Basic Concepts and Applications of Experimental Designs and Analysis

Felix Kutsanedzie
Sylvester Achio
Edmund Ameko

SciencePG
Science Publishing Group

Basic Concepts and Applications of Experimental Designs and Analysis

Felix Kutsanedzie

Sylvester Achio

Edmund Ameko

SciencePG
Science Publishing Group

Science Publishing Group

548 Fashion Avenue
New York, NY 10018
www.sciencepublishinggroup.com

Published by Science Publishing Group 2015

First Edition

ISBN: 978-1-940366-50-0

To order additional copies of this book, please contact:
Science Publishing Group
book@sciencepublishinggroup.com
www.sciencepublishinggroup.com

Printed and bound in India

Dedication

This book is dedicated to our families and friends for their constant support in course of writing this book. It is time to celebrate the successes we have chalked together.

Acknowledgement

We wish to acknowledge our colleague lecturers, research fellows of Accra Polytechnic and all well wishers.

Preface

Researchers, academics and students are engaged in one form of research or the other that requires designing. However be it as it may, most of these individuals are not conversant with selecting the appropriate experimental designs that should best suit their respective researches or studies. To some understanding the concepts and basis of these designs are quite a challenge. Still others have a huge challenge handling these designs because of the complex mathematics underpinning these designs.

In addition there is no one stop book that treats in details the various designs and the mathematical principles underlying them. One might advance the argument that computer applications aid analyzing of these designs but the same cannot be said of these designing and selecting the appropriate designs for individuals in the conduct researches.

This book covers thoroughly the explanations of the concepts and basic terms in almost all the known experimental designs; the mathematics underlying these; how to select the appropriate designs for a study and a logical sequence of analyzing these designs. For each of these designs, hypothetical examples of experiments have been provided with stepwise approaches towards analyzing them. It treats complex designs in a simplified way to enhance the understanding of its readership. The designs are arranged in a systematic order of increasing complexity.

Albeit the underlying principles of experimental designs and analysis are based on mathematic, the other aspect of designs, which is the actualization or practising is equally challenging. Oftentimes when it comes to mounting of an experiment using a particular design, people understand the mathematical basis, however identifying and allocating the different treatments or levels or factors of treatments to each plot or experimental units (as to whether they are

homogeneous or heterogeneous) become a dilemma to them. This book does not only deal with the mathematics behind each design but also how to identify the treatments; levels of treatments; whether a plot or a unit is homogenous or heterogeneous; and explains and demonstrate with practical examples on how to identify and allocate factor in the design as main factors or sub factors based when applicable.

Each chapter dwells extensively and exhaustively on a particular design with hypothetical data analysed and interpretations to aid the readers understanding of the design.

This is thus a must read book for all involved in designing of experiments in diverse fields.

Contents

Chapter 1

Concepts and Basis of Experimental Design

Concepts and Basis of Experimental Design

Felix Kutsanedzie[1 *]; Sylvester Achio[1]; Edmund Ameko[1]; Victoria Ofori[2]; Paul Goddey[1]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Agricultural Engineering Department, KNUST, Ghana

Abstract

The concepts and basic terms underlying experimental designs are not well understood by students and some researchers. For experimental designs to be understood, the various terms used and applied in designing an experiment must be well explained. Some of the terms used and applied in the field of experimental designs are quite wrongly used and applied by students and many others involved in research. This paper defines and explains the terms comprehensively.

Keywords

Experimental Design, Treatment, Bias, Randomization, Experiment, Variance, Variable

1.1 Introduction

There is variability in all living and non-living things that exist in nature. Even though things may be categorized or classified into a particular group, there still could exist variability between things placed in the same group. Variations exist more in living things than in non-living things because of their characteristics – growth, reproduction, excretion, irritability, movement, respiration. The variability in living organisms may be linked to special functions they play. Some plants have narrower leaves than others; others have their leaves reduced to spine as a way of conserving water loss through respiration and transpiration. Some human beings are ambidextrous, thus they are able to do so many things at a goal. It is very important as a matter of fact when researching on things to be able to ascertain and quantify the variability between them. In order to do this, there is a need for one to design experiments to prove statistically how variables differ or relate to each other.

Oftentimes people complain of the mathematical principles on which the concept rests as it is tedious working such calculations manually. In the advent of computers, computations have been made quite easy but the underlying principles need to be understood before using computers to do the various calculations. It will also explain and define the terms used in the design of experiments as well as the types of design and when they are supposed to be used in experiment.

1.2 Terms Used in Experimental Designs

There are so many terms one to know and understand in the designing of experiment. When these terms are not properly understood and used by the

designer, it would lead to confusion and wrong results being churned out at the end. Some of the terms are explained below.

1.3 Experimental Design

It refers to the allocation of treatments to experimental units or materials or plots. Normally when an experiment is to be mounted, it must first be designed. Thus experimental design encapsulates how an experiment must be conducted, and how data collected is to be collected, analysed and interpreted. It therefore suggests that every experiment (not survey) must have its unique and appropriate design.

1.4 Experiment

It is an investigation where an investigator imposes treatment(s) on experimental units to ascertain the effects on the unit by the measurement of response variables on the unit due to the imposition.

When one wants to conduct an experiment, the individual considers what to perform the experiment on (experiment unit or material) and what to use on the experiment unit to be able to study their effects on the unit (treatment) and the measurements to take from the units (response variable) in order to arrive at conclusion. However, a researcher can have a test and a controlled experiment.

1.5 Test Experiment

In a test experiment, all conditions or treatment are made available except the treatment the researcher is interested in investigating for effects on the experimental material. For instance, if the researcher wants to know the effects

of nitrogen on a plant growth, he / she provides all the conditions for plant growth with the exception of nitrogen. When all other need conditions all provided, whatever effects or defects that are observed are considered to be due to the condition lacked in the experiment.

1.6 Control Experiment

It is an experiment in which all treatments or conditions are necessary for the experimental material to allow the needed response variables to be observed and measured. This is the converse of the test experiment. The researcher would therefore compare the results from a test and a control experiment to do interpretation for a conclusion to be made.

1.7 Observatory Study

It is also an investigation in which no treatment is imposed on an experimental unit by the investigator but he or she observes and measures response variables on the unit.

1.8 Treatments

It is the set of conditions or circumstances created for an experiment or investigation. For instance an investigator might want to study the effect of acids on a particular experimental unit or material. The treatment one can then impose on the experimental unit to create and acid condition could be the introduction of an acid. Hence the acid becomes the treatment in this case.

1.9 Factor

It is an explanatory variable to be studied in an experiment and can be set at different values.

Once the treatment has been selected as acid, the experimenter might be interested in looking at more than one form or type of an acid. Each form or type of an acid thus becomes a factor in the experiment.

1.10 Level

The level of a factor refers to the different values of factors under consideration or to be studied. For example in the case of a treatment (acid condition), factors such as sulphuric acid, carbonic acid, hydrogen sulphide can all be considered as factors because they are all different kinds of acids. Now, the concentrations or the quantities of these acids used in creating the acidic conditions referred to as the levels of the factors

1.11 Replication

It refers to the number of times a complete set of treatment is repeated in an experiment. When a researcher applies the same treatment to the same experimental unit or materials, provided the conditions are the same, the response variables would yield the same measurement (the results should be the same). To ensure consistency in results, a researcher repeats the same treatments over the experimental units in order to minimize errors and biases which are likely to be overlooked when repetition or cross checking is not done.

1.12 Experimental Unit

It is physical entity or subject on which the treatment is applied independent of other units. In other words, it is what generates the response variable that the investigator needs to measure and observe to address the objective of the experiment. It should be noted that an experiment can have homogeneous or heterogeneous experimental units and thus might give different response variables. It is therefore important for the research to know if his or her units are homogeneous or not in order to apply the appropriate design.

1.13 A Plot

It is an example of experimental units - a smallest unit of land that a treatment is applied to. The conditions or compositions of a piece of land differ in chemically, physically and biologically. If an experiment needs to be conducted on a piece of land, these chemical, physical and biological are likely to affect the results. Thus a plot is just a small piece of land, where the variations in its chemical, physical and biological properties are expected to lesser.

1.14 A Block

It is a large area or experimental unit consisting of several identical units on which all or most of the treatments under consideration are applied. Thus plots can be blocked – can be classified into blocks in an experiment. This allows the experimenter to compare variations in all the treatments when considering heterogeneous experimental units.

1.15 Response Variable

A characteristic of an experimental unit observed or measured after a treatment has been applied to it. In other words, it is the reaction observable or measurable reactions generated by experimental units as a result of application or imposition of a treatment to/on it. These observations and measures are what the investigator analyzes to address the objectives of the experiment.

1.16 Explanatory Variable

It is the characteristics of a treatment (factor) that induces the experimental unit to generate a response variable.

1.17 Experimental Error

It is a measure of the differences between experimental units on which the same treatment is applied. It seeks to establish the variation or variances in the experimental units. These variations may stem from the units, the lack of uniformity in the way the investigator applied the treatment, uncontrolled external influences and others that cannot be explained (natural). For example, if the same treatment (seed of pepper) is applied on an experimental units (plots).

1.18 Randomization

It is the act of allocating treatments to plots in an experimental design such that each has equal chance of receiving each treatment – plots or experimental units are not favoured or discriminated against. Randomization reduces the incidence of biases in allocation of treatment. An experimenter can decide to

treat a particular unit differently due to personal beliefs or ideals thereby introducing errors. However, when treatments are randomized, this can never happen. Randomization can be done by using random numbers generated using the computer or random number tables. To give different ration of feed (treatments) to rats (experimental units), the quantity feeds can be written on a paper, folded, placed in a cup, mixed up, then picked randomly and them applied to the experimental units.

1.19 Single-Factor Experiment

It is an experiment in which the investigator varies only one factor while all the others are kept constant. For an experiment in which the treatment is an acid, the factors can be hydrochloric acid and sulphuric acid. The experimenter can choose to vary the concentration of one of the factors while maintaining the other. When this happens, the experiment becomes a single-factor experiment.

1.20 Multi–Factorial Experiment

It is an experiment in which all the factors involved in the experiment are varied unlike the single-factor experiment.

1.21 Full Factorial Treatment Design

In a full factorial treatment design, the treatments involve all possible combinations of the levels of the factors of interest.

1.22 Observational Unit

An observational unit is a unit on which the response variable is observed and measured. This unit can either be the same as the experimental unit in some cases and in other cases not. For instance, if seeds are sown on different types of soil to ascertain the yields of plants in those soils, the soil is the experimental units but the fruits on the plant become observation units.

1.23 Conducting an Experiment

To conduct an experiment, one needs to consider the following:

- *Identify, define and state the problem*

  One cannot investigate a problem without identifying it. The first and foremost thing to do is to identify and define the problem such that all could understand it as a problem that needs to be tackled and solved. The problem statement should be precise and concise. It should not be ambiguous.

- *State the objectives and develop a hypothesis of the study*

  The objectives of conducting the experiment must be clearly stated. They are basically the reasons for conducting the experiment. The problem can be woven to develop a hypothesis – a statement which is neither considered as true or false but needs to be investigated and proven to otherwise. It is therefore the data collected from an experiment carried out that can provide evidence for or against the hypothesis.

- *Designing and conducting the experiment*

Having identified, defined and stated the problem, the research needs to design appropriate experiment that would help him or her conduct relevant data to prove the hypothesis. The designing of the experiment therefore plays a crucial role in proving the hypothesis and hence achieving the stated objectives of the experiment. Research therefore needs to be well baked in experimental designs to be able to design good experiments.

- *Collecting data*

The execution of the experiment allows the researcher to collect data on the response variables from experimental units induced by explanatory variables from factors of the treatments being considered in the experiment or to observe and measure the response variables in the case of observatory studies.

Bibliography

[1] Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs.* John Willey & Sons, N.Y.

[2] Adams, W. K, Wieman, C. E. (2010). Development and validation of instruments to measure learning of expert-like thinking. *Int J. Sci. Educ.* 33:1289–1312.

[3] American Association for the Advancement of Science (2011). Vision and Change in Undergraduate Biology Education: A Call to Action, Washington, DC. http://visionand change.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf (accessed 9 June 2013).

[4] Anderson-Cook, C. M., Doraj-Raj, S. (2001). An active learning in class demonstration of good experimental design. J Stat Educ *9*(1). http://amstat.org/publications/jse/v9n1/anderson-cook.html (accessed 6 November 2013).

[5] Dasgupta, A. P., Anderson T. R., Pelaez, N. (2014). Development and validation of a rubric for diagnosing students' experimental design knowledge and difficulties. *CBE Life Sci Educ.* 13:265–284.

[6] Doran, R. (1980). Basic Measurement and Evaluation of Science Instruction, Washington, DC: National Science Teachers Association.

[7] Festing, M. F. W. (2003). Principles: the need for better experimental design. Trends *Pharmacol Sci.* 24:341–345.

[8] Green, D. S., Bozzone, D. M. (2001). A test of hypotheses about random mutation: using classic experiments to teach experimental design. *Am Biol Teach*. 63:54–58.

[9] Hiebert, S. M. (2007). Teaching simple experimental design to undergraduates: do your students know the basics? *Adv Physiol Educ.* 31:82–92.

[10] Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecol Monogr.* 54:187–211.

[11] Lambert, C. G., Black, L. J. (2012). Learning from our GWAS mistakes: from experimental design to scientific method. *Biostat* 13:195–203.

[12] Pollack, A. E. (2010). Exploring the complexities of experimental design: using an on-line reaction time program as a teaching tool for diverse student populations. *J Undergrad Neurosci Educ.* 9: A47–A50.

[13] Shi, J., Power, J. M., Klymkowsky, M. W. (2011). Revealing student thinking about experimental design and the roles of control experiments. *Int J Schol Teach Learn. 5*:1–19.

[14] Sirum, K., Humburg, J. (2011). The Experimental Design Ability Test (EDAT). *Bioscience* 37:8–16.

[15] Zolman, J. F. (1999). Teaching experimental design to biologists. *Adv Physiol Educ.* 277:111–118.

[16] Brownlee, K. A. *Statistical theory and methodology in science and engineering*. New York: Wiley, 1960.

[17] Campbell, D., Stanley, J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), *Handbook on research on teaching*. Chicago: Rand McNally & Co.

[18] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.

[19] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs*. John Willey & Sons, N.Y.

[20] Le Clerg, E. L., Leonard, W. H., Clark, A. G. (1966). *Field Plot Technique*. Burgess Pub. Co., Minn., USA.

Chapter 2

Complete Randomized Design (CRD)

2

Complete Randomized Design (CRD)

Felix Kutsanedzie[1] *; Sylvester Achio[1]; Edmund Ameko[1]; Victoria Ofori[2]; Edith Mensah[1]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Agricultural Engineering Department, KNUST, Ghana

Abstract

Complete Randomized Design is one of the design analysis that is not well understood by researchers. However the need to understand such a design is very vital in the case where the experimental units being considered are homogeneous or uniform. This paper uses examples to explain the underlying principles and how the design is used for analysis.

Keywords

Randomization, Homogeneous, Treatment, Experimental Units, Factors, Levels, Replications

2.1 Introduction

Complete Randomized Design (CRD) is the simplest form of design used in experimental analysis. It involves randomization of treatments on homogeneous or uniform experimental units /plots uniform. Since in nature or reality, most experimental units are not homogeneous, it makes this design suited for large experiments vis-à-vis the homogeneity requirement of the experimental units. This presupposes that it is not suitable for the analysis of large field experiments. It is thus suited for small experiments.

In allocation or assignment of treatments to experimental units, randomization is used. Randomization involves the assigning of treatments to experiment units such that each treatment has equal chance of being assigned to units available. The randomization uses random tables and computer programmes to generate random numbers for the allocation of treatments to units. For instance, when one has five treatments – chloroquine (A), malarex (B), paracetamol (C), chamoquine (D) and panadol (E) are given to sterile rabbits(experimental units) to study their effects on them. Assuming that this experiment is to be replicated (repeated) five times, it means 5 (treatments) x 5 (replicates) totaling 25 (experimental units) from which data will be collected. The table below indicates how the randomization is done:

Table 2.1 *Randomization of Five Treatments in a Complete Randomized Design (CRD).*

A	B	C	D	E
B	A	D	C	D
C	D	E	A	C
D	E	B	B	A
E	C	A	E	B

Each of the boxes in the tables represent an experimental unit or plots (sterile rabbits) on which the various treatments (the medicines) denoted in the boxes are to be subjected to. This means that the researcher can table twenty (25) pieces of papers and label them with the letters A to E in turns until the twenty (25) pieces of papers are exhausted. The labeled pieces of papers are then folded, dropped in a cup and, shuffled. The pieces of papers are then picked randomly one after the order to denote the various plots or experimental units that the picked treatments would be applied to i.e. from 1^{st} to the 25^{th} experimental units. When this procedure is followed religiously, treatments are said to be randomized on the experimental units.

Once the experiment design is set, the treatment can then be allocated randomly to the experimental units according to the design. During this time onwards are the researcher is supposed to observe the observation units and to measure and determine responses of the experimental units to the treatments they have been subjected to.

2.2 Analysis of Data Obtained from the CRD

Assuming the treatments were subjected to the experiment units in order to determine their efficacies based on the time taken in days for a named symptom to be corrected, the researcher would have to observe the various treatments and record the time taken in days as summarized in the table below:

Data collected on time (measured in days) taken for a named symptoms on experimental units for each the respective replicates of each treatment considered in the CRD.

Table 2.2 *Results Collected on responses of treatment to experimental units in CRD.*

A = 3	**B = 3**	**C = 3**	**D = 6**	**E = 3**
B = 4	A = 4	D = 3	C = 4	D = 4
C = 5	D = 4	E = 2	A = 3	C = 4
D = 6	E = 3	B = 4	B = 3	A = 5
E = 3	C = 2	A = 5	E = 2	B = 3

In analyzing CRD designs after that had been collected from an experiment that have been carried, the must approach the Analysis of Variance (ANOVA) in terms of the following classification: designs having treatments with equal replication; designs having treatments with unequal replications, designs having treatments with equal number of samples per experimental unit; designs having number of subsamples that are unequal.

Design and Analysis of Variance (ANOVA) for treatments with equal number of replication

For the researcher to subject the data recorded in the Table 2.2, it must be re-summarized as below:

No/Replications	**A**	**B**	**C**	**D**	**E**
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4	5	3	4	4	3
5	5	3	2	6	2

A = 3	B = 3	C = 3	D = 6	E = 3
B = 4	A = 4	D = 3	C = 4	D = 4
C = 5	D = 4	E = 2	A = 3	C = 4
D = 6	E = 3	B = 4	B = 3	A = 5
E = 3	C = 2	A = 5	E = 2	B = 3

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypothesis

H_o: $\mu_A = \mu_B = \mu_C = \mu_D = \mu_E$

H_A: $\mu_A \neq \mu_B = \mu_C = \mu_D = \mu_E$

H_o: All the means of the treatments are equal

H_A: At least one of the means of the treatments is different

H_o=Null hypothesis, H_A=Alternate hypothesis, μ= means of treatments

No/Replications	**A**	**B**	**C**	**D**	**E**
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4	5	3	4	4	3
5	5	3	2	6	2
Tt =	20	17	18	23	13
ΣTt =	91				
CF = (ΣTt)^2/(r x t)	3.64				
TSS = Σ((3)^2 + (3)^2 +...+ (2)^2) –CF	357.36				
TtSS = Σ((20)^2 +(17)^2 + (13)^2) /r) – CF	338.56				
ESS = TSS –TtSS	18.8				

$$\mathrm{Tt}_A = \sum(3 + 4 + \cdots + 5) = 20$$

$$\mathrm{Tt}_B = \sum(3 + 4 + \cdots + 3) = 17$$

$$\mathrm{Tt}_C = \sum(3 + 4 + \cdots + 2) = 18$$

$$\mathrm{Tt_D} = \sum(6 + 3 + \cdots + 6) = 23$$

$$\mathrm{Tt_E} = \sum(3 + 2 + \cdots + 2) = 13$$

$$\mathrm{Tt_T} = \sum(\mathrm{Tt_A} + \mathrm{Tt_B} + \mathrm{Tt_C} + \mathrm{Tt_D} + \mathrm{Tt_E}) = 91$$

Tt_T=Total treatment, ESS=Error Sum of Squares, TSS= Total Sum of Squares, TtSS=Treatment Sum of Squares, CF= Correction Factor.

The second step is calculating of the Correction Factor:

$$\mathrm{CF} = \frac{\sum(\mathrm{Tt_T})^2}{(\mathrm{r\ x\ t})} = 331.24$$

The third step is calculating of the Total Sum of Squares:

$$\mathrm{TSS} = \sum((3)^2 + (3)^2 + (3)^2 + (6)^2 + (3)^2 + (4)^2 \ldots\ldots + (2)^2) - \mathrm{CF} = 357.36$$

The fourth step is calculating of the Treatment Sum of Squares:

$$\mathrm{TtSS} = \sum \frac{((20)^2 + (17)^2 + (18)^2 + (23)^2 + (13)^2)}{5} - \mathrm{CF} = 338.56$$

The fifth step is calculating of the Error Sum of Squares:

$$\mathrm{ESS} = \mathrm{TSS} - \mathrm{TtSS} = 357.36 - 338.56 = 18.8$$

Tt_T=Treatment Total, TSS=Total Sum of Squares, TtSS=Treatment Sum of Squares, ESS=Error Sum of Squares, CF=Correction Factor, t=number of treatment, r = number of replication.

The sixth step is completing the ANOVA table:

ANOVA TABLE					
Sources of Variation	**df**	**SS**	**MS**	**Fcal.**	**Fcrit. (5%)**
Treatment	t-1 = 5 - 1 = 4	338.56	84.64	90.04255	2.87
Error	[((r x t)-1)-(t-1))] = 24 -4 = 20	18.8	0.94		
Total	(r x t) -1 = (5 x 5) - 1 =24	357.36			
Fcrit (5%) @ df of 4,20	= 2.87				
Fcrit (1%) @ df of 4,20	= 4.43				

df=degree of freedom, SS=Sum of Squares, MS=Mean Square.

The degree of freedom for the all items under the sources of variations is one minus the item: df of treatment is 5 – 1; df for the total number of observation is 25-1; the df of freedom for the error is 24 - 4.

The Mean Square (MS) is the ratio of the SS to df: Treatment MS is calculated as:

$$\text{Treatment Mean Square (MS)} = \frac{338.56}{4} = 84.64$$

$$\text{Error Mean Square (MS)} = \frac{18.8}{20} = 0.94$$

The F calculated value is the ratio of Treatment Mean Square to the Error Mean Square:

$$\text{Fcal.} = \frac{\text{Treatment Mean Square}}{\text{Error Mean Square}} = \frac{84.64}{0.94} = 90.04255$$

The seventh step looking up the F-critical table to the F- critical values:

In looking at the F-critical table for the critical value, the level of significance or level of confidence is used. If the level of significance used is 1%, then the level of confidence is 99%; if the level of significance is 5%, the level of

confidence is 95%. Normally the 1% and 5% levels of significances are used depending on the precision of the work being undertaken.

Whether the 1% or 5% level of significance is used or not, the degree of freedom of the treatment variation is used against the degree of freedom of the error variation as (x, y), where the degree of freedom of the treatment is located horizontally (x-axis) against the degree of freedom of the error variation on the vertically (y-axis) and the intercept of these two gives the corresponding value for the F-critical or tabulated value.

Thus using the F-critical value from 5% (0.05) level of significance table for instance, one has to locate the degree of freedom of the treatment i.e. 4 horizontally on the 5% F- table, and the degree of freedom of the error i.e. 20 on the vertically. Where the two intercepts, the F- critical or tabulated value can be read as 2.87. When the same procedure is followed on the 1% F –table, at using the degree of freedom of the treatment as 4 against the degree of freedom of the error as 20, the F- critical or tabulated value of 4.43 is obtained. This is summarized as follows:

$$\text{Fcrit or tab. at } 5\%\ (0.05) = 2.87$$

$$\text{Fcrit or tab. at } 1\%\ (0.01) = 4.43$$

The eighth step is making the decision or conclusion:

The decision rule is summarized as follows:

If F calculated at the level of significance or level of confidence and degrees of freedoms stated is greater or equal to (≥) the F critical or tabulated, the null hypothesis is rejected and therefore concluded that there is enough evidence from the data to support the rejection of the null hypothesis because the mean of the various treatments are significantly different. This is expressed mathematically as:

$$\text{Fcal} \geq \text{Fcrit. or tab.}(5\%)\text{at } 4,20$$

$$\text{Fcal} \geq \text{Fcrit. or tab.}(1\%)\text{at } 4,20$$

For instance in the case being considered:

$$\text{Fcal}\,(90.04) \geq 2.87, \text{Fcrit. or tab.}(5\%)\text{at } 4,20$$

$$\text{Fcal}\,(90.04) \geq 4.43, \text{Fcrit. or tab.}(1\%)\text{at } 4,20$$

Therefore the means of the various treatments being considered are significantly different at both 5% (0.05) and 1% (0.01)

However if the F calculated at the level of significance or level of confidence and degrees of freedoms stated is less than to (<) the F critical or tabulated. We fail to reject the null hypothesis and conclude that there is enough evidence supporting the fact that the means of the various treatments are not significantly different. This is expressed as follows:

$$\text{Fcal} < Fcrit.\,or\ tab.(5\%)\text{at } 4,20$$

$$\text{Fcal} < Fcrit.\,or\ tab.(1\%)\text{at } 4,20$$

It should be noted whenever a level of significance is chosen; the p-value (probability value) is equal to the level of significance. Thus for 5%, it means the p-value = 0.05, at 1%, p-value = 0.01. This implies that when the F calculated ≥ F critical, it is concluded that means are significantly different; hence the p-value will be less than the level of significance chosen i.e. < 0.05 and < 0.01 at 5% and 1% respectively.

Table 2.3 *F critical table at 5% (0.05).*

Critical values of F for the 0.05 significance						
	1	2	3	4	5	
1	161.45	199.50	215.71	224.58	230.16	2
2	18.51	19.00	19.16	19.25	19.30	
3	10.13	9.55	9.28	9.12	9.01	
4	7.71	6.94	6.59	6.39	6.26	
5	6.61	5.79	5.41	5.19	5.05	
6	5.99	5.14	4.76	4.53	4.39	
7	5.59	4.74	4.35	4.12	3.97	
8	5.32	4.46	4.07	3.84	3.69	
9	5.12	4.26	3.86	3.63	3.48	
10	4.97	4.10	3.71	3.48	3.33	
11	4.84	3.98	3.59	3.36	3.20	
12	4.75	3.89	3.49	3.26	3.11	
13	4.67	3.81	3.41	3.18	3.03	
14	4.60	3.74	3.34	3.11	2.96	
15	4.54	3.68	3.29	3.06	2.90	
16	4.49	3.63	3.24	3.01	2.85	
17	4.45	3.59	3.20	2.97	2.81	
18	4.41	3.56	3.16	2.93	2.77	
19	4.38	3.52	3.13	2.90	2.74	
20	4.35	3.49	3.10	2.87	2.71	
21	4.33	3.47	3.07	2.84	2.69	

Table 2.4 *F critical table at 1% (0.01).*

Critical values of F for the 0.01 significance						
	1	2	3	4	5	
1	4052.19	4999.52	5403.34	5624.62	5763.65	5
2	98.50	99.00	99.17	99.25	99.30	
3	34.12	30.82	29.46	28.71	28.24	
4	21.20	18.00	16.69	15.98	15.52	
5	16.26	13.27	12.06	11.39	10.97	
6	13.75	10.93	9.78	9.15	8.75	
7	12.25	9.55	8.45	7.85	7.46	
8	11.26	8.65	7.59	7.01	6.63	
9	10.56	8.02	6.99	6.42	6.06	
10	10.04	7.56	6.55	5.99	5.64	
11	9.65	7.21	6.22	5.67	5.32	
12	9.33	6.93	5.95	5.41	5.06	
13	9.07	6.70	5.74	5.21	4.86	
14	8.86	6.52	5.56	5.04	4.70	
15	8.68	6.36	5.42	4.89	4.56	
16	8.53	6.23	5.29	4.77	4.44	
17	8.40	6.11	5.19	4.67	4.34	
18	8.29	6.01	5.09	4.58	4.25	
19	8.19	5.93	5.01	4.50	4.17	
20	8.10	5.85	4.94	4.43	4.10	
21	8.02	5.78	4.87	4.37	4.04	

The ninth step is calculating the Coefficient of Variation:

$$\%CV = \frac{S}{\bar{X}} \text{ X } 100$$

$$\text{Error MS} = S^2$$

$$S = \sqrt{\text{Error MS}}$$

$$\%CV = \frac{\sqrt{\text{Error MS}}}{\bar{X}}$$

$$\%CV = \frac{\sqrt{0.94}}{3.64} \text{ x } 100 = 26.64$$

$$CV = 26.64\%$$

$$CV = coefficient\ of\ variation,\ S = Standard\ deviation$$

$$Error\ MS = Error\ Mean\ Square,\ \bar{x} = mean\ of\ treatment$$

The Coefficient of Variation is the ratio of the standard deviation to the mean of the treatments expressed in percentage. It is a measure of the consistency of the mean of the treatment or the variations in the mean of the treatment. It

implies that when the standard deviation increases the variations in the means of the treatment increases. The CV can also be taken as the measure of variations in the means of the treatments. Thus the lower the CV value, the more consistent or uniform the means of the treatment.

Design and Analysis of Variance (ANOVA) for treatments with unequal number of replication

No/Replications	A	B	C	D	E
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4		3		4	3
5		3			2

This analysis is done when dealing with unequal number of replications of the treatment. In cases like this, the same preambles are used but with variations in the various formulae used in the computation. So we proceed to follow the same steps used earlier.

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypothesis

H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_D = \mu_E$

H_1: $\mu_1 \neq \mu_2 = \mu_3 = \mu_D = \mu_E$ (At least one of the treatment means differ from the others).

The second step is calculating of the Correction Factor:

No/Replications	A	B	C	D	E
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4		3		4	3
5		3			2
Tt	10	17	12	17	13
∑Tt	69				
CF = (ΣTt)^2/(r))	238.05				
TSS	16.95				
TtSS	7.13				
ESS	9.82				
Tt mean	3.45				

$$CF = \frac{\sum(Tt_T)^2}{r}$$

$$CF = \frac{(69)^2}{20} = 238.05$$

where $r = (r_A + r_B + r_C + r_D + r_E)$, *r= total number of observations for the treatments.*

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum((3)^2 + (3)^2 + (3)^2 + (6)^2 + (3)^2 + (4)^2 \ldots\ldots + (2)^2) - 238.05 = 16.95$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TtSS = \sum\left(\frac{(Tt_A)^2}{r_A} + \frac{(Tt_B)^2}{r_B} + \frac{(Tt_C)^2}{r_C} + \frac{(Tt_D)^2}{r_D} + \frac{(Tt_E)^2}{r_E}\right) - CF$$

$$TtSS = \sum\left(\frac{(10)^2}{3} + \frac{(17)^2}{5} + \frac{(12)^2}{3} + \frac{(17)^2}{4} + \frac{(13)^2}{5}\right) - 238.05 = 7.13$$

$$\text{TtSS} = \sum 245.18 - 238.05 = 7.13$$

The fifth step is calculating of the Error Sum of Squares:

$$\text{ESS} = \text{TSS} - \text{TtSS} = 16.95 - 7.13 = 9.82$$

The sixth step is completing the ANOVA table:

ANOVA TABLE

Sources of Variation	df	SS	MS	Fcal.	Fcrit. (5%)
Treatment	t-1 = 5 - 1 = 4	7.13	1.78	2.72	3.06
Error	[((r -1)-(t-1))] = 19-4 = 15	9.82	0.65		
Total	r -1 = (20) - 1 =19	16.95			

The seventh step looking up the F-critical Table to the F-critical values:

This follows the earlier procedures used in looking up the critical values from the F critical Table

$$\text{Fcrit. or tab. } (5\%) \text{at df of } 4,15 = 3.06$$

$$\text{Fcrit. or tab. } (1\%) \text{at df of } 4,15 = 4.89$$

Critical values of F for the 0.05 significanc

	1	2	3	4	5
1	161.45	199.50	215.71	224.58	230.16
2	18.51	19.00	19.16	19.25	19.30
3	10.13	9.55	9.28	9.12	9.01
4	7.71	6.94	6.59	6.39	6.26
5	6.61	5.79	5.41	5.19	5.05
6	5.99	5.14	4.76	4.53	4.39
7	5.59	4.74	4.35	4.12	3.97
8	5.32	4.46	4.07	3.84	3.69
9	5.12	4.26	3.86	3.63	3.48
10	4.97	4.10	3.71	3.48	3.33
11	4.84	3.98	3.59	3.36	3.20
12	4.75	3.89	3.49	3.26	3.11
13	4.67	3.81	3.41	3.18	3.03
14	4.60	3.74	3.34	3.11	2.96
15	4.54	3.68	3.29	3.06	2.90
16	4.49	3.63	3.24	3.01	2.85

Critical values of F for the 0.01 significanc

	1	2	3	4	5
1	4052.19	4999.52	5403.34	5624.62	5763.65
2	98.50	99.00	99.17	99.25	99.30
3	34.12	30.82	29.46	28.71	28.24
4	21.20	18.00	16.69	15.98	15.52
5	16.26	13.27	12.06	11.39	10.97
6	13.75	10.93	9.78	9.15	8.75
7	12.25	9.55	8.45	7.85	7.46
8	11.26	8.65	7.59	7.01	6.63
9	10.56	8.02	6.99	6.42	6.06
10	10.04	7.56	6.55	5.99	5.64
11	9.65	7.21	6.22	5.67	5.32
12	9.33	6.93	5.95	5.41	5.06
13	9.07	6.70	5.74	5.21	4.86
14	8.86	6.52	5.56	5.04	4.70
15	8.68	6.36	5.42	4.89	4.56
16	8.53	6.23	5.29	4.77	4.44

The eighth step is making the decision or conclusion:

$$\text{Fcal. } (2.72) < (3.06), \text{Fcrit. or tab. } (5\%) \text{at df of } 4,15$$

$$\text{Fcal.}(2.72) < (4.89), \text{Fcrit. or tab.}(1\%) \text{at df of } 4,15$$

We therefore fail to reject the null hypothesis at both 5% (0.05) and 1% (0.01) level of significance and conclude that the means are not significantly different based on the data available.

The ninth step is calculating the Coefficient of Variation:

$$\%\text{CV} = \frac{\sqrt{\text{Error MS}}}{\bar{\text{X}}}$$

$$\%\text{CV} = \frac{\sqrt{0.65}}{3.45} \text{ x } 100 = 23.45$$

$$\text{CV} = 23.45\%$$

Design and Analysis of Variance (ANOVA) with equal number of samples per experimental units with treatments of equal number of replication

Consider an experimental design where three different concentrations (10M, 20M and 30M) of acetone are applied to dissolve a Perspex material chopped into five different thicknesses and replicated four times alongside taking the records of the time for the dissolution of chopped to be realized in minutes. This is an example of CRD design that requires ANOVA with equal number of samples per experimental unit with treatments of equal number of replication. The table below represents the summary of this information:

Conc. (M)	**10M**				**20M**				**30M**			
	Perspex″				**Perspex″**				**Perspex″**			
Thickness (inches)	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1	4	5	6	4	5	6	7	8	9	4	5	5
2	5	4	6	4	3	6	7	7	4	6	5	5
3	6	7	7	5	6	6	7	5	5	5	4	7
4	4	7	8	6	4	5	5	6	6	6	4	6
5	7	8	9	4	5	6	7	4	7	7	8	9

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypothesis

H_o: $\mu_1 = \mu_2 = \mu_3$

H_1: $\mu_1 \neq \mu_2 = \mu_3$ (At least one of the treatment means differ from the others)

Conc. (M)		**10M**				**20M**				**30M**			
		Perspex″				**Perspex″**				**Perspex″**			
Thickness (inches)	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1	4	5	6	4	5	6	7	8	9	4	5	5
2	5	4	6	4	3	6	7	7	4	6	5	5
3	6	7	7	5	6	6	7	5	5	5	4	7
4	4	7	8	6	4	5	5	6	6	6	4	6
5	7	8	9	4	5	6	7	4	7	7	8	9
EU	26	31	36	23	23	29	33	30	31	28	26	32
Tt_T		116				115				117		
CF	2018											
Tt mean	5.8											
TSS	123.6											
TtSS	0.1											
EUSS	34.8											
EESS	34.7											
SESS	88.8											

The second step is calculating of the Correction Factor:

$$CF = \frac{\sum(Tt_T)^2}{rts}$$

$$CF = \frac{(348)^2}{(4)(3)(5)} = 2018.40$$

r= number of replication, t=number of treatment, s=number of samples.

The third step is calculating of the Total Sum of Squares:

$$\text{TSS} = \sum((4)^2 + (5)^2 + (6)^2 + (4)^2 + (5)^2 + (6)^2 \ldots\ldots + (9)^2) - 2018.40 = 123.60$$

The fourth step is calculating of the Treatment Sum of Squares:

$$\text{TtSS} = \sum\left(\frac{(\text{Tt}_{t1})^2}{\text{rs}} + \frac{(\text{Tt}_{t2})^2}{\text{rs}} + \frac{(\text{Tt}_{t3})^2}{\text{rs}}\right) - \text{CF}$$

$$\text{TtSS} = \sum\left(\frac{(116)^2}{(4)(5)} + \frac{(115)^2}{(4)(5)} + \frac{(117)^2}{(4)(5)}\right) - 2018.40$$

$$\text{TtSS} = \sum 2018.5 - 2018.4 = 0.1$$

The fifth step is calculating the Sum of Squares among the Experimental Units:

$$EUSS = \sum\left(\frac{(EU_1)^2}{s} + \frac{(EU_2)^2}{s} + \frac{(EU_3)^2}{s} + \ldots\ldots \frac{(EU_{12})^2}{s}\right) - CF$$

$$EUSS = \sum\left(\frac{(26)^2}{5} + \frac{(23)^2}{5} + \frac{(31)}{5}\right) - 2018.40$$

$$\text{EUSS} = 2053.2 - 2018.40 = 34.8$$

The sixth step is calculating the Experimental Error Sum of Squares:

$$\text{Experimental Error SS(EESS)} = \text{EUSS} - \text{TtSS}$$

$$\text{Experimental Error SS (EESS)} = 34.8 - 0.1 = 34.70$$

The seventh step is calculating the Sampling Error Sum of Squares:

$$\text{Sampling Error SS (SESS)} = \text{TSS} - \text{EUSS}$$

$$\text{SamplingError SS(SESS)} = 123.60 - 34.80 = 88.8$$

The eighth step is completing the ANOVA table:

ANOVA TABLE					
Sources of Variations	**df**	**SS**	**MS**	**Fcal**	**Fcrit.**
Treatment	t - 1 = 3 -1 = 2	0.1	0.05	0.013	4.26(5%)
Experimental error	(tr-1) - (t - 1) = (3x4 -1) -2 = 9	34.7	3.86		8.02(1%)
Sampling error	(trs -1) - (tr - 1) = (3x4x5 -1) -9 = 50	88.8			
Total	trs - 1 = 60 -1 =59	123.6			

The seventh step looking up the F-critical Table to the F-critical values:

This follows the earlier procedures used in looking up the critical values from the F critical Table

$$\text{Fcrit. or tab.}(5\%)\text{at df of } 2,9 = 4.26$$

$$\text{Fcrit. or tab.}(1\%)\text{at df of } 2,9 = 8.02$$

The eighth step is making the decision or conclusion:

$$\text{Fcal.}(0.013) < (4.26), \text{Fcrit. or tab.}(5\%)\text{at df of } 2,9$$

$$\text{Fcal.}(0.013) < (8.02), \text{Fcrit. or tab.}(1\%)\text{at df of } 2,9$$

We therefore fail to reject the null hypothesis at both 5% (0.05) and 1% (0.01) level of significance and conclude that the means are not significantly different based on the data available.

The ninth step is calculating the Coefficient of Variation:

$$\%\text{CV} = \frac{\sqrt{\text{Error MS}}}{\bar{\text{X}}}$$

$$\%\text{CV} = \frac{\sqrt{3.86}}{5.8} \text{ x } 100$$

$$\%\text{CV} = \frac{1.96}{5.8} \text{ x } 100 = 33.87$$

$$CV = 33.87\%$$

Design and Analysis of Variance (ANOVA) with unequal number of subsamples

Taking for instance a design similar to the previously treated one where this time around the number of subsamples are unequal as indicated in the table, the following approach is used in analyzing it.

Conc. (M)	**10M**				**20M**				**30M**			
	Perspex″				**Perspex″**				**Perspex″**			
Thickness (inches)	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1	4	5	6	4	5	6	7	8	9	4	5	5
2	5	4	6	4	3	6	7	7	4	6	5	5
3	6	3	7	5	6	7	7	5	5	4	4	7
4					4	5	5	6	4	4	4	6
5									6	5	8	5
TEU	15	12	19	13	18	24	26	26	28	23	26	28
Tt_T			59			94				105		
CF	1387											
Tt mean	5.37											
TSS	85											
TtSS	7											
EUSS	30											
EESS	23											
SESS	55											

The first step in the analysis is to put forth the hypothesis:

H_o: $\mu_1 = \mu_2 = \mu_3$

H_1: $\mu_1 \neq \mu_2 = \mu_3$ (At least one of the treatment means differ from the others)

The second step is calculating of the Correction Factor:

$$\text{CF} = \frac{(\text{sum of observations})^2}{\text{number of observations}}$$

$$\text{CF} = \frac{(258)^2}{(48)}$$

$$\text{CF} = \frac{66564}{48} = 1387$$

The third step is calculating of the Total Sum of Squares:

$$\text{TSS} = \sum((4)^2 + (5)^2 + (6)^2 + (4)^2 + (5)^2 + (6)^2 \ldots\ldots + (5)^2) - 1387$$

$$\text{TSS} = 1472 - 1387 = 85$$

The fourth step is calculating of the Treatment Sum of Squares:

$$\text{TtSS} = \sum\left(\frac{(\text{Tt}_{t1})^2}{\text{rs}} + \frac{(\text{Tt}_{t2})^2}{\text{rs}} + \frac{(\text{Tt}_{t3})^2}{\text{rs}}\right) - \text{CF}$$

$$\text{TtSS} = \sum\left(\frac{(59)^2}{(4)(3)} + \frac{(94)^2}{(4)(4)} + \frac{(105)^2}{(4)(5)}\right) - 1387$$

$$\text{TtSS} = \sum 1394 - 1387 = 7$$

The fifth step is calculating the Sum of Squares among the Experimental Units:

$$\text{EUSS} = \sum\left(\frac{(\text{EU}_1)^2}{\text{s}_1} + \frac{(\text{EU}_2)^2}{\text{s}_2} + \frac{(\text{EU}_3)^2}{\text{s}_3} + \ldots\ldots \frac{(\text{EU}_{12})^2}{\text{s}_i}\right) - \text{CF}$$

$$\text{EUSS} = \sum\left(\frac{(15)^2}{3} + \frac{(18)^2}{4} + \frac{(28)^2}{5}\right) - 1387$$

$$\text{EUSS} = 1417 - 1387 = 30$$

$$\text{EUSS} = \text{Sum of Squares Among Experimental Units}$$

$$\text{TEU} = \text{Total treatment on Experiemental Units}$$

The sixth step is calculating the Experimental Error Sum of Squares:

$$\text{Experimental Error SS(EESS)} = \text{EUSS} - \text{TtSS}$$

$$\text{Experimental Error SS (EESS)} = 30 - 7 = 23$$

The seventh step is calculating the Sampling Error Sum of Squares:

$$\text{Sampling Error SS (SESS)} = \text{TSS} - \text{EUSS}$$

$$\text{SamplingError SS(SESS)} = 85 - 30 = 55$$

ANOVA TABLE

Sources of Variations	df	SS	MS	Fcal	Fcrit. (5%)	Fcrit (1%)
Treatment	t - 1 = 3 -1 = 2	7	3.5	1.3696	4.26	8.02
Experimental error	df(EU) -df(t) = (11-2) = 9	23	2.56	2.5556		
Sampling error	df(T) - df(EU) = (47 - 11) = 36	55				
Total	Total observ. – 1 = (48 -1)= 47	85				

df(EU)= degree of freedom of experimental unit.
df(T)= degree of freedom of total observation.

The seventh step looking up the F-critical Table to the F-critical values:

This follows the earlier procedures used in looking up the critical values from the F critical Table

$$\text{Fcrit. or tab. (5\%) at df of 2,9} = 4.26$$

$$\text{Fcrit. or tab. (1\%) at df of 2,9} = 8.02$$

The eighth step is making the decision or conclusion:

$$\text{Fcal. (1.37)} < \text{(4.26), Fcrit. or tab. (5\%) at df of 2,9}$$

$$\text{Fcal. (1.37)} < \text{(8.02), Fcrit. or tab. (1\%) at df of 2,9}$$

We therefore fail to reject the null hypothesis at both 5% (0.05) and 1% (0.01) level of significance and conclude that the means are not significantly different based on the data available.

Bibliography

[1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.

[2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), Handbook on research on teaching. Chicago: Rand McNally & Co.

[3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

[4] Cochran, W. G., G. M. (1957). Experimental Designs, John Willey & Sons, N.Y.

[5] Le Clerg, E. L., Leonard, W. H., Clark A. G. (1966). *Field Plot Technique*. Burgess Pub. Co., Minn., USA.

Chapter 3

Complete Randomised Block Design

3

Complete Randomised Block Design

Felix Kutsanedzie[1] *; Sylvester Achio[1];

Edmund Ameko[1]; Victoria Ofori[2];

Diaba Kwasi Selassie[3]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Agricultural Engineering Department, KNUST, Ghana

[3]Anglican University College of Technology, Nkorazan, Sunyani, Ghana

Abstract

Complete Randomised Design (CRD) is used for experiments in which the experimental materials or units are homogeneous. However most experimental materials in life are heterogeneous in nature, hence the need to employ Complete Randomized Block Designs (CRBD) in designing and analysis of experiments with such nature. It is the most used of all the types of design, hence this chapter explains the underlying conditions and how it is used in analyzing experiments comprehensively for researchers and would-be experiment designers.

Keywords

Block, Homogeneous, Heterogeneous, Analysis, Randomized

3.1 Introduction

So long as researchers or individuals continue to design experiments to explain mechanisms, phenomenon etc. using heterogeneous experimental materials, the Complete Randomized Block Design (RCBD) would be required for analyses. Whereas Complete Randomized Design (CRD) is used appropriate for the analysis for experiments which involves the use of homogenous experimental materials, RCBD is used for experiments that require the use of heterogeneous materials. Generally when using animals as experimental materials in an experiment, when they are of the same species, they can be considered as homogeneous but when they are of different species then they are considered as heterogeneous. Assuming a researcher wants to use a plot of land to serve an experimental material for a study which involves the sowing of seeds to ascertain the germination rate, the plot of land needs to be divided into blocks, because it cannot be taken as a homogeneous unit due to the variability of nutrients at each point of the land. Thus the use of RCBD requires that when experimental materials which are not homogenous are to be used in an experiment, they must be divided into subgroups which are similar in nature and referred to as blocks or replicates before the design is employed for its analysis. The import of reducing the heterogeneous experimental materials to blocks or replicates is to make sure variations are minimized or reduced as much as possible so that all variations existing would be due to variability in the treatments applied.

In RCBD, randomization of treatments is done such that every block is restricted to a single treatment. However it should be noted that based on randomization designs are partially classified as follows: homogeneity of experimental material – Complete Randomized Design (CRD); heterogeneity of experimental material – Randomized Complete Block Design, that is a single restriction of treatment; Latin Square Design (LSD) and Cross Over Designs, designs with two restrictions of

treatment; Graeco-Latin Square Design, design with more than two restriction of treatment allocation; Incomplete Block Design (IBD) includes those not grouped into replications and those grouped into replications.

With RCBD, there is at least a single restriction of treatments per block, the treatment are randomly allocated at least once for each replicate or block. Also treatments are randomized separately for each block and have equal probability of being allocated to any experimental unit per block or replicate.

3.2 Merits of RCBD over CRD

There is preciseness of RCBD over CRD.

The species or objects experimental materials generally thought to be homogeneous though it may not be necessarily as they may differ in one way or the other when carefully examined. This exposes the idea of homogeneity of material in the case of CRD as a flaw should differences exist between the experimental materials or units. However the RCBD reduces this flaw as it is intended to cater for the heterogeneity of experimental materials. As regard restriction of treatment, there is at least a single restriction, thus is every treatment is expected to be allocated at least once per block or replicate.

3.3 Illustration of Randomized Complete Block Design

Assuming a researcher wants to design an experiment to manage waste materials such coconut fruit waste (shell and fibre), palm nut shells, waste plastic bottles, waste plastic water sachets and waste plastic packaging bags through pyrolysis for the recovery of other usable products and by-products. A design of such nature is to ascertain whether the amount of products such as oil,

gas or char produced subjected to different treatments (different weights of the various materials used) on equal weight per weight basis are significant or not. The design below is an illustration RCBD.

	Treatments (weights of the waste types) to be Pyrolysed			
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**
Coconut waste fruits (shell and fibre) [A]	2.3	5.0	11.2	15.7
Palm nut shell [B]	1.6	4.8	9.2	14.0
Waste plastic bottles [C]	4.4	9.2	14.4	19.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2
Waste plastic bags [E]	4.1	8.5	11.9	16.9

The table above is a design for five different treatments in terms of weights of the various waste types being considered in the study. The data being considered is the weights of the oil generated from these waste types. It should however be noted that the values fielded in the table are all being assumed for explanation and not real.

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	2.3	5.0	11.2	15.7	34.2
Palm nut shell [B]	1.6	4.8	9.2	14.0	29.2
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	11.9	16.9	41.4
Treatment Totals	16.6	36.4	60	84.1	197.1

The first step: Stating of hypothesis

For Blocks:

H_o: All the means of the blocks are equal

H_A: At least one of the means of the blocks (blk) is different

H_o: $\mu blk_A = \mu blk_B = \mu blk_C = \mu blk_D = \mu blk_E$

H_1: $\mu blk_A \neq \mu blk_B = \mu blk_C = \mu blk_D = \mu blk_E$

For Treatments:

H_o: All the means of the treatments are equal

H_A: At least one of the means of the treatments (t) is different

H_o: $\mu trt_A = \mu trt_B = \mu trt_C = \mu trt_D = \mu trt_E$

H_1: $\mu trt_A \neq \mu trt_B = \mu trt_C = \mu trt_D = \mu trt_E$

The second step: Calculating the Correction Factor (CF)

$$CF = \frac{Y^2}{TB}$$

where Y^2=square of the sum of all observations

where TB is the product of the number of treatments and number of blocks or number of replicates (r)when replicates are used rather instead of the blocks

$$CF = \frac{(2.3 + 5 + 11.2 \ldots\ldots\ldots. + 16.9)^2}{4 \times 5}$$

$$CF = \frac{(197.1)^2}{4 \times 5}$$

$$\mathrm{CF} = \frac{38848.41}{20}$$

$$\mathrm{CF} = 1942.42$$

The third step: calculating the Total Sum of Squares (TSS)

$$\mathrm{TSS} = \sum \mathrm{Y}_{\mathrm{IJ}}^2 - \mathrm{CF}$$

$$\mathrm{TSS} = \sum (2.3^2 + 5.0^2 + 11.2^2 \ldots\ldots + 16.9^2) - 1942.42$$

$$\mathrm{TSS} = 2513.57 - 1942.42 = 571.15$$

Y_{IJ}^2 *is the square of all observations.*

The fourth step: calculating the Block Sum of Squares (BLKSS)

$$\mathrm{BLKSS} = \sum \frac{\mathrm{BLKSS}^2}{\mathrm{nTrt}} - \mathrm{CF}$$

$$\mathrm{BLKSS} = \sum \frac{34.2^2 + 29.6^2 + 47.3^2 \ldots + 41.4^2}{4} - 1942.42$$

$$\mathrm{BLKSS} = \frac{1169.64 + 876.16 + 2237.29 + \cdots + 1713.96}{5} - 1942.42$$

$$\mathrm{BLKSS} = \frac{7986.21}{4} - 1942.42$$

$$\mathrm{BLKSS} = 1996.55 - 1942.42$$

$$\mathrm{BLKSS} = 54.13$$

The fifth step: calculating the Treatment Sum of Squares (TRTSS)

$$\mathrm{TRTSS} = \sum \frac{\mathrm{TRTSS}^2}{\mathrm{nBLK}} - \mathrm{CF}$$

$$\mathrm{TRTSS} = \sum \frac{16.6^2 + 36.4^2 + \ldots\ldots + 84.1^2}{5} - 1942.42$$

$$\mathrm{TRTSS} = \frac{275.56 + 1324.96 + 3600 \ldots\ldots\ldots + 7072.81}{5} - 1942.42$$

$$\text{TRTSS} = \frac{12273.33}{5} - 1942.42$$

$$\text{TRTSS} = 2454.67 - 1942.42$$

$$\text{TRTSS} = 512.25$$

The sixth step: calculating the Error Sum of Squares (ERRSS)

$$\text{ERRSS} = \text{TSS} - (\text{BLKS} + \text{TRTSS})$$

$$\text{ERRSS} = 571.15 - (54.13 + 512.25)$$

$$\text{ERRSS} = 571.15 - 566.38$$

$$\text{ERRSS} = 4.77$$

Where nTrt =Number of Treatment, nBLK = Number of Blocks.

The seventh step: completing the ANOVA Table

ANOVA TABLE					
Sources of Variation	**df**	**SS**	**MS**	**Fcal.**	**Fcrit.**
block	blk - 1 = 5-1 = 4	54.13	13.53	33.83	
treatment	trt - 1 = 4 -1 = 3	512.25	170.75	426.88	
error	(blk-1)(trt -1) = 4 x 3 = 12	4.77	33.67		
total	blk x trt - 1 = 20 -1 = 19	571.15			

For block: Fcrit (5%) @ df of 4, 12 = 3.26; Fcrit (1%) @ df of 4, 12 =5.41.
For treatment: Fcrit (5%) @ df of 3, 12 =3.49; Fcrit (1%) @ df of 3, 12 = 5.95.

The tenth step: looking up the F-critical table to find the F-critical values:

For the block:

$$Fcrit.or\ tab.(5\%) at\ df\ of\ 4,12 = 3.26$$

$$Fcrit.or\ tab.(1\%) at\ df\ of\ 4,12 = 5.41$$

For the treatment:

$$Fcrit.or\ tab.(5\%) at\ df\ of\ 3,12 = 3.49$$

$$Fcrit.or\ tab.(1\%)at\ df\ of\ 3{,}12 = 5.95$$

The eighth step: making the decision or conclusion:

Decision on block:

$$Fcal.(33.83) > (3.26), Fcrit.or\ tab.(5\%)at\ df\ of\ 4{,}12$$

$$Fcal.(33.83) > (5.41), Fcrit.or\ tab.(1\%)at\ df\ of\ 4{,}12$$

At both 5% and 1% levels of significance for the block, we reject the null hypotheses because the F-calculated values are greater than the F-critical values; thus we conclude that there is enough evidence provided by the data collected in the experiment to reject the null hypotheses. It thus implies that the blocks are significantly different, hence there was a need to block or in other words blocking can be justified. It means RCBD is the right design adopted for the study.

Decision on treatment:

$$Fcal.(426.88) > (5.95), Fcrit.or\ tab.(1\%)at\ df\ of\ 3{,}12$$

At 1%, since the F-calculated value is greater than the F-critical value, we reject the null hypothesis and conclude that the data collected provides enough evidence that the treatments are significantly different.

$$Fcal.(426.88) > (3.49), Fcrit.or\ tab.(5\%)at\ df\ of\ 3{,}12$$

Again at 5% level of significance for the treatment, we reject the null hypothesis on the basis that the F-calculated value is greater than the F-critical value and thus conclude that the treatments are significantly different, therefore the need to find out the treatments that are significantly different.

3.3.1 Finding the Treatments that are Significantly Different

Once the test reveals that there exist significance differences between the treatments, we proceed to find out which of the treatments significantly differ. This procedure of finding out is referred to as the *pair comparisons*.

Since there are four treatments and we are to do a pair comparison, it presupposes that we will have C_2^4 i.e. 4 combination 2 which means there would be 6 paired comparisons possible. These are as follow:

AB, AC, AD, BC, BD and CD. There are two statistical methods used in doing the pair comparisons, these are namely: the Least Significance Difference Test (LSD) and the Duncan Multiple Range Test (DMRT).

3.3.2 The Fisher's LSD Test

With regards to the LSD (Least Significant Difference) test, an LSD value is calculated at a prescribed level of significance either 5% or 1%, which serves as a boundary for the classification between whether one treatment is significantly or not significantly different from another when their means are compared. This means that if the means difference of any two treatments compared exceed the LSD computed at a prescribed significance level, then we conclude that the two treatments are significantly different or otherwise not. It is used or valid when used for independent (orthogonal) comparison and used when the treatment size is less i.e. less than six treatments. The generalized version of LSD is given by the formula:

$$LSD = \frac{\left|\overline{x_i} - \overline{x_j}\right|}{\sqrt{wms\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}$$

$$However\ sd = \left(\sqrt{WMS^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}\right)$$

Since $n_i = n_j$,

$$sd = \left(\sqrt{WMS^2\left(\frac{2}{n}\right)}\right)$$

$$sd = \sqrt{\frac{2WMS^2}{n}}$$

where WMS =Within (Residual or unexplained)Mean Square, n=number of sample per each treatment or replication; sd=standard deviation.

$$LSD_{\propto} = t_{\propto}\ x\ sd$$

However, $t_{\frac{\propto}{2}}$ is used to in order to a two sided hypothesis, therefore

$$LSD_{\propto} = t_{\frac{\propto}{2}}\ x\ sd$$

Taking the example being handled under the RCBD design for instance, we used the following data:

	Treatments (weights of the waste types) to be Pyrolysed			
Blocks	**A**	**B**	**C**	**D**
Coconut waste fruits (shell and fibre)	2.3	5	11.2	15.7
Palm nut shell	1.6	4.8	9.2	14
Waste plastic bottles	4.4	9.2	14.4	19.3
Waste plastic satchets	4.2	8.9	13.3	18.2
Waste plastic bags	4.1	8.5	11.9	16.9
Treatment Means	3.32	7.28	12	16.82

$$\mathrm{LSD}_{\propto} = \mathrm{t}_{\propto}\ \mathrm{x}\ \mathrm{sd}$$

$$\text{sd} = \sqrt{\frac{2\text{WMS}^2}{\text{n}}}$$

$WMS\ from\ the\ ANOVA\ table =$ 33.67; n = 4; Error or residual or unexplained df = 12

$$sd = \sqrt{\frac{2(33.67)^2}{5}}$$

$$sd = \sqrt{\frac{2\ x\ 1133.89}{5}}$$

$$sd = \sqrt{\frac{2267.79}{5}}$$

$$sd = \sqrt{453.56}$$

$$sd = 21.30$$

We therefore proceed to read the t-critical value from the two-tailed table at the significance level at which the design revealed that the treatments were significant (5% or 0.05).

Therefore from the two tailed t-critical at 5% level of significance table we obtain:

$$t_{\frac{0.05}{2}} = t_{o.o25}, df\ 12\ of\ the\ residual\ or\ unexplained\ = 2.18$$

$$t - crit(two - tailed) at\ df 12 = 2.18$$

$$LSD_{\propto} = 2.18\ x\ 21.30$$

$$LSD_{0.05} = 46.43$$

The value t-crit (two-tailed) at df 12 can be read from the $t_{\frac{0.05}{2}} = t_{o.o25}$ critical table.

α (1 tail)	0.05	0.025
α (2 tail)	0.1	0.05
df		
1	6.3138	12.7065
2	2.9200	4.3026
3	2.3534	3.1824
4	2.1319	2.7764
5	2.0150	2.5706
6	1.9432	2.4469
7	1.8946	2.3646
8	1.8595	2.3060
9	1.8331	2.2621
10	1.8124	2.2282
11	1.7959	2.2010
12	1.7823	2.1788
13	1.7709	2.1604

Means Difference Table				
	A($\overline{x_A}$)	**B($\overline{x_B}$)**	**C($\overline{x_C}$)**	**D($\overline{x_D}$)**
A ($\overline{x_A}$)	-			
B ($\overline{x_B}$)	($\overline{x_B} - \overline{x_A}$)	-		
C($\overline{x_C}$)	($\overline{x_C} - \overline{x_A}$)	($\overline{x_C} - \overline{x_B}$)	-	
D($\overline{x_D}$)	($\overline{x_D} - \overline{x_A}$)	($\overline{x_D} - \overline{x_B}$)	($\overline{x_D} - \overline{x_C}$)	-

Means difference table				
	A (3.32)	**B (7.28)**	**C(12.00)**	**D(16.82)**
A (3.32)	-			
B(7.28)	3.96	-		
C(12.00)	8.68	4.72	-	
D(16.82)	13.5	9.54	4.82	-

Since all the mean differences for all the treatments are (A, B, C, D) are less than (<) the LSD value obtained (45.39), hence the LSD test confirms they are not significantly different.

$$(\overline{x_B} - \overline{x_A}), (\overline{x_C} - \overline{x_A}), (\overline{x_D} - \overline{x_A}), (\overline{x_C} - \overline{x_B}), (\overline{x_D} - \overline{x_B}), (\overline{x_D} - \overline{x_C}) < LSD_{0.05}$$

The Duncan Multiple Range Test

This test is also used for paired comparisons for a larger treatment size than the LSD. With the DMRT, the sample means of the treatments are ranked from the lowest to the highest and then the steps apart denoted by (r) is derived and used with the total degree of freedom (df) to read the q tabulated value from the studentized range table.

When using the DMRT, two population means are significantly different if the absolute value of their sample differences exceed W, where W is defined as below:

$$W = q\ (r, df\ of\ residual\ or\ unexplianed)\ x\ \sqrt{\frac{WMS}{n}}$$

Where n=number of samples or observations per treatment group, r=the number of steps from the lowest treatment mean to the highest treatment mean when the treatment means are arranged in ascending order, df=degree of freedom of the residual or unexplained or error, WMS=within residual or unexplained or error mean square derived from the ANOVA table.

Arrangement of the treatment means in ascending order to determine the steps is done as follows:

$$\overline{x_A} = 3.32 < \overline{x_B} = 7.28 < \overline{x_C} = 12.0 < \overline{x_D} = 16.82$$

Since the distance between $\overline{x_A}\ and\ \overline{x_D}$ is 4, thus moving from $\overline{x_A}\ to\ \overline{x_D}\ is\ 4\ steps\ then\ \overline{x_D}$ has r = 4, $\overline{x_c}\ has\ r = 3, \overline{x_B}\ has\ r = 2$

Therefore the table below can be constructed to aid the calculation of W:

	$\overline{x_B}$	$\overline{x_C}$	$\overline{x_D}$
	$r = 2$	$r = 3$	$r = 4$
q(r, df of error = 12)	3.082	3.773	4.199
$W = q\ (r, df\ of\ error = 12)\ x\ \sqrt{\frac{WMS}{n}}$	$3.082\ x\ \sqrt{\frac{33.67}{5}}$	$3.773\ x\ \sqrt{\frac{33.67}{5}}$	$4.199\ x\ \sqrt{\frac{33.67}{5}}$
W	7.98	9.77	10.88

The values of q can be read from the critical values of the Studentized Range (0.05) presented.

Critical Values of the Studentuzed Range (0.05 level)

dfe								
	2	3	4	5	6	7	8	9
2	6.0849	8.3308	9.7980	10.8810	11.7340	12.4345	13.0266	13.5381
3	4.5007	5.9096	6.8245	7.5016	8.0370	8.4780	8.8521	9.1766
4	3.9265	5.0403	5.7571	6.2870	6.7065	7.0528	7.3465	7.6015
5	3.6354	4.6017	5.2185	5.6731	6.0329	6.3299	6.5823	6.8014
6	3.4605	4.3390	4.8956	5.3049	5.6285	5.8953	6.1222	6.3192
7	3.3439	4.1648	4.6812	5.0601	5.3591	5.6058	5.8154	5.9975
8	3.2612	4.0410	4.5288	4.8858	5.1672	5.3991	5.5962	5.7673
9	3.1991	3.9485	4.4149	4.7554	5.0235	5.2444	5.4319	5.5947
10	3.1511	3.8768	4.3266	4.6543	4.9120	5.1242	5.3042	5.4605
11	3.1127	3.8195	4.2561	4.5736	4.8229	5.0281	5.2021	5.3531
12	3.0813	3.4428	4.1985	4.5076	4.7477	4.9469	5.1159	5.2625
13	3.0553	3.7341	4.1509	4.4529	4.6897	4.8841	5.0490	5.1920
14	3.0332	3.7014	4.1105	4.4066	4.6385	4.8290	4.9903	5.1300

We can now proceed to conclude based on the computation of the W-value and the mean difference:

Means difference table				
	A (3.32)	**B (7.28)**	**C(12.00)**	**D(16.82)**
A (3.32)	-			
B(7.28)	3.96	-		
C(12.00)	8.68	4.72	-	
D(16.82)	13.5*	9.54	4.82	-

$(\overline{x_B} - \overline{x_A})$*=3.96<7.98, there is no significant difference between A and B.*

$(\overline{x_C} - \overline{x_A})$*=8.68<9.77, there is no significant difference between C and A.*

$(\overline{x_C} - \overline{x_B})$*=4.72<9.77, there is no significant difference between C and B.*

$(\overline{x_D} - \overline{x_A})$*=13.5>10.88, there is significant difference between D and A.*

$(\overline{x_D} - \overline{x_B})$ *= 9.54 < 10.88, there is no significant difference between D and B.*

$(\overline{x_D} - \overline{x_C})$ *= 4.82 < 10.88, there is no significant difference between D and C.*

3.4 Missing Data Handling

There is bound to be some data missing while conducting the experiment based on accidents such as breakage, death of an animal, spilling of a substance, destruction of a treatment on the experimental material by any extraneous subject or object (human or animal). When the experiment is started and these accidents occur, then the data on some experimental units cannot be obtained and thus referred to as 'missing data'. The experiment cannot be halted but continued and the missing data estimated after the experiment. In order to estimate the missing data the formula below can be used:

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

where Y_{ij}=the missing data value, r=number of replicates or blocks, B=block or replicate total for block or replicate with the missing data, t=number of treatment, T=treatment total for treatment with missing data, $G = \sum_{ij}^{n} Y_{ij}$ = grand total of experimental units or observation units.

3.4.1 Handling a Single Missing Data

Therefore using the same experiment presented for the RCBD design with one of the data taken out as a missing data in the table below, the missing data value is estimated as follows:

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	2.3	5.0	11.2	15.7	34.2
Palm nut shell [B]	1.6	4.8	9.2	14.0	29.2
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	Y_{ij}	13.3	18.2	35.7
Waste plastic bags [E]	4.1	8.5	11.9	16.9	41.4
Treatment Totals	16.6	27.5	60	84.1	188.2

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$r = 5, t = 4, B = 44.6, T = 36.4, G = 188.2$$

$$Y_{ij} = \frac{(5 * 35.7 + 4 * 27.5 - 188.2)}{(5-1)(4-1)}$$

$$Y_{ij} = \frac{(178.5 + 110 - 188.2)}{(4)(3)}$$

$$Y_{ij} = \frac{288.5 - 188.2}{12}$$

$$Y_{ij} = \frac{100.3}{12}$$

$$Y_{ij} = 8.4$$

where r=number of replication, t=number of treatments, B=Block or replicate total with missing data, T=Treatment total with missing data, G= grand total of all observations.

After the estimation of the missing data, the bias is computed using the formulae:

$$\beta_1 = \frac{[B_0 - (t-1)x]^2}{t(t-1)}$$

$$\beta_1 = \frac{[35.7 - (4-1)8.4]^2}{4(4-1)}$$

$$\beta_1 = \frac{[35.7 - (3)8.4]^2}{4(3)}$$

$$\beta_1 = \frac{[35.7 - 25.2]^2}{12}$$

$$\beta_1 = \frac{[10.5]^2}{12}$$

$$\beta_1 = \frac{(110.25)}{12}$$

$$\beta_1 = 9.2$$

Therefore the estimated bias $(\beta_1) = 9.2$

Adjustment in the Analysis of using the bias:

When the estimated bias has been computed, the adjustment of the analysis is done by subtracting the bias value from only the TSS (Total Sum of Squares) i.e.

[571.15 – 9.2 =561.95] and TrtSS (Treatment Sum of Squares) i.e. [512.25 – 9.2 = 503.05]

Once the missing data value is estimated, it can be put in the table in order to complete the ANOVA table. It must however be noted that one df (degree of freedom) is lost from the error i.e. [((5-1)(4-1) – 1) = (12 -1) = 11] and also the total degrees of freedom because of one missing data. Therefore the ANOVA would look as below:

Complete Anova Table					
Sources of Variation	**df**	**SS**	**MS**	**Fcal.**	**Fcrit.**
Block	blk - 1 = 5-1 = 4	54.13	13.53	31.21	
Treatment	trt - 1 = 4 -1 = 3	503.05	167.68	386.69	
Error	(blk-1)(trt -1) -1 =[4 x 3]-1 = 11	4.77	0.43		
Total	blk x trt – 1-1 = 20 -1-1 = 18	561.95			

From the ANOVA table, the decisions to be taken on the blocks and treatments remains the same, after the missing data had been computed and the ANOVA table completed for the data.

3.4.2 Handling More than One Missing Data Under RCBD

Assuming there are more than one missing data or values to be estimated under the RCBD, how this can be done is explained using the RCBD table below:

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	Y_{43}	16.9	29.5
Treatment Totals	14.3	36.4	48.1	70.1	168.9

From the above table, it can be seen that three data values (Y_{11}, Y_{24} and Y_{43}) are conspicuously missing and which must be estimated in order to complete the table for the analysis of the data recorded.

The first procedure adopted in estimating the values is to use the formula given below to find the average estimated values of the missing data in the design:

$$Y_{ij} = \frac{(T_m + B_m)}{2}$$

This formula can apply only to RCBD when estimating average missing data.

Where Y_{ij} =estimated average missing data value, T_m =the mean value of treatment with the missing data, B_m =the mean value of treatment with the missing data.

Thus applying this formula, one can obtain the estimated average values of the missing data: Y_{11}, Y_{24} and Y_{43}.

$$Y_{ij} = \frac{(T_m + B_m)}{2}$$

$$For\ Y_{11} = \frac{(T_m + B_m)}{2}$$

$$T_m = \frac{14.3}{5} = 2.86,\ B_m = \frac{31.9}{4} = 7.98$$

$$Y_{11} = \frac{(2.86 + 7.98)}{2} = \frac{10.84}{2} = 5.42$$

$$For\ Y_{24} = \frac{(T_m + B_m)}{2}$$

$$T_m = \frac{70.1}{5} = 14.02,\ B_m = \frac{15}{4} = 3.75$$

$$Y_{24} = \frac{(14.02 + 3.75)}{2} = \frac{17.77}{2} = 8.89$$

$$For\ Y_{43} = \frac{(T_m + B_m)}{2}$$

$$T_m = \frac{48.1}{5} = 9.62, B_m = \frac{29.5}{4} = 7.38$$

$$Y_{43} = \frac{(9.62 + 7.38)}{2} = \frac{17.0}{2} = 8.50$$

Now that we have estimated the averages of missing data for Y_{11}, Y_{24} and Y_{43}, two of these estimated values can be substituted into the table for the formula given below to be used:

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

for computing the estimated values of each of the missing data in turns, and replacing the average estimated values with the new computed values until no change in the new computed values occur. When no change occurs in the estimated values, it means the accurate missing data values have been found.

Using the stated procedures, we proceed to apply them.

Complete the table with the average estimated values (italicized in the table) leaving only the one whose estimated value one wants to determine first and in this case we start with Y_{11}

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	*8.89*	24.49
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*8.50*	16.9	38.0
Treatment Totals	14.3	36.4	56.6	78.99	186.29

Compute for the estimated value of Y_{11} using the formula given as:

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{11} = \frac{(5(31.9) + 4(14.3) - 186.29)}{(5-1)(4-1)}$$

$$Y_{11} = \frac{((159.5 + 57.2) - 186.29)}{(4)(3)}$$

$$Y_{11} = \frac{(216.70 - 186.29)}{12}$$

$$Y_{11} = \frac{30.41}{12} = 2.53$$

Therefore substitute the value of $Y_{11} = 2.53$ into the table and compute for the estimated value of the next missing data, Y_{24}. The new table is given below:

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	*2.53*	5.0	11.2	15.7	34.43
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*8.50*	16.9	38.0
Treatment Totals	16.83	36.4	56.6	70.1	179.93

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 179.93)}{(5-1)(4-1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 179.93)}{(4)(3)}$$

$$Y_{24} = \frac{((78 + 280.4) - 179.93)}{12}$$

$$Y_{24} = \frac{((358.4 - 179.93)}{12}$$

$$Y_{24} = \frac{178.47}{12} = 14.87$$

The estimated missing data value for $Y_{24} = 14.87$. This value is then be substituted into the table for the subsequent estimation of the value for the missing data (Y_{43}).

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	*2.53*	5.0	11.2	15.7	34.43
Palm nut shell [B]	1.6	4.8	9.2	*14.9*	30.47
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	Y_{43}	16.9	29.5
Treatment Totals	16.83	36.4	48.1	84.97	186.3

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{43} = \frac{(5(29.5) + 4(48.1) - 186.3)}{(5-1)(4-1)}$$

$$Y_{43} = \frac{((147.5 + 192.4) - 186.3)}{(4)(3)}$$

$$Y_{43} = \frac{(339.9 - 186.3)}{12}$$

$$Y_{43} = \frac{(339.9 - 186.3)}{12}$$

$$Y_{43} = \frac{(153.6)}{12} = 12.8$$

The estimated value for the missing data $Y_{43} = 12.8$. This is therefore substituted in the table and the value of the first estimated value Y_{11} taken out from the table and freshly computed for to ascertain whether the value will change or not.

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	*14.9*	30.47
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*12.8*	16.9	42.3
Treatment Totals	14.3	36.4	60.9	84.97	196.57

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{11} = \frac{(5(31.9) + 4(14.3) - 196.57)}{(5-1)(4-1)}$$

$$Y_{11} = \frac{((159.5 + 57.2) - 196.57)}{12}$$

$$Y_{11} = \frac{(216.70 - 196.57)}{12}$$

$$Y_{11} = \frac{20.13}{12} = 1.7$$

Since Y_{11} has changed from 2.53 to 1.7, it means one must continue computing until constant values are obtained.

So we continue to compute for Y_{24} by replacing the value of $Y_{11} = 1.7$

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	*1.7*	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*12.8*	16.9	42.3
Treatment Totals	16.0	36.4	60.9	70.1	183.4

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 183.4)}{(5-1)(4-1)}$$

$$Y_{24} = \frac{((78 + 280.4) - 183.4)}{(4)(3)}$$

$$Y_{24} = \frac{(358.4 - 183.4)}{12}$$

$$Y_{24} = \frac{175}{12} = 14.58$$

The new value for $Y_{24} = 14.58$, which has also changed from 14.9 to 14.58, so we substitute this value in the table:

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	*1.7*	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	*14.58*	30.18
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	Y_{43}	16.9	29.5
Treatment Totals	16.0	36.4	48.1	84.68	185.18

$$Y_{ij} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{43} = \frac{(5(29.5) + 4(48.10) - 185.18)}{(5-1)(4-1)}$$

$$Y_{43} = \frac{((147.5 + 192.4) - 185.18)}{(4)(3)}$$

$$Y_{43} = \frac{(339.9 - 185.18)}{12}$$

$$Y_{43} = \frac{(154.72)}{12} = 12.9$$

Therefore the missing data for $Y_{43} = 12.89$, which is change from 12.8 to 12.9 appropriately equal.

We proceed to compute for the Y_{11} by substituting the value of $Y_{43} = 12.9$

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	*14.58*	30.18
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*12.9*	16.9	42.4
Treatment Totals	14.3	36.4	61	84.68	196.38

$$Y_{11} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{11} = \frac{(5(31.9) + 4(14.3) - 196.38)}{(5-1)(4-1)}$$

$$Y_{11} = \frac{((159.5 + 57.2) - 196.38)}{(4)(3)}$$

$$Y_{11} = \frac{(216.7 - 196.38)}{12}$$

$$Y_{11} = \frac{20.32}{12} = 1.7$$

Therefore $Y_{11} = 1.7$, This value equals the previous hence suggest the right value for $Y_{11} = 1.7$, we then check for Y_{24} to see whether the estimated missing value remains the same as the previous value computed by substituting the value $Y_{11} = 1.7$.

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	*1.7*	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*12.9*	16.9	42.4
Treatment Totals	16	36.4	61	70.1	183.5

$$Y_{24} = \frac{(rB + tT - G)}{(r-1)(t-1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 183.5)}{(5-1)(4-1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 183.5)}{(4)(3)}$$

$$Y_{24} = \frac{((78 + 280.4) - 183.5)}{12}$$

$$Y_{24} = \frac{((358.4 - 183.5)}{12}$$

$$Y_{24} = \frac{(174.9)}{12} = 14.58$$

Since the value of the missing data $Y_{24} = 14.58$, same as the previous computed value for Y_{24}. It thus confirms that all the missing data values have been accurately estimated hence the table can now be completed and used for the Analysis of Variance (ANOVA).

Thus the completed table is shown below:

	Treatments (weights of the waste types) to be Pyrolysed				
Blocks	**10 kg [A]**	**20kg [B]**	**30kg [C]**	**40kg [D]**	**Block Totals**
Coconut waste fruits (shell and fibre) [A]	*1.7*	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	*14.58*	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	*12.9*	16.9	42.4
Treatment Totals	16	36.4	61	70.1	183.5

Bibliography

[1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.

[2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), Handbook on research on teaching. Chicago: Rand McNally & Co.

[3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.

[4] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs*. John Willey & Sons, N.Y..

[5] Le Clerg, E. L., Leonard, W. H., Clark A. G. (1966). *Field Plot Technique*. Burgess Pub. Co., Minn., USA.

Chapter 4

Latin Square Design

4

Latin Square Design

Felix Kutsanedzie[1] *; Sylvester Achio[1]; Edmund Ameko[1]; George Kutsanedzie[2]; Diaba Kwasi Selassie[3]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Project Office of Electricity Company of Ghana, Accra, Ghana

[3]Anglican University College of Technology, Nkorazan, Sunyani, Ghana

Abstract

The Latin Square Design (LSD) is one of the known experimental designs used in analysis designed experiments. Though this design exists, most researchers are unfamiliar with its usage. The prime aim of analysis experiment using the various known experimental designs is to reduce error as much as possible, or eliminate them altogether. This paper thus explain the concept of how the LSD is used for analysing experiment as well as how it helps to reduce errors due to the differences that exist in two directions (rows and columns) often referred to as double blocking.

Keywords

Design, Double Blocking, Experiment, Analysis

4.1 Introduction

In an RCBD, blocking is used to place all treatments in groups that are similar or homogenous so as to reduce the error due to variations as much as possible among the various treatments being handled. In RCBD there is single blocking adopted – only treatments are blocked. However with the Latin Square Design (LSD) blocking is done – blocking by row and column. The LSD is set such that each treatment is found in both a row and a column. Thus the design typically assumes a square shape. The number of experimental units is denoted by n^2, where n is the number of treatments or blocks or replications. It presupposes that in LSD, the number of blocks or replication used must be equal to the number of treatments being handled in order for each treatment to be found in the each block.

If a researcher needs to analysis an experiment using the Latin Square Design, the treatments must be arranged in the rows and columns in such a way that the major sources come from them.

Below are the types of LSD, and they are limited by the number of treatments.

3 X 3 LSD		
A	C	B
B	A	C
C	B	A

4X4 LSD			
A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

5X5 LSD				
A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A

It can be seen from the designs above that the treatments are all arranged in such a way that each column and row contain all the treatments be experimented on. Thus the number and type of treatments in a row are equal to those in a column.

Assuming a researcher wants to extract milk from five different plants such Tiger nut (T), coconut (Co), soyabean (So), cashew nut (Ca) and shea nut (Sh) to ascertain whether there are significance differences between the volume of milk in cubic centimetres that can be extracted from a kilogramme of each of the fruits; the following 5 x 5 Latin Square Design can be drawn:

5X5 LSD					
	Milk Extracted in cm^3/kg of treatment				
	Column1	**Column2**	**Column3**	**Column4**	**Column5**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3
Row2	Co = 5.5	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6

The first step: Stating of hypothesis

For Treatments:

H_o: All the means of the treatments are equal

H_A: At least one of the means of the treatments (t) is different

H_o: $\mu trt_A = \mu trt_B = \mu trt_C = \mu trt_D = \mu trt_E$

H_1: $\mu trt_A \neq \mu trt_B = \mu trt_C = \mu trt_D = \mu trt_E$

For Columns:

H_o: All the means of the Columns (C) are equal

H_A: At least one of the means of the Columns (C) is different

H_0: $\mu C1 = \mu C2 = \mu C3 = \mu C4 = \mu C5$

H_1: $\mu C1 \neq \mu C2 = \mu C3 = \mu C4 = \mu C5$

For Rows:

H_o: All the means of the Rows (R) are equal

H_A: At least one of the means of the Rows (R) is different

H_0: $\mu R1 = \mu R2 = \mu R3 = \mu R4 = \mu R5$

H_1: $\mu R1 \neq \mu R2 = \mu R3 = \mu R4 = \mu R5$

The second step: Calculating the Correction Factor (CF)

$$GT = \sum_{i=1,j=1}^{n} (T, Co, Ca, So, Sh) = \text{sum of all observations on experimental units}$$

$$GT = \sum_{i=1,j=1}^{n=5} (5.3 + 5.0 + \ldots 5.6) = 108.6$$

$$CF = \frac{(GT)^2}{N}$$

$$CF = \frac{(108.6)^2}{25} = \frac{11793.96}{25} = 471.76$$

where GT= grand Total, N=total number of observation, CF=Correction factor.

The third step: Calculating the Total Sum of Squares (TSS)

$$TSS = \sum_{i=1,j=1}^{n=5} [(T_{1,1})^2 + (Co_{1,2})^2 + \ldots . (T_{5,5})^2] - CF$$

$$TSS = \sum_{i=1,j=1}^{n=5} [(5.3)^2 + (5.0)^2 + \ldots . (5.6)^2] - 471.96$$

$$TSS = 494.74 - 471.96 = 22.98$$

CT	22	21.3	22.1	20.1	23.1	108.6
RT	21.6	21.6	20.7	21.8	22.9	108.6
T_T	5.3	5	4.9	4.5	5.6	25.3
Co_T	5.5	5	6	5.4	5.9	27.8
Ca_T	3.9	4.8	4.1	3.9	4.2	20.9
So_T	4.2	3.6	4	3.9	4.1	19.8
Sh_T	3.1	2.9	3.1	2.4	3.3	14.8

The fourth step: Calculating the Treatment Sum of Squares (TrtSS)

$$Trt\ SS = \sum (\frac{(T_T)^2 + (Co_T)^2 + (Ca_T)^2 + (So_T)^2 + (Sh_T)^2}{t}) - CF$$

$$Trt\ SS = \sum (\frac{(25.3)^2 + (27.8)^2 + (20.9)^2 + (19.8)^2 + (14.8)^2}{5}) - 471.96$$

$$Trt\ SS = \sum \left(\frac{640.09 + 772.84 + 436.81 + 392.04 + 219.04}{5}\right) - 471.96$$

$$Trt\ SS = \sum (\frac{2460.82}{5}) - 471.96$$

$$Trt\ SS = 492.16 - 471.96 = 20.41$$

where t=number of treatments, T_T=Treatment T total, Sh_T=Treatment Sh total, Co_T=Treatment Co total, Ca_T=Treatment Ca total, So_T=Treatment So total.

The sixth step: Calculating the Rows Sum of Squares (RSS)

$$RSS = \sum\left(\frac{(R1_T)^2 + (R2_T)^2 + (R3_T)^2 + (R4_T)^2 + (R5_T)^2}{R}\right) - CF$$

$$RSS = \sum\left(\frac{(21.6)^2 + (21.6)^2 + (20.7)^2 + (20.1)^2 + (23.1)^2}{5}\right) - 471.96$$

$$RSS = \sum\left(\frac{466.56 + 466.56 + 428.49 + 475.24 + 524.41}{5}\right) - 471.96$$

$$RSS = \sum\left(\frac{2361.26}{5}\right) - 471.96$$

$$RSS = 472.25 - 471.96$$

$$RSS = 0.49$$

where R=number of rows, $R1_T$=Treatment Row 1 total, $R2_T$=Treatment Row 2 total, $R3_T$=Treatment Row 3 total, $R4_T$=Treatment Row 4 total, $R5_T$=Treatment Row 5 total.

The seventh step: Calculating the Column Sum of Squares (CSS)

$$CSS = \sum\left(\frac{(C1_T)^2 + (C2_T)^2 + (C3_T)^2 + (C4_T)^2 + (C5_T)^2}{R}\right) - CF$$

$$CSS = \sum\left(\frac{(22)^2 + (21.3)^2 + (22.1)^2 + (20.1)^2 + (23.1)^2}{R}\right) - 471.96$$

$$CSS = \sum\left(\frac{(484) + (453.69) + (488.41) + (404.01) + (533.61)}{5}\right) - 471.96$$

$$CSS = \sum\left(\frac{2363.72}{5}\right) - 471.96$$

$$CSS = 472.74 - 471.96$$

$$CSS = 0.99$$

where C=number of rows, $C1_T$=Treatment Column 1 total, $C2_T$=Treatment Column 2 total, $C3_T$=Treatment Column 3 total, $C4_T$=Treatment Column 4 total, $C5_T$=Treatment Column 5 total.

The eighth step: Calculating the Error Sum of Squares (ESS)

$$ESS = TSS - (TrtSS + RSS + CSS)$$

$$ESS = 22.98 - (20.41 + 0.49 + 0.99)$$

$$ESS = (22.98 - 21.89)$$

$$ESS = 1.1$$

where ESS= Error Sum of Squares or unexplained sum of squares.

The ninth step: Completing the ANOVA table

ANOVA TABLE

Sources of Variation	*df*	*SS*	*MS*	*Fcal.*	*Fcrit. (5%)*
Treatment	t - 1 = 5 - 1 = 4	20.41	5.1025	46.39	3.48
Row	R – 1 = 5 – 1 = 4	0.49	0.1225	1.11	3.48
Column	C - 1 = 5 – 1 = 4	0.99	0.2475	2.25	3.48
Error	[C x R -1] –[(t-1)+(R-1)+(C-1)] =24-16 =10	1.1	0.11		
Total	(C X R) – 1 = 25 -1 =24	22.98			
Fcrit (5%) @ df of 4,10	= 3.48				
Fcrit (1%) @ df of 4,10	= 5.99				

Tenth step: looking up the F-critical table for the critical values

$$Fcrit(1\%) at\ df\ of\ 4,10\ for\ Treatments = 5.99$$

$$Fcrit(1\%) at\ df\ of\ 4,10\ for\ Rows = 5.99$$

$$Fcrit(1\%) at\ df\ of\ 4,10\ for\ Columns = 5.99$$

$$Fcrit(5\%) at\ df\ of\ 4,10\ for\ Treatments = 3.48$$

$$Fcrit(5\%) at\ df\ of\ 4{,}10\ for\ Rows = 3.48$$

$$Fcrit(5\%) at\ df\ of\ 4{,}10\ for\ Columns = 3.48$$

These values for the treatments, rows and columns can be obtained from the F-critical tables below and the values highlighted.

Critical values of F for the 0.05 significance level:

	1	2	3	4	5	6
1	161.45	199.50	215.71	224.58	230.16	233.99
2	18.51	19.00	19.16	19.25	19.30	19.33
3	10.13	9.55	9.28	9.12	9.01	8.94
4	7.71	6.94	6.59	6.39	6.26	6.16
5	6.61	5.79	5.41	5.19	5.05	4.95
6	5.99	5.14	4.76	4.53	4.39	4.28
7	5.59	4.74	4.35	4.12	3.97	3.87
8	5.32	4.46	4.07	3.84	3.69	3.58
9	5.12	4.26	3.86	3.63	3.48	3.37
10	4.97	4.10	3.71	3.48	3.33	3.22
11	4.84	3.98	3.59	3.36	3.20	3.10
12	4.75	3.89	3.49	3.26	3.11	3.00
13	4.67	3.81	3.41	3.18	3.03	2.92
14	4.60	3.74	3.34	3.11	2.96	2.85

Critical values of F for the 0.01 significance level:

	1	2	3	4	5	6
1	4052.19	4999.52	5403.34	5624.62	5763.65	5858.97
2	98.50	99.00	99.17	99.25	99.30	99.33
3	34.12	30.82	29.46	28.71	28.24	27.91
4	21.20	18.00	16.69	15.98	15.52	15.21
5	16.26	13.27	12.06	11.39	10.97	10.67
6	13.75	10.93	9.78	9.15	8.75	8.47
7	12.25	9.55	8.45	7.85	7.46	7.19
8	11.26	8.65	7.59	7.01	6.63	6.37
9	10.56	8.02	6.99	6.42	6.06	5.80
10	10.04	7.56	6.55	5.99	5.64	5.39
11	9.65	7.21	6.22	5.67	5.32	5.07
12	9.33	6.93	5.95	5.41	5.06	4.82
13	9.07	6.70	5.74	5.21	4.86	4.62
14	8.86	6.52	5.56	5.04	4.70	4.46

Eleventh Step: Making the decisions and Conclusions.

Treatments

$$Fcal\ (4{,}10) = 46.39 > Fcrit\ (3.48)\ at\ 5\%$$

$$Fcal\ (4{,}10)\ = 46.39 > Fcrit\ (5.99)\ at\ 1\%$$

Since the F-calculated value (46.39) for the treatments at both 1% and 5% level of significance respectively are greater that the F-critical value (5.99) and

(3.48), the null hypothesis is rejected on the grounds that there is enough evidence to suggest that there is significant difference between the treatments. Hence we fail to reject the alternate hypothesis.

Columns

$$Fcal\ (4{,}10)\ = 2.25 < Fcrit\ (3.48)\ at\ 5\%$$

$$Fcal\ (4{,}10) = 2.25 < Fcrit\ (5.99)\ at\ 1\%$$

F calculated value (2.25) for the columns at both 1% and 5% level of significance respectively lesser than the f critical value (3.48) and (5.99), we fail to reject the null hypothesis and say there is no significant difference in the columns, hence no justification for blocking by columns.

Rows

$$Fcal\ (4{,}10) = 1.11 < Fcrit\ (3.48)\ at\ 5\%$$

$$Fcal\ (4{,}10) = 1.11 < Fcrit\ (5.99)\ at\ 1\%$$

F calculated value (1.11) for the rows at both 1% and 5% level of significance respectively lesser than the F-critical value (3.48) and (5.99), we fail to reject the null hypothesis and say there is no significant difference in the rows, hence no justification for blocking by rows.

Since there is no justification for blocking by rows and columns, it presupposes that the LSD is not the right design for the experiment in question. Since the treatments are significantly different, there is the need to perform the Lsd (Least significance difference) test or the Duncan Multiple Range Test (DMRT).

Handling Missing Data under Latin Square Design (LSD).

In every experimental design data can be missing destruction or damage of living things being used as experimental materials, improper allocation of

treatments, loss of materials due to harvesting and processing of materials and illogical data that cannot be considered as reliable results due to how extreme the values are.

For single missing data under the LSD, the formula below can be used:

$$X_O = \left[\left(\frac{t(R_O + C_O + T_O) - 2G_O}{(t-1)(t-2)}\right)\right]$$

X_O=missing data observed, R_O=total observed row values with missing data, C_O=total observed column values with missing data, T_O=total observed treatment values with missing data, G_O=grand total of all observed value, t=number of treatments.

Formula for the Computation of Bias Value

$$\beta_0 = \frac{[G_O - R_O - C_O - T_O(t-1)]^2}{[(t-1)(t-2)]^2}$$

β_0= bias value

For adjustment of the analysis based on the estimation of the missing data the following is done

1. The value one (1) is substracted from the total degree of freedom and the unexplained or error degree of freedom.

2. The biased value must be computed and subtracted from the Total Sum of Squares and Treatment Sum of Squares and not that of the column and the row.

Taking the 5 x 5 LSD table below, the missing that can be computed for as follows:

5X5 LSD					
Milk Extracted in cm^3/kg of treatment					
	Column1	**Column2**	**Column3**	**Column4**	**Column5**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3
Row2	$Co = Y_{21}$	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	$Co = Y_{21}$	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	16.1
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9	20.7
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6	22.9
Column Total	16.5	21.3	22.1	20.1	23.1	103.1

$$T_o(Co) = Co_{12} + Co_{35} + Co_{43} + Co_{54}$$

$$T_o(Co) = 5.0 + 6.0 + 5.4 + 5.9 = 22.3$$

$$G_o = \sum_{ij}^{n=5} Y_{ij} = Y_{11} + Y_{12} + \ldots Y_{55} = 5.3 + 5.0 + \ldots 5.6 = 103.1$$

$$R_o = Y_{22} + Y_{23} + \ldots Y_{25} = 5.0 + 3.1 + 3.9 + 4.1 = 16.1$$

$$C_o = Y_{11} + Y_{31} + \ldots Y_{51} = 5.3 + 3.9 + 4.2 + 3.1 = 16.5$$

$$Y_{21} = \left[\left(\frac{t(R_O + C_O + T_O) - 2G_O}{(t-1)(t-2)}\right)\right]$$

$$Y_{21} = \left[\left(\frac{5(16.1 + 16.5 + 22.3) - 206.2)}{(5-1)(5-2)}\right)\right]$$

$$Y_{21} = \left[\left(\frac{5(54.4) - 206.2}{12}\right)\right]$$

$$Y_{21} = \left[\left(\frac{274.5 - 206.2}{12}\right)\right]$$

$$X_0 = \left[\left(\frac{68.3}{12}\right)\right]$$

$$X_0 = [5.69]$$

The bias is now computed using the formula:

$$\beta_0 = \frac{[G_0 - R_0 - C_0 - T_0(t-1)]^2}{[(t-1)(t-2)]^2}$$

$$\beta_0 = \frac{[103.1 - 16.1 - 16.5 - 22.3(5-1)]^2}{[(5-1)(5-2)]^2}$$

$$\beta_0 = \frac{[103.1 - 16.1 - 16.5 - 22.3(4)]^2}{[(4)(3)]^2}$$

$$\beta_0 = \frac{[103.1 - 16.1 - 16.5 - 89.2]^2}{[12]^2}$$

$$\beta_0 = \frac{[-18.7]^2}{[12]^2}$$

$$\beta_0 = \frac{[349.69]^2}{[144]^2}$$

$$\beta_0 = \frac{[349.69]}{[144]}$$

$$\beta_0 = 2.43$$

Now the table can be completed with the estimated missing data in order to do the Analysis of Variance (Anova) table

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	$Co = 5.69$	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9	20.7
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6	22.9
Column Total	22.19	21.3	22.1	20.1	23.1	108.79
CT	**22.19**	**21.3**	**22.1**	**20.1**	**23.1**	**108.79**
RT	21.6	21.79	20.7	21.8	22.9	108.79
T_T	5.3	5	4.9	4.5	5.6	25.3
Co_T	5.69	5	6	5.4	5.9	27.99
Ca_T	3.9	4.8	4.1	3.9	4.2	20.9
So_T	4.2	3.6	4	3.9	4.1	19.8
Sh_T	3.1	2.9	3.1	2.4	3.3	14.8

The first step: Stating of the Hypothesis.

The hypothesis stated remains the same.

The second step: Calculating the Correction Factor (CF)

$GT = \sum_{i=1,j=1}^{n}(T, Co, Ca, So, Sh) =$ sum of all observations on experimental units

$$GT = \sum_{i=1,j=1}^{n=5} (5.3 + 5.0 + \ldots 5.6) = 108.79$$

$$CF = \frac{(GT)^2}{N}$$

$$CF = \frac{(108.79)^2}{25} = \frac{11835.26}{25} = 473.41$$

where GT= Grand Total, N=total number of observation, CF=Correction factor.

The third step: Calculating the Total Sum of Squares (TSS)

$$TSS = \sum_{i=1,j=1}^{n=5} [(T_{1,1})^2 + (Co_{1,2})^2 + \ldots . (T_{5,5})^2] - CF$$

$$TSS = \sum_{i=1,j=1}^{n=5} [(5.3)^2 + (5.0)^2 + \ldots . (5.6)^2] - 473.41$$

$$TSS = 496.87 - 473.41 = 23.46$$

The fourth step: Calculating the Treatment Sum of Squares (TrtSS)

$$Trt\ SS = \sum (\frac{(T_T)^2 + (Co_T)^2 + (Ca_T)^2 + (So_T)^2 + (Sh_T)^2}{t}) - CF$$

$$Trt\ SS = \sum \left(\frac{(25.3)^2 + (27.99)^2 + (20.9)^2 + (19.8)^2 + (14.8)^2}{5}\right) - 473.41$$

$$Trt\ SS = \sum \left(\frac{640.09 + 783.44 + 436.81 + 392.04 + 219.04}{5}\right) - 473.41$$

$$Trt\ SS = \sum (\frac{2471.42}{5}) - 473.41$$

$$Trt\ SS = 494.28 - 473.41 = 20.87$$

where t=number of treatments, T_T=Treatment T total, Sh_T=Treatment Sh total, Co_T=Treatment Co total, Ca_T=Treatment Ca total, So_T=Treatment So total.

The sixth step: Calculating the Rows Sum of Squares (RSS)

$$RSS = \sum (\frac{(R1_T)^2 + (R2_T)^2 + (R3_T)^2 + (R4_T)^2 + (R5_T)^2}{R}) - CF$$

$$RSS = \sum (\frac{(21.6)^2 + (21.79)^2 + (20.7)^2 + (21.8)^2 + (22.9)^2}{5}) - 473.41$$

$$RSS = \sum (\frac{466.56 + 474.80 + 428.49 + 475.24 + 524.41}{5}) - 473.41$$

$$RSS = \sum(\frac{2369.50}{5}) - 473.41$$

$$RSS = 473.90 - 473.41$$

$$RSS = 0.49$$

where R=number of rows, $R1_T$=Treatment Row 1 total, $R2_T$=Treatment Row 2 total, $R3_T$=Treatment Row 3 total, $R4_T$=Treatment Row 4 total, $R5_T$=Treatment Row 5 total.

The seventh step: Calculating the Column Sum of Squares (CSS)

$$CSS = \sum(\frac{(C1_T)^2 + (C2_T)^2 + (C3_T)^2 + (C4_T)^2 + (C5_T)^2}{R}) - CF$$

$$CSS = \sum(\frac{(22.19)^2 + (21.3)^2 + (22.1)^2 + (20.1)^2 + (23.1)^2}{R}) - 473.41$$

$$CSS = \sum(\frac{(492.40) + (453.69) + (488.41) + (404.01) + (533.61)}{5}) - 473.41$$

$$CSS = \sum(\frac{2372.12}{5}) - 473.41$$

$$CSS = 474.42 - 473.41$$

$$CSS = 1.01$$

where C=number of rows, $C1_T$=Treatment Column 1 total, $C2_T$=Treatment Column 2 total, $C3_T$=Treatment Column 3 total, $C4_T$=Treatment Column 4 total, $C5_T$=Treatment Column 5 total.

The eighth step: Calculating the Error Sum of Squares (ESS)

$$ESS = TSS - (TrtSS + RSS + CSS)$$

$$ESS = 23.46 - (20.87 + 0.49 + 1.01)$$

$$ESS = (23.46 - 22.38)$$

$$ESS = 1.1$$

where ESS=Error Sum of Squares or unexplained sum of squares.

The ninth step: Making the necessary Adjustments in the Analysis and Completing the ANOVA table

Making the necessary Adjustments in the Analysis

The following adjustments must be done in completing the Anova table:

1. Deduct the value (1) from the total degree of freedom and the unexplained or error degree of freedom
2. Subtract the computed bias from the Total Sum of Squares (TSS) and Treatment Sum of Squares (TrtSS)

Adjustments

$$Total\ df = [\mathrm{C\ x\ R}] - 1 - 1 = 25 - 1 - 1 = 23$$

$$Unexplained\ or\ error = [\mathrm{C\ x\ R}\ - 1] - [(\mathrm{t} - 1) + (\mathrm{R} - 1) + (\mathrm{C} - 1)] - 1$$

$$= 24 - 16 - 1\ = 9$$

$$\text{Adjusted TSS} = 23.45 - \beta_0$$

$$where\ \beta_0 = 2.43$$

$$\text{Adjusted TSS} = 23.45 - 2.43 = 21.02$$

$$\text{Adjusted TrtSS} = 20.87 - \beta_0$$

$$\text{Adjusted TrtSS} = 20.87 - 2.43 = 18.44$$

ANOVA TABLE					
Sources of Variation	*df*	*SS*	*MS*	*Fcal.*	*Fcrit. (5%)*
Treatment	4	18.44	4.61	38.78	3.63
Row	4	0.5	0.13	1.05	3.63
Column	4	1.01	0.25	2.12	3.63
Error	9	1.07	0.12		
Total	23	21.02			
Fcrit (5%) @ df of 4,9	3.63				
Fcrit (1%) @ df of 4,9	6.42				

Tenth step: looking up the F-critical table for the critical values

$$Fcrit(1\%) at\ df\ of\ 4{,}9\ for\ Treatments = 6.42$$

$$Fcrit(1\%) at\ df\ of\ 4{,}9\ for\ Rows = 6.42$$

$$Fcrit(1\%) at\ df\ of\ 4{,}9\ for\ Columns = 6.42$$

$$Fcrit(5\%) at\ df\ of\ 4{,}9\ for\ Treatments = 3.63$$

$$Fcrit(5\%) at\ df\ of\ 4{,}9\ for\ Rows = 3.63$$

$$Fcrit(5\%) at\ df\ of\ 4{,}9\ for\ Columns = 3.63$$

The Eleventh step: Making the Decision and Conclusions

The values obtained from the F-critical table compared with the F-values calculated for the Treatments, Rows and Columns reveals the decisions and conclusions remain the same as the table without the missing data. This proves the missing data value has been accurately estimated.

4.2 Handling More than One Missing Data Under LSD

The procedure is the same as done in the case of RCBD but the formula for estimating the missing data values differ. For LSD, the following formula are used:

$$Y_{ij} = \frac{(T_m + C_m + R_m)}{2}$$

This formula can apply only to LSD when estimating average missing data

where Y_{ij}=estimated average missing data value, T_m=the mean value of treatment with the missing data, C_m=the mean value of Column with the missing data, R_m=the mean value of Row with the missing data.

$$Y_{ij} = \left[\left(\frac{t(R_O + C_O + T_O) - 2G_O}{(t-1)(t-2)}\right)\right]$$

Y_{ij}=missing data observed, R_O=total observed row values with missing data, C_O=total observed column values with missing data, T_O=total observed treatment values with missing data, G_O=grand total of all observed value, t=number of treatments.

$$\beta_0 = \frac{[G_O - R_O - C_O - T_O(t-1)]^2}{[(t-1)(t-2)]^2}$$

β_0= bias value

For example to compute for the missing data under LSD as seen in the table above, the approach below is used:

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	*Ca = 3.9*	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co = Y_{35}*	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = Y_{52}*	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	17.2	98.09

We first find the average estimated missing data using the formula:

$$Y_{ij} = \frac{(T_m + C_m + R_m)}{3}$$

For $Y_{52} = \frac{(T_m + C_m + R_m)}{3}$

$$C_m = \frac{\sum Y_{11} + Y_{22} + Y_{32} + Y_{42}}{5} = \frac{16.5}{5} = 3.3$$

$$R_m = \frac{\sum Y_{51} + Y_{53} + Y_{54} + Y_{55}}{5} = \frac{18.1}{5} = 3.62$$

$$T_m = \frac{\sum Y_{31} + Y_{13} + Y_{24} + Y_{45}}{5}$$

$$T_m = \frac{\sum(3.9 + 4.1 + 3.9 + 4.2)}{5} = \frac{16.1}{5}\ 3.22$$

$$Y_{52} = \frac{(3.22 + 3.3 + 3.62)}{3} = \frac{10.14}{3} = 3.38$$

For $Y_{35} = \frac{(T_m + C_m + R_m)}{2}$

$$C_m = \frac{\sum Y_{15} + Y_{25} + Y_{45} + Y_{55}}{5} = \frac{17.2}{5} = 3.44$$

$$R_m = \frac{\sum Y_{31} + Y_{32} + Y_{33} + Y_{34}}{5} = \frac{14.8}{5} = 2.96$$

$$T_m = \frac{\sum Y_{21} + Y_{12} + Y_{43} + Y_{54}}{5} = \frac{5.69 + 5.0 + 6.0 + 5.4}{5} = 4.42$$

$$Y_{35} = \frac{(4.42 + 3.44 + 2.96)}{2} = \frac{10.82}{2} = 3.61$$

Since the estimated average values of the missing data Y_{52} and Y_{35} have been computed, the values of each of this missing data can now be estimated by completing the table with one of the average estimated value where finding the estimated value of the other one.

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co = 3.61*	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca* = Y_{52}	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	20.81	101.7

In order to estimate the value of Y_{52}, the average estimated value of Y_{35} is substituted in the table as seen.

$$Y_{52} = \left[\left(\frac{t(R_O + C_O + T_O) - 2G_O}{(t-1)(t-2)}\right)\right]$$

$$R_O = Y_{51} + Y_{53} + \ldots + Y_{55}$$

$$R_O = 3.1 + 4.0 + 5.4 + 5.6 = 18.1$$

$$C_O = Y_{12} + Y_{22} + \ldots + Y_{52}$$

$$C_O = 5.0 + 5.0 + 3.6 + 2.9 = 16.5$$

$$T_O(Ca) = Y_{51} + Y_{53} + \ldots + Y_{55}$$

$$T_O(Ca) = 3.9 + 4.1 + 3.9 + 4.2 = 16.1$$

$$Y_{52} = \left[\left(\frac{5(18.1 + 16.5 + 16.1) - 2(101.7)}{(5-1)(5-2)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{5(50.7) - 203.4}{(4)(3)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{253.5 - 203.4}{12}\right)\right]$$

$$Y_{52} = \left[\left(\frac{50.1}{12}\right)\right]$$

$$Y_{52} = [4.18]$$

To compute the estimate for Y_{35}, substitute the value of $Y_{52} = 4.18$ into the table as shown:

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	$Co = Y_{35}$	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = 4.18*	So = 4.0	Co = 5.4	T = 5.6	22.28
Column Total	22.19	20.68	22.1	20.1	17.2	102.27

$$Y_{35} = \left[\left(\frac{t(R_O + C_O + T_O) - 2G_O}{(t-1)(t-2)}\right)\right]$$

$$R_O = Y_{31} + Y_{32} + \ldots + Y_{34}$$

$$R_O = 3.9 + 3.6 + 4.9 + 2.4 = 14.8$$

$$C_O = Y_{15} + Y_{25} + \ldots + Y_{55}$$

$$C_O = 3.3 + 4.1 + 4.2 + 5.6 = 17.2$$

$$T_O(Ca) = Y_{21} + Y_{12} + \ldots + Y_{54}$$

$$T_O(Ca) = 5.69 + 5.0 + 6.0 + 5.4 = 22.09$$

$$Y_{35} = \left[\left(\frac{5(14.8 + 17.2 + 22.09) - 2(102.27)}{(5-1)(5-2)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{5(54.09) - 2(102.27)}{(4)(3)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{270.45 - 204.54}{12}\right)\right]$$

$$Y_{35} = \left[\left(\frac{65.91}{12}\right)\right]$$

$$Y_{35} = [5.50]$$

The estimated missing data for $Y_{35} = 5.50$. This is replaced in the table and the value of Y_{52} is estimated again. This step is repeated until a constant value is obtained for all the missing data before the accurate missing data values are obtained.

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co = 5.50*	20.30
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = Y_{52}*	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	22.70	103.9

$$Y_{52} = \left[\left(\frac{t(R_O + C_O + T_O) - 2G_O}{(t-1)(t-2)}\right)\right]$$

$$R_O = 16.5$$

$$C_O = 18.1$$

$$T_O(Ca) = 3.9 + 4.1 + 3.9 + 4.2 = 16.1$$

$$Y_{52} = \left[\left(\frac{5(16.5 + 18.1 + 16.1) - 2(103.9)}{(5-1)(5-2)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{5(50.7) - 2(103.9)}{(4)(3)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{253.5 - 207.8}{12}\right)\right]$$

$$Y_{52} = \left[\left(\frac{45.70}{12}\right)\right]$$

$$Y_{52} = [3.81]$$

Since the value for the previous estimated of $Y_{52} = 4.18$ is not the same the 3.81 as obtained now, the process must be repeated. We then substitute the value of $Y_{52} = 3.81$ into the table to estimate for Y_{35}

$$Y_{35} = \left[\left(\frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)}\right)\right]$$

$$R_0 = 14.8$$

$$C_0 = 17.2$$

$$T_0(Co) = 22.09$$

$$Y_{35} = \left[\left(\frac{5(14.8 + 17.2 + 22.09) - 2(101.27)}{(5-1)(5-2)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{5(54.09) - 2(101.27)}{(4)(3)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{270.45 - 207.8}{12}\right)\right]$$

$$Y_{35} = \left[\left(\frac{67.91}{12}\right)\right]$$

$$Y_{35} = [5.66]$$

Substitute $Y_{35} = 5.66$ into the table and compute for Y_{52}

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co = 5.66*	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	$Ca = Y_{52}$	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	17.2	103.75

$$Y_{52} = \left[\left(\frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)}\right)\right]$$

$$R_0 = 18.1$$

$$C_0 = 16.5$$

$$T_0(Ca) = 16.1$$

$$Y_{52} = \left[\left(\frac{5(18.1 + 16.5 + 16.1) - 2(103.75)}{(5-1)(5-2)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{5(50.7) - 2(103.75)}{(4)(3)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{253.5 - 207.5}{12}\right)\right]$$

$$Y_{52} = \left[\left(\frac{46}{12}\right)\right]$$

$$Y_{52} = [3.83]$$

Substitute $Y_{52} = 3.83$ into the table and compute for Y_{35}

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co* = Y_{35}	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = 3.83*	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	20.33	22.1	20.1	17.2	103.9

$$Y_{35} = \left[\left(\frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)}\right)\right]$$

$$R_0 = 14.8$$

$$C_0 = 17.2$$

$$T_0(Co) = 22.09$$

$$Y_{35} = \left[\left(\frac{5(14.8 + 17.2 + 22.09) - 2(101.92)}{(5-1)(5-2)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{5(54.09) - 2(101.92)}{(4)(3)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{270.45 - 203.84}{12}\right)\right]$$

$$Y_{35} = \left[\left(\frac{66.61}{12}\right)\right]$$

$$Y_{35} = [5.55]$$

Substitute $Y_{35} = 5.55$ into the table and compute for Y_{52}

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co = 5.55*	20.35
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = Y_{52}*	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	22.75	103.64

$$Y_{52} = \left[\left(\frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)}\right)\right]$$

$$R_0 = 18.1$$

$$C_0 = 16.5$$

$$T_0(Co) = 16.1$$

$$Y_{52} = \left[\left(\frac{5(18.1 + 16.5 + 16.1) - 2(103.64)}{(5-1)(5-2)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{5(50.7) - 2(103.64)}{(4)(3)}\right)\right]$$

$$Y_{52} = \left[\left(\frac{253.50 - 207.28}{12}\right)\right]$$

$$Y_{52} = \left[\left(\frac{46.22}{12}\right)\right]$$

$$Y_{52} = [3.85]$$

Substitute $Y_{52} = 5.55$ into the table and compute for Y_{35}

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	$Co = Y_{52}$	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = 3.85*	So = 4.0	Co = 5.4	T = 5.6	21.95
Column Total	22.19	20.35	22.1	20.1	22.75	101.94

$$Y_{35} = \left[\left(\frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)}\right)\right]$$

$$R_0 = 14.8$$

$$C_0 = 17.2$$

$$T_0(Co) = 22.09$$

$$Y_{35} = \left[\left(\frac{5(14.8 + 17.2 + 22.09) - 2(103.64)}{(5-1)(5-2)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{5(54.09) - 2(101.94)}{(4)(3)}\right)\right]$$

$$Y_{35} = \left[\left(\frac{270.45 - 203.88}{12}\right)\right]$$

$$Y_{35} = \left[\left(\frac{66.57}{12}\right)\right]$$

$$Y_{35} = [5.55]$$

Since the $Y_{35} = 5.55$ is the same as obtained previously, then it presupposes that the right missing data has been estimated for Y_{35}. Therefore the estimated missing data values for Y_{52} and Y_{35} are 3.85 and 5.55 respectively.

5X5 LSD						
Milk Extracted in cm^3/kg of treatment						
	Column1	**Column2**	**Column3**	**Column4**	**Column5**	**Row Total**
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	*Co = 5.69*	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	*Co = 5.55*	20.35
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	*Ca = 3.85*	So = 4.0	Co = 5.4	T = 5.6	21.95
Column Total	22.19	20.35	22.1	20.1	22.75	107.49
CT	**22.19**	**20.35**	**22.1**	**20.1**	**22.75**	**107.49**
RT	21.6	21.79	20.35	21.8	21.95	107.49
T_T	5.3	5	4.9	4.5	5.6	25.3
Co_T	5.69	5	6	5.4	5.55	27.64
Ca_T	3.9	3.85	4.1	3.9	4.2	19.95
So_T	4.2	3.6	4	3.9	4.1	19.8
Sh_T	3.1	2.9	3.1	2.4	3.3	14.8

The first step: Stating of the Hypothesis

The hypothesis stated remains the same.

The second step: Calculating the Correction Factor (CF)

$GT = \sum_{i=1,j=1}^{n}(T, Co, Ca, So, Sh)$ =sum of all observations on experimental units

$$GT = \sum_{i=1,j=1}^{n=5} (5.3 + 5.0 + \ldots 5.6) = 107.49$$

$$CF = \frac{(GT)^2}{N}$$

$$CF = \frac{(107.49)^2}{25} = \frac{11554.1}{25} = 462.16$$

where GT= Grand Total, N=total number of observation, CF=Correction factor.

The third step: Calculating the Total Sum of Squares (TSS)

$$TSS = \sum_{i=1,j=1}^{n=5} [(T_{1,1})^2 + (Co_{1,2})^2 + \ldots . (T_{5,5})^2] - CF$$

$$TSS = \sum_{i=1,j=1}^{n=5} [(5.3)^2 + (5.0)^2 + \ldots . (5.6)^2] - 462.16$$

$$TSS = 484.64 - 462.16 = 22.48$$

The fourth step: Calculating the Treatment Sum of Squares (TrtSS)

$$Trt\ SS = \sum(\frac{(T_T)^2+(Co_T)^2+(Ca_T)^2+(So_T)^2+(Sh_T)^2}{t}) - CF$$

$$Trt\ SS = \sum(\frac{(25.3)^2+(27.64)^2+(19.95)^2+(19.8)^2+(14.8)^2}{5}) - 462.16$$

$$Trt\ SS = \sum\left(\frac{640.09 + 763.97 + 398.0 + 392.04 + 219.04}{5}\right) - 462.16$$

$$Trt\ SS = \sum(\frac{2413.14}{5}) - 462.16$$

$$Trt\ SS = 482.63 - 462.16 = 20.47$$

where t=number of treatments, T_T=Treatment T total, Sh_T=Treatment Sh total, Co_T=Treatment Co total, Ca_T=Treatment Ca total, So_T=Treatment So total.

The sixth step: Calculating the Rows Sum of Squares (RSS)

$$RSS = \sum(\frac{(R1_T)^2+(R2_T)^2+(R3_T)^2+(R4_T)^2+(R5_T)^2}{R}) - CF$$

$$RSS = \sum(\frac{(21.6)^2+(21.79)^2+(20.35)^2+(21.8)^2+(21.95)^2}{5}) - 462.16$$

$$RSS = \sum(\frac{466.56+474.80+414.12+475.24+481.80}{5}) - 462.16$$

$$RSS = \sum(\frac{2312.53}{5}) - 462.16$$

$$RSS = 462.51 - 462.16$$

$$RSS = 0.35$$

where R=number of rows, $R1_T$=Treatment Row 1 total, $R2_T$=Treatment Row 2 total, $R3_T$=Treatment Row 3 total, $R4_T$=Treatment Row 4 total, $R5_T$=Treatment Row 5 total.

The seventh step: Calculating the Column Sum of Squares (CSS)

$$CSS = \sum\left(\frac{(C1_T)^2 + (C2_T)^2 + (C3_T)^2 + (C4_T)^2 + (C5_T)^2}{R}\right) - CF$$

$$CSS = \sum\left(\frac{(22.19)^2 + (20.35)^2 + (22.1)^2 + (20.1)^2 + (22.75)^2}{R}\right) - 462.16$$

$$CSS = \sum\left(\frac{(492.40) + (414.12) + (488.41) + (404.01) + (517.56)}{5}\right) - 462.16$$

$$CSS = \sum\left(\frac{2316.50}{5}\right) - 462.16$$

$$CSS = 463.30 - 462.16$$

$$CSS = 1.14$$

where C=number of rows, $C1_T$=Treatment Column 1 total, $C2_T$=Treatment Column 2 total, $C3_T$=Treatment Column 3 total, $C4_T$=Treatment Column 4 total, $C5_T$=Treatment Column 5 total.

The eighth step: Calculating the Error Sum of Squares (ESS)

$$ESS = TSS - (TrtSS + RSS + CSS)$$

$$ESS = 22.48 - (20.47 + 0.35 + 1.14)$$

$$ESS = 22.48 - 21.96)$$

$$ESS = 0.52$$

where ESS= Error Sum of Squares or unexplained sum of squares.

Now the bias can be computed using the formula:

Calculating bias using Ca as the missing data

$$\beta_0 = \frac{[G_0 - R_0 - C_0 - T_0(t-1)]^2}{[(t-1)(t-2)]^2}$$

β_O= bias value

$$G_0 = 103.64,\ C_0 = 16.5,\ R_0 = 18.1,\ T_0 = 16.1$$

$$\beta_0 = \frac{[103.64 - 18.1 - 16.5 - 16.1(5-1)]^2}{[(5-1)(5-2)]^2}$$

$$\beta_0 = \frac{[103.64 - 34.6 - 64.4)]^2}{[(4)(3)]^2}$$

$$\beta_0 = \frac{[103.64 - 99)]^2}{[(12)]^2}$$

$$\beta_0 = \frac{[4.64]^2}{[12]^2}$$

$$\beta_0 = \frac{[21.53]}{[144]} = 0.15$$

Calculating bias using Co as the missing data:

$$\beta_0 = \frac{[G_0 - R_0 - C_0 - T_0(t-1)]^2}{[(t-1)(t-2)]^2}$$

β_O= bias value

$$G_0 = 101.94,\ C_0 = 17.2,\ R_0 = 14.8,\ T_0 = 22.09$$

$$\beta_0 = \frac{[101.94 - 14.8 - 17.2 - 22.09(5-1)]^2}{[(5-1)(5-2)]^2}$$

$$\beta_0 = \frac{[101.94 - 32 - 88.36]^2}{[(4)(3)]^2}$$

$$\beta_0 = \frac{[101.94 - 20.36)]^2}{[(12)]^2}$$

$$\beta_0 = \frac{[-18.42]^2}{[12]^2}$$

$$\beta_0 = \frac{[339.30]}{[144]} = 2.36$$

The ninth step: Making the necessary Adjustments in the Analysis and Completing the ANOVA table.

Making the necessary Adjustments in the Analysis.

The following adjustments must be done in completing the Anova table:

1. Deduct the value (1) from the total degree of freedom and the unexplained or error degree of freedom.
2. Subtract the computed bias from the Total Sum of Squares (TSS) and Treatment Sum of Squares (TrtSS).

Adjustments

$$Total\ df = [\mathrm{C\ x\ R}] - 1 - 1 = 25 - 1 - 1 = 23$$

$$Unexplained\ or\ error = [\mathrm{C\ x\ R} - 1] - [(t-1) + (R-1) + (C-1)] - 1$$

$$= 24 - 16 - 1 = 9$$

Using Ca as treatment having the missing data

$$\text{Adjusted TSS} = 22.48 - \beta_0$$

$$where\ \beta_0 = 0.15$$

$$\text{Adjusted TSS} = 22.48 - 0.15 = 22.33$$

$$\text{Adjusted TrtSS} = 20.47 - \beta_0$$

$$\text{Adjusted TrtSS} = 20.47 - 0.15 = 20.32$$

Using Co as treatment having the missing data

$$\text{Adjusted TSS} = 22.48 - \beta_0$$

$$where\ \beta_0 = 2.36$$

$$\text{Adjusted TSS} = 22.48 - 2.36 = 20.12$$

$$\text{Adjusted TrtSS} = 20.47 - \beta_0$$

$$\text{Adjusted TrtSS} = 20.47 - 2.36 = 18.11$$

Anova table when as missing data is from Ca treatment:

ANOVA TABLE					
Sources of Variation	***df***	***SS***	***MS***	***Fcal.***	***Fcrit. (5%)***
Treatment	4	20.32	5.08	9.77	3.63
Row	4	0.35	0.09	0.17	3.63
Column	4	1.14	0.29	0.55	3.63
Error	9	0.52	0.06		
Total	23	22.33	0.97		
Fcrit (5%) @ df of 4,9	3.63				
Fcrit (1%) @ df of 4,9	6.42				

Anova table when as missing data is from Co treatment:

ANOVA TABLE					
Sources of Variation	***df***	***SS***	***MS***	***Fcal.***	***Fcrit. (5%)***
Treatment	4	18.11	4.53	8.71	3.63
Row	4	0.35	0.09	0.17	3.63
Column	4	1.14	0.29	0.55	3.63
Error	9	0.52	0.06		
Total	23	20.12	0.87		
Fcrit (5%) @ df of 4,9	3.63				
Fcrit (1%) @ df of 4,9	6.42				

Tenth step: looking up the F-critical table for the critical values

$$Fcrit(1\%) at\ df\ of\ 4,9\ for\ Treatments = 6.42$$

$$Fcrit(1\%) at\ df\ of\ 4,9\ for\ Rows = 6.42$$

$$Fcrit(1\%) at\ df\ of\ 4,9\ for\ Columns = 6.42$$

$$Fcrit(5\%) at\ df\ of\ 4,9\ for\ Treatments = 3.63$$

$$Fcrit(5\%) at\ df\ of\ 4{,}9\ for\ Rows = 3.63$$

$$Fcrit(5\%) at\ df\ of\ 4{,}9\ for\ Columns = 3.63$$

The Eleventh step: Making the Decision and Conclusions.

The values obtained from the F-critical table compared with the F-values calculated for the Treatments, Rows and Columns reveals the decisions and conclusions remain the same as the table without the missing data. This proves the missing data value has been accurately estimated.

Bibliography

[1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.

[2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), Handbook on research on teaching. Chicago: Rand McNally & Co.

[3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.

[4] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs*. John Willey & Sons, N.Y.

[5] Le Clerg, E. L., Leonard, W. H., Clark A. G. 1966. *Field Plot Technique*. Burgess Pub. Co., Minn., USA.

[6] Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

Chapter 5

Multifactorial Design

5

Multifactorial Design

Felix Kutsanedzie[1] *; Sylvester Achio[1]; Edmund Ameko[1]; Gyekye Appiah Lewis[1]; Diaba Kwasi Selassie[2]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Anglican University College of Technology, Nkorazan, Sunyani–Ghana

Abstract

A researcher may have two or more treatments to handle at a time and these treatments to be handled may or may not interact with each other at certain levels. The effects of these interactions cannot be established with the types of designs such as CRD, RCBD and LSD. The multifactorial design thus helps to establish whether the variations between each type of factor and their interactions are significantly different or not. This paper thus explains how the multifactorial design is used to establish whether there are any significant differences between the treatment types and their interactions.

Keywords

Treatments, Factorial, Factor, Interaction, Variations

5.1 Introduction

Designs like CRD, RCBD and LSD have been used variously under appropriate situations for the design and analysis of experiments. In every experiment, a researcher is bound to handle one or more treatments – that which is expected to be subjected to an experimental material or unit or plot. A factor is defined as a basic treatment and in terms of a multifactorial or factorial designs, a treatment in a sense is a combination of two or more levels of factors. This means that in considering two levels of factors or treatments like the amount of sugar used in the preparation of porridge (A) – half cup of sugar (A_1) and two cups of sugar (A_2); and the quantity of porridge prepared (Q) – twenty (20) litres (Q_1) and twenty-five (25) litres (Q_2) prepared. Thus in handling of factors and treatments at different levels, the researcher would be faced with how to deal or cope with multiple factors. However it must be noted that CRD, RCBD and LSD can be designed and contained in multifactorial designs. It is the appropriate design used in handling treatments with different levels in order to ascertain or establish whether there exist significant differences between the various factor levels or treatment levels as well as the interactions between the treatments.

The factorial experimental designs are more complicated because they are used for designing and analyzing many factors which can be observed on large experimental units and thus not suitable for designing simple experiments. It allows for greater precision when estimating the overall effects of factors as it helps to expose what is referred to as hidden replications. For instance when studying two factors say A and B with two levels A_1, A_2 and B_1, B_2 respectively, the researcher would observe the normal required plots such as A, B and AB and extra plots or units. In all, the researcher would observe these plots or units: A_1B_1, A_1B_2, A_2B_1, A_2B_2, (A_3B_3, this represents interaction between AB).

It must however be noted that the objectives of factorial experimental designs include testing of the main effects or factors and their interactions, and not the treatments. Hence the treatment Mean Squares (MSQ) are not found.

To consider a CRD, let us assume a researcher wants to study the time it takes for the same quantity of milk from cow (A1) and soyabean (A2) to ferment into yoghurt under two temperature conditions B1 and B2. When this experiment is replicated four times, it can be represented below:

5.2 CRD

	A1		A2	
REP	**B1**	**B2**	**B1**	**B2**
1	5.4	5.3	5.1	4.7
2	4.6	5.6	4.5	4.9
3	4.1	4.8	5.6	5.2
4	3.9	4.8	4.5	5.0

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypotheses

H_o: $\mu_{A1B1} = \mu_{A1B2} = \mu_{A2B2} = \mu_{A2B2}$

H_1: $\mu_{A1B1} \neq \mu_{A1B2} = \mu_{A2B2} = \mu_{A2B2}$

H_o: All the means of the treatments are equal

H_A: At least one of the means of the treatments is different

H_o: $\mu_{A1} = \mu_{A2}$

H_1: $\mu_{A1} \neq \mu_{A2}$

H_o: The means of the 2 levels of A are equal

H_A: The means of the 2 levels of A are unequal

H_o: $\mu_{B1} = \mu_{B2}$

H_1: $\mu_{B1} \neq \mu_{B2}$

H_o: The means of the 2 levels of B are equal

H_A: The means of the 2 levels of B are unequal

H_o: $\mu_{A \times B} = \mu_{A \times B}$

H_1: $\mu_{A \times B} \neq \mu_{A \times B}$

H_o: The effects of the interaction of two factors at the 2 levels are the same

H_A: The effects of the interaction of two factors at the 2 levels are different

	A_1		A_2		
REP	**B_1**	**B_2**	**B_1**	**B_2**	
1	5.4	5.3	5.1	4.7	20.5
2	4.6	5.6	4.5	4.9	19.6
3	4.1	4.8	5.6	5.2	19.7
4	3.9	4.8	4.5	5	18.2
	18	20.5	19.7	19.8	78

$$Tt_{A_1B_1} = \sum(5.4 + 4.6 + \cdots + 3.9) = 18$$

$$Tt_{A_1B_2} = \sum(5.3 + 5.6 + \cdots + 4.8) = 20.5$$

$$Tt_{A_2B_1} = \sum(5.1 + 4.5 + \cdots + 4.5) = 19.7$$

$$Tt_{A_2B_2} = \sum(4.7 + 4.9 + \cdots + 5) = 19.8$$

$$Tt_T = \sum\left(Tt_{A_1B_1} + Tt_{A_1B_2} + Tt_{A_2B_2} + Tt_{A_2B_2}\right) = 78$$

The second step is calculating of the Correction Factor:

$$CF = \frac{\sum(Tt_T)^2}{(r\ x\ t)} = \frac{(78)^2}{4\ x\ 4} = \frac{(78)^2}{16} = \frac{6084}{16} = 380.25$$

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum((5.4 + (5.3)^2 + (5.1)^2 + (4.7)^2 + (4.6)^2 + (5.6)^2 \ldots\ldots + (5)^2) - 380.25 = 3.63$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TrtSS = \sum\left(\frac{(Tt_{A_1B_1})^2 + (Tt_{A_1B_2})^2 + (Tt_{A_2B_2})^2 + (Tt_{A_2B_2})^2}{4}\right) - 380.25 = 0.85$$

$$TrtSS = \sum\left(\frac{(18)^2 + (20.5)^2 + (19.7)^2 + (19.8)^2}{4}\right) - 380.25 = 0.85$$

$$TrtSS = \sum\left(\frac{1524.38}{4}\right) - 380.25 = 0.85$$

$$TrtSS = 381.10 - 380.25 = 0.85$$

The fifth step is calculating of the Factor "A" Sum of Squares (ASS):

$$A = A_1 + A_2$$

$$A_1 = 18 + 20.5 = 38.5$$

$$A_2 = 19.7 + 19.8 = 39.5$$

$$ASS = \sum\left(\frac{(A_1)^2 + (A_2)^2}{r\ x\ n}\right) - CF$$

$$ASS = \sum\left(\frac{(38.5)^2 + (39.5)^2}{4\ x\ 2}\right) - 380.25$$

$$ASS = \sum\left(\frac{1482.25 + 1560.25}{8}\right) - 380.25$$

$$ASS = \sum\left(\frac{3042.5}{8}\right) - 380.25$$

$$ASS = \sum 380.31 - 380.25 = 0.063$$

The sixth step is calculating of the Factor "B" Sum of Squares (BSS):

$$B = B_1 + B_2$$

$$B_1 = 18 + 19.7 = 37.7$$

$$B_2 = 20.5 + 19.8 = 40.3$$

$$BSS = \sum (\frac{(B_1)^2 + (B_2)^2}{r\ x\ n}) - CF$$

$$BSS = \sum (\frac{(37.7)^2 + (40.3)^2}{4\ x\ 2}) - 380.25$$

$$BSS = \sum (\frac{1421.29 + 1624.09}{8}) - 380.25$$

$$BSS = \sum (\frac{3045.38}{8}) - 380.25$$

$$BSS = \sum 380.67 - 380.25 = 0.42$$

The seventh step is calculating of the interaction between "A x B" Sum of Squares (ABSS):

$$ABSS = TrtSS - (ASS + BSS)$$

$$ABSS = 0.85 - (0.063 + 0.42)$$

$$ABSS = 0.85 - 0.48 = 0.37$$

The eighth step is calculating of the Error Sum of Squares (ESS):

$$ESS = TSS - TrtSS = 3.63 - 0.85 = 2.78$$

Tt_T=Treatment Total, ASS=Factor "A" Sum of Squares, TSS=Total Sum of Squares, BSS=Factor "B" Sum of Squares, TrtSS=Treatment Sum of Squares, ABSS=Interaction between Factors "A x B" Sum *of Squares, ESS=Error Sum of Squares, CF=Correction Factor, n = number of levels of factors, r =* number of replication.

The ninth step looking up the F-critical table to the F-critical values:

In looking for the critical value, the degree of the variable of the variable whose critical value is being read from the table is located on the outer most row and matched with the error degree of freedom on the outer most column on the left of the table. Following the given procedure, the following readings would be obtained from the 5% and 1% F-critical table:

Treatment: F crit at df 3, 9 for 5% = 3.86; F crit at df 3, 9 for 1% = 6.99

Factor A: F crit at df 1, 9 for 5% = 5.12; F crit at df 1, 9 for 1% = 10.56

Factor B: F crit at df 1, 9 for 5% = 5.12; F crit at df 1, 9 for 1% = 10.56

Interaction (A x B): F crit at df 1, 9 for 5% = 5.12; F crit at df 1, 9 for 1% = 10.56.

The tenth step is completing the ANOVA table:

ANOVA TABLE

Sources of Variation	*df*	*SS*	*MS*	*Fcal.*	*Fcrit. (5%)*	*Fcrit (1%)*
Treatment	4 - 1= 3	0.85	0.28	0.92	3.86	6.99
Factor A	2 - 1 = 1	0.063	0.06	0.20	5.12	10.56
Factor B	2 - 1 = 1	0.42	0.42	1.36	5.12	10.56
Interaction (A x B)	2 – 1 =1	0.37	0.37	1.20	5.12	10.56
Error	15 – 6 = 9	2.78	0.31			
Total	16 – 1=15	3.63				

The eleventh step is making the decision or conclusion:

Decision on Treatment

$$Fcal\ (0.92) < 3.86, Fcrit.or\ tab.(5\%) at\ 3, 9$$

$$Fcal\ (\ 0.92) < 6.99, Fcrit.or\ tab.(1\%) at\ 3, 9$$

Since *Fcal < Fcrit. or tab. at both 1% and 5% levels of significance*, it can be concluded that there are no significant differences between the various treatments and then fail to reject the null hypothesis.

Decision of Factor A

$$Fcal\ (0.20) < 5.12, Fcrit.or\ tab.(5\%)at\ 1,9$$

$$Fcal\ (\ 0.20) < 10.56, Fcrit.or\ tab.(1\%)at\ 1,9$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence there are no significant differences between the two different levels of factor A.

Decision of Factor B

$$Fcal\ (1.36) < 5.12, Fcrit.or\ tab.(5\%)at\ 1,9$$

$$Fcal\ (\ 1.36) < 10.56, Fcrit.or\ tab.(1\%)at\ 1,9$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence there are no significant differences between the two different levels of factor B.

Decision on the interactions (A x B)

$$Fcal\ (1.20) < 5.12, Fcrit.or\ tab.(5\%)at\ 1,9$$

$$Fcal\ (\ 1.20) < 10.56, Fcrit.or\ tab.(1\%)at\ 1,9$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence the two levels of factor A do not vary significantly at the two levels of B when applied together.

5.3 RCBD

With regards to RCBD, the researcher can study the time it takes for the same quantity of milk from cow (A_1) and (A_2) to ferment into yoghurt under two

temperature conditions B_1 and B_2 blocked under the mode of storage: block 1(fridge), block 2 (air condition) and block 3 (normal room condition).

	A_1		A_2	
BLK	**B_1**	**B_2**	**B_1**	**B_2**
1	5.4	5.3	5.1	4.7
2	4.6	5.6	4.5	4.9
3	4.1	4.8	5.6	5.2
4	3.9	4.8	4.5	5.0

The first step in the analysis is to put forth the hypothesis:

For RCBD the following hypotheses can be put forth:

H_o: $\mu BLK_1 = \mu BLK_2 = \mu BLK_3 = \mu BLK_4$

H_1: $\mu BLK_1 \neq \mu BLK_2 = \mu BLK_3 = \mu BLK_4$

H_o: All blocks have equal treatment means

H_A: At least one block's treatment means differ from the others

H_o: $\mu A_1B_1 = \mu A_1B_2 = \mu_{A2B2} = \mu_{A2B2}$

H_1: $\mu_{A1B1} \neq \mu_{A1B2} = \mu_{A2B2} = \mu_{A2B2}$

H_o: All the means of the treatments are equal

H_A: At least one of the means of the treatments is different

H_o: $\mu A_1 = \mu A_2$

H_1: $\mu_{A1} \neq \mu_{A2}$

H_o: The means of the 2 levels of A are equal

H_A: The means of the 2 levels of A are unequal

H_o: $\mu_{B1} = \mu_{B2}$

H_1: $\mu_{B1} \neq \mu_{B2}$

H_o: The means of the 2 levels of B are equal

H_A: The means of the 2 levels of B are unequal

H_o: $\mu_{A \times B} = \mu_{A \times B}$

H_1: $\mu_{A \times B} \neq \mu_{A \times B}$

H_o: The effects of the interaction of two factors at the 2 levels are the same

H_A: The effects of the interaction of two factors at the 2 levels are different

	A_1		A_2		
BLK	B_1	B_2	B_1	B_2	
1	5.4	5.3	5.1	4.7	20.5
2	4.6	5.6	4.5	4.9	19.6
3	4.1	4.8	5.6	5.2	19.7
4	3.9	4.8	4.5	5	18.2
	18	20.5	19.7	19.8	78

$$Tt_{A_1B_1} = \sum(5.4 + 4.6 + \cdots + 3.9) = 18$$

$$Tt_{A_1B_2} = \sum(5.3 + 5.6 + \cdots + 4.8) = 20.5$$

$$Tt_{A_2B_1} = \sum(5.1 + 4.5 + \cdots + 4.5) = 19.7$$

$$Tt_{A_2B_2} = \sum(4.7 + 4.9 + \cdots + 5) = 19.8$$

$$Tt_T = \sum\left(Tt_{A_1B_1} + Tt_{A_1B_2} + Tt_{A_2B_2} + Tt_{A_2B_2}\right) = 78$$

The second step is calculating of the Correction Factor:

$$CF = \frac{\sum(Tt_T)^2}{t} = \frac{\sum(Tt_T)^2}{4\ x\ 2\ x\ 2} = \frac{(78)^2}{4\ x\ 4} = \frac{(78)^2}{16} = \frac{6084}{16} = 380.25$$

where t=number of blocks x 2 levels of A x 2 levels of B=4 x 2 x 2

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum((5.4 + (5.3)^2 + (5.1)^2 + (4.7)^2 + (4.6)^2 + (5.6)^2 \ldots\ldots + (5)^2) - 380.25 = 3.63$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TrtSS = \sum\left(\frac{(Tt_{A_1B_1})^2 + (Tt_{A_1B_2})^2 + (Tt_{A_2B_2})^2 + (Tt_{A_2B_2})^2}{4}\right) - 380.25 = 0.85$$

$$TrtSS = \sum\left(\frac{(18)^2 + (20.5)^2 + (19.7)^2 + (19.8)^2}{4}\right) - 380.25 = 0.85$$

$$TrtSS = \sum\left(\frac{1524.38}{4}\right) - 380.25 = 0.85$$

$$TrtSS = 381.10 - 380.25 = 0.85$$

The fifth step is calculating of the Factor "A" Sum of Squares (ASS):

$$A = A_1 + A_2$$

$$A_1 = 18 + 20.5 = 38.5$$

$$A_2 = 19.7 + 19.8 = 39.5$$

$$ASS = \sum(\frac{(A_1)^2 + (A_2)^2}{r\ x\ n}) - CF$$

$$ASS = \sum(\frac{(38.5)^2 + (39.5)^2}{4\ x\ 2}) - 380.25$$

$$ASS = \sum(\frac{1482.25 + 1560.25}{8}) - 380.25$$

$$ASS = \sum(\frac{3042.5}{8}) - 380.25$$

$$ASS = \sum 380.31 - 380.25 = 0.063$$

The sixth step is calculating of the Factor "A" Sum of Squares (ASS):

$$A = A_1 + A_2$$

$$A_1 = 18 + 20.5 = 38.5$$

$$A_2 = 19.7 + 19.8 = 39.5$$

$$ASS = \sum \left(\frac{(A_1)^2 + (A_2)^2}{r\ x\ n}\right) - CF$$

$$ASS = \sum \left(\frac{(38.5)^2 + (39.5)^2}{4\ x\ 2}\right) - 380.25$$

$$ASS = \sum \left(\frac{1482.25 + 1560.25}{8}\right) - 380.25$$

$$ASS = \sum \left(\frac{3042.5}{8}\right) - 380.25$$

$$ASS = \sum 380.31 - 380.25 = 0.063$$

The seventh step is calculating of the Factor "B" Sum of Squares (BSS):

$$B = B_1 + B_2$$

$$B_1 = 18 + 19.7 = 37.7$$

$$B_2 = 20.5 + 19.8 = 40.3$$

$$BSS = \sum \left(\frac{(B_1)^2 + (B_2)^2}{r\ x\ n}\right) - CF$$

$$BSS = \sum \left(\frac{(37.7)^2 + (40.3)^2}{4\ x\ 2}\right) - 380.25$$

$$BSS = \sum \left(\frac{1421.29 + 1624.09}{8}\right) - 380.25$$

$$BSS = \sum \left(\frac{3045.38}{8}\right) - 380.25$$

$$BSS = \sum 380.67 - 380.25 = 0.42$$

The eighth step is calculating the Block Sum of Squares (BLKSS):

$$BLKSS = \sum\left(\frac{(BLK1_T)^2+(BLK2_T)^2+(BLK3_T)^2+(BLK4_T)^2}{number\ \ of\ blocks}\right) - CF$$

$$BLKSS = \sum\left(\frac{(20.5)^2+(19.6)^2+(19.7)^2+(18.2)^2}{4}\right) - 380.25$$

$$BLKSS = \sum\left(\frac{(420.25)+(384.16)+(388.09)+(331.24)}{4}\right) - 380.25$$

$$BLKSS = \sum\left(\frac{(1523.74)}{4}\right) - 380.25$$

$$BLKSS = \sum 380.94 - 380.25 = 0.69$$

The ninth step is calculating of the interaction between "A x B" Sum of Squares (ABSS):

$$ABSS = TrtSS - (ASS + BSS)$$

$$ABSS = 0.85 - (0.063 + 0.42)$$

$$ABSS = 0.85 - 0.48 = 0.37$$

The tenth step is calculating of the Error Sum of Squares (ESS):

$$ESS = TSS - (BLKSS + TrtSS)$$

$$ESS = 3.63 - (0.85 + 0.69)$$

$$ESS = 3.63 - 1.54 = 2.09$$

Tt_T=Treatment Total, ASS=Factor "A" Sum *of Squares, TSS=Total Sum of Squares, BSS=Factor* "B" Su *m of Squares, TrtSS=Treatment Sum of Squares, BLKSS=Block Sum of Squares, ABSS=Interaction between Factors* "A x B" Sum *of Squares, ESS=Error Sum of Squares, CF=Correction Factor, n=number levels of factors, r = number of replication.*

The eleventh step looking up the f-critical table to the f- critical values:

Block: f crit at df 3, 6 for 5% = 4.76; f crit at df 3, 6 for 1% = 9.78

Treatment: F crit at df 3, 6 for 5% = 4.76; F crit at df 3, 6 for 1% = 9.78

Factor A: F crit at df 1, 6 for 5% = 5.99; F crit at df 1, 6 for 1% = 13.75

Factor B: F crit at df 1, 6 for 5% = 5.99; F crit at df 1, 6 for 1% = 13.75

Interaction (A x B): F crit at df 1, 6 for 5% = 5.99; F crit at df 1, 6 for 1% = 13.75.

The twelfth step is completing the ANOVA table:

ANOVA TABLE

Sources of Variation	*df*	*SS*	*MS*	*Fcal.*	*Fcrit. (5%)*	*Fcrit (1%)*
Block	4 – 1= 3	0.69	0.23	0.66	4.76	9.78
Treatment	4 - 1= 3	0.85	0.28	0.81	4.76	9.78
Factor A	2 - 1 = 1	0.063	0.06	0.18	5.99	13.75
Factor B	2 - 1 = 1	0.42	0.42	1.21	5.99	13.75
Interaction (A x B)	2 – 1 =1	0.37	0.37	1.06	5.99	13.75
Error	15 – 9 = 6	2.09	0.35			
Total	16 - 1=15	3.63				

The thirteenth step is making the decision or conclusion:

Decision on Block

$$Fcal\ (0.66) < 4.76, Fcrit.or\ tab.(5\%)at\ 3,6$$

$$Fcal\ (\ 0.66) < 9.78, Fcrit.or\ tab.(1\%)at\ 3,6$$

The F cal < F crit for the block at both 1% and 5% levels of significance, hence the researcher must fail to reject the null hypothesis for the block and conclude that there is no significant difference between the blocks. It there for suggests that there is no justification for blocking and the RCBD is therefore not the appropriate design for analysis of this experiment.

Decision on Treatment

$$Fcal\ (0.81) < 4.76, Fcrit.or\ tab.(5\%) at\ 3,6$$

$$Fcal\ (\ 0.81) < 9.78, Fcrit.or\ tab.(1\%) at\ 3,6$$

Since *Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, it can* be concluded that there are no significant differences between the various treatments and then fail to reject the null hypothesis.

Decision of Factor A

$$Fcal\ (0.18) < 5.99, Fcrit.or\ tab.(5\%) at\ 1,6$$

$$Fcal\ (\ 0.18) < 13.75, Fcrit.or\ tab.(1\%) at\ 1,6$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence there are no significant differences between the two different levels of factor A.

Decision of Factor B

$$Fcal\ (1.21) < 5.99, Fcrit.or\ tab.(5\%) at\ 1,6$$

$$Fcal\ (\ 1.21) < 13.75, Fcrit.or\ tab.(1\%) at\ 1,6$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence there are no significant differences between the two different levels of factor B.

Decision on the interactions (A x B)

$$Fcal\ (1.06) < 5.99, Fcrit.or\ tab.(5\%) at\ 1,6$$

$$Fcal\ (\ 1.06) < 13.75, Fcrit.or\ tab.(1\%) at\ 1,6$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence the two levels of factor A do not vary significantly at the two levels of B when applied together.

5.4 LSD

For the LSD, the same experiment can be used as in the case of CRD and RCBD, but this time around there is no replication and blocking as done respectively under the two mentioned designs.

		A_1		A_2	
		B_1	B_2	B_1	B_2
A_1	B_1	5.4	5.3	5.1	4.7
	B_2	4.6	5.6	4.5	4.9
A_2	B_1	4.1	4.8	5.6	5.2
	B2	3.9	4.8	4.5	5.0

The first step in the analysis is to put forth the hypothesis:

The following hypotheses can be stated for the LSD:

H_o: $\mu COL_1 = \mu COL_2 = \mu COL_3 = \mu COL_4$

H_1: $\mu COL_1 \neq \mu COL_2 = \mu COL_3 = \mu COL_4$

H_o: All columns have equal treatment means

H_A: At least one column's treatment means differ from the others

H_o: $\mu ROW_1 = \mu ROW_2 = \mu ROW_3 = \mu ROW_4$

H_1: $\mu ROW_1 \neq \mu ROW_2 = \mu ROW_3 = \mu ROW_4$

H_o: All rows have equal treatment means

H_A: At least one row's treatment means differ from the others

H_o: $\mu A_1B_1 = \mu A_1B_2 = \mu_{A2B2} = \mu_{A2B2}$

H_1: $\mu_{A1B1} \neq \mu_{A1B2} = \mu_{A2B2} = \mu_{A2B2}$

H_o: All the means of the treatments are equal

H_A: At least one of the means of the treatments is different

H_o: $\mu A_1 = \mu A_2$

H_1: $\mu_{A1} \neq \mu_{A2}$

H_o: The means of the 2 levels of A are equal

H_A: The means of the 2 levels of A are unequal

H_o: $\mu_{B1} = \mu_{B2}$

H_1: $\mu_{B1} \neq \mu_{B2}$

H_o: The means of the 2 levels of B are equal

H_A: The means of the 2 levels of B are unequal

H_o: $\mu_{A \times B} = \mu_{A \times B}$

H_1: $\mu_{A \times B} \neq \mu_{A \times B}$

H_o: The effects of the interaction of two factors at the 2 levels are the same

H_A: The effects of the interaction of two factors at the 2 levels are different

		A_1		**A_2**		
		B_1	**B_2**	**B_1**	**B_2**	
A_1	B_1	5.4	5.3	5.1	4.7	20.5
	B_2	4.6	5.6	4.5	4.9	19.6
A_2	B_1	4.1	4.8	5.6	5.2	19.7
	B_2	3.9	4.8	4.5	5	18.2
		18	20.5	19.7	19.8	78

The second step is calculating of the Correction Factor:

$$CF = \frac{\sum(Tt_T)^2}{t} = \frac{\sum(Tt_T)^2}{4\ x\ 2\ x\ 2} = \frac{(78)^2}{4\ x\ 4} = \frac{(78)^2}{16} = \frac{6084}{16} = 380.25$$

where t=number of blocks x 2 levels of A x 2 levels of B=4 x 2 x 2

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum((5.4 + (5.3)^2 + (5.1)^2 + (4.7)^2 + (4.6)^2 + (5.6)^2 \ldots\ldots + (5)^2) - 380.25 = 3.63$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TrtSS = \sum\left(\frac{(Tt_{A_1B_1})^2 + (Tt_{A_1B_2})^2 + (Tt_{A_2B_2})^2 + (Tt_{A_2B_2})^2}{4}\right) - 380.25 = 0.85$$

$$TrtSS = \sum\left(\frac{(18)^2 + (20.5)^2 + (19.7)^2 + (19.8)^2}{4}\right) - 380.25 = 0.85$$

$$TrtSS = \sum\left(\frac{1524.38}{4}\right) - 380.25 = 0.85$$

$$TrtSS = 381.10 - 380.25 = 0.85$$

The fifth step is calculating of the Column Sum of Squares (CSS):

$$CSS = \sum\left(\frac{(COL1_T)^2 + (COL2_T)^2 + (COL3_T)^2 + (COL4_T)^2}{number\ of\ columns}\right) - CF$$

$$CSS = \sum\left(\frac{(18)^2 + (20.5)^2 + (19.7)^2 + (19.8)^2}{4}\right) - 380.25$$

$$CSS = \sum\left(\frac{(324) + (420.25) + (388.09) + (392.04)}{4}\right) - 380.25$$

$$CSS = \sum 381.10 - 380.25 = 0.85$$

The sixth step is calculating of the Row Sum of Squares (RSS):

$$RSS = \sum\left(\frac{(ROW1_T)^2 + (ROW2_T)^2 + (ROW3_T)^2 + (ROW4_T)^2}{number\ of\ rows}\right) - CF$$

$$RSS = \sum \left(\frac{(20.5)^2 + (19.6)^2 + (19.7)^2 + (18.2)^2}{4} \right) - 380.25$$

$$RSS = \sum \left(\frac{(420.25) + (384.16) + (388.09) + (331.24)}{4} \right) - 380.25$$

$$RSS = \sum 380.94 - 380.25 = 0.69$$

The seventh step is calculating of the Factor "A" Sum of Squares (ASS):

$$A = A_1 + A_2$$

$$A_1 = 18 + 20.5 = 38.5$$

$$A_2 = 19.7 + 19.8 = 39.5$$

$$ASS = \sum \left(\frac{(A_1)^2 + (A_2)^2}{r \; x \; n} \right) - CF$$

$$ASS = \sum \left(\frac{(38.5)^2 + (39.5)^2}{4 \; x \; 2} \right) - 380.25$$

$$ASS = \sum \left(\frac{1482.25 + 1560.25}{8} \right) - 380.25$$

$$ASS = \sum \left(\frac{3042.5}{8} \right) - 380.25$$

$$ASS = \sum 380.31 - 380.25 = 0.063$$

The eighth step is calculating of the Factor "A" Sum of Squares (ASS):

$$A = A_1 + A_2$$

$$A_1 = 18 + 20.5 = 38.5$$

$$A_2 = 19.7 + 19.8 = 39.5$$

$$ASS = \sum \left(\frac{(A_1)^2 + (A_2)^2}{r \; x \; n} \right) - CF$$

$$ASS = \sum \left(\frac{(38.5)^2 + (39.5)^2}{4 \; x \; 2} \right) - 380.25$$

$$ASS = \sum \left(\frac{1482.25 + 1560.25}{8} \right) - 380.25$$

$$ASS = \sum \frac{3042.5}{8} - 380.25$$

$$ASS = \sum 380.31 - 380.25 = 0.063$$

The ninth step is calculating of the Factor "B" Sum of Squares (BSS):

$$B = B_1 + B_2$$

$$B_1 = 18 + 19.7 = 37.7$$

$$B_2 = 20.5 + 19.8 = 40.3$$

$$BSS = \sum \left(\frac{(B_1)^2 + (B_2)^2}{r\ x\ n}\right) - CF$$

$$BSS = \sum \left(\frac{(37.7)^2 + (40.3)^2}{4\ x\ 2}\right) - 380.25$$

$$BSS = \sum \left(\frac{1421.29 + 1624.09}{8}\right) - 380.25$$

$$BSS = \sum \left(\frac{3045.38}{8}\right) - 380.25$$

$$BSS = \sum 380.67 - 380.25 = 0.42$$

The tenth step is calculating of the interaction between "A x B" Sum of Squares (ABSS):

$$ABSS = TrtSS - (ASS + BSS)$$

$$ABSS = 0.85 - (0.063 + 0.42)$$

$$ABSS = 0.85 - 0.48 = 0.37$$

The eleventh step is calculating of the Error Sum of Squares (ESS):

$$ESS = TSS - (TrtSS + RSS + CSS)$$

$$ESS = 3.63 - (0.85 + 0.69 + 0.85)$$

$$ESS = 3.63 - 2.39 = 1.24$$

Tt_T=Treatment Total, ASS=Factor "A" Sum of Squares, TSS=Total Sum of Squares, BSS=Factor "B" Sum of Squares, TrtSS=Treatment Sum of Squares, BLKSS=Block Sum of Squares, ABSS=Interaction between Factors "A x B" Sum of Squares, ESS=Error Sum of Squares, CF=Correction Factor, n=number levels of factors, r = number of replication.

The twelfth step looking up the F-critical table to the F-critical values:

Column: F crit at df 3, 3 for 5% =9.28; F crit at df 3, 3 for 1% = 29.46

Row: F crit at df 3, 3 for 5% = 9.28; F crit at df 3, 3 for 1% = 29.46

Treatment: F crit at df 3, 3 for 5% = 9.28; F crit at df 3, 3 for 1% = 29.46

Factor A: F crit at df 1, 3 for 5% = 10.13; F crit at df 1, 3 for 1% = 34.12

Factor B: F crit at df 1, 3 for 5% = 10.13; F crit at df 1, 3 for 1% = 34.12

Interaction (A x B): F crit at df 1, 3 for 5% = 10.13; F crit at df 1, 3 for 1% = 34.12.

The thirteenth step is completing the ANOVA table:

Sources of Variation	df	SS	MS	Fcal.	Fcrit. (5%)	Fcrit (1%)
Column	4 – 1= 3	0.69	0.23	0.56	9.28	29.46
Row	4 -1 = 3	0.85	0.28	0.69	9.28	29.46
Treatment	4 - 1= 3	0.85	0.28	0.69	9.28	29.46
Factor A	2 - 1 = 1	0.063	0.06	0.15	10.13	34.12
Factor B	2 - 1 = 1	0.42	0.42	1.02	10.13	34.12
Interaction (A x B)	2 – 1 =1	0.37	0.37	0.90	10.13	34.12
Error	15 – 12 = 3	1.24	0.41			
Total	16 -1=15	3.63				

The fourteenth step is making the decision or conclusion:

Decision on Column

$$Fcal\ (0.56) < 9.28, Fcrit.or\ tab.(5\%)at\ 3,3$$

$$Fcal\ (\ 0.56) < 29.46, Fcrit.or\ tab.(1\%)at\ 3,3$$

The F cal < F crit for the block at both 1% and 5% levels of significance, hence the researcher must fail to reject the null hypothesis for the column and conclude that there is no significance difference between the columns.

Decision on Row

$$Fcal\ (0.69) < 9.28, Fcrit.or\ tab.(5\%)at\ 3,3$$

$$Fcal\ (\ 0.69) < 29.46, Fcrit.or\ tab.(1\%)at\ 3,3$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, it can be concluded that there are no significant differences between rows hence the researcher must fail to reject the null hypothesis for the row.

Once it has been established that there are no significant differences between the columns and rows respectively in the design, one can therefore conclude that there is no justification for the use of LSD, it is thus inappropriate for analysis of the experiment in question.

Decision on Treatment

$$Fcal\ (0.69) < 9.28, Fcrit.or\ tab.(5\%)at\ 3,3$$

$$Fcal\ (\ 0.69) < 29.46, Fcrit.or\ tab.(1\%)at\ 3,3$$

Since *Fcal < Fcrit. or tab. at both 1% and 5% levels of significance*, it can be concluded that there are no significant differences between the various treatments and then fail to reject the null hypothesis for the treatment.

Decision of Factor A

$$Fcal\ (0.15) < 10.13, Fcrit.or\ tab.(5\%) at\ 1,3$$

$$Fcal\ (\ 0.15) < 34.12, Fcrit.or\ tab.(1\%) at\ 1,3$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence there are no significant differences between the two different levels of factor A.

Decision of Factor B

$$Fcal\ (1.02) < 10.13, Fcrit.or\ tab.(5\%) at\ 1,3$$

$$Fcal\ (\ 1.02) < 34.12, Fcrit.or\ tab.(1\%) at\ 1,3$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence there are no significant differences between the two different levels of factor B.

Decision on the interactions (A x B)

$$Fcal\ (0.90) < 10.13, Fcrit.or\ tab.(5\%) at\ 1,3$$

$$Fcal\ (\ 0.90) < 34.12, Fcrit.or\ tab.(1\%) at\ 1,6$$

Fcal < Fcrit. or tab. at both 1% and 5% levels of significance, hence the two levels of factor A do not vary significantly at the two levels of B when applied together.

Bibliography

[1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.

[2] Campbell, D., Stanley, J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), Handbook on research on teaching. Chicago: Rand McNally & Co.

[3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.

[4] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs.* John Willey & Sons, N.Y.

[5] Le Clerg, E. L., Leonard, W. H., Clark A. G. (1966). *Field Plot Technique.* Burgess Pub. Co., Minn., USA.

[6] Daniel, C. (1959). Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments. *Technometrics* 1:311–341.

[7] Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

[8] Ju, H. L., Lucas, J. M. (2002). L*k* Factorial Experiments with Hard-To-Change and Easy-To-Change Factors. *Journal of Quality Technology* 34:411–421.

[9] Montgomery, D. C. (2008). *Design and Analysis of Experiments*. 7th ed. John Wiley & Sons. New York, NY.

[10] Taguchi, G. (1987). *System of Experimental Design*. Kraus International Publications, White Plains, NY.

[11] Webb, D. F., Lucas, J. M., Borkowski, J. J. (2004). Factorial Experiments when Factor Levels Are Not Necessarily Reset. *Journal of Quality Technology* 36:1–11.

[12] Yates, F. (1935). Complex Experiments with Discussion. *Journal of the Royal Statistical Society, Series B* 2:181–223.

Chapter 6

Split Plot Experimental Design

6

Split Plot Experimental Design

Felix Kutsanedzie[1] *; Sylvester Achio[1]; Edmund Ameko[1]; Victoria Ofori[2]; Edith Mensah[1]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Agricultural Engineering Department, KNUST, Ghana

Abstract

Spilt plot experimental design analysis is quite complicated because it involves the analysis of the main plot and then the sub plot. Due to its complicated nature, analysis of such a design is quite difficult for the understanding of students and most researchers. It is the appropriate design used for the analysis of a two factor experiment where all the treatments or factors cannot be contained in complete block design. This chapter explains how this design can be used and handled when analysing experiments.

Keywords

Split Plot, Design, Two Factor Experiment, Complete Block Design

6.1 Introduction

Every experimental design has its peculiar usefulness and cannot be relegated to the background when the situation calls for it. Whenever a researcher or an experimenter is faced handling two factor experiment which cannot be contained in a complete block design, then the split plot designs are used. The split plot design is thus suited for two factor experiment where the main factor is assigned to the main plot and the second factor assigned to the subplot which emanates from the division of the main plot. Thus the main plot becomes a block on its own and has subplot which holds treatment. The *main factor* is the factor which the researcher is very familiar with its characteristics. In a split plot design, the precision of measurement of effects of the main plot factors is sacrificed to improve that of the sub-plot factor. The relative size of the main effects and the precision of measurement of effects should not be the same for both factors – main factor and the sub factor. Assignment of factors to the main plot and the sub- plot is important and the guidelines to make this choice or assign the factors to the plot is determined by the relative size of the main effects and the precision of its measurement in relation to the researcher or experimenter's interest. As such it is often referred to as an experiment of convenience.

When the experimenter is considering greater precision for one factor as opposed to the other, the factor that requires a greater precision is assigned to the sub-plot and that with less precision assigned to the main plot. Taking for example a machine designer and a chemist considering a split plot design that involves the sizes of a machine type (S) and the chemical composition of food processed using these machines (C). The machine designer and the chemist are likely to assign these two factors as follows as per the precision required from each of on the measurement of the factors:

Machine Designer			
Factors	Factor type	Precision required	Plot type the factor must be assigned
Sizes of machines	Sub factor	More precision	sub plot -
Chemical composition of processed food	Main factor	Less precision	Main factor to be assigned to main plot

Chemist			
Factors	Factor type	Precision required	Plot type the factor must be assigned
Chemical composition of processed food	Sub factor	More precision	sub plot -
Sizes of machines	Main factor	Less precision	Main factor to be assigned to main plot

As regard the relative size of the main effects, the experimenter assigns factors to the main and subplot according to the relative size of their expected effects. For instance, if the main effect of one factor is larger and easier to detect than the other it is assigned to the main plot and the other assigned to the sub-plot. Taking for instance in an experiment to test the effects of different methods of processing food (P) and storing food (A) on nutrient loss; the factors in this experiment can be assigned in the design based on the relative size of their effects as shown below:

Assignment of factors based on the relative size of their effects			
Factors	Factor type	Size of effects factors	Plot type the factor must be assigned
Food processing methods (P)	Main factor	Larger effects	Main plot
Food storage methods (A)	Sub-factor	Lesser effects	Sub-plot

Another important factor in assigning factors is the management practices to be adopted in handling the factors under the design. For instance if the researcher is experimenting on the effect of a particular food item on the growth in terms of height on humans above twenty (20) year old (H) and humans below

one (1) to twenty (20) years (G), since it has been established that humans cease to grow in height at age 21, those who are above twenty (H) needs to be assigned to the sub-plot while those between 1 and 20 be assigned to the main plot to minimize the effect of the food to those who already have the potential or capable of growing in height. This last factor applies to agricultural experiments. However with other industrial experiments, there are the hard-to-change factors which are assigned to the main plots and the easy-to-change factors which are assigned to the sub-plots. For instance assuming a sapele board is to be subjected to three different treatments (A, B, C) and then painted with three different paints (X, Y, Z) to ascertain its acceptability by users. These can be achieved in two ways: the first is to treat each sapele board in the three different conditions, divide each of them into three different portions and then paint them with the three selected paints; the other way is to divide each of the sapele board into three portions first, treat each of them and then paint them.

Therefore in a split plot experimental designs, the levels of the main plot factor multiplied by the levels of the sub-plot factor gives the number of treatments. It presupposes that if one is considering three levels of H (H_1, H_2, H_3) as sub-plot factor and G (G_1, G_2, G_3) as the main plot factor, then the number of treatments equals nine (9), 3 levels of H x 3 levels of G. if these are replicated for four (4) times, then there will be 36 treatments or observable units - 4 replicates x 3 levels of H x 3 levels of G.

To explain the step-by-step procedures used to analyse the split plot design, we can consider a hypothetical situation of designing an experiment involving three machines (M_1, M_2, M_3) made out of different materials such stainless stain, Aluminium and Iron respectively and their respective wear in terms of particles size and quantity (W_1, W_2, W_3) into the flour produced when used in milling the same quantity of maize.

In this particular case the main plot factors would be the machine types (M_1, M_2, M_3) and the subplot factors would be the different quantity of wear (W_1, W_2, W_3). Thus the split-plot design is shown below:

REPLICATION I		
M_1	M_2	M_3
W_1	W_1	W_1
W_2	W_2	W_2
W_3	W_3	W_3

REPLICATION II		
M_2	M_3	M_1
W_2	W_2	W_2
W_3	W_3	W_3
W_1	W_1	W_1

REPLICATION III		
M_3	M_1	M_2
W_3	W_3	W_3
W_1	W_1	W_1
W_2	W_2	W_2

In this experiment, the first replication shows how the split plot design looks like and randomization can be achieved as done in the other replications II and III. The design and the replications show how the main plots factors and the subplot factors would be arranged in the experiment. However when the experiment is conducted, the data obtained can be represented as follows:

REPLICATION I		
M_1	M_2	M_3
1.3	1.9	2.2
1.1	2.1	1.5
1.5	1.6	1.4

REPLICATION II		
M_2	M_3	M_1
1.2	2.2	1.9
1.4	1.5	2.1
1.3	1.4	1.6

REPLICATION III		
M_3	M_1	M_2
1.9	2.2	2.5
2.1	1.5	1.6
1.6	1.4	2.1

In order to do the calculation to complete the ANOVA table, the data obtained from the experiment is summarized and shown in the table below:

		W_1	W_2	W_3
REPLICATION I	M_1	1.3	2.1	1.5
	M_2	1.9	2.1	1.6
	M_3	2.2	1.5	1.4
REPLICATION II	M_2	1.3	1.2	1.4
	M_3	1.4	2.2	1.5
	M_1	1.6	1.9	2.1
REPLICATION III	M_3	2.1	1.6	1.9
	M_1	1.5	1.4	2.2
	M_2	1.6	2.1	2.5

6.2 Analysis of the Split Plot Design

For the analysis of the split plot design, one needs to understand that the size of the main plot is W times the size of the of the sub plot; the number of times the main plot treatment is tested is equal to the number of replications (r) used; there are three degrees of precision where the main plot factor is associated with the lowest degree of precision and the sub plot is associated with the highest degree of precision.

In doing the analysis for the split plot, the whole plot and the sub plot analyses must be done.

6.3 Whole Plot Analysis

The main plot treatments M_1, M_2 and M_3 within the blocks and handled as randomized complete block design. Once the main plot treatments are randomized as in the RCBD, no adjustment is required as regard the sum of squares for the main plot treatments (M).

In other to do the analysis the following steps should be followed:

		W_1	W_2	W_3
REPLICATION I	M_1	1.3	2.1	1.5
	M_2	1.9	2.1	1.6
	M_3	2.2	1.5	1.4
REPLICATION II	M_2	1.3	1.2	1.4
	M_3	1.4	2.2	1.5
	M_1	1.6	1.9	2.1
REPLICATION III	M_3	2.1	1.6	1.9
	M_1	1.5	1.4	2.2
	M_2	1.6	2.1	2.5

The first step is to calculate the replication total and the grand total.

In other to do these calculations the data has to be arranged in the tables below:

A cross tabulation of replication and machine type – main plot treatment (RM)

		W_1	W_2	W_3
REPLICATION I	M_1	1.3	2.1	1.5
	M_2	1.9	2.1	1.6
	M_3	2.2	1.5	1.4
REPLICATION II	M_2	1.3	1.2	1.4
	M_3	1.4	2.2	1.5
	M_1	1.6	1.9	2.1
REPLICATION III	M_3	2.1	1.6	1.9
	M_1	1.5	1.4	2.2
	M_2	1.6	2.1	2.5

Cross tabulation of Replication and Machine type (RM)

	M_1	M_2	M_3	Rep Totals
REPLICATION I	4.9	5.6	5.1	15.6
REPLICATION II	5.6	3.9	5.1	14.6
REPLICATION III	5.1	6.2	5.6	16.9
Machine type Totals	15.6	15.7	15.8	
Grand Total				47.1

Calculating the Grand Total (GT)

$$GT = \sum_{i=1,j=1}^{n=9} (4.9 + 5.6 + 5.1 + 5.6 + \cdots + 5.6) = 47.1$$

Calculating the Correction Factor (CF)

$$CF = \frac{(GT)^2}{rmw}$$

$$CF = \frac{(47.1)^2}{3\ x\ 3\ x\ 3}$$

$$CF = \frac{2218.41}{27}$$

$$CF = 82.16$$

where r=no. of replications, m=no. of levels of machine types, w=no. of levels of machine wears.

At this point one must compute the sum of squares of the main plot and these are done and illustrated below:

Calculating the Total Sum of Squares (TSS)

$$TSS = \sum sum\ of\ square\ of\ all\ observations - CF$$

$$TSS = \sum_{i=1,j=1,k=1}^{n=9} (m_1 w_1 r_1)^2 + \ldots\ldots\ldots\ldots\ldots + (m_2 w_3 r_3)^2 - CF$$

$$TSS = \sum [(1.3)^2 + (2.1)^2 + (1.5)^2 + (2.2)^2 + \cdots + (2.5)^2] - 82.16$$

$$TSS = \sum [1.69 + 4.41 + 2.25 + 4.84 + \cdots + 6.25] - CF$$

$$TSS = 85.55 - 82.16 = 3.39$$

$$TSS = 3.39$$

Calculating the Replication Sum of Squares (RSS)

$$RSS = \sum \frac{[(R_{1_T})^2 + (R_{2_T})^2 + (R_{3_T})^2]}{mw} - CF$$

$$RSS = \sum \frac{[(15.6)^2 + (14.6)^2 + (16.9)^2]}{3\ x\ 3} - 82.16$$

$$RSS = \sum \frac{[\ 243.36 + 213.16 + 285.61]}{9} - 82.16$$

$$RSS = \sum (\frac{[\ 742.13]}{9} - 82.16)$$

$$RSS = 82.46 - 82.16$$

$$RSS = 0.30$$

Calculating the Machine Type Sum of Squares (MSS).

It should be noted here that the main plot factor is the machine type.

$$MSS = \sum \frac{[(M_{1_T})^2 + (M_{2_T})^2 + (M_{3_T})^2]}{rw} - CF$$

$$MSS = \sum \frac{[(15.6)^2 + (15.7)^2 + (15.8)^2]}{3\ x\ 3} - 82.16$$

$$MSS = \sum \frac{[\ 243.36 + 246.49 + 249.64]}{9} - 82.16$$

$$MSS = \sum (\frac{[\ 739.49]}{9} - 82.16)$$

$$MSS = 82.17 - 82.16$$

$$MSS = 0.01$$

Calculating the Replication and Machine Type Sum of Squares (RMSS).

Note: one is expected to use the cross tabulation table for replication and machine type

$$RMSS = \sum_{i=1,j=1}^{n=9} \frac{(m_1 r_1)^2 + \ldots\ldots\ldots\ldots\ldots + (m_2 r_3)^2}{w} - CF$$

$$RMSS = \sum \frac{[(4.9)^2 + (5.6)^2 + (5.1)^2 + \cdots + (5.6)^2]}{w} - 82.16$$

$$RMSS = \sum \frac{[24.01 + 31.36 + 26.01 + \cdots + 31.36]}{3} - 82.16$$

$$RMSS = \sum \frac{[249.77\ ]}{3} - 82.16$$

$$RMSS = 83.26 - 82.16)$$

$$RMSS = 1.10$$

$$\text{Main plot Error Sum of Squares (MPESS)} = RMSS - RSS - MSS$$

$$MPESS = RMSS - RSS - MSS$$

$$MPESS = 1.10 - 0.30 - 0.01$$

$$MPESS = 0.79$$

6.4 Sub Plot Analysis

This is where the sum of squares of the sub plot factors are computed and these have been illustrated as below:

Cross tabulation of Machine type and machine wear (MW)

	M_1	M_2	M_3	W Totals
W_1	4.4	4.8	5.7	14.9
W_2	5.4	5.4	5.3	16.1
W_3	5.8	5.5	4.8	16.1

Calculating the Machine Wear Sum of Squares (WSS)

It should be noted that the machine wear is the sub plot factor for this particular experiment being considered for analysis.

$$WSS = \sum \frac{[(W_{1T})^2 + (W_{2T})^2 + (W_{3T})^2]}{rm} - CF$$

$$WSS = \sum \frac{[(14.9)^2 + (16.1)^2 + (16.1)^2]}{3\ x\ 3} - 82.16$$

$$WSS = \sum \frac{[\ 222.01 + 259.21 + 259.21]}{9} - 82.16$$

$$WSS = \sum (\frac{[\ 740.43]}{9} - 82.16)$$

$$WSS = 82.27 - 82.16$$

$$WSS = 0.11$$

Calculating the Sum of Squares of the interaction between Machine Type and Machine Wear

$$(M\ xW)SS = \sum_{i=1,j=1}^{n=9} \frac{(m_1w_1)^2 + \ldots\ldots\ldots\ldots\ldots + (m_3w_3)^2}{r} - CF$$

$$(MxW)SS = \sum \frac{[(4.4)^2 + (4.8)^2 + (5.7)^2 + \cdots + (4.8)^2]}{r} - 82.16$$

$$(MxW)SS = \sum \frac{[19.36 + 23.04 + 32.49 + \cdots + 23.04]}{3} - 82.16$$

$$(MxW)SS = \sum \frac{[248.23\ ]}{3} - 82.16$$

$$(MxW)SS = 82.74 - 82.16$$

$$(MxW)SS = 0.58$$

$$Sub\ Plot\ Sum\ of\ Squares\ (SPSS) = TSS - All\ Sums\ of\ Squares$$

$$Error\ Sub\ Plot\ Sum\ of\ Squares\ (ESPSS) = TSS - (RSS + MSS + RMSS + MPESS + WSS + (MxW)SS$$

$$SPSS = 3.39 - (0.30 + 0.01 + 1.10 + 0.79 + 0.11 + 0.58)$$

$$ESPSS = 3.39 - 2.89$$

$$SPSS = 0.50$$

6.5 Completing the ANOVA Table

Sources of Variation	df	Sum of Squares	Mean Sum of Squares	Fcal	Fcrit (5%)	Fcrit (1%)
Replication	`r -1 = 3 – 1 = 2	0.30	$\frac{0.30}{2} = 0.15$			
Main Plot factor (M)	m -1 = 3-1 = 2	0.01	$\frac{0.01}{2} = 0.005$	$\frac{0.005}{0.395} = 0.01$		
Error (M)	(r-1)(m-1) = 2 x 2 =4	0.79	$\frac{0.79}{4} = 0.395$			
Sub Plot factor (W)	w -1 = 3 -1 =2	0.11	$\frac{0.11}{2} = 0.055$	$\frac{0.055}{0.25} = 0.22$		
Interaction between (M x W)	(m-1)(w-1) = 2 x 2 =4	0.58	$\frac{0.58}{4} = 0.29$	$\frac{0.29}{0.25} = 1.16$		
Error (z)	m(r-1)(w-1) = 3 x 2 x 2 = 12	0.50	$\frac{0.50}{12} = 0.25$			
Total	rmw-1 = 27 -1 = 26					

Reading of the F-critical or tabulated values from the F-table at the various assigned levels of significance allowed.

For the case being considered, 1% and 5% levels of significance would be used.

Critical values of F for the 0.05 significance level:

	1	2	3	4	5	6
1	161.45	199.50	215.71	224.58	230.16	233.99
2	18.51	19.00	19.16	19.25	19.30	19.33
3	10.13	9.55	9.28	9.12	9.01	8.94
4	7.71	6.94	6.59	6.39	6.26	6.16
5	6.61	5.79	5.41	5.19	5.05	4.95
6	5.99	5.14	4.76	4.53	4.39	4.28
7	5.59	4.74	4.35	4.12	3.97	3.87
8	5.32	4.46	4.07	3.84	3.69	3.58
9	5.12	4.26	3.86	3.63	3.48	3.37
10	4.97	4.10	3.71	3.48	3.33	3.22
11	4.84	3.98	3.59	3.36	3.20	3.10
12	4.75	3.89	3.49	3.26	3.11	3.00
13	4.67	3.81	3.41	3.18	3.03	2.92
14	4.60	3.74	3.34	3.11	2.96	2.85

Critical values of F for the 0.01 significance level:

	1	2	3	4	5	6
1	4052.19	4999.52	5403.34	5624.62	5763.65	5858.97
2	98.50	99.00	99.17	99.25	99.30	99.33
3	34.12	30.82	29.46	28.71	28.24	27.91
4	21.20	18.00	16.69	15.98	15.52	15.21
5	16.26	13.27	12.06	11.39	10.97	10.67
6	13.75	10.93	9.78	9.15	8.75	8.47
7	12.25	9.55	8.45	7.85	7.46	7.19
8	11.26	8.65	7.59	7.01	6.63	6.37
9	10.56	8.02	6.99	6.42	6.06	5.80
10	10.04	7.56	6.55	5.99	5.64	5.39
11	9.65	7.21	6.22	5.67	5.32	5.07
12	9.33	6.93	5.95	5.41	5.06	4.82
13	9.07	6.70	5.74	5.21	4.86	4.62
14	8.86	6.52	5.56	5.04	4.70	4.46

Sources of Variations	Fcal	Fcrit (5%)	Fcrit (1%)
Main Plot Factor (M)	$\frac{0.005}{0.25} = 0.01$	df(2,4) = 6.94	df(2,4) = 18.00
Sub Plot Factor (W)	$\frac{0.055}{0.25} = 0.22$	df (2,12) = 3.89	df(2,12) = 6.93
Interactions (M x W)	$\frac{0.29}{0.25} = 1.16$	df (2,12) = 3.89	df(2,12) = 6.93

*Note in reading the f tabulated value for the main plot factor, the degree of freedom of the main plot factor is used against the degree of freedom of the error (M). However for the subplot factor and the interactions their respective degrees of freedom are used against the degree of freedom of error (z).

Taking the Decision and Making the Conclusion

Main Plot Factor (Machine Type):

$$Fcal\ (0.01) < Fcrit\ (\ df\ 2{,}4\ @\ 5\% = 6.94)$$

$$Fcal\ (0.01) < Fcrit\ (\ df\ 2{,}4\ @\ 1\% = 18.00)$$

Conclusion on Main Plot Factor (Machine Type):

There is no significance difference between the machine types used in the experiment at both levels of significance. This means the machine types have a similar effect.

Subplot Factor (Wear Type):

$$Fcal\ (0.22) < Fcrit\ (\ df\ 2{,}12\ @\ 5\% = 3.89)$$

$$Fcal\ (0.22) < Fcrit\ (\ df\ 2{,}12\ @\ 1\% = 6.93)$$

Conclusion on Subplot (Wear Type).

There is no significant difference between the various wear types. Meaning the wear types did not differ statistically

Interactions (M x W):

$$Fcal\ (1.16) < Fcrit\ (\ df\ 2{,}12\ @\ 5\% = 3.89)$$

$$Fcal\ (1.16) < Fcrit\ (\ df\ 2{,}12\ @\ 1\% = 6.93)$$

Conclusion on Interactions (M x W).

There exists no significance difference. Hence one can conclude that the interactions between the two did not differ statistically.

Bibliography

[1] Goos, P. and Donev, A. (2007). Tailor-Made Split-Plot Designs for Mixture and Process Variables. *Journal of Quality Technology.* 39:326–339.

[2] Goos, P., Vandebroek, M. (2003). D-Optimal Split-Plot Designs with Given Numbers and Sizes of Whole Plots. *Technometrics* 45:235–245.

[3] Huang, P., Dechang, C., Voelkel, J. O. (1998). Minimum-Aberration Two-Level Split-Plot Designs. *Technometrics* 40:314–326.

[4] Jones, B., Goos, P. (2009). D-Optimal Split-Split-Plot Designs. *Biometrika* 96:67–82.

[5] Kowalski, S. M., Cornell, J. A., Vining, G. G. (2002). Split-Plot Designs and Estimation Methods for Mixture Experiments with Process Variables. *Technometrics* 44:72–79.

[6] Loeppky, J. L., Sitter, R. R. (2002). Analyzing Unreplicated Blocked or Split-Plot Fractional Factorial Designs. *Journal of Quality Technology*. 43:229–243.

[7] McLeod, R. G., Brewster, J. F. (2004). The Design of Blocked Fractional Factorial Split-Plot Experiments. *Technometrics.* 46:135–146.

[8] McLeod, R. G., Brewster, J. F. (2008). Optimal Foldover Plans for Two-Level Fractional Factorial Split-Plot Designs. *Journal of Quality Technology*. 40:227–240.

Chapter 7

Strip Plot Design

7

Strip Plot Design

Felix Kutsanedzie[1] *; Sylvester Achio[1]; Edmund Ameko[1]; George Kutsanedzie[2]; Gyekye Appiah Lewis[1]

[1]Accra Polytechnic, GP 561, Accra, Ghana

[2]Project Office of Electricity Company of Ghana, Accra, Ghana

Abstract

The strip plot design is one of the uncommon experimental designs and thus most researchers had little or no knowledge of it. Though it has semblance of the split plot design, it is used differently. This paper examines conditions or situations that necessitate the use of the strip plot designs and explains comprehensively with designed examples of experiments on how to handle such designs. The chapter examines the factor that is mainly used when greater precision is given to the interactions between the two factors. It also shows how the two factors are arranged in the design.

Keywords

Horizontal Strips, Vertical Strips, Interaction Strips, Design

7.1 Introduction

The strip plot design is most uncommon and less used design. However its use becomes necessary when the experimenter or the researcher needs to handle certain experiments. It used for analysing two factor experiments in which the factors to be handled are so large that they cannot be accommodated in a split plot design or would bring about a condition of heterogeneity in terms of the factors being considered. For instance in a two factor (subplot factor and main factor) experiment which involves testing the effects of four (4) tillage methods - *subplot factor*; and four (4) soil types - *main plot factor* on productivity of crops. In the design of such an experiment since large space is involved, it cannot be accommodated in a split plot design because large land space is needed and this would bring about heterogeneity. As a result the appropriate design for such experiment would be the strip plot design. In the strip plot design, the factors are handled by arranging them in strips. The factors are arranged and placed in a horizontal strip, vertical strip and interaction strip. One factor is placed in the horizontal strip and the second factor in the vertical strip; and the interactions between the factors placed in the vertical and horizontal strips are observed in the interaction strip, which is diagonal to the vertical and the horizontal strips.

While in the split plot designs, greater precision is given to the subplot factors while sacrificing that of the main plot factors; in strip plot designs, the precision for both the main plot factors and the subplot factors are sacrificed to give greater precision to only their interactions. Thus this design is only used when the experimenter needs to give greater precision to the interactions between the two factors being considered in an experiment. Usually in a strip plot design, the subplot factors are arranged by randomizing and placed within the horizontal strips while the main plot factors are placed in the vertical strips. The interactions between both factors are then seen or observed diagonally

between the spaces or field between the vertical and horizontal strips as in the diagrammatic illustration given below:

REPLICATION 1				
B1	B1A1	B1A2	B1A3	B1A4
B2	B2A1	B2A2	B2A3	B2A4
B3	B3A1	B3A2	B3A3	B3A4
B4	B4A1	B4A2	B4A3	B4A4
	A1	A2	A3	A4

REPLICATION 2				
B2	B2A2	B2A3	B2A4	B2A1
B3	B3A2	B3A3	B3A4	B3A1
B4	B4A2	B4A3	B4A4	B4A1
B1	B1A2	B1A3	B1A4	B1A1
	A2	A3	A4	A1

REPLICATION 3				
B3	B3A3	B3A4	B3A1	B3A2
B4	B4A3	B4A4	B4A1	B4A2
B1	B1A3	B1A4	B1A1	B1A2
B2	B2A3	B2A4	B2A1	B2A2
	A3	A4	A1	A2

For the purpose of illustrating and explaining the strip plot design, we proceed to design a strip plot experiment and learn how to analyze it. Assuming we are considering a strip plot experiment to study the effect on four (4) tillage methods on four (4) different soil types, we will have a similar design as given above, where the variables B are the sub plot factors arranged in the vertical strip and A, the main plot factors arranged the horizontal whereas the interaction between the subplot factor and the main plot factor.

Thus in the experimental layout, the vertical strip must be divided into four to contain the four different levels of the main factor A; the horizontal strip must

also be divided into four to contain the four different levels of the second factor B. The interaction between the two factors A and B (A x B) is observed in the interaction strip. This is done for every replication. Therefore if three (3) replications are to be considered, then the arrangement is done three times.

Analysis of this design is done in three-fold: analyzing the vertical strip, the horizontal strip and the interaction strip. The vertical analysis is done by first creating a cross table of the replication and the main factor which in this case is A. The analysis of the design involves computing the Sum of Squares (SS) due to the main factor (A); SS due replication; SS due to the interactions between the main factor (A) and replication (R) – A x R; and then the error SS (x). The horizontal strip analysis involves creating a cross table of the second factor and the replication and using it to compute the SS due to the second factor (B); SS due to the replications; SS due to the interaction between B x R; and the SS due to error (y). For the analysis of the interaction strip, the cross table of the main factor and the second factor is generated and then used in computing the SS due to the interaction of two factors (A x B). Thus the Total SS can be computed by considering every observation of the experiment while the error SS (z) can be obtained by subtracting all various variations from the Total SS variations.

Now let us assume a researcher used strip plot design to investigate the effect of four tillage methods (A) and four different soil types (B) with the experiment replicated three (3) times and obtained the results as shown in the table below:

REPLICATION 1				
B2	1.3	1.4	1.3	1.1
B3	1.5	1.4	1.5	1.3
B4	1.4	1.6	1.5	1.6
B1	1.4	1.2	1.7	1.3
	A2	A3	A4	A1

REPLICATION 2				
B1	1.3	1.4	1.4	2.3
B2	1.4	1.5	1.5	1.3
B3	1.5	1.3	1.3	1.4
B4	1.8	1.3	1.4	1.4
	A1	A2	A3	A4

REPLICATION 3				
B2	2.4	1.6	1.3	2.1
B3	1.6	2.0	1.5	2.0
B1	2.4	1.6	1.4	1.9
B4	1.4	1.5	1.7	1.8
	A4	A3	A1	A2

		A_1	A_2	A_3	A_4
	B_2	1.1	1.3	1.4	1.3
	B_3	1.3	1.5	1.4	1.5
REPLICATION 1	B_4	1.6	1.4	1.6	1.5
	B_1	1.3	1.4	1.2	1.7
	B_1	1.3	1.4	1.4	2.3
	B_2	1.4	1.5	1.5	1.3
REPLICATION 2	B_3	1.5	1.3	1.3	1.4
	B_4	1.8	1.3	1.4	1.4
	B_2	1.3	2.1	1.6	2.4
	B_3	1.5	2.0	2.0	1.6
REPLICATION 3	B_1	1.4	1.9	1.6	2.4
	B_4	1.7	1.8	1.5	1.4

7.2 For the Vertical Strip Analysis

Cross tabulation of Replication and Tillage Method (R x A)

	A1	A2	A3	A4	Rep Totals
REPLICATION I	5.3	5.6	5.6	6.0	22.5
REPLICATION II	6.0	5.5	5.6	6.4	23.5
REPLICATION III	5.9	7.8	6.7	7.8	28.2
Tillage Method Totals	17.2	18.9	17.9	20.2	
Grand Total					74.2

Calculating the Grand Total (GT)

$$GT = \sum_{i=1,j=1}^{n=12} (5.3 + 5.6 + 5.6 + 6.0 + \cdots + 7.8) = 74.2$$

Calculating the Correction Factor (CF)

$$CF = \frac{(GT)^2}{rab}$$

$$CF = \frac{(74.2)^2}{3\ x\ 4\ x\ 4}$$

$$CF = \frac{5505.64}{48}$$

$$CF = 114.70$$

where r=no. of replications a=no. of levels of tillage methods, b=no. of levels of soil.

At this point one must compute the sum of squares for items or factor within the vertical strip and these are done be done and illustrated below:

Calculating the Total Sum of Squares (TSS)

$$TSS = \sum sum\ of\ square\ of\ all\ observations - CF$$

$$TSS = \sum_{i=1,j=1,k=1}^{n=48} (a_1 b_1 r_1)^2 + \ldots\ldots\ldots\ldots\ldots + (a_2 b_3 r_3)^2 - CF$$

$$TSS = \sum[(1.1)^2 + (1.3)^2 + (1.4)^2 + (1.3)^2 + \cdots + (1.4)^2] - 114.70$$

$$TSS = \sum[1.21 + 1.69 + 1.96 + 1.69 + \cdots + 1.96] - 114.70$$

$$TSS = 118.96 - 114.70$$

$$TSS = 4.26$$

Calculating the Replication Sum of Squares (RSS)

$$RSS = \sum\frac{[(R_{1_T})^2 + (R_{2_T})^2 + (R_{3_T})^2]}{ab} - CF$$

$$RSS = \sum\frac{[(22.5)^2 + (23.5)^2 + (28.2)^2]}{4\ x\ 4} - 114.70$$

$$RSS = \sum\frac{[506.25 + 552.25 + 795.24]}{16} - 114.70$$

$$RSS = \sum(\frac{[1853.74]}{16} - 114.70)$$

$$RSS = 115.86 - 114.70$$

$$RSS = 1.16$$

Calculating the Tillage Method Sum of Squares (ASS)

$$ASS = \sum\frac{[(A_{1_T})^2 + (A_{2_T})^2 + (A_{3_T})^2 + (A_{4_T})^2]}{br} - CF$$

$$ASS = \sum\frac{[(17.2)^2 + (18.9)^2 + (17.9)^2 + (20.2)^2]}{4\ x\ 3} - 114.70$$

$$ASS = \sum\frac{[295.84 + 357.21 + 320.41 + 408.04]}{12} - 114.70$$

$$ASS = \sum(\frac{[1381.5]}{12} - 114.70)$$

$$ASS = 115.13 - 114.70$$

$$ASS = 0.42$$

Calculating the Replication and Tillage Method Sum of Squares (RASS)

Note: one is expected to use the cross tabulation table for replication and machine type

$$RASS = \sum_{i=1,j=1}^{n=12} \frac{(A_1r_1)^2 + \ldots\ldots\ldots\ldots\ldots + (A_2r_3)^2}{b} - CF$$

$$RASS = \sum \frac{[(5.3)^2 + (5.6)^2 + (5.6)^2 + \cdots + (7.8)^2]}{4} - 114.70$$

$$RASS = \sum \frac{[28.09 + 31.36 + 31.36 + \cdots + 60.84]}{4} - 114.70$$

$$RASS = \sum \frac{[466.76\ ]}{4} - 114.70$$

$$RASS = (116.69 - 114.70)$$

$$RASS = 1.99$$

Error (x)Sum of Squares (for the vertical strip) $= RASS - RSS - ASS$

$$E(x)SS = RASS - RSS - ASS$$

$$E(x)SS = 1.99 - 1.16 - 0.42$$

$$E(x)SS = 0.41$$

7.3 For the Horizontal Strip Analysis

Cross tabulation of Replication and Soil Type (R x B)

	B1	B2	B3	B4	Rep Totals
REPLICATION I	5.6	5.1	5.7	6.1	22.5
REPLICATION II	6.4	5.7	5.5	5.9	23.5
REPLICATION III	7.3	7.4	7.1	6.4	28.2
Soil Type Totals	19.3	18.2	18.3	18.4	
Grand Total					74.2

Calculating the Replication Sum of Squares (RSS)

$$RSS = \sum \frac{[(R_{1_T})^2 + (R_{2_T})^2 + (R_{3_T})^2]}{ab} - CF$$

$$RSS = \sum \frac{[(22.5)^2 + (23.5)^2 + (28.2)^2]}{4\ x\ 4} - 114.70$$

$$RSS = \sum \frac{[\ 506.25 + 552.25 + 795.24]}{16} - 114.70$$

$$RSS = \sum (\frac{[1853.74]}{16} - 114.70)$$

$$RSS = 115.86 - 114.70$$

$$RSS = 1.16$$

Calculating the Soil Type Sum of Squares (BSS)

$$BSS = \sum \frac{[(B_{1_T})^2 + (B_{2_T})^2 + (B_{3_T})^2 + (B_{4_T})^2]}{ar} - CF$$

$$BSS = \sum \frac{[(19.3)^2 + (18.2)^2 + (18.3)^2 + (18.4)^2]}{4\ x\ 3} - 114.70$$

$$BSS = \sum \frac{[\ 372.49 + 331.24 + 334.89 + 338.56]}{12} - 114.70$$

$$BSS = \sum (\frac{[\ 1377.18]}{12} - 114.70)$$

$$BSS = 114.77 - 114.70$$

$$BSS = 0.07$$

Calculating the Replication and Tillage Method Sum of Squares (RBSS)

Note: one is expected to use the cross tabulation table for replication and machine type

$$RBSS = \sum_{i=1,j=1}^{n=12} \frac{(B_1 r_1)^2 + \ldots\ldots\ldots\ldots\ldots + (B_2 r_3)^2}{a} - CF$$

$$RBSS = \sum \frac{[(5.6)^2 + (5.1)^2 + (5.7)^2 + \cdots + (6.4)^2]}{4} - 114.70$$

$$RBSS = \sum \frac{[31.36 + 26.01 + 32.49 + \cdots + 40.96]}{4} - 114.70$$

$$RBSS = \sum \frac{[465\ ]}{4} - 114.70$$

$$RBSS = (116.25 - 114.70)$$

$$RBSS = 1.55$$

Error (y) Sum of Squares (for the horizontal strip) $= RBSS - RSS - BSS$

$$E(y)SS = RBSS - RSS - BSS$$

$$E(y)SS = 1.55 - 1.16 - 0.07$$

$$E(y)SS = 0.32$$

7.4 For the Interaction Strip Analysis

Cross tabulation of Tillage Method (A) and Soil Type (B) (A x B)

	A1	A2	A3	A4	B Totals
B1	4.0	4.7	4.2	6.4	19.3
B2	3.8	4.9	4.5	5.0	18.2
B3	4.3	4.8	4.7	4.5	18.3
B4	5.1	4.5	4.5	4.3	18.4
A TOTAL	17.2	18.9	17.9	20.2	74.2

Calculating the Sum of Squares of the interaction between Tillage Method and Soil Type

$$(AxB)SS = \sum_{i=1,j=1}^{n=16} \frac{(A_1B_1)^2 + \ldots\ldots\ldots\ldots\ldots + (A_4B_4)^2}{r} - CF$$

$$(AxB)SS = \sum \frac{[(4.0)^2 + (4.7)^2 + (4.2)^2 + \cdots + (4.3)^2]}{r} - 114.70$$

$$(AxB)SS = \sum \frac{[16 + 22.09 + 17.64 + \cdots + 18.49]}{3} - 114.70$$

$$(AxB)SS = \sum \frac{[248.23\ ]}{3} - 114.70$$

$$(AxB)SS = 116.42 - 114.70$$

$$(AxB)SS = 1.72$$

$$Error\ (z)\ Sum\ of\ Squares\ (E(z)SS) = TSS - (RSS + ASS + E(x)SS + BSS + (AxB)SS + E(y)SS$$

$$E(z)SS = 4.26 - (1.16 + 0.42 + 0.41 + 0.07 + 1.72 + 0.32)$$

$$E(z)SS = 4.26 - 4.1$$

$$E(z)SS = 0.16$$

7.5 Completing the ANOVA Table

Sources of Variation	df	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Fcal
Replication	`r -1 = 3 – 1 = 2	1.16	$\frac{1.16}{2} = 0.58$	
Main factor - Tillage Method (A)	a -1 = 4-1 = 3	0.42	$\frac{0.42}{3} = 0.14$	$\frac{MSS(A)}{MSS(error\ x)} = \frac{0.14}{0.07} = 2$
Error (x)	(r-1)(a-1) = 2 x 3 = 6	0.41	$\frac{0.41}{6} = 0.07$	
Sub factor – Soil type (B)	b -1 = 4 -1 =3	0.07	$\frac{0.07}{3} = 0.02$	$\frac{MSS(A)}{MSS(errory)} = \frac{0.02}{0.53} = 0.04$
Error (y)	(r -1)(b -1) = (3-1)(4-1) =2 x 3 = 6	0.32	$\frac{0.32}{6} = 0.53$	
Interaction between – Tillage Method x Soil Type (A x B)	(a-1)(b-1) = (4-1)(4-1) = 3 x 3 = 9	1.72	$\frac{1.72}{9} = 0.19$	$\frac{MSS(A)}{MSS(error\ z)} = \frac{0.19}{0.01} = 19$
Error (z)	(r-1)(a-1)(b-1) =(3-1)(4-1)(4-1) = 2 x 3 x 3 = 18	0.16	$\frac{0.16}{18} = 0.01$	
Total	rab-1 = 48 -1 = 47	4.26		

Reading of the F-critical or tabulated values from the F- table at the various assigned levels of significance allowed.

For the case being considered, 1% and 5% levels of significance would be used.

F-critical table at 5% (0.05)

Critical values of F for the 0.05 significance level:

	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.39	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.97	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.97	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38

F-critical table at 1% (0.01)

Critical values of F for the 0.01 significance level:

	1	2	3	4	5	6	7	8	9	10
1	4052.19	4999.52	5403.34	5624.62	5763.65	5858.97	5928.33	5981.10	6022.50	6055.85
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05
6	13.75	10.93	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	8.86	6.52	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.90	3.81
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.25	4.02	3.84	3.71	3.60	3.51
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43

Note: The read values from the table have been highlighted to show how the F-critical values read from the F-table. The values are always read using the degree of freedom of a particular source of variation against the degree of freedom of a particular error.

Sources of Variations	Fcal	Fcrit (5%)	Fcrit (1%)
Main Plot Factor (A)	$\frac{0.14}{0.07} = 2$	df(3,6) = 4.76	df(3,6) = 9.78
Sub Plot Factor (B)	$\frac{0.02}{0.53} = 0.04$	df (3,6) = 4.76	df(3,6) = 9.78
Interactions (A x B)	$\frac{0.19}{0.01} = 1.16$	df (9,18) =2.46	df(9,18) = 3.60

7.6 Taking the Decision and Making the Conclusion

Main Plot Factor (Tillage Method):

$$Fcal\ (2) < Fcrit\ (\ df\ 3{,}6\ @\ 5\% = 4.76)$$

$$Fcal\ (2) < Fcrit\ (\ df\ 3{,}6\ @\ 1\% = 9.78)$$

Conclusion on Main Plot Factor (Soil Type):

There is no significance difference exists between the tillage methods (A) used in the experiment at both levels of significance. Thus the tillage methods are similar.

Subplot Factor or Second Factor (Soil Type):

$$Fcal\ (0.04) < Fcrit\ (\ df\ 3{,}6\ @\ 5\% = 4.76)$$

$$Fcal\ (0.04) < Fcrit\ (\ df\ 3{,}6\ @\ 1\% = 9.78)$$

Conclusion on Subplot (Soil Type)

No significant difference exists between the various soil types. Meaning the soil types did not differ statistically

Interactions (M x W):

$$Fcal\ (1.16) < Fcrit\ (\ df\ 9{,}18\ @\ 5\% = 2.46)$$

$$Fcal\ (1.16) < Fcrit\ (\ df\ 9{,}18\ @\ 1\% = 3.60)$$

Conclusion on Interactions (A x B)

There exists no significance difference. Hence one can conclude that the interactions between the main and sub factors did not differ statistically.

Bibliography

[1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.

[2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), Handbook on research on teaching. Chicago: Rand McNally & Co.

[3] Cornfield, J., Tukey, J. W. (1956) Average values of mean squares in factorials. *Annals of Mathematical Statistics*, 27:907-949.

[4] Cox, D. R. (1935). *Planning of experiments*. New York: Wiley.

[5] Fisher, R. A. (1935). *The design of experiments*. 1st ed. Oliver & Boyd, London.

[6] Winer, B. J. (1962). *Statistical principles in experimental design*. McGraw-Hill, New York.

[7] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs*. John Willey & Sons, N.Y.

[8] Le Clerg, E. L., Leonard, W. H., Clark, A. G. (1966). *Field Plot Technique*. Burgess Pub. Co., Minn., USA.

[9] Gomez, K. A. and Gomez, A. A. (1984). *Statistical Procedures for Agricultural Research*. John Wiley & Sons, New York.

[10] Mead, R., Curnow, R. N. (1983). *Statistical Methods in Agriculture and Experimental Biology*. Chapman and Hall, London.

[11] Pearce, S. C., Clarke, G. M., Dyke, G. V., Kempson, R. E. (1988). *A Manual of Crop Experimentation*. Charles Griffin & Co. Ltd., Oxford.

Mr. Felix Kutsanedzie

Mr. Felix Kutsanedzie, the lead author of this book, is a Senior Research Fellow / lecturer and currently the Head of the Accra Polytechnic Research and Innovations Centre. He holds a MSc. degree in Bio-Engineering; a BSc. (Agric-Mech); and an Adv. Dip. (Project Management). He is an Associate Editor of Directory Open Access Journals (DOAJ) and a prolific writer with several peer-reviewed publications to his credit.

Professor Sylvester Achio

Professor Sylvester Achio, is currently the Rector of Accra Polytechnic and a professor at the Department of Science Laboratory Technology of Accra Polytechnic, who holds MSc. Hons. (Research) Agric (Agronomy), PGC Ed. (Russian Language), PhD. Bio. Sci. (Microbiology). He has authored several technical books and also has presented papers at many national and international conferences, workshops, seminars as well as several peer-reviewed journal publications to his credit.

Mr. Edmund Ameko

Mr. Edmund Ameko is a Senior Lecturer at the Department of Science Laboratory Technology of Accra Polytechnic. He holds a BSc. in Biochemistry and a MSc. degree in Food Science and Technology, and currently the Dean of School of Applied Sciences at Accra Polytechnic. He is a prolific writer with several peer-reviewed publications.

To order additional copies of this book, please contact:
Science Publishing Group
book@sciencepublishinggroup.com
www.sciencepublishinggroup.com

ISBN 978-1-940366-50-0

Price: US $120