1 Design of experiments:

It means **how** to **design** an experiment in the sense that how the observations or measurements should be obtained to answer a **valid**, **efficient** and **economical** way. If the experiment is **designed properly,** then the data generated is **valid** and **properly analyzed.** If the experiment is **not well designed**, the validity of the statistical analysis is **questionable** and may be **invalid**.

- How does the feed formulation affect the body weight of animals?
- Which variety of crop species would be good for particular region?
- Does the date of sowing affect the crop yield?
- How does the water availability and its quality influence the crop yield?
- How does greenhouse gases emission influence the global warming?
- Does the use of pesticide in crops affect the health of farmers as well as the people consuming the produce?

2 Brief history of design of experiments:

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

3 Basic terminology in Experiment Design.

3.1 Experiment:

An **e**xperiment is a **controlled study** in which the researcher attempts to **understand** *cause and effect relationships*. Based on the analysis, the researcher draws a conclusion about whether the treatment (**independent** variable) had a causal effect on the **dependent** variable.

3.2 Treatment

In experiments, a **treatment** is something that researchers administer to experimental units. Different **objects** or **procedures** which are to be **compared** in an experiment are called treatments. Some examples of treatments are:

- Different kinds of fertilizer in agronomic experiments,
- Different irrigation methods or levels of irrigation

- Different fungicides in pest management experiments
- Different varieties of crops,
- Different pesticides
- Grazing systems for animals
- Different tree species in agro-forestry experiments

3.3 Control treatment

Is a standard treatment that is used as a baseline or **basis of comparison** for the other treatments? This control treatment might be the treatment which is **currently in use**, or it might be a **no treatment** at all. For example:

- A study of **new pesticides** could use a standard pesticide as a **control treatment**
- An experiment involving **fertilizers** may have one treatment as **no fertilizers** at all.

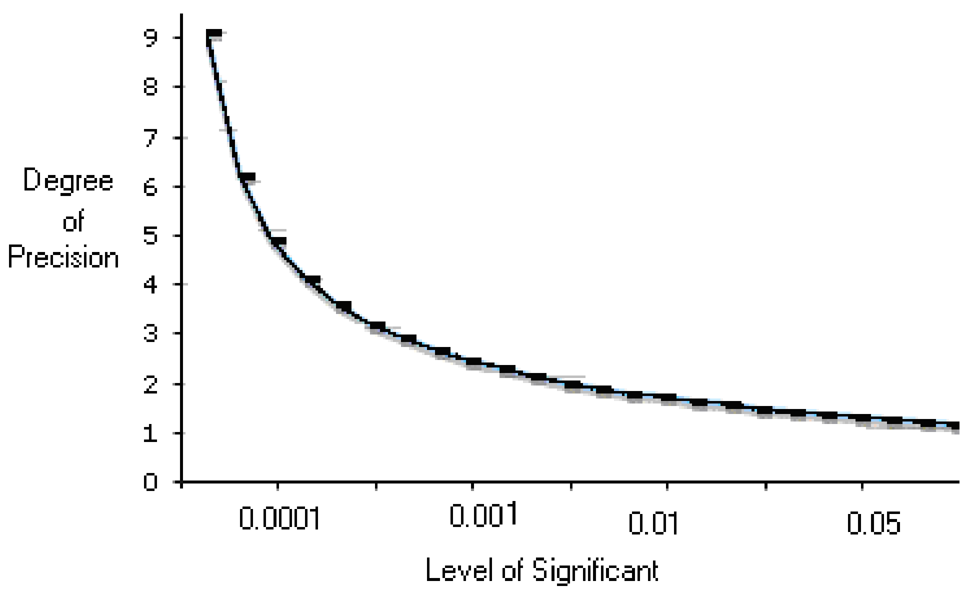
3.4 Experimental unit:

Experimental units are the subjects or objects on which the treatments are applied. For example:

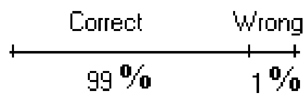
- Plots of land receiving fertilizer.
- Groups of animals receiving different feeds.
- Pots in glasshouse experiments.
- Petri dishes or tissues to culture bacteria or micro-organisms in laboratory experiments.

3.5 Levels of Significance:

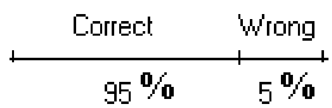
It is the probability level below which we reject Ho, and it is range between **0.01 - 0.05**. also denoted as **alpha** or **α**. As the level is low the precision (accuracy) is high, because you minimize the area of study and repeat the experiment for several times, so the precision will be high. Thus the errors will be low as the number of the observations is high and vice versa. Also as the S.E is low, the level of significant is low, i.e the precision is high, and this is called. It will be done by **increasing sample size (n)**.



1) 0.01 Or 1% means your experiment is **99%** correct only **1%** is wrong. It used in **Lab experiments**, and it is **highly significant** or very accurate, because the surrounding environment is under your control & mistake is unknown means there is a little deviation.



2) 0.05 Or 5% means your experiment is **95%** correct but **5%** is wrong which considered inaccurate when compared with 0.01, it is used when the experiment is done in the field the surrounding environment is not under your control, in 0.05 the mistake is known by neglecting of your result sometimes is unknown because of conditions of the outside environment.



4 Testing Hypothesis:

Any **study** about the parameters is called **Hypothesis**. Before doing anything, we have to set a hypothesis for our study. These hypotheses or probabilities will be done **before doing the experiment**. Then **after** achieving the experiment, the result will be the judge that decides which one of the **hypotheses** is **true** and which one is not.

There are two types of hypothesis:

1) Null hypothesis: **(H_0)**
It is denoted by **H_0**. It states that there is **no statistical significance** between the two variables. Here;

Ho; m1 = m2 = m3 = …. m_n.

E.g. we have to assume that the effect of the recently invented pesticide is **equal** to the effect of the conventional pesticide found before.

Thus; m1 (mean of the new pesticide) = m2 (mean of the conventional pesticide)

2) Alternative hypothesis: (H_i, or H_A)
It is denoted by (Hi, or HA). It states that there is statistical significance between two variables. It means we reject Ho and accept the alternative hypothesis.

m1 ≠ m2 ≠ m3 ≠ …. m_n.

E.g. Also there may be another person who assumes that the effect of the recently invented pesticide is not equal to the effect of the other pesticide.

Thus; m1 (mean of the new pesticide) ≠ m2 (mean of the conventional pesticide)

5 Principles of experimental design

There are three basic principles of design which were developed by Sir Ronald A. Fisher:

5.1 Randomization

The principle of randomization involves the **allocation** of **treatment** to **experimental units** at **random** to **avoid** any **bias** in the experiment resulting from the influence of some extraneous unknown factor that may affect the experiment.
The **random assignment** of experimental units to treatments results in the following outcomes:
a) It eliminates the systematic bias.
b) It is needed to obtain a **representative sample** from the population.
c) It helps in **distributing** the **unknown variation** due to confounded variables throughout the experiment and breaks the confounding influence.

If the **randomization process** is such that every experimental unit has an **equal chance** of receiving each treatment, it is called a **complete randomization.**

5.2 Replication:

In the replication principle, any treatment is **repeated** a number of times to obtain a **valid** and more reliable **estimate** than which is possible with one observation only. Replication **provides** an **efficient way** of increasing the **precision** of an experiment. The precision increases with the increase in the number of observations. Replication provides more observations when the same treatment is used, so it increases precision.

$$SE = \sqrt{\frac{s^2}{n}} = \frac{s}{\sqrt{n}}$$

5.3 Local control (Blocking)

Local control or blocking is a technique to account for the variability in response because of the variability in the experimental units. To block an experiment is to divide the experimental units into groups or blocks of similar units in such a way that the observations in each block are collected under relatively similar experimental conditions.

6 Independent and Dependent variables:

When an experiment is conducted, some variables are manipulated by the experimenter "independent” and others are measured from the subjects “dependent”.

- **Independent variable** is the **cause**. Its value is *independent* of other variables in your study.
- **Dependent variable** is the **effect**. Its value *depends* on changes in the independent variable.

Research Question	Independent variable(s)	Dependent variable(s)
Do tomatoes grow fastest under fluorescent, incandescent, or natural light?	The type of light the tomato plant is grown under	The rate of growth of the tomato plant
How well do different plant species tolerate salt water?	The amount of salt added to the plants' water	• Plant growth • Plant wilting • Plant survival rate