DESIGN OF EXPERIMENTS AND BASIC DESIGN

Introduction

The design of an experiment is just like construction of a house. Based on the basic requirements, cost, etc., different house plans may be drawn up. We may choose one which has potential advantages compared to others; similarly, an experiment is designed based on the objectives, availability of experimental material, cost of the experiment, etc. In order to test hypothesis with acceptable degree of precision, we should have carefully designed experiments.

Modern concepts of experimental design are due primarily to R.A. Fisher. He developed them in the planning of agricultural field experiments.

Basic Concepts

The choice of treatments, the method of assigning treatments to experimental units and arrangement of experimental units in various patterns to suit the requirement of particular problems, are combined known as the design of experiment.

Through experimentation, we wish to study the effect of changes in one variable (such as application of fertilizer) on another variable (such as grain yield of a crop). The variable whose change we wish to study may be termed as a dependent or response variable. The variable whose effect on the response variable we wish to study may be termed as an independent variable or a factor. Thus fertilizer, spacing irrigation schedule, pesticide etc., are known as factors. The crop yield, mortality of pests, etc. are known as responses.

Treatments and Level

Any specific experimental conditions applied to the experimental units are termed as treatments. A treatment is usually a combination of specific values, called levels of each of the experimental factors.

Experimental Unit:

The basic objects on which the experiment is carried out are known as experimental units. The experimental unit may be a plot of land, a plant, an insect, an animal etc.

Experimental Error

A major problem in experimentation is that the responses of the experimental units are influenced not only by treatments but also by the extraneous factors. The variation in responses caused by the extraneous factors is termed as experimental error.

Basic Principles of Experimental Design

The basic principles of the experimental design are replication, randomization and local control.

Replication

The repeated application of the treatments under investigation is known as replication. The major functions of the replication are,

1) to provide an estimate of experimental errors and
2) to reduce the experimental error

Randomization

When all the treatments have equal chances of being allocated to different experimental units it is known as randomization. Since for valid conclusions about experimental results, we should have not merely an estimate of experimental error but it should be an unbiased estimate. For this purpose we use the technique of randomization.

Local Control (or) Error Control

We know that the estimate of experimental error is based on the variation from experimental unit to experimental unit. In other words, the error in an experiment is a measure of 'within block' variation. This suggests that if we group the homogeneous experimental units in to blocks, the experimental error will be reduced considerably. Grouping of homogeneous experimental units into blocks is known as local control or error control.

Analysis of Variance (ANOVA)

The analysis of variance is the systematic algebraic procedure of decomposing the overall variation in the responses observed in an experiment in to different components. Each component is attributed to an identifiable cause or source of variation.

In other words, the ANOVA is a technique of partitioning the total variation of the response into between group variation (such as treatment variance) and within group variation (such as error variance). The structure of these component parts is determined by the design of experiments.

Completely Randomized Design

When the treatments are arranged randomly over the whole of a previously determined set of experimental units, the design is known as completely randomized design (CRD).

Layout of CRD

Suppose that there are t treatments and r replication, the layout of CRD is given as follows.

Replication	Treatment				Total
	1	2	...	t	
1	Y_{11}	Y_{12}	...	Y_{1t}	
2	Y_{21}	Y_{22}	...	Y_{2t}	
⋮	⋮	⋮	...	⋮	
r	Y_{r1}	Y_{r2}	...	Y_{rt}	
Total	T_1	T_2	...	T_t	G.T

Here we have equal number of replications, in some of the cases we may have unequal number of replication.

Analysis and Interpretation of Results

The ANOVA model for CRD is given by

$$Y_{ij} = \mu + t_i + e_{ij}; \quad i = 1,2,\cdots t, \quad j = 1,2,\cdots r,$$

Where Y_{ij} – Response from the j^{th} plot receiving i^{th} treatment.

μ – overall mean effect

t_i – effect due to i^{th} treatment

e_{ij} – error term iid $N(0, \sigma^2)$

The null hypothesis to be tested here is

H_0: $t_1 = t_2 = \ldots = t_t$ (the treatment effect is equal)

Vs

H_1: $t_1 \neq t_2 = \ldots = t_t$ (atleast one of the treatment effect is not equal)

The various sum of squares to be calculated are as follows.

Correction Factor (CF) = $\frac{(GT)^2}{rxt}$ G.T: Grand Total

Total Sum of Squares (TSS) = $\Sigma\Sigma Y_{ij}^2 - CF$

Treatment Sum of Squares (TrSS) = $\frac{1}{r}\Sigma T_i^2 - CF$

T_i = i^{th} treatment total

Error Sum of Squares (ESS) = TSS – Tr SS

The calculated sums of squares are summarized in the analysis of variance (ANOVA) Table.

ANOVA Table

Sources of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F – ratio	F – Table Value
Treatment	t – 1	TrSS	$T_rMS = \frac{TrSS}{t-1}$ – (a)	(a) / (b)	$F_\alpha(t-1, t(r-1))$
Error	t (r – 1)	ESS	$EMS = \frac{ESS}{t(r-1)}$ – (b)		
Total	rt – 1	TSS			

Inference:

If the calculated F – ratio corresponds to treatment is greater than the Table F value of specified level of significance and degrees of Freedom, we reject the null hypothesis H_0. Otherwise we do not reject the null hypothesis H_0.

Advantages and Disadvantages of CRD

Advantages:

1) The statistical analysis of data from CRD is very simple.

2) CRD provides the maximum number of degrees of freedom for the estimation of experimental error.

3) The number of replication may be unequal for treatment to treatment. Because of this, all the available experimental material can be utilized without any wastage.

Disadvantages:

1) CRD, it is less accurate than other designs.

2) The importance given to the treatments, not for the replication.

3) If we increase the number of treatments, the heterogeneity of experimental materials will be increased. This will result in increased experimental error and reduced precision.

Randomized Complete Block Design (RBD)

Suppose the experimental material is divided into r blocks. Let there be t treatments. Each block is then divided into t units and the treatments are allocated within a block at random. The resulting design called randomized complete block design.

Layout and Analysis of RBD

The results from RBD can be arranged in two–way tables according to the replications and treatments. There will be rt observation in total. The layout of RBD is as follows.

Treatment	Replication				Total
	1	2	...	r	
1	Y_{11}	Y_{12}	...	Y_{1t}	T_1
2	Y_{21}	Y_{22}	...	Y_{2t}	T_2
⋮	⋮	⋮	...	⋮	⋮
t	Y_{t1}	Y_{t2}	...	Y_{tr}	T_t
Total	B_1	B_2	...	B_r	G.T

The statistical linear model for RBD is given by

$$Y_{ij} = \mu + t_i + r_j + e_{ij}; \quad i = 1,2,\cdots t, \quad j = 1,2,\cdots r,$$

Where Y_{ij} – Response from the j^{th} plot receiving i^{th} treatment.

μ – overall mean effect
t_i – effect due to i^{th} treatment
β_j – effect due to j^{th} replicate

e_{ij} – error term iid $N(0, \sigma^2)$

The null hypothesis to be tested are

H_{01}: The treatment effects are equal

and

H_{02}: The replication effects are equal.

The sum of squares for different source of variation is calculated as follows.

Correction Factor (CF) = $\frac{(GT)^2}{r \times t}$ where GT= Grand Total

Total Sum of Squares (TSS) = $\Sigma\Sigma Y_{ij}^2 - CF$

Treatment Sum of Squares (TrSS) = $\frac{1}{r}\Sigma T_i^2 - CF$ $\quad T_i = i^{th}$ treatment total

Block Sum of Squares (BSS) = $\frac{1}{r}\Sigma B_j^2 - CF$ $\quad B_j = j^{th}$ block total

Error Sum of Squares (ESS) = TSS – TrSS – BSS

The calculated sums of squares are summarized in the analysis of variance (ANOVA) Table.

ANOVA Table

Sources of Variation	**Degrees of Freedom**	**Sum of Squares**	**Mean Sum of Squares**	**F – ratio**	**F – Table Value**
Treatment	t – 1	TrSS	$T_rMS = \frac{TrSS}{t-1}$ – (a)	(a) /(b)	$F_\alpha(t-1,(r-1)(t-1))$
Block	r – 1	BSS	$BMS = \frac{BSS}{r-1}$ – (b)	(b) /(c)	$F_\alpha((r-1),(t-1)(r-1$
Error	$\frac{t(r-1)}{(t-1)(r-1)}$	ESS	$EMS = \frac{ESS}{(r-1)(t-1}$ – (c)		
Total	rt – 1	TSS			

Inference:

Due to treatments, if the calculated F – ratio is greater than the table F – value of specified level of significance and degrees of freedom, we reject H_{01} otherwise we do not reject H_{01}.

Similarly the decision will be made for the blocks.

Advantages:

1) It increases the precision of the experiment.
2) The amount of information got in RBD is more as compared in CRD.
3) The statistical analysis simple and easy.

Disadvantages:

1) The number of treatments is increased, the block size will increase. Due to this, it may be difficult to maintain homogeneity in within blocks.

2) The RBD may not be suitable for large number of treatments.

LATIN SQUARE DESIGN (LSD)

When the experimental material is divided into rows and the treatments are allocated such that each treatment occurs only once in a row and once is a column, the design in known as Latin Square Design (LSD). In LSD the number of rows and columns are equal.

Example: The following in the layout of 3×3 LSD with three treatments.

A	B	C
B	C	A
C	A	B

Model and Analysis of LSD

The linear model for LSD is given by

$$Y_{ijk} = \mu + \alpha_i + \beta_j; + \tau_k + e_{ijk} \quad i = 1,2,\cdots m, \quad j = 1,2,\cdots m, \quad k = 1,2,\cdots m,$$

Where, α_i – i^{th} row effect

β_j – j^{th} column effect

τ_k – k^{th} treatment effect

Other notations have usual meaning.

Let R_i, C_j and T_k be the row total, column total and treatment total respectively.

Correction Factor (CF) = $\frac{(GT)^2}{r \times t}$ G.T: Grand Total

Total Sum of Squares (TSS) = $\Sigma\Sigma Y_{ijk}^2 - CF$

Row Sum of Squares (RSS) = $\frac{1}{r}\Sigma R_i^2 - CF$

Column Sum of Squares (CSS) = $\frac{1}{r}\Sigma C_j^2 - CF$

Treatment Sum of Squares (TrSS) = $\frac{1}{r}\Sigma T_k^2 - CF$

Error Sum of Squares (ESS) = TSS – RSS – CSS - TrSS

The computed sums of squares are summarized in the ANOVA Table.

H_{01} : The row effects are equal

H_{02} : The column effects are equal

H_{03} : The treatment effects are equal

ANOVA Table

Sources of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F – ratio	F – Table Value
Row	m – 1	RSS	RMS = $\frac{RSS}{m-1}$	RMS / EMS	F(m - 1,(m - 1)(m - 2) d.f
Column	m – 1	TrSS	CMS = $\frac{CSS}{m-1}$	CMS / EMS	,,

Treatment	m– 1	TrSS	$T_rMS = \frac{TrSS}{m-1}$	TrMS / EMS	,,
Error	(m–1)(m-2)	ESS	$EMS = \frac{ESS}{(-)}$ –		
Total	$m^2 - 1$	TSS			

Inference:

Due to Rows

If cal F ≤ table F – value of specified level of significance, we do not subject H_{01}, otherwise we reject H_{01}.

Similarly the decision to be taken for column and treatment.

****