm1,"treatment", console=TRUE)
```

III.2.5 Tukey method

Tukey's test is performed for multiple comparisons and the test also reports the confidence intervals. Tukey's multiple comparison test is also called as Tukey's honest significant difference test or Tukey's HSD. If the researcher wants to compare every treatment effect with every other treatment effect (all the possible pairwise treatment comparisons), then use the Tukey's test. If the interest of the researcher is to compare every treatment effect with a control treatment effect, then use the Dunnett's test. The Tukey's method applies simultaneously to the set of all the pairwise treatment comparisons $(\tau_i - \tau_l),\ i \neq l = 1, 2, \cdots, v$. This method is used for all pair-wise treatment contrasts and is based on distribution statistic Q, given by

$$Q = \frac{Max(T_1, T_2, \ldots, T_v) - Min(T_1, T_2, \ldots, T_v)}{\sqrt{MSE / r}}$$

The distribution of Q is called Studentized range distribution. For CRD and the one way ANOVA, a set of overall simultaneous $100(1-\alpha)\%$ confidence interval for all pairwise treatment contrasts $(\tau_i - \tau_l; i \neq l)$ is given by

$$\left(\hat{B} \pm q_{v, n-v, \alpha} / \sqrt{2} \ \sqrt{MSE\left(\frac{1}{r_i} + \frac{1}{r_l} \right)} \right)$$

For equal replication ($r_i = r, \forall i = 1, 2, \cdots, v$), the overall confidence level is exactly $100(1-\alpha)\%$ and for unequal replications ($r_i \neq r, \forall i$), the overall confidence level is at least $100(1-\alpha)\%$.

The SAS code for multiple comparisons with Tukey's method is given in the sequence. TUKEY is the SAS code that performs multiple comparison tests by Tukey's method and it is used with LSMEANS statement as follows.

```
DATA mult_comp;
INPUT treatment yield;
CARDS;
. . . Enter Data Here. . .
;
PROC GLM;
CLASS treatment;
MODEL yield = treatment;
LSMEANS treatment/ PDIFF = ALL ADJUST = TUKEY LINES;
RUN;
```

The R code for comparing treatments using Tukey's method is given below.

```
library(agricolae)
treatment=as.factor(treatment)
lm1=lm(yield~treatment)
anova(lm1)
HSD.test(lm1,"treatment",group=FALSE, console=TRUE)
HSD.test(lm1,"treatment", console=TRUE)
```

III.2.6 Dunnett's Method

Dunnett's method is a kind of special method that provides a set of simultaneous $100(1 − α)\%$ confidence intervals for preplanned "treatments *vs* control" contrasts of the type $τ_2 − τ_1, τ_3 − τ_1, \dots, τ_v − τ_1$ (in general $τ_i − τ_1 \ \forall i = 2, 3, \cdots, v$), given that $τ_1$ is the effect of control treatment, such that the overall confidence levels is at least $100(1 − α)\%$. This test is used only for treatments *vs* control comparisons and not for all the possible pairwise treatment comparisons. This test provides confidence intervals shorter in length compared to the ones provided by the Bonferroni, Scheffe and Tukey methods.

This test obtains the simultaneous confidence intervals by using the joint distribution of the estimators $\overline{T}_i − \overline{T}_1$ of $τ_i − τ_1$, where $\overline{T}_i$ $(i = 1, 2, \cdots, v − 1)$ is the mean of the *i*th treatment. The joint distribution of $\overline{T}_i − \overline{T}_1$ is a special case of multivariate *t* distribution. The general formula for Dunnett's two-sided overall simultaneous $100(1 − α)\%$ confidence interval for "treatment *vs* control" contrast $τ_i − τ_1$ is given by

$$\left(\hat{d}_{i1} \pm \{t_α (v−1, n−v)\} \sqrt{\hat{V}ar(\hat{d}_{i1})} \right)$$

where $t_α (v−1, n−v)$ is the Dunnett's *t* at $α$ level of significance.

The SAS code for multiple comparisons with Dunnett's method is given in the sequence. DUNNETT is the SAS code that performs multiple comparisons test by Dunnett's method and it is used with LSMEANS statement as follows.

```
DATA mult_comp;
INPUT treatment yield;
CARDS;
. . .Enter Data Here. . .
;
PROC GLM;
CLASS treatment;
MODEL yield = treatment;
MEANS treatment / DUNNETT('t') CLDIFF;
```

/* Here in DUNNETT('t'), treatment t is control treatment against which all the comparisons are made. t could be 1, or 2, or, *v**/

RUN;

The R code for comparing treatments using Dunnett's test is given below.
```
library(multcomp)
treatment=as.factor(treatment)
lm1=lm(yield~treatment)
anova(lm1)
```

test = glht(lm1,linfct=mcp(trt="Dunnett"))
#Here the first treatment is by default the reference treatment
summary(test)

Remark III.1 Hsu's method is used for multiple comparisons where each treatment is compared with the 'best' treatment. This is similar to the multiple comparisons of treatments with a control except that in this case the control treatment is not designated before the actual start of the experiment. The control treatment in this case is the best treatment. Hsu's method makes comparison of a treatment effect with the best of the other $v-1$ treatments, *i.e.*, tests contrasts of the type $\left\{\tau_i - \underset{j \neq i = 1,2,...,v}{max}(\tau_j)\right\}$ for all $i = 1,2,...,v$. Obviously, when the effect of τ_i is the best treatment, then max (τ_j) is the effect of second best treatment in terms of response.

Thus, $\left\{\tau_i - \underset{j \neq i = 1,2,...,v}{max}(\tau_j)\right\}$ will be positive if the effect of τ_i is the best treatment, zero if the effect of τ_i is tied with max (τ_j), the effect of second best treatment in terms of response and negative if the effect of τ_i is worse than the best of max (τ_j). Hsu's method is, therefore, called as RSMCB (Ranking, Selection and Multiple Comparisons with the Best treatment). For equi-replicate C.R.D., Hsu's formula for overall simultaneous $100(1-\alpha)\%$ confidence interval is given

by $\tau_i - \underset{j \neq i}{max} \tau_j \in \left(\bar{T}_i - \underset{j \neq i}{max} \bar{T}_j \pm \{t_\alpha(v-1, n-v)\}\sqrt{MSE(2/r)} \right)$.

If the lower bound comes out to be positive, then it is set to zero and the ith treatment is selected as the best, and if the lower bound comes out to be negative, then it is set to zero and the ith treatment is declared worse treatment. On the other hand if lower responses are favoured instead of higher responses, then the criterion of selecting the best or the worse treatment is just the opposite of what has been mentioned for the high favoured response.

III.3 Example 1

{Nigam, A.K. and Gupta V.K., 1979, Handbook on Analysis of Agricultural experiments, First Edition, I.A.S.R.I. Publication, New Delhi, pp16-20}

A feeding trial with 3 feeds namely (i) Pasture (control), (ii) Pasture and Concentrates and (iii) Pasture, Concentrates and Minerals was conducted at the Yellachihalli Sheep Farm, Mysore, to study their effect on wool yield of Sheep. For this purpose twenty five ewe lambs were allotted at random to each of the three treatments and the three treatments and the weight records of the total wool yield (in gms) of first two clipping were obtained. The data for two lambs for feed 1, three for feed 2 and one for feed 3 are missing. The details of the experiment are given in Table III.1.

Table III.1: Wool yield (in gms)

Feed 1	Feed 2	Feed 3		Feed 1	Feed 2	Feed 3
850.50	510.30	992.25		1020.60	623.70	1077.30
453.60	963.90	850.50		708.75	538.65	850.50
878.85	652.05	1474.20		652.05	737.10	680.40
623.70	1020.60	510.30		623.70	453.60	737.10
510.30	878.85	850.50		396.90	481.95	737.10
765.45	567.00	793.80		822.15	368.55	708.75
680.40	680.40	453.60		680.40	567.00	708.75
595.35	538.65	935.55		652.05	595.35	652.05
538.65	567.00	1190.70		538.65	567.00	567.00
850.50	510.30	481.95		850.50	595.35	453.60
850.50	425.25	623.70		680.40		652.05
793.80	567.00	878.85				567.00

The purpose of the experiment is to know if the three feeds differ statistically so far as the wool yield is concerned. The purpose of experimentation is also to identify the best feed, *i.e.* the treatment giving the highest yield.

III.3.1 Analysis of data

In the sequel are given the commands of SAS for analysis of data.

DATA crd;
INPUT trt woolyield;
CARDS;
1   850.50
1   453.60
1   878.85
1   623.70
1   510.30
1   765.45
1   680.40
1   595.35
1   538.65
1   850.50
1   850.50
1   793.80
1   1020.60
1   708.75
1   652.05

1	623.70
1	396.90
1	822.15
1	680.40
1	652.05
1	538.65
1	850.50
1	680.40
2	510.30
2	963.90
2	652.05
2	1020.60
2	878.85
2	567.00
2	680.40
2	538.65
2	567.00
2	510.30
2	425.25
2	567.00
2	623.70
2	538.65
2	737.10
2	453.60
2	481.95
2	368.55
2	567.00
2	595.35
2	567.00
2	595.35
3	992.25
3	850.50
3	1474.20
3	510.30
3	850.50
3	793.80
3	453.60
3	935.55
3	1190.70
3	481.95
3	623.70
3	878.85
3	1077.30

3	850.50
3	680.40
3	737.10
3	737.10
3	708.75
3	708.75
3	652.05
3	567.00
3	453.60
3	652.05
3	567.00

```
;
PROC GLM;
CLASS trt;
MODEL woolyield = trt;
MEANS trt / LSD Lines ; /*Lines is optional and only helpful in group letters to treatments,
however, the Standrad error of difference beteen treatment means is obtained with harmonic
mean of replications of treatments*/
LSMEANS trt / pdiff lines;
LSMEANS trt / PDIFF ADJUST = BON LINES;
LSMEANS trt / PDIFF ADJUST = SCHEFFE LINES;
LSMEANS trt / PDIFF ADJUST = TUKEY LINES;
MEANS trt / DUNNET ('1') CLDIFF;
/*DUNNET ('1') will make comparisons of feed 2 and feed 3 with feed 1, given that feed 1 is a
control treatment*/
RUN;
```

III.3.2 Output of analysis

It is seen from the analysis that the model used is able to explain only about 11 per cent of the total variability on the data. The CV is also very high (27.84). The treatment differences are significant though (p-value = 0.0258).

Table III.2: Analysis results using SAS commands

ANOVA

Source	DF	SS	MS	F- Value	Prob > F
Model	2	288208.648	144104.324	3.87	0.0258
Error	66	2460243.443	37276.416		
Corrected Total	68	2748452.091			

R-Square	CV	Root MSE	woolyield Mean
0.105	27.839	193.071	693.519

371

The pairwise treatment comparisons are made using different methods. These are given in the sequence.

Fisher's LSD

The pairwise treatment comparisons are made using the Fisher's Least Significant Difference. This method is the most popular for making the pairwise treatment comparisons. The confidence intervals are also obtained for all the pairwise differences between two treatment effects, $(\tau_i - \tau_l)$, $i \neq l = 1,2,3$. The 95 percent confidence coefficient lengths of the confidence intervals are 229.92, 224.96 and 227.56 gms for the difference of the treatment effects of the pairs of treatments (1, 2), (1, 3) and (2, 3), respectively.

Table III.3: Pairwise comparison using LSD

Alpha	0.05
Error Degrees of Freedom	66
Error Mean Square	37276.42
Critical Value of t	1.997

Comparisons significant at the 0.05 level are indicated by ***			
Treatment Comparison	**Difference Between Means**	**95% Confidence Limits**	**Significance**
3 - 1	71.39	−41.09 183.87	
3 - 2	158.38	44.60 272.16	***
1 - 3	−71.39	−183.87 41.09	
1 - 2	86.99	−27.97 201.95	
2 - 3	−158.38	−272.16 −44.60	***
2 - 1	−86.99	−201.95 27.97	

The above output is without using the option Lines in Means Statement. If we make use of Lines statement, the output would be as given in Table III.4.

Note: This test controls the Type I comparison wise error rate, not the experiment wise error rate

Table III.4: Analysis results of multiple comparison using LSD with LINES statement

Alpha	0.05
Error Degrees of Freedom	66
Error Mean Square	37275.5
Critical Value of t	1.997
Least Significant Difference	113.74
Harmonic Mean of Cell Sizes	22.971

Note: Cell sizes are not equal.

t Grouping		Mean	N	Treatment
		Means with the same letter are not significantly different		
	A	767.81	24	3
B	A	696.42	23	1
B		609.53	22	2

It may also be seen that only treatments 2 and 3 are significantly different. Treatments 1 and 2 and treatments 1 and 3 are not significantly different. The output using LSMEANS statement with default option of LSD is given in Table III.5.

Table III.5: Analysis results of multiple comparison using LSD with LSMEANS STATEMENT

i/j	1	2	3
	Least squares means for treatment effects Pr > |t| for H_0: LSMean(i) = LSMean(j)		
1		0.1356	0.2096
2	0.1356		0.0071
3	0.2096	0.0071	

		woolyield LSMEAN	Treatment	LSMEAN Number
		***t* Comparison Lines for Least Squares Means of Treatment Effects**		
		LS-means with the same letter are not significantly different		
	A	767.813	3	3
B	A	696.424	1	1
B		609.434	2	2

Bonferroni method

Bonferroni method produces similar results as obtained using LSD. Treatments 2 and 3 are significantly different while treatments 1 and 2 and 1 and 3 are not significantly different. The probability levels are higher than those obtained using LSD. This is so because the significance levels have been adjusted for multiple comparisons.

Table III.6: Analysis results of multiple comparison using Bonferroni method

i/j	1	2	3
	Least squares means for treatment effects Pr > |t| for H0: LSMean(i) = LSMean(j)		
1		0.4068	0.6287
2	0.4068		0.0213
3	0.6287	0.0213	

Bonferroni Comparison Lines for Least Squares Means of treatment effects				
LS-means with the same letter are not significantly different				
		woolyield LSMEAN	Treatment	LSMEAN Number
	A	767.813	3	3
B	A	696.424	1	1
B		609.434	2	2

Least Squares Means for treatment effects				
i	*j*	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i) - LSMean(j)	
1	2	86.99	−54.45	228.43
1	3	−71.39	−209.78	67.01
2	3	−158.38	−298.37	−18.39

Scheffe's method

Scheffe's method produces similar results as obtained using LSD and Bonferroni method. Treatments 2 and 3 are significantly different while treatments 1 and 2 and 1 and 3 are not significantly different. The probability levels are higher than those obtained using LSD but are almost similar to those obtained using Bonferroni method.

Table III.7: Analysis results of multiple comparison using Scheffe's method

Least Squares Means for treatment effects Pr > |t| for H0: LSMean(i) = LSMean(j)			
i/j	1	2	3
1		0.3256	0.4524
2	0.3256		0.0259
3	0.4524	0.0259	

Scheffe Comparison Lines for Least Squares Means of treatment effects				
LS-means with the same letter are not significantly different				
		woolyield LSMEAN	Treatment	LSMEAN Number
	A	767.813	3	3
B	A	696.424	1	1
B		609.434	2	2

Least squares means for treatment effect				
i	*j*	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	86.90	−57.20	231.18
1	3	−71.39	−212.48	69.70
2	3	−158.38	−301.10	−15.66

Tukey - Kramer method

Tukey-Kramer method, commonly known as Tukey's method is the most popular method for making multiple comparisons. Tukey's method produces similar results as obtained using LSD and Bonferroni methods. Treatments 2 and 3 are significantly different while treatments 1 and 2 and 1 and 3 are not significantly different. The probability levels are higher than those obtained using LSD but are almost similar to those obtained using Bonferroni method.

Table III.8: Analysis results of multiple comparison using Tukey-Kramer method

i/j	1	2	3
	Least squares means for treatment effects Pr > |t| for H0: LSMean(i) = LSMean(j)		
1		0.2925	0.4186
2	0.2925		0.0192
3	0.4186	0.0192	

		woolyield LSMEAN	Treatment	LSMEAN Number
		Tukey-Kramer Comparison Lines for Least Squares Means of treatment effects		
		LS-means with the same letter are not significantly different		
	A	767.813	3	3
B	A	696.424	1	1
B		609.434	2	2

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
			Least squares means for effect trt	
1	2	86.99	−51.06	225.04
1	3	−71.39	−206.46	63.69
2	3	−158.38	−295.01	−21.74

Dunnett's test

Dunnett's test has been used for obtaining a set of simultaneous confidence intervals for a preplanned "treatments *vs* control" comparisons of the type $\tau_2 - \tau_1$ and $\tau_3 - \tau_1$ (in general $\tau_i - \tau_1$ $\forall i = 2,3,\cdots, v$, given that τ_1 is the effect of control treatment). The length of the confidence intervals are 254.69 and 260.90, respectively for the comparisons of the type $\tau_3 - \tau_1$ and $\tau_2 - \tau_1$. However, the corresponding length of the confidence intervals obtained using *LSD* (Student's *t*) are 224.96 and 229.92, respectively. The major difference arises because the critical value of *t* in case of LSD method is 1.997 while the critical value of Dunnett's *t* is 2.260.

Table III.9: Analysis results of multiple comparison using Dunnett's test

Alpha	0.05
Error Degrees of Freedom	66
Error Mean Square	37276.42
Critical Value of Dunnett's *t*	2.26038

Comparisons significant at the 0.05 level are indicated by ***				
Treatment Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		**Significance**
3 - 1	71.39	−55.96	198.73	-
2 - 1	−86.99	−217.14	43.16	-

The comparison of the length of the confidence intervals is also made with those obtained by other methods. The lengths of the confidence intervals of various contrasts are given in Table III.10.

Table III.10: Length of confidence intervals by different multiple procedure methods

Comparison	LSD	Bonferroni	Scheffe	Tukey	Dunnett
|1, 2|	229.92	282.88	288.38	276.10	260.30
|1, 3|	224.96	276.79	282.18	270.15	254.69
|2, 3|	227.56	279.98	285.44	273.27	-

Table III.10 reveals that LSD produces smallest length confidence intervals for the difference of treatment effects, because it has same significance level for each comparison. This method does not adjust the significance levels for multiple comparisons. Among the other methods, Dunnett's method produces the smallest length simultaneous confidence intervals. But this method is restricted to making only treatment *vs* control comparisons, while other methods are applicable to all the pairwise treatment comparisons. In fact these can be applied to any other set of contrasts. If there are v treatments, then the Dunnett's method makes only $v -$ 1 comparisons (treatment *vs* control), whereas the other methods make $v(v - 1)/2$ pairwise treatment comparisons. Thus, if the problem is not of making treatment *vs* control comparisons, then Tukey's method is most popular to be used.

Alternatively, one can also represent the pairwise comparisons of treatments without arranging them in ascending or descending order. Table III.11 gives the pairwise comparison of treatment effects. Any two treatments with at least one letter common are not statistically significant using Tukey's Honest Significant Difference. This Table once again reveals that treatments 2 and 3 are significantly different while treatments 1 and 2 and treatments 1 and 3 are not significantly different.

<div align="center">**Table III.11: Treatment comparison using Tukey's method**</div>

Treatment	Treatment Mean of 'woolyield'	Rank of Treatment
1	696.42[A,B]	2
2	609.43[B]	3
3	767.81[A]	1
General Mean	693.52	.
p-Value		.
CV(%)	27.84	.
SE(*d*)	56.969	.
Tukey HSD at 5%	136.59	.

In the sequel we shall present an example, wherein null hypothesis regarding equality of treatment effects through ANOVA is not rejected and if one use the multiple comparison tests based on individual error rate (such as LSD or Duncan's Multiple Comparison test), one may be able to see that some treatment pairs are significantly different, which should not be interpreted so, as this leads to false discovery. However, if one makes use of multiple comparison procedure that controls family error rate (such as Bonferroni correction or Tukey's HSD), the treatment pairs would be found to be statistically at par.

III.3.3 Analysis using R code

The analysis of data in Example 1 can be performed in R using the code given below.

```
d36=read.table("woolyield.txt",header=TRUE)
attach(d36)
names(d36)
trt=as.factor(trt)
lm1=lm(woolyield~trt)
anova(lm1)
LSD.test(lm1,"trt",group=FALSE,console=TRUE)
#Bonferroni method
LSD.test(lm1,"trt",group=FALSE,p.adj="bonferroni",console=TRUE)
#Scheffe method
scheffe.test(lm1,"trt",group=FALSE,console=TRUE)
#Tukey's method
HSD.test(lm1,"trt",group=FALSE,console=TRUE)
#Dunnett's test
library(multcomp)
test <- glht(lm1,linfct=mcp(trt="Dunnett"))
summary(test)
detach(d36)
```

III.4 Example 2

An experiment was conducted during 2004 at IARI, New Delhi to study the bio-efficacy of controlled release formulations of Carbofuran against the rice-leaf-folder (*Cnaphalocrocis medinalis*) on rice cultivar 'Pusa Sugandh 3' susceptible to leaf folder. The experiment was conducted using a randomized complete block designs in 3 replications for comparing the 11 treatments which are: Carbofuran 3G (10g); Sodium Carboxy Methyl Cellulose- Kaolinite-3 (10g); Sodium Carboxy Methyl Cellulose- Kaolinite-3 (5g); Sodium Carboxy Methyl Cellulose-Kaolinite-3 (2.5g);   Sodium Carboxy Methyl Cellulose-3 (10g); Sodium Carboxy Methyl Cellulose- 3 (5g); Sodium Carboxy Methyl Cellulose-3 (2g); Polyvinyl Chloride-3 (10g); Polyvinyl Chloride- 3(5g); Polyvinyl Chloride-3 (2.5g); Control. The plot size was 2m × 2m. The rows and plants were spaced 20cm and 15cm apart, respectively. All the formulations were broadcast in standing crop 20 days after transplanting. Percent leaf folder damage, 27 days after treatment, was recorded by counting the number of damaged leaves per hill, before and after treatment, from 10 randomly selected hills in each plot. Observations were similarly recoded at 42, 57 and 72 days after treatment. The leaf folder damage (%) was calculated as ratio of damaged leaves per hill to total leaves per hill and is given as Leaf folder damage (%)

$\dfrac{\text{Damaged leaves/10 hills}}{\text{Total leaves/10 hills}}$ 100 . The data obtained on leaf folder damage (%) before pre-treatment

is given in Table III.12.

Table III.12: Leaf folder damage (%)

Treatment	block	Pre-treatment	Treatment	block	Pre-treatment
T1	1	0	T7	2	0.4
T2	1	0	T8	2	0
T3	1	0	T9	2	0
T4	1	0	T10	2	0.23
T5	1	0	T11	2	0.83
T6	1	0	T1	3	0
T7	1	0	T2	3	0
T8	1	0	T3	3	0
T9	1	0.9	T4	3	0
T10	1	0	T5	3	2.9
T11	1	0.83	T6	3	0
T1	2	0	T7	3	0.8
T2	2	0.21	T8	3	0.93
T3	2	0	T9	3	0.43
T4	2	0	T10	3	0.49
T5	2	0.4	T11	3	0.68
T6	2	0			

The objective of the experiment is to analyse leaf folder damage (%) before pre-treatment. Eventually, the objective is to identify the best treatment that gives minimum leaf folder damage or maximum grain yield using appropriate multiple comparison procedure.

III.4.1 Analysis of data

Since the data is in percentages based on numbers, an angular transformation was performed on the data and SAS steps are given in the sequel.

DATA leaf;

INPUT trt $ rep st;

x2=ARSIN(SQRT(st/100))*180*7/22;

x3=100-1/400;

IF st=0 THEN x2=ARSIN(SQRT(1/400))*180*7/22;

IF st=100 THEN x2= ARSIN(SQRT(x3/100))*180*7/22;

CARDS;

trt	rep	st
T1	1	0
T2	1	0
T3	1	0
T4	1	0
T5	1	0
T6	1	0
T7	1	0
T8	1	0
T9	1	0.9
T10	1	0
T11	1	0.83
T1	2	0
T2	2	0.21
T3	2	0
T4	2	0
T5	2	0.4
T6	2	0
T7	2	0.4
T8	2	0
T9	2	0
T10	2	0.23
T11	2	0.83
T1	3	0

T2	3	0
T3	3	0
T4	3	0
T5	3	2.9
T6	3	0
T7	3	0.8
T8	3	0.93
T9	3	0.43
T10	3	0.49
T11	3	0.68

```
;
PROC GLM;
CLASS rep trt;
MODEL x2  = rep trt;
MEANS trt/LSD;
MEANS trt/DUNCAN;
MEANS trt/BON;
MEANS trt/TUKEY;
RUN;
```

Table III.13: Analysis results using SAS commands

Source	DF	SS	MS	F Value	Prob > F
Model	12	34.369	2.864	1.75	0.1302
Error	20	32.782	1.639		
Corrected Total	32	67.150			

R-Square	CV	Root MSE	x2 Mean
0.512	35.571	1.280	3.599

Source	DF	Type III SS	MS	F Value	Prob > F
Replication	2	8.165	4.082	2.49	0.1082
Treatment	10	26.204	2.620	1.60	0.1783

From Table III.13, one can easily see that the *p*-value for testing the hypothesis for equality of treatment effects is 0.1783 and at 5% level of significance null hypothesis regarding equality of treatment effects is not rejected. The results from different multiple comparison procedures are given in the sequel.

Fisher's Least Significant Difference

Table III.14: Analysis results of multiple comparison using LSD

Note: This test controls the Type I comparison wise error rate, not the experiment wise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	1.639
Critical Value of t	2.086
Least Significant Difference	2.181

Means with the same letter are not significantly different.					
t Grouping			**Mean**	**N**	**Treatment**
	A		5.430	3	T5
B	A		5.059	3	T11
B	A	C	4.022	3	T9
B	A	C	3.873	3	T7
B	A	C	3.754	3	T8
B		C	3.208	3	T10
		C	2.865	3	T6
		C	2.865	3	T4
		C	2.865	3	T3
		C	2.865	3	T1
		C	2.785	3	T2

Duncan's Multiple Comparison Test

Table III.15: Analysis results of multiple comparison using DMRT

Note: This test controls the Type I comparison wise error rate, not the experiment wise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	1.639

p (difference between ranks of treatments+1)	2	3	4	5	6	7	8	9	10	11
Critical Range	2.181	2.289	2.358	2.406	2.441	2.468	2.489	2.506	2.520	2.530

Means with the same letter are not significantly different				
Duncan Grouping		Mean	N	Treatment
	A	5.430	3	T5
B	A	5.059	3	T11
B	A	4.022	3	T9
B	A	3.873	3	T7
B	A	3.754	3	T8
B	A	3.208	3	T10
B		2.865	3	T6
B		2.865	3	T4
B		2.865	3	T3
B		2.865	3	T1
B		2.785	3	T2

Bonferroni Method

Table III.16: Analysis results of multiple comparison using Bonferroni method

Note: This test controls the Type I experiment wise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	1.639
Critical Value of t	3.890
Minimum Significant Difference	4.067

Means with the same letter are not significantly different			
Bon Grouping	Mean	N	**Treatment**
A	5.430	3	T5
A	5.059	3	T11
A	4.022	3	T9
A	3.873	3	T7
A	3.754	3	T8
A	3.208	3	T10
A	2.865	3	T6
A	2.865	3	T4
A	2.865	3	T3
A	2.865	3	T1
A	2.785	3	T2

Tukey's HSD

Table III.17: Analysis results of multiple comparison using Tukey's method

Note: This test controls the Type I experiment wise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	1.639
Critical Value of Studentized Range	5.108
Minimum Significant Difference	3.776

Means with the same letter are not significantly different			
Tukey Grouping	Mean	N	**trt**
A	5.430	3	T5
A	5.059	3	T11
A	4.022	3	T9
A	3.873	3	T7
A	3.754	3	T8
A	3.208	3	T10
A	2.865	3	T6
A	2.865	3	T4
A	2.865	3	T3
A	2.865	3	T1
A	2.785	3	T2

III.4.2 Analysis using R

The data in Example 2 can be analysed using R code given below. We do not produce the output for the sake of duplication.

```
d37=read.table("leaf.txt",header=TRUE)
attach(d37)
names(d37)
x2=asin(sqrt(st/100))*180*7/22
x3=100-1/400
for (i in 1:length(st))
{
if(st[i]==0) x2[i]=asin(sqrt(1/400))*180*7/22
if(st[i]==100) x2[i]=asin(sqrt(x3[i]/100))*180*7/22
}
trt=factor(trt)
rep=factor(rep)
```

```
lm1=lm(x2~rep+trt)
library(agricole)
anova(lm1)
LSD.test(lm1,"trt",console=TRUE)
#Duncan test
duncan.test(lm1,"trt",console=TRUE)
#Bonferroni method
LSD.test(lm1,"trt",p.adj="bonferroni",console=TRUE)
#Scheffe method
scheffe.test(lm1,"trt",console=TRUE)
#Tukey's method
HSD.test(lm1,"trt",console=TRUE)
#Dunnett's test
library(multcomp)
test <- glht(lm1,linfct=mcp(trt="Dunnett"))
summary(test)
detach(d37)
```

III.5 Conclusions

The purpose of this Appendix has been to describe various methods of treatment comparisons particularly when the researcher has to make multiple comparisons. In this case, there is a need to adjust the significance level. One can easily see that all the tests for multiple comparisons involve a minimum significant difference which is product of a table value (chosen based on level of significance and test statistic used) and the estimated standard error of the estimated difference of treatment effects. Contrast analysis is an easy way of making any type of comparison. The researcher should be able to convert the problem of treatment comparison to that of a contrast analysis and then the analysis could be done easily. LS MEANS and PDIFF make all the possible pairwise treatment comparisons using Student's *t* statistic by default. Other methods of multiple comparisons described in this Appendix can also be used. One may also do the analysis online by visiting Design Resources Server at www.iasri.res.in/sscnars/ and then using IP Authenticated Indian NARS Statistical Computing portal (for Indian National Agricultural Research System users). It is user friendly and menu driven. It uses 5 per cent and 1 percent level of significance. Various methods of treatment comparisons are listed and the researcher can use one depending upon the requirement. It is suggested that if null hypothesis regarding equality of treatment effects through ANOVA is not rejected then one should not make use of multiple comparisons tests based on individual error rate. In those cases, it is better to make use of multiple comparisons procedures based on family error rate. Generally, Tukey's HSD method is the most popular method.

Design Resources Server

IV.1 Introduction

Design Resources Server is a web resource which has been created as an International Public Good to disseminate research in Design of Experiments among the scientists of NARES in particular and researchers all over the globe in general. The web resource is hosted at home page of IASRI so as to make available design theory and actual randomized layout of the designs through web. The URL of the resource is www.iasri.res.in/design. This web resource is open to everyone all over the globe and anyone can join the resource and add information to the site to strengthen it further with the permission of the developers. The goal of this web resource is to help experimenters in agricultural, biological and social sciences, industry, etc. in planning and designing their experiments and then analyzing the data generated. It also targets at spreading advances in theoretical, analytical and application of design of experiments among mathematicians and statisticians both in academia and also involved in advisory and consultancy services. This is also a very useful resource for faculty and students to disseminate and learn theory and application of design of experiments.

The server is matter-of-factly mobile library on Design of Experiments. It is dynamic in nature and new additions are posted on it from time to time. Ultimate objective of this server is to provide E-advisory services and become an E-learning platform. The contents of the server are divided into four broad categories *viz.* (a) Useful for Experimenters, (b) Useful for Statisticians, (c) Other Useful Links, (d) Site Information. We discuss the resources available in the subsequent sections.

IV.2 Resources for experimenters

The design resources server contains electronic Books and online generation facility of a number of designs, analysis steps using SAS, SPSS and Excel and a section on statistical genomics.

IV.2.1 Electronic books

There are two electronic books namely "Design and Analysis of Agricultural Experiments" and "Advances in Data Analytical Techniques", which are available on the server and they act as basic learning material for understanding the fundamental concepts of design and analysis of experiments. The electronic book "Design and Analysis of Agricultural Experiments" has five modules namely computer usage, designs for one- and two- way elimination of heterogeneity, block designs with nested factors, factorial experiments and other important considerations.

The electronic book "Advances in Data Analytical Techniques" has six modules, which deal with computer usage and statistical software packages, basic statistical techniques, diagnostics and remedial measures, application of multivariate techniques, modeling and forecasting techniques in agriculture and other useful techniques.

IV.2.2 Online design generation-I (facilities for experimenters)

The server provides facilities for online generation of a number of designs which include randomized layout of designs including completely randomized designs, randomized complete block designs both for single and multi-factor experiments, Latin Square designs for single factor experiments, augmented designs, alpha designs and square lattice designs.

IV.2.3 Online analysis of data

Under online analysis of data, the server provides analysis of data from augmented designs. There is step by step detailed guidance on analysis of data which includes descriptive statistics, test of significance, correlation, regression and data analysis from designed experiments using statistical software SAS and SPSS. An important feature of this link is that all the data used are from real experiments and can be downloaded.

IV.2.4 Statistical genomics

The server has a small section on statistical genomics which contains designs for microarray experiments and useful material on QTL mapping. It provides a facility for online generation of row-column designs for microarray experiments and for generation of block designs with baseline parameterization.

IV.3 Resources for statisticians

The server contains literature on a number of topics of experimental design and catalogues of a number of designs which may be very useful for statisticians engaged in research in the design of experiments. The various classes of designs which are available are described below.

IV.3.1 Block designs

Three broad classes of designs are covered in the server, *viz.*, (i) binary balanced block designs, (ii) block designs for test *vs* control comparisons, and (iii) efficient incomplete block designs. The server contains a catalogue for binary balanced designs, balanced incomplete block designs, efficient incomplete block designs and balanced treatment incomplete block designs. Layouts of the designs are given for parametric combinations of v (number of treatments), b (number of blocks) and k (block size).

IV.3.2 Designs for bioassays

The server contains a very useful resource for conducting bioassays. Both the two types of bioassays, namely parallel line assays and slope ratio assays, are covered and currently there is a catalogue on parallel line assays are available. The catalogue gives a number of designs for various parametric combinations.

IV.3.3 Designs for factorial experiments

Factorial experiments are very common in various scientific disciplines including agricultural sciences. Out of the various classes of designs used for factorial experiments, the server provides a catalogue of orthogonal arrays, their online generation along with a detailed bibliography on such arrays. A very useful class of designs called block designs with factorial structures are discussed in detailed including their method of construction, a catalogue of such designs and bibliography. A class of designs supersaturated designs have recently received a lot of attention because of their economical run size for screening experiments. The server contains catalogue of two-level, multi-level, balanced and unbalanced mixed-level, extended two-level, extended multi-level, k-circulant multi-level and k-circulant mixed-level supersaturated designs. It also has a bibliography on supersaturated designs. Another useful content is row-column designs for factorial experiments based on orthogonal parameterization in two rows.

IV.3.4 Response surface designs

Response surface designs are particularly useful in process and product experimentations. On response surface designs, the Design Resources Server has a very elaborate content which includes response surface methodology, response surface designs, robustness against one missing observation, response surface designs for quantitative and qualitative factors, a catalogue of designs along with design layout. It also contains references and extensive bibilography.

IV.3.5 Mixture experiments

Mixture experiments are often used in a number of experiments. Some real life situations have been described where experiment with mixtures methodology is applicable. The server has facilities for online generation of simplex centroid designs and simplex lattice designs for mixture experiments and also provides a bibliography on experiments with mixtures.

IV.3.6 Online design generation-II (facility for statisticians)

It is well known that Hadamard matrices, mutually orthogonal Latin squares and orthogonal arrays are very useful combinatorial structures for constructing a number of other classes of designs. The server provides facilities for online generation of each of these combinatorial structures. Resolvable mixed orthogonal arrays provide fractional factorials with blocking for asymmetrical factorial experiments. These plans allow orthogonal estimation of all the main effects and the general mean. Recently, balanced incomplete Latin square designs are introduced in literature. The server is also has provision for generation of such designs.

IV.4 Other useful links

The server has some other useful links which include a discussion board, ask a question facility, an archive of questions, a collection of who-is-where in the field of design of experiments, some important links and a list of books on design of experiments.

"Ask a Question" is an important feature of the server through which any user can ask a question on design and / or analysis of data or seek any other clarification. E-mails automatically

go to developers of the site who in turn answer the query. On an average two questions are asked every week. The server also provides an archive of questions asked and answers given for the benefit of stakeholders. The "Discussion Board" is for sharing any useful piece of research or idea with any other scientist over the globe. Under who-is-where, the server provides a list of experts in design of experiments all over the globe. The list includes 84 experts from USA, UK, Mexico, Syria, Canada, China, Australia, Japan, New Zealand, Oman, Taiwan, India and Vietnam.

One can also provide feedback and suggestion to the developers using the "Feedback" link available on the server. Important links provide users a platform to connect to some very useful resources for the stake holders. Some important ones are Design resources, GENDEX, free encyclopedia on design of experiments, access to some journals, some software, etc. Books on design of experiments is a useful link for faculty and students.

IV.5 Concluding remarks

Design Resources Server is a virtual library on Design of Experiments in particular and Statistics in general. The server is dynamic in nature and new links on various topics are added to it regularly. It is very useful for the agricultural researchers and statisticians across NARES for e-learning and e-advisory. The server is a copyright of IASRI (ICAR): L46452/2013. This web resource has become very popular with scientists of NARES and is also being used in the 6 continents throughout the globe.

The server is updated on a regular basis with addition of new contents on design of experiments. The developers also wish to add contents on field book of all the designs generated, labels generation for preparing seed packets, online analysis of data, steps for analyzing data using GENSTAT, R, SYSTAT, etc.

Several other web resources on designed experiments developed and hosted at www.iasri.res.in are (i) web generation of experimental designs balanced for indirect effects of treatments, (ii) web based generation and analysis of partial diallel crosses, (iii) general block design analysis, (iv) analysis of row column designs, besides several other resources.

Bibliography

Agrawal, R. C., and Sapra, R. L. (1995). Augment 1: A microcomputer based program to analyze augmented randomized complete block design. *Indian Journal of Plant Genetic Resources*, **8**, 61-69.

Anderson, V. L. and McLean, R. A. (1974). *Design of Experiments: A Realistic Approach.* Marcel Dekker, Inc., New York.

Andrews, D. F., and Pregibon, D. (1978). Finding the outliers that matter. *Journal of the Royal Statistical Society. Series B (Methodological)*, 85-93.

Atiqullah, M. and Cox, D. R. (1962). The use of control observations as an alternative to incomplete block designs. *Journal of the Royal Statistical Society*, Series B (Methodological), **24**,464-471.

Atkinson, A. C. (1985). *Plots, Transformation and Regression.* Oxford University Press.

Atkinson, A. C. and Donev, A. N. (1992). *Optimal Experimental Designs.* Oxford University Press.

Atkinson, A. C. and Donev, A. N. (1996). Experimental designs optimally balanced against trend. *Technometrics*, **38**, 333-341.

Bailey, R. A. and Rowley, C. A. (1987). Valid randomization. *Proceedings of Royal Society, London*, A, **410**, 105-124.

Bartlett, M. S. (1938). The approximate recovery of information from field experiments with large blocks. *Journal of Agricultural Sciences*, **28**, 418-427.

Bartlett, M. S.(1947). The use of transformation. *Biometrics*, **3**, 39-52.

Beckman, R. J., and Cook, R. D. (1983). Outlier………. s. *Technometrics*, **25**, 119-149.

Bhar, L., and Gupta, V. K. (2001). A useful statistic for studying outliers in experimental designs. *Sankhyā: The Indian Journal of Statistics, Series B* **63**, 338-350.

Bhar, L., Parsad, R. and Gupta, V.K. (2008). Outliers in Designed Experiments. IASRI, New Delhi. I.A.S.R.I./ P.R.-05/2008

Bose, M., and Dey, A. (2009). *Optimal Crossover Designs*. World Scientific.

Bose, R. C. (1938). On the application of Galois fields to the problem of the construction of Hyper-Graeco Latin squares. *Sankhyā*, **3**, 323-338.

Bose, R.C. (1942). A note on resolvability of balanced incomplete block designs. *Sankhya*,**6,** 105-110.

Bose, R. C. and Bush, K. A. (1952). Orthogonal arrays of strength two and three. *Annals of Mathematical Statistics*, **23**, 508-524.

Bose, R. C., Shrikhande, S. S. and Parker, E. T. (1960). Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture. *Canadian Journal of Mathematics*, **12**, 189-203.

Box, G. E. P. (1953). Non-normality and tests on variances. *Biometrika*, **40**, 318-335.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, *Series B* **26**, 211–243.

Brien, C. J. and Payne, R. W. (1999). Tiers, structure formulae and the analysis of complicated experiments. *Journal of the Royal Statistical Society*, *series D* **48**, 41-52.

Brown, M. B. and Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, **69**, 364-367.

Chao, S. C. and Shao, J. (1997). Statistical methods for two-sequence three-period cross-over trials with incomplete data. *Statistics in Medicine*, **16**, 1071-1039.

Clatworthy, W. H. (1973). Tables of two-associate-class partially balanced designs, *National Bureau of Standards, Applied Mathematics Series* **63**.

Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs*, Second Edition. Wiley, New York.

Conover, W. J. and Iman, R. L. (1976). In Some Alternative Procedure Using Ranks for the Analysis of Experimental Designs. *Comm. Statist.: Theory & Methods,* **A5**, 1349-1368.

Conover, W. J. and Iman, R. L.(1981). Rank transformations as a bridge between parametric and non-parametric statistics (with discussion). *American Statistician*, **35**, 124-133.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, **19**, 15-18.

Cook, R. D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association*, **74**, 169-174.

Cook, R. D. and Weisberg, S. (1982). *Residuals and Inference in Regression.* London: Chapman and Hall.

Cox, D. R. (1951). Some systematic experimental designs. *Biometrika*, **38**, 310-315.

Cox, D. R. (1954). The design of an experiment in which certain treatment arrangements are inadmissible. *Biometrika*, **41**, 287-295.

Cox, D. R. (1957). The use of a concomitant variable in selecting an experimental design. *Biometrika*, **44**, 150-158.

Cox, D. R. (1958). *Planning of Experiments.* New York: Wiley.

Cox, D. R. (1982). Randomization and concomitant variables in the design of experiments. In *Statistics and Probability*: Essays in honour of C.R. Rao. Editors G. Kallianpur, P.R. Krishnaiah and J.K. Ghosh. Amsterdam: North Holland, pp. 197-202.

Cox, D. R. and McCullagh, P. (1982). Some aspects of analysis of covariance (with discussion). *Biometrics*, **38**, 541-561.

D'Agostino, R.B. and Stephens, M. A.(1986). *Goodness-of-fit Techniques*. Marcel Dekkar, Inc., New York.

Daniel, C. (1960). Locating outliers in factorial experiments. *Technometrics*, **2**(**2**), 149-156.

Daniel, C. (1994). Factorial one-factor-at-a-time experiments. *American Statistician*, **48**, 132-135.

Das, M. N. and Kulkarni, G. A. (1966). Incomplete block designs for bio-assays. *Biometrics,* **22**, 706–729.

Das, M. N. and Giri, N. C. (1986): *Design and Analysis of Experiments*, 2nd edition, New Age International (P) Limited Publishers.

Davies, O. L., editor (1963). *Design and Analysis of Industrial Experiments.* Second Edition, Oliver and Boyd, London.

Dean, A., and Voss, D. (1999). *Design and Analysis of Experiments.* Springer-Verlag, New York.

Denʾes, J. and Keedwell, A. D. (1974). *Latin Squares and Their Applications.* English Universities Press, London.

Design Resources Server (2005). Indian Agricultural Statistics Research Institute (ICAR), New Delhi 110 012, India. www.iasri.res.in/design.

Deshpande, J. V., Gore, A. P. and Shanubhogue, A. (1995). *Statistical Analysis of Non-normal Data.* Wiley Eastern Limited, New Delhi.

Dey, A. (1986). *Theory of Block Designs.* John Wiley.

Dey, A. (2010). *Incomplete Block Designs* (Vol. 57). World Scientific, Hackensack NJ.

Dolby, J. L.(1963). A quick method for choosing a transformation. *Technometrics*, **5**, 317-326.

Draper, N. R. and Hunter, W. G.(1969). Transformations: Some examples revisited. *Technometrics*, **11**, 23-40.

Draper, N. and Smith, H. (1998). *Applied Regression Analysis.* 3rd edition. Wiley, New York.

Duncan, D. B. (1955). Multiple range and multiple F tests. *Biometrics,* 11(1), 1-42.

Dunnett, C. W. (1955). A multiple comparisons procedure for comparing several treatments with a control. *Journal of the American Statistical Association* **50**, 1096–1121.

Dunnett, C. W. (1980). Pairwise multiple comparisons in the unequal variance case. *Journal of the American Statistical Association* **75**, 796–800.

Durbin, J. (1951). Incomplete blocks in ranking experiments. *British Journal of Psychology,* **4**, 85–90.

Federer, W. T. (1956). Augmented (or hoonuiaku) designs. *Hawaiian Planters' Record*, **55**, 191-208.

Federer, W. T. (1961). Augmented designs with one-way elimination of heterogeneity. *Biometrics*, **17**, 447-473.

Federer, W. T. (1963). Procedures and designs useful for screening material in selection and allocation, with a bibliography. *Biometrics*, **19**, 553-87.

Federer, W. T., and Raghavarao, D. (1975). On augmented designs. *Biometrics*, **31**, 29-35.

Federer, W. T., Nair, R. C., and Raghavarao, D. (1975). Some augmented row-column designs. *Biometrics*, **6**, 361-373.

Fisher, R. A. (1925). *Statistical Methods for Research Workers.* Genesis Publishing Pvt Ltd.

Fisher, R. A. and Yates, F. (1973). *Statistical Tables for Biological, Agricultural and Medical Research*, Oliver and Boyd, Edinburgh.

Fisher, R. A. (1935). *Design of Experiments.* Oliver and Boyd, Edinburgh.

Fisher, R. A. and Mackenzie, W. A. (1923). Studies in crop variation II. The manurial response of different potato varieties. *Journal of Agricultural Sciences*, **13**, 311-320.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, **32**, 675-701.

Gentleman, J. F., and Wilk, M. B. (1975). Detecting outliers in a two-way table: I. Statistical behavior of residuals. *Technometrics*, **17**, 1-14.

Gentleman, J. F., and Wilk, M. B. (1975). Detecting outliers. II. Supplementing the direct analysis of residuals. *Biometrics*, 33, 387-410.

Gilmour, A.R., Cullis, B. R. and Verbyla, A. P. (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural Biological and Environmental Statistics*, **2**, 269-273.

Gomez, K. A., and Gomez, A. A. (1984). *Statistical Procedures for Agricultural Research*. John Wiley and Sons.

Graybill, F. A. (1976). *Theory and Application of the Linear Model*. Duxbury Press.

Grundy, P. M. and Healy, M. J. R. (1950). Restricted randomization and quasi-Latin squares. *Journal of Royal Statistical Society*, *series B* **12**, 286-291.

Harshabarger, B. (1949). Triple rectangular lattices. *Biometrics*, **5**, 1-13.

Hartley, H. O. and Smith, C. A. B. (1948). The construction of Youden squares. *Journal of Royal Statistical Society*, s*eries B* **10**, 262-263.

Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications.* New York: Springer.

Hicks, C. R. (1956). Fundamentals of Analysis of Variance, Part III—Nested designs in analysis of variance. *Industrial Quality Control*, **13**, part 4, 13–16.

Hicks, C. R. (1965). The analysis of covariance. *Industrial Quality Control*, **22**, 282–286.

Hicks, C. R. (1993). *Fundamental Concepts in the Design of Experiments*, Fourth Edition, Oxford University Press Inc., New York.

Hochberg, Y. and Tamhane, A. C. (1987). *Multiple Comparison Procedures*. John Wiley and Sons, New York.

Hocking, R. R. (1996). *Methods and Applications of Linear Models; Regression and the Analysis of Variance*. John Wiley and Sons, New York.

Hollander, M. and Wolfe, D. A. (1973). *Non-parametric Statistical Methods*. John Wiley and Sons, New York.

Hollander, M. and Wolfe, D.A. (1999). *Nonparametric Statistical Methods*, 2nd Ed, John Wiley, New York.

Hsu, J. C. (1984). Ranking and selection and multiple comparisons with the best. In *Design of Experiments: Ranking and Selection (Essays in Honor of Robert E. Bechhofer)*. Editors: T. J. Santner and A. C. Tamhane). 23–33, Marcel Dekker, New York.

Huber, P. J. (1964). Robust Estimation of a Location Parameter. *Annals of Mathematical Statistics*, **35**, 73-101.

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological monographs*, **54**, 187-211.

J. A. (1987). *Cyclic Designs*. Chapman and Hall, London.

John, J. A. and Turner, G. (1977). Some new group divisible designs. *Journal of Statistical Planning and Inference*, **1**, 103–107.

John, J. A., Wolock, F. W., and David, H. A. (1972). Cyclic Designs, *National Bureau of Standards, Applied Mathematics Series*, **62**.

John, J. A. and Quenouille, M. H. (1977). *Experiments: Design and Analysis.* Griffin, London.

John, J. A. and Williams, E. R. (1995). *Cyclic and Computer Generated Designs.* 2nd edition. Chapman and Hall, London.

John, J. A., Russell, K. G., Williams, E. R. and Whitaker, D. (1999). Resolvable designs with unequal block sizes. *Australian and New Zeland Journal of Statistics*, **41**, 111-116.

John, P. W. M. (1961). An application of a balanced incomplete design, *Technometrics*, **3**, 51–54.

John, P. W. M. (1971). *Statistical Design and Analysis of Experiments*, Macmillan, New York.

John, P. W. M. (1980). *Incomplete Block Designs*, Lecture Notes in Statistics, Volume I, Marcel Dekker, New York.

John, P. W. M. (1971). *Statistical Design and Analysis of Experiments.* New York: Macmillan.

Johnson, R. A. and Wichern, D. W. (1988). *Applied Multivariate Statistical Analysis*, 2nd Edition. Prentice-Hall International, Inc., London

Jones, B. and Kenward, M. G. (1989). *Design and Analysis of Crossover Trials.* Chapman and Hall, London.

Kempthorne, O. (1952). *Design of Experiments.* Wiley, New York.

Jones, B. and Kenward, M. G. (2014). *Design and Analysis of Cross-over Trials*. CRC Press.

Kempthorne, O. (1976). *Design and Analysis of Experiments.* John Wiley and Sons, New York.

Kempthorne, O. (1977). Why Randomize? *Journal of Statistical Planning and Inference,* **1**, 1-25.

Kempthorne, O. (1994). *Design and Analysis of Experiments.* New York: Wiley.

Kiefer, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. *Annals of Mathematical Statistics*, **29**, 675- 699.

Kiefer, J. (1959). Optimum experimental design (with discussion). *Journal of the Royal Statistical Society*, *series B* **21**, 272-319.

Kiefer, J. (1975). Optimal design: variation in structure and performance under change of criterion. *Biometrika*, **62**, 277-288.

Kleczkouski, A. (1960). Interpreting relationship between the concentrations of plants viruses and number of local lesions. *Journal of General Microbiology*, **4**, 53-69.

Kuehl, R. O. (1994). *Statistical Principles of Research Design and Analysis.* Duxbury Press, Belmont, California.

Levene, H. (1960). Robust Tests for Equality of Variances. **I** Olkin, ed., *Contributions to Probability and Statistics*, Palo Alto, Calif.: Stanford University Press, 278-292.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models.* 2nd edition. Chapman and Hall, London.

Montgomery, D. C. (1997). *Design and Analysis of Experiments.* 4th edition. Wiley, New York.

Nair, K. R. (1951). Rectangular lattices and partially balanced incomplete block designs. *Biometrics*, **7**, 145-154.

Nandi, P. K. (2007). *Design and analysis for multi-response experiments.* Unpublished Ph. D. thesis, Indian Agricultural Research Institute, New Delhi.

Nelder, J. A. (1965a). The analysis of experiments with orthogonal block structure. I Block structure and the null analysis of variance. *Proceedings of Royal Society of London*, series A **283**, 147-162.

Nelder, J. A. (1965b). The analysis of experiments with orthogonal block structure. II Treatment structure and the general analysis of variance. *Proceedings of Royal Society of London*, **A 283**, 163-178.

Neter, J., Kutner, M. H., Nachtsheim, C. J. and Wasserman, W. (1996). *Applied Linear Statistical Models*, Fourth Edition, Richard D. Irwin, Inc., Chicago.

Nigam, A. K., Puri, P. D., and Gupta, V. K. (1988). *Characterizations and Analysis of Block Designs*. John Wiley and Sons.

Nigam, A. K., Parsad, R., and Gupta, V. K. (2003). *Design and analysis of on-station and on-farm agricultural experiments: a revisit.* Workshop proceedings, IASRI, New Delhi.

Parsad, R. and Gupta, V. K. (2000). A note on augmented designs. *Indian Journal of Plant Genetic Resources*, **13**, 53-58.

Parsad, R., Gupta, V. K. and Prasad, N. S. G. (2003). Structurally incomplete row-column designs. *Communications in statistics. Theory and methods*, **32**, 239-261.

Parsad, R., Gupta, V. K., Batra, P.K., Srivastava, R., Kaur, R., Kaur, A. and Arya, P. (2004). A diagnostic study of design and analysis of field experiments. Project Report, IASRI, New Delhi. I.A.S.R.I./P.R.-01/2004

Parsad, R., Gupta, V. K., Batra, P. K., Satpati, S. K. and Biswas, P. (2007a). *α-Designs*. IASRI, New Delhi.

Parsad, R., Gupta, V. K. and Srivastava, R. (2007b). Designs for cropping systems research. *Journal of Statistical Planning and Inference*, **137**, 1687-1703.

Parsad, R., Mandal, B. N., Dhandapani, A., Dash, S, Varghese, E. and Mishra, D. C. (2015). R-Scripts For Statistical Analyses using R-Studio. http://www.iasri.res.in/sscnars/content_rmanual.htm

Patterson, H. D., and Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika*, **63**, 83-92.

Pearce, S. C. (1960). Supplemented balance. *Biometrika*, **47**, 263-271.

Pearce, S. C. (1983). *The Agricultural Field Experiment. A Statistical Examination of Theory and Practice*. John Wiley and Sons.

Pearce, S. C. (1970). The efficiency of block designs in general. *Biometrika*, **57**, 339-346.

Pena, D. and Yohai, V. J. (1995). The detection of influential subsets in linear regression by using an influence matrix, *Journal of Royal Statistical Society, Series B*, **57**, 145-156.

Prasad, R., and Gupta, V. K. (2000). A Note on augmented designs. *Indian Journal of Plant Genetic Resources*, **13**(**1**), 53-58.

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

Preece, D. A. (1983). Latin squares, Latin cubes, Latin rectangles, etc. In *Encyclopaedia of Statistical Sciences, Vol.* **4**. S. Kotz and N.L. Johnson, Eds, 504-510.

Preece, D. A. (1988). Semi-Latin squares. In *Encyclopaedia of Statistical Sciences, Vol.* **8**. S. Kotz and N.L. Johnson, Eds, 359-361.

Rao, C. R.(1973). *Linear Statistical Inference and Application*. Wiley Eastern Ltd., New Delhi.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raghavarao, D. (1971). *Construction and Combinatorial Problems in Design of Experiments.* Wiley, New York.

Rai, S.C. and Rao, P. P. (1980). Use of ranks in groups of experiments. *Journal of Indian Society of Agricultural Statistics*, **32**, 25-32.

Rai, S.C. and Rao, P. P. (1984). Rank analysis of group of split-plot experiments. *Journal of Indian Society of Agricultural Statistics*, **36**, 105-113.

Reeves, G. K. (1991). Estimation of contrast variances in linear models. *Biometrika*, **78**, 7-14.

Ridout, M. S. (1989). Summarizing the results of fitting generalized linear models from designed experiments. In *Statistical Modelling: Proceedings of GLIM89.* A. Decarli et al. editors, pp. 262-269. New York: Springer.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, **79**, 871-880.

Royston, P. (1992). Approximating the Shapiro-Wilk W-Test for non-normality. *Statistics and Computing*, **2**, 117 -119.

SAS/STAT User's Guide (1990). Volume 1, Version 6, Fourth Edition, SAS Institute Inc., Cary, NC.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components, *Biometrics*, **2**, 110-114.

Scheff´e, H. (1959). *The Analysis of Variance*. John Wiley and Sons, New York.

Schilling, E. G. (1973). A systematic approach to the analysis of means. Part II. Analysis of contrasts. *Journal of Quality Technology*, **5**, 147–155.

Seber, G. A. F.(1983). *Multivariate Observations*. Wiley series in Probability and Statistics.

Senn, S. J. (1993). *Cross-over Trials in Clinical Research.* Chichester: Wiley.

Sen, P. K. (1996). *Design and Analysis of Experiments: Non-parametric Methods with Applications to Clinical Trials*. Handbook of Statistics, Vol 13, eds. S. Ghosh and C. R. Rao.

Shah, K. R. and Sinha, B. K. (1989). *Theory of Optimal Designs.* Springer, Berlin.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variances test for normality (Complete Samples). *Biometrika*, **52**, 591-611.

Siegel, S. and Castellan, N. J. Jr. (1988*). Non parametric statistics for the Behavioural Sciences*, 2nd Ed., McGraw Hill, New York.

Silvey, S. D. (1980). *Optimal Design.* Chapman and Hall, London.

Singh, M. and Dey, A. (1979). Block designs with nested rows and columns. *Biometrika*, **66**, 321-326.

Skillings J. H. and Mack G. A. (1981). On the use of a Friedman-type statistic in balanced and unbalanced block designs. *Technometrics*, **23**, 171-177.

Smith, H. F. (1938). An empirical law describing heterogeneity in the yields of agricultural crops. *The Journal of Agricultural Science*, **28**, 1-23.

Smith, H. Jr. (1969). The analysis of data from a designed experiment. *Journal of Quality Technology*, **1**, 259–263.

Snedecor, G. W. and Cochran, W. G. (1967). *Statistical Methods*. Ames.

Speed, T. P. (1987). What is an analysis of variance? (with discussion). *Annals of Statistics*, **15**, 885-941.

Street, A. P. and Street, D. J. (1987). *Combinatorics of Experimental Design*. Oxford University Press.

Thompson, W. A. Jr. and Moore, J. R. (1963). Non-negative estimates of variance components. *Technometrics*, **5**, 441–449.

Tukey, J. W. (1949). One degree of freedom for non-additivity. *Biometrics*, **5**, 232-242.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Publishing Company, Reading.

Ture, T. E. (1994). Optimal row-column designs for multiple comparisons with a control: A complete catalogue. *Technometrics*, **36**, 292-299.

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, **29**, 350–362.

Williams, E. J. (1950). Experimental designs balanced for pairs of residual effects. *Australian Journal of Scientific Research*, **A 3**, 351-363.

Wooding, W. M. (1973). The split-plot design. *Journal of Quality Technology,* **5**, 16–33.

Yates, F. (1935). Complex experiments (with discussion). *Supplement to the Journal of the Royal Statistical Society*, II, 181-247.

Yates, F. (1936). A new method of arranging variety trials involving a large number of varieties. *Journal of Agricultural Science*, **26**, 424-455.

Youden, W. J. (1956). Randomization and experimentation (abstract). *Annals of Mathematical Statistics*, **27**, 1185-1186.

NOTE

NOTE

Designing an experiment properly is very important in any scientific endeavour because a carefully designed experiment is able to answer all the queries by making efficient use of available resources with the researcher. For successful experimentation, it is highly desirable that scientists and researchers in agricultural sciences understand the basic principles of designing an experiment and analysis of data generated from the completed experiment. Most of the books available on Design of Experiments treat the subject through rigorous algebraic theories, which often the agricultural scientists and researchers find difficult to understand and adopt in their experimentation. More often than not, the interaction of agricultural scientist with statistician is absent. This creates problems in a proper choice of a design for experimentation and then the analysis of data to answer the questions of the experimenter. The purpose of this book is to build a bridge between the agricultural sciences and statistical sciences so as to help make the agricultural research globally competitive, visible and acceptable by proper use of experimental designs and analysis of data generated.

The subject of experimental designs would be treated in two parts: Part I would focus on fundamentals of design and analysis of experiments related to mostly single factor experiments and Part II would feature essentially multifactor experiments and advanced topics. The present book is the Part I, which covers basic concepts of various designs for single factor experiments along with enough examples, easy to follow analysis steps using SAS and R software and interpretation of analysis results. The Examples used in this book are, by and large, the actual experiments conducted by the scientists. Though intended principally for agricultural researchers, none-the-less this book will also serve as a text book for students undergoing graduation and post graduation in the field of Statistics or Agricultural Statistics. The authors fervently expect that the book would be enormously useful to the agricultural scientists in planning and designing their experiments and analyzing the data generated from their experiments.