DESIGN OF EXPERIMENT

Experiment

A device or a means of getting an answer to the problem under consideration. Experiments Experiment can be classified into two.

- Absolute:- designed to calculate certain measures of relationships. Eg. Correlation, averages etc.

- Comparative:- designed to compare the effect of two more characteristics. Eg. Comparison of effect of two or medicines, fertilizers etc.

Design of Experiment

The logical construction of the experiment in which the degree of uncertainty with which the inference is drawn may be well defined.

The components of experimental design are

- Planning of Experiment
- Obtaining the relevant information from it regarding the Statistical Hypothesis under study
- Making Statistical Analysis of Data

Important Terms and Definitions

Treatments: Various object of comparison, in a comparative experiment, are termed as treatments.

Eg: In agricultural experiment, different fertilizers, different verities of crops or different method of cultivation are treatments

Experimental Unit: Smallest division of the experimental material to which we apply the treatments and on which we makes observation on the variable under study the treatments and on which we makes observations on the variable under study is termed as experimental unit.

Eg: In the agricultural experiments the *plot* or *land* is the experimental unit. In medical experiments the unit may be a patient in a hospital, a group of pig etc.

Blocks: In agricultural experiments, most of the times we divide the whole experimental unit (field) into relatively homogeneous sub groups or strata These strata, which are more uniform amongst themselves than the field as a whole are known as blocks.

Yield: The measurement of the variable under study on different experimental units are termed as yield.

Experimental Error: Extraneous or random (or chance or non assignable error) variation in the yield due to

- the inherent variability in the experimental material to which treatment are applied
- the lack of uniformity in the methodology of conducting the experiment or in other words failure to standardize the experiment technique and
- the lack of representative ness of the sample to the population under study

are called experimental error.

Principle of Experimental Design

The basic principle of the design of experiments are

- Replication
- Randomisations and
- Local Control

Replication

Replication means the execution of an experiment more than once. In other words the replication of the treatments under investigation is known as replication. By replication the experimenter tries to average out as far as possible the effect of uncontrolled factors.

Randomisation

A process of assigning the treatments to the various experimental unit in a purely chance manner. The purpose of randomization is to assure that the source of variation, not controlled in the experiment, operate randomly so that the average effect of any group of unit is zero.

Local Control

The process of reducing the experimental error by dividing the relatively heterogeneous experimental area (field) homogenous blocks is known as Local control.

Types of Experimental Designs

1. Completely Randomised Design (CRD)
2. Randomised Block Design (RBD)
3. Latin Square Design (LSD)

Completely Randomized Design

The CRD is the simplest of all the designs based on principle of randomization and replication. In this design treatments are allotted at random to the experimental units over the entire experimental units over the entire experimental materials. Let us suppose that we have k treatments, the i^{th} treatment being replicated r_i times $i = 1, 2, \ldots, k$. Then the whole experimental material is divided into $n = \sum r_i$ experimental units and the treatments are distributed completely at random over the units subject to the condition that the i^{th} treatment occurs r_i times.

ANOVA Table

Source of variation	d.f	Sum of Squares (SS)	Mean Sum of Squares (MSS)	F ratio
Treatment	k-1	S_t^2	$MS_t^2 = S_t^2/(k-1)$	$F_t = \frac{MS_t^2}{MS_E^2}$
Error	n-k	S_E^2	$MS_E^2 = S_E^2/(n-k)$	
Total	n-1	$S_T^2 = S_t^2 + S_E^2$		

Advantages

1. CRD results in the maximum use of the experimental units since all the experimental material can be used.
2. The design is very flexible. In this case any number of treatment can be used and different treatment can be used unequal number of times.
3. The statistical analysis remains simple even if some or all observations for any treatment are rejected or lost or missing for some purely random accidental reasons.
4. It provides the maximum number of degree of freedom for the estimation of the error variance, which increases the sensitivity or the precision of the experiments when the number of treatment are small.

Disadvantages

In certain case (i.e. when the experimental materials are not homogenous) the design suffers from the disadvantage of being inherently less informative than other more sophisticated designs.

Applications

1. CRD is most useful in laboratory technique and methodological studies.
2. CRD is also recommended in situations where an appreciable fraction of units is likely to be destroyed or fail to respond.

Randomized Block Design

In Randomized block design we divide the experimental area into smaller homogenous blocks and the treatment are applied to this block at random and is is replicated over all the blocks.

1. Since all the treatments are to be applied within each block, in each block we take as many units (or plots) as the number of treatments. With this design each treatment will have the same number of replications. If we want additional replications for some treatments each of them must be applied to more than one unit in a block.
2. It is assumed that though the general level of results is different in the different blocks, the relative effects of treatments are same in all the blocks apart from experimental error. In other words, there is no interaction between treatments and blocks. In practice we interpret this as meaning that the interaction, if any exist, are not appreciable compared with the treatment effects. Interactions will be separable from experimental error and if the interactions are large, the experiment may yield misleading results.
3. R.B.D. has become extremely popular in a large number of fields. It is flexible, readily adaptable and easy to analyze and these points have made it the most popular of all designs with Latin Square Design (L.S.D) being its closest rival.
4. R.B.D. provides a method of eliminating or reducing the effects of trends. The design is not limited to experiments in fields or experiments in industrial plants alone but applies whenever the trials are spread over a period of time or space and the possibility of systematic variation or trends exists.

Analysis of variance

Let us suppose that we have k treatments, and they are applied to r blocks. Then the whole experimental material is replicated r times (i.e. treatment occurs r times).

ANOVA Table

Source of variation	d.f	Sum of Squares (SS)	Mean Sum of Squares (MSS)	F ratio

Treatment	k-1	S_t^2	$MS_t^2 = S_t^2/(k-1)$	$F_t = \frac{MS_t^2}{MS_E^2}$
Blocks or replication	r-1	S_b^2	$MS_b^2 = S_b^2/(r-1)$	$F_b = \frac{MS_b^2}{MS_E^2}$
Error	(r-1)(k-1)	S_E^2	$MS_E^2 = S_E^2/(r-1)(k-1)$	
Total	nk-1	$S_T^2 = S_t^2 + S_b^2 + S_E^2$		

Lay out of R.B.D.

In agricultural experimentation, the lay out of R.B.D. can be illustrated as follows.

Let us consider five treatments A, B, C, D, E each replicated four times. We divide the whole experimental area into four relatively homogeneous strata or blocks and each block into five units or plots. Treatments are then allocated at random to the plots of a block, fresh randomization being done for each block. A particular layout may be as follows.

Block I	A	E	B	D	C
Block II	E	D	C	B	A
Block III	C	B	A	E	D
Block IV	A	D	E	C	B

Advantages of R.B.D.

1. **Accuracy**. This design has been shown to be more efficient or accurate than C.R.D for most types of experimental work. The elimination of between S.S, usually results in a decrease of error mean S.S.
2. **Flexibility**. In R.B.D. no restrictions are placed on the number of treatments or the number of replicates. In general, at least two replicates are required to carry out the test of significance (Factorial design is an exception). In addition, control (check) or some other treatments may be included more than once without complications in the analysis.
3. **Ease of analysis**. Statistical analysis is simple and rapid. Moreover the error of any treatment can be isolated and any number of treatments may be omitted from the analysis without complicating it.

Disadvantages of R.B.D.

It is not suitable for large number of treatments or for cases in which complete block contains considerable variability.

Latin Square Design

In R.B.D. whole of the experimental area is divided into relatively homogeneous groups (blocks) and treatments are allocated at random to units within each block, i.e., randomization was subjected to one restriction, i.e., within blocks. But in field experimentation, it may happen that experimental area (field) exhibits fertility in strips, eg, cultivation might result in alternative strips of high or low fertility. R.B.D. will be effective if the blocks happen to be parallel to these strips and would be extremely inefficient if the blocks are across the strips. Initially fertility gradient is seldom known. A useful method of eliminating fertility variations consists in an experimental layout which will control variation in two perpendicular directions. Such a layout is a *Latin Square Design* (L.S.D.)

Lay out of design. In this design the number of treatments is equal to the number of replications. Thus in case of m treatments, there have to be $m \times m = m^2$ experimental units. The whole of experimental area is divided into m^2 experimental units (plots) arranged in a square so that each row as well as each column contains m units. The m treatments are then allocated at random to these rows and columns in such a way that every treatment occurs once and only once in each row and in each column. Such a layout is known as $m \times m$ *Latin Square Design* (L.S.D.) and is extensively used in agricultural experiments, eg, if we are interested in studying the effects of m types of fertilizers on the yield of a certain variety of wheat;, it is customary to conduct the experiments on a square field, with m^2 plots of equal area and to associate treatments with different fertilizers and row and column effects with variations in fertility of soil.

Obviously there can be many arrangements for an $m \times m$ L.S.D. and a particular layout in an experiment must be determined randomly, eg, with four treatments A, B, C and D, one typical arrangement of 4 × 4 L.S.D. is given below.

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

Fischer and Yates have tabulated Latin Squares up to 12 × 12.

Advantages of Latin Square Design

1. With two way grouping or stratification L.S.D. controls more of the variation than C.R.D. or R.B.D. The two way elimination of result of cross grouping often results in small error mean sum of squares. Thus in field experimentation if the fertility gradient is in two directions at right angles to each other (i.e., if there is a diagonal trend in fertility) or in one unknown direction then L.S.D. is likely to be more efficient than R.B.D. In fact L.S.D. can be used with advantage in those cases where the variation in experimental material is from two orthogonal sources.
2. L.S.D. is an incomplete 3-way layout. Its advantage over the complete 3-way layout is that instead of m^3 experimental units only m^2 units are needed. Thus a 4 ×4 L.S.D. results in saving of 64-16=48 observations over a complete 3-way layout.
3. The statistical analysis is simple though slightly complicated than fro R.B.D. Even with missing data the analysis remains relatively simple.
4. More than one factor can be investigated simultaneously and with fewer trials than more complicated designs.

Disadvantages of L.S.D.

1. The fundamental assumption that there is no interaction between different factors (i.e., the factors act independently) may not be true in general.
2. Unlike R.B.D. in L.S.D. the number of treatments is restricted to the number of replications and this limits its field of application. L.S.D. is suitable for the number of treatments between 5 and 10 and for more than 10 to 12 treatments the design is seldom used since in that case the square becomes too large and does not remain homogeneous.
3. In case of missing plots, when several units are missing the statistical analysis becomes quite complex. If one or two blocks in a field are attacked by some disease

or pest then in R.B.D. we can easily omit the data for these blocks without complicating the analysis at all whereas a much more complicated analysis is necessitated in L.S experiment under similar conditions.

4. In the field layout, R.B.D. is much easy to manage than L.S.D. since the former can be performed equally well on a square or rectangular field or a field of any shape whereas for the latter approximately a square field is necessary.

Latin Square Design

ANOVA Table

Source of variation	d.f	Sum of Squares (SS)	Mean Sum of Squares (MSS)	F ratio
Treatment	k-1	S_t^2	$MS_t^2 = S_t^2/(k-1)$	$F_t = \frac{MS_t^2}{MS_E^2}$
Rows	k-1	S_R^2	$MS_b^2 = S_R^2/(k-1)$	$F_R = \frac{MS_R^2}{MS_E^2}$
Column	k-1	S_C^2	$MS_b^2 = S_C^2/(k-1)$	$F_C = \frac{MS_C^2}{MS_E^2}$
Error	(k-1)(k-2)	S_E^2	$MS_E^2 = S_E^2/(k-1)(k-2)$	
Total	$k^2 - 1$	$S_T^2 = S_t^2 + S_R^2 + S_C^2 + S_E^2$		