Unit 2
Research Design

Research Design

- Task of defining the research problem is the preparation of the research project, popularly known as the “research design".
- Decisions regarding what, where, when, how much, by what means concerning an inquiry or a research study constitute a research design.
- A plan or strategy for conducting the research.
- A research design is one that minimizes bias and maximizes the reliability of the data.
- It also yields maximum information, gives minimum experimental error, and provides different aspects of a single problem.
- A research design depends on the purpose and nature of the research problem. Thus, one single design cannot be used to solve all types of research problem, i.e., a particular design is suitable for a particular problem.

Need of research design

- Smooth sailing of research operations.
- Stands for advance planning of the methods.
- Great bearing on reliability of the results.
- Helps to give directions.
- Helps in decision making.
- Research design prevent blind searching.
- Helps researchers to anticipate potential problems in collecting data etc.

Features of a Good Design

- The means of obtaining information
- The availability and skills of the researcher and his staff, if any;
- The availability of time and money for the research work.
- It should be flexible enough to consider different aspects of the study in case of exploratory.
- The design should be accurate with minimum bias in case of accurate description
- Control of extraneous variables
- Statistical correctness for testing hypothesis

Concepts relating to research design

- Dependent and independent variables
- Extraneous variable
- Control
- Experimental and non-experimental hypothesis- testing research
- Experimental and control groups
- Treatments

Variables

- Any characteristic which is subject to change and can have more than one value such as age, intelligence, motivation, gender, etc.
- Types of variables
- **Independent vs. Dependent vs. Controlled Variables**
- **Categorical vs. Continuous Variables**
- **Quantitative vs. Qualitative Variables**

You design a study to test whether changes in room temperature have an effect on math test scores.

Your independent variable is the temperature of the room. You vary the room temperature by making it cooler for half the participants, and warmer for the other half.

Your dependent variable is math test scores. You measure the math skills of all participants using a standardized test and check whether they differ based on room temperature.

Identifying independent vs. dependent variables

Distinguishing between independent and dependent variables can be tricky when designing a complex study or reading an academic research paper.

A dependent variable from one study can be the independent variable in another study, so it’s important to pay attention to research design.

Here are some tips for identifying each variable type.

Recognizing independent variables

Use this list of questions to check whether you’re dealing with an independent variable:

- Is the variable manipulated, controlled, or used as a subject grouping method by the researcher?
- Does this variable come before the other variable in time?
- Is the researcher trying to understand whether or how this variable affects another variable?

Recognizing dependent variables

Check whether you’re dealing with a dependent variable:

- Is this variable measured as an outcome of the study?
- Is this variable dependent on another variable in the study?
- Does this variable get measured only after other variables are altered?

Dependent Variable

- Variable affected by the independent variable
- It responds to the independent variable.
- In an experiment that which is supposed to be changed by the independent.

Independent Variable

- Variable that is presumed to influence other variable
- It is the presumed cause, whereas the dependent variable is the presumed effect.
- In an experiment that which is supposed to be manipulated by you.
- The variable manipulated by the experimenter.

Extraneous variable

- Independent variable that are not related to the purpose of the study, but may affect the dependent variable are termed as extraneous variables.
- Whatever effect is notices on dependent variable as a result of extraneous variable is technically described as an 'experimental error'.

Control

- One of the important characteristics of a good research design is to minimize the influence or effect of extraneous variable.

Experimental and non experimental hypothesis testing

- Research in which the independent variable is manipulated is termed experimental hypothesis –testing research.
- Research in which an independent variable is not manipulated is called non-experimental hypothesis.

Confounded Relationship

Confounding relationship means the distortion of the association between the independent and dependent variables because a third variable is independently associated with both.

A causal relationship between two variables is often described as the way in which the independent variable affects the dependent variable.

Experimental and control group

- In an experimental hypothesis –testing research when a group is exposed to usual conditions, it is termed as control group.
- When a group is exposed to some novel or special condition it is termed as experimental group.

Treatments

- The different conditions under which experimental and control groups are put are usually referred to as treatments.

Categorical vs. Continuous Variables

- *Categorical variables are variables that can take on* specific values only within a defined range of values like gender, marital status
- consisting of discrete, mutually exclusive categories, such as “male/female,” “White/Black,” etc
- *Continuous variables are variables that can theoretically take on any* value along a continuum like age, income weight, height etc..
- When compared with categorical variables, continuous variables can be measured with a greater degree of precision.
- The choice of which statistical tests will be used to analyze the data is partially dependent on whether the researcher uses categorical or continuous variables.
- Certain statistical tests are appropriate for categorical variables, while other statistical tests are appropriate for continuous variables.
- As with many decisions in the research-planning process, the choice of which type of variable to use is partially dependent on the question that the researcher is attempting to answer.

Quantitative vs. Qualitative Variables

- *Qualitative variables are variables that vary in kind, like* "attractive" or "not attractive," "helpful" or "not helpful," or "consistent" or "not consistent"

- *Quantitative variables are those that vary in amount like height, weight, salary etc*

Hypothesis

- The research hypothesis is a predictive statement that relates an independent variable to dependent variable.
- A hypothesis may be defined as a proposition or set of proposition set forth as an explanation for occurrence of some specified group of phenomena either asserted merely as a provisional conjucture to guide some investigation or accepted as highly probable in the light of established facts

Purpose

- Guides/gives direction to the study/investigation
- Defines Facts that are relevant and not relevant
- Suggests which form of research design is likely to be the most appropriate
- Provides a framework for organizing the conclusions of the findings
- Limits the research to specific area
- Offers explanations for the relationships between those variables that can be empirically tested
- Furnishes proof that the researcher has sufficient background knowledge to enable her/him to make suggestions in order to extend existing knowledge
- Structures the next phase in the investigation and therefore furnishes continuity to the examination of the problem

Experimental and non-experimental hypothesis testing

- When a group is exposed to usual conditions, it is termed as a control group.
- But when the group is exposed to be some special condition, it is termed as Experimental group

Categorizing Hypotheses

Can be categorized in different ways

1. Based on their *formulation*

- Null Hypotheses and Alternate Hypotheses

2. Based on direction

- Directional and Non-directional Hypothesis

1. Null Hypotheses and Alternate Hypotheses

- Null hypothesis always predicts that
  - no differences between the groups being studied (e.g., experimental vs. control group) or
  - no relationship between the variables being studied

- By contrast, the alternate hypothesis always predicts that there will be a difference between the groups being studied (or a relationship between the variables being studied)

Problem Statement 1: Does eating an apple daily ensure weight loss? State both Null and Alternative hypotheses.

Answer:
Null Hypothesis (H0): Eating apples daily does not affect weight loss.

Alternative Hypothesis (H1): Eating apples affects weight loss.

2. Directional Hypothesis and Non-directional Hypothesis

- Simply based on the wording of the hypotheses we can tell the difference between directional and non-directional

  – If the hypothesis simply predicts that there will be a difference between the two groups, then it is a ***non-directional hypothesis***. It is non-directional because it predicts that there will be a difference but does not specify how the groups will differ.

  – If, however, the hypothesis uses so-called comparison terms, such as "greater," "less," "better," or "worse," then it is a ***directional hypothesis***. It is directional because it predicts that there will be a difference between the two groups and it specifies how the two groups will differ

Directional hypothesis: A directional (or one tailed hypothesis) states which way you think the results are going to go, for example in an experimental study we might say…"Participants who have been deprived of sleep for 24 hours will have more cold symptoms in the following week after exposure to a virus than participants who have not been sleep deprived"; the hypothesis compares the two groups/conditions and states which one will ….have more/less, be quicker/slower, etc.

Non-directional hypothesis: A non-directional (or two tailed hypothesis) simply states that there will be a difference between the two groups/conditions but does not say which will be greater/smaller, quicker/slower etc. Using our example above we would say "There will be a difference between the number of cold symptoms experienced in the following week after exposure to a virus for those participants who have been sleep deprived for 24 hours compared with those who have not been sleep deprived for 24 hours."

Experimental Design

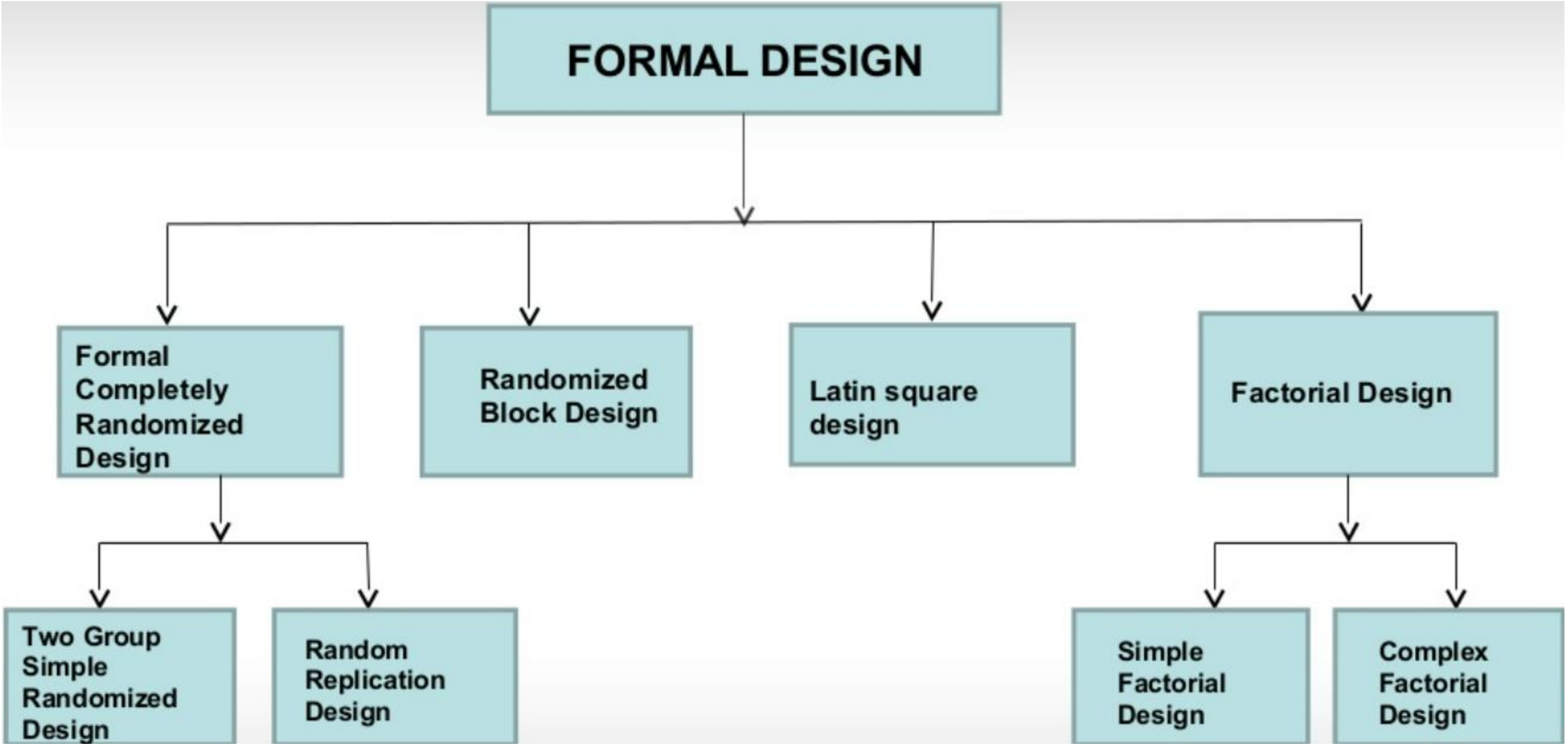
Experimental design methods allow the experimenter to understand better and evaluate the factors that influence a particular system by means of statistical approaches. Such approaches combine theoretical knowledge of experimental designs and a working knowledge of the particular factors to be studied.

Principles of Experimental Design

- **Control** – make conditions as similar as possible for all treatment groups (aside from the actual treatments).
  - *If we observe a difference between groups, we want to know that it is a result of the treatment(s)!*

- **Randomization** – the use of chance to assign subjects/units to treatments
  - *This helps create roughly equivalent groups of experimental units by balancing the effects of lurking variables that aren't controlled on the treatment groups.*

- **Replication** of the experiment on many subjects/in different locations/etc.



Single factor/Single independent variable

Multi factors/Multiple independent variables

Single Factorial Design

- Single factor experiments are those experiments in which only a single factor varies while all others are kept constant.

- Here the treatments consist exclusively of the different levels of the single variable factor.

- All other factors are applied uniformly to all plots.

- Single-factor experiments are useful for understanding the effect of a single independent variable on a dependent variable.
- They are often used as a first step in understanding the relationships between variables and can help to identify potential relationships or patterns that can be further explored in more complex experiments.
- A single-factor experiment, also known as a one-factor experiment or a one-way experiment, is a type of experimental design in which the effect of a single independent variable on a dependent variable is studied.
- In a single-factor experiment, the experimenter manipulates the level of the independent variable and measures the effect on the dependent variable.

An example of a single-factor experiment might be a study examining the effect of different levels of light intensity on the growth of a particular type of plant. In this case, the independent variable would be the level of light intensity, and the dependent variable would be the growth of the plant.

All other parameters are kept constant like type of fertilizer applied, amount of fertilizer, the amount of water applied to the plant, etc

Completely Randomised Design

- It is commonly called as CRD.
- CRD is a statistical experimental design where the treatments are assigned completely at random so that each treatment unit has the same chance of receiving any one treatment.
- In CRD, any difference among experimental units receiving the same treatment is considered as an experimental error.
- CRD is applicable only when the experimental material is homogenous (Example: Homogenous soil condition in the field).
- Usually in the field, the soil will be HETEROGENOUS.
- Thus, CRD is not a preferable method in field experiments.
- CRD is generally applicable to the lab experimental conditions.
- In labs, the environmental conditions can be easily controlled.
- The concept of 'Local-control' is not used in CRD.

- In CRD, as the name suggests, treatments are assigned completely randomly so that each treatment unit gets the same chance of receiving any one treatment.
- This is suitable only for the experiments such as laboratory experiments or greenhouse studies etc, where the experiment material is homogeneous and not for heterogeneous studies.
- CRD experiment For example, an agricultural scientists wants to study the effect of 4 different fertilizers (A,B,C,D) on corn productivity. 4 replicates of the 4 treatments are assigned at random to the 16 experimental units ! ➢Treatment : Types of fertilizer (A,B,C,D) ➢Experimental unit : Corn tree ➢Dependent variable : Production of corn

Advantages of CRD

- Ø CRD is easy to understand and calculate the variance.
- Ø The number of replications can vary from treatment to treatment.
- Ø CRD has high flexibility and thus any number of treatments can be used.
- Ø Simple statistical analysis is required in the analysis of CRD.
- Ø CRD provides maximum number of degree of freedom.

Disadvantages of CRD

- Ø CRD can be applied only to homogenous experiments.
- Ø The principle of 'Local-control' is not used in CRD.

Randomised Complete Block Design

- Ø It is also called RBD.
- Ø RBD is the most commonly used experimental design in agriculture.
- Ø Here the 'local-control' is adopted and the experimental material is grouped into homogenous subgroups.
- Ø Each such sub-group is called blocks.
- Ø A block contains the entire set of treatments and thus a block is equivalent to a replication.
- Ø Characteristic of RBD:
- $ Presence of blocks of equal sizes, each of which contains all the treatments.
- $ In RBD, we have to use all the three principles of design of experiments (replication, randomization and local-control).

To illustrate how the RCBD works, consider a study that is testing the effectiveness of two different fertilizers on the growth of tomato plants. The experimental units in this case are individual tomato plants, and the blocks are created based on factors that can affect plant growth, such as soil quality, temperature, and light exposure.

The study would involve randomly assigning each treatment (fertilizer) to each plant within each block. This ensures that each treatment is tested under the same conditions, and any differences observed between the treatments are likely to be due to the treatment itself rather than extraneous variables.

How to implement the RCBD in your research

If you are interested in using the RCBD in your research, here are the steps you should follow:

- Identify the factor(s) that may affect the outcome of your study.
- Divide your experimental units into blocks based on these factors.
- Randomly assign each treatment to the experimental units within each block.
- Collect and record the data from the experiment.
- Analyze the data using statistical methods to determine if the differences observed between treatments are significant.

Examples of RBCD

Clinical trial: A pharmaceutical company wants to test the effectiveness of a new drug for treating a particular medical condition. The company recruits patients with the same condition and divides them into blocks based on their age or severity of the condition. Then, the patients within each block are randomly assigned to receive either the new drug or a placebo. This design allows the company to control for the effect of age or severity on the treatment outcome.

Educational study: A researcher wants to investigate the effectiveness of two teaching methods on student performance in a particular subject. The researcher divides the students into blocks based on their prior academic performance and randomly assigns each teaching method to different classes within each block. This design allows the researcher to control for the effect of prior academic performance on the outcome and determine which teaching method is more effective.

Differences between CRD and RCBD

When the treatments are arranged randomly over the whole of a previously determined set of experimental units, the design is known as Complete Randomized Design(CRD).
In Randomized Complete Block Design (RCBD) the entire experimental material is divided into blocks and the complete set of treatments are allocated randomly in each block
CRD is used where the experimental units are homogeneous in nature like in controlled environment or chemical environment but RCBD is used in case of heterogeneous experimental material like agricultural experiments conducted in field where the soil fertility is heterogeneous.
In case of RCBD all the 3 principles of experimental design is used i.e. randomisation, replication and local control but local control is not used in CRD.
There are 2 sources of variation in case of ANOVA analysis of CRD viz. Treatments and error. But in RCBD there are 3 sources of variation in case of ANOVA analysis viz. Treatments, replication and error.

Advantages of RBD

Ø RBD is more efficient and accurate when compared to CRD.

Ø Chance of error in RBD is comparatively less.

Ø Flexibility is also very high in RBD and thus any number of treatments and any number of replications can be used.

Ø Statistical analysis is relatively simple and easy.

Ø Statistical analysis simple when one value is missing.

Ø Errors of any treatment can be isolated.

Disadvantages of RBD

Ø RBD is not advised for very large number of treatments.

Ø If the heterogeneity of the plot is very high, RBD cannot be applied. When the number of treatments is very large then the size of each block will be increased so that there may be heterogeneous blocks within.

Ø With large number of treatments, the possibility of experimental errors will be high.

Latin Square Design

- Commonly called as LSD.
- LSD is a design where the experimental material is divided into ‘m’ rows, ‘m’ columns and ‘m’ treatments – assigned by randomization method to rows and columns.
- The randomization is in such a way that each treatment occurs only once in each row and in each column.

Latin Square Design

- Allows the researcher to statistically control two **non-interacting** external variables as well as to manipulate the independent variable.
- Each external or blocking variable is divided into an **equal** number of blocks, or levels.
- The independent variable is also divided into the same number of levels.
- A Latin square is conceptualized as a table, with the rows and columns representing the blocks in the two external variables.
- The levels of the independent variable are assigned to the cells in the table.
- The **assignment rule** is that each level of the independent variable should appear only once in each row and each column, as shown in Table

Advantages of LSD

- Statistical analysis is relatively simple (complicated than CRD and RBD).
- Statistical analysis is simple if one value is missing.
- Most efficient design when compared to CRD and RBD.

Disadvantages of LSD

- LSD is not suitable for agricultural experiments.
- Statistical analysis is complicated when two or more values are missing.
- Difficult when treatments are more than ten.

Factorial Design

- Is used to measure the effects of two or more independent variables at various levels.
- A factorial design may also be conceptualized as a table.
- In a two-factor design, each level of one variable represents a row and each level of another variable represents a column.

Simple factorial designs

- In case of simple factorial designs, we consider the effects of varying two factors on the dependent variable, but when an experiment is done with more than two factors, we use complex factorial designs.
- Simple factorial design is also termed as a ‘two-factor factorial design’.
- Simple factorial design may either be a 2×2 simple factorial design, or it may be, say, 3×4 or 5×3 or the like type of simple factorial design

- In this design the extraneous variable to be controlled by homogeneity is called the control variable and the independent variable, which is manipulated, is called the experimental variable.
- Then there are two treatments of the experimental variable and two levels of the control variable. As such there are four cells into which the sample is divided. Each of the four combinations would provide one treatment or experimental condition.
- Subjects are assigned at random to each treatment in the same manner as in a randomized group design. The means for different cells may be obtained along with the means for different rows and columns.
- Means of different cells represent the mean scores for the dependent variable and the column means in the given design are termed the main effect for treatments without taking into account any differential effect that is due to the level of the control variable. Similarly, the row means in the said design are termed the main effects for levels without regard to treatment.
- This design also facilitates the testing of several hypothesis at a single time.
- Typical factorial design incorporates 2X2 or 2X3 factorial, but it can be in any combination.
- The first number (α) refers to the independent variables or the type of experimental treatments, and the second number (β) refers to the level or frequency of the treatment.

- This shows that a 2 × 2 simple factorial design can be generalised to any number of treatments and levels. Accordingly we can name it as such and such (–×–) design. In such a design the means for the columns provide the researcher with an estimate of the main effects for treatments and the means for rows provide an estimate of the main effects for the levels. Such a design also enables the researcher to determine the interaction between treatments and levels.
- Thus, through this design we can study the main effects of treatments as well as the main effects of levels. An additional merit of this design is that one can examine the interaction between treatments and levels, through which one may say whether the treatment and levels are independent of each other or they are not so. The following examples make clear the interaction effect between treatments and levels. The data obtained in case of two (2 × 2) simple factorial studies may be as given below.

Advantages of factorial design

- There are many advantages of the factorial experiments:

1. Factorial experiments are advantageous to study the combined effect of two or more factors simultaneously and analyze their interrelationships. Such factorial experiments are economic in nature and provide a lot of relevant information about the phenomenon under study. It also increases the efficiency of the experiment.
2. It is an advantageous because a wide range of factor combination are used. This will give us an idea to predict about what will happen when two or more factors are used in combination.
3. The factorial approach will result in considerable saving of the time and the experimental materials. It is because the time required for the combined experiment is less than that required for the separate experiments.
4. In single factor experiments, the results may not be satisfactory because of the changes in environmental conditions. However, in factorial experiments such type of difficulties will not arise even after several factors are investigated simultaneously.
5. Information may obtained from factorial experiments is more complete than that obtained from a series of single factor experiments, in the sense that factorial experiments permit the evaluation of interaction effects.

Disadvantages of factorial design

➢ The disadvantages of the factorial experiments are:

- It is disadvantageous because the execution of the experiment and the statistical analysis becomes more complex when several treatments combinations or factors are involved simultaneously.
- It is also disadvantageous in cases where may not be interested in certain treatment combinations but we are forced to include them in the experiment. This will lead to wastage of time and also the experimental material.
- In factorial experiments, the number of treatment combinations will increase if the factors are increased. This will also lead to the increase in block size, which in turn will increase the heterogenicity in the experimental material. Because of this it will lead to the increased experimental error and will decrease the precision in the experiment. Appropriate block size must be maintained.

Thank You