UNIT-IV

Design of experiments

- **Prepared by Dr.N.PADMAPRIYA**

In statistics, experiment means collection of data for a scientific investigation, according to certain specified sampling procedures.

The **design of experiment** may be defined as the logical construction of the experiment in which the degree of uncertainty with which the reference is drawn may be well defined i.e., planning an experiment to get a deliberate or purposeful outcome.

Experiments can be classified into 2 categories: Absolute and Comparative

Absolute experiments consist in determining the absolute value of some characteristic like

(1) Obtaining the average intelligence quotient (IQ) of a group of people.
(2) Finding the correlation coefficient between two variables in a bivariate distribution etc.

Comparative experiments are designed to compare the effect of two or more objects on some population characteristic.

Example: Comparison of different fertilizers on yield.

Treatments – Various objects of comparison in a comparative experiment are termed as treatments.

Example: different varieties of crop or different methods of cultivation are the treatments

Experimental Unit The smallest division of the experimental material to which we apply the treatments and on which we make observations on the variable under study, is termed as experimental unit. Example: Patient in a hospital, a batch of seeds.

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.

Principle of experimental design

There are three basic principles in the design of experiments

(1) Replication (2) Randomization (3) Local Control

Replication

Performing the same experiments more than once under same condition is known as replication. The advantage is that it will average out the chance factor.

Randomization

If the treatments are allotted to the experimental unit purely on chance basis, then it is known as randomization. Randomization is needed in order to remove bias and for making valid estimates of standard errors.

Local control

The chances of reducing experimental error by dividing the whole experimental unit into a relatively homogeneous group is known as local control

Some basic designs of experiment

- Completely Randomized Design(CRD)
- Randomized Block Design (RBD)
- Latin Square Design (LSD)

We use the method of ANOVA for studying the relationship between factors.

Analysis of Variance (ANOVA)

ANOVA is a technique that will enable us to test for the significance of the difference among more than two sample means. That is to test the homogeneity of several means. t-test is used for testing the hypothesis that two population means are equal . ANOVA is used for testing the significance of more than two population means and therefore considered as an extension of t- test.

ANOVA may be classified into 2 categories as

- One- way Classification
- Two- way Classification

One Way Classification (Completely Randomized Design) (CRD)

The completely randomized design is simplest of all designs and is based on the principle of replication and randomization. In this design, treatments are allocated randomly to the entire experimental unit with replication.

Advantage

- It is easy to layout the design
- It allows complete flexibility. Any number of factor classes and replications may be used.
- It is usually suited only for small number of treatments
- It provides the maximum number of degrees of freedom for the estimation of the error variance, which increases the sensitivity or the precision of the experiment for small experiments, i.e., for experiments with small number of treatments.

ANOVA TABLE

Source of Variation	Sum of squares	Degree of freedom	Mean of square	F-ratio
Between samples(Treatments)	SSC	$v_1 = k - 1$	MSC = $\frac{SSC}{k-1}$	$F_c = \frac{MSC}{MSE}$
Within Samples (Treatments)	SSE	$v_1 = N - k$	MSE = $\frac{SSE}{N-k}$	
Total	TSS	N-1		

Procedure

1. Find the total number of observations N
2. Find the total value of all observations $T = \sum X_1 + \sum X_2 + \sum X_3 + \cdots + \sum X_k$
3. Find the correction factor (C.F) $= \frac{T^2}{N}$
4. Calculate the total sum of squares TSS = Sum of squares all the items – C.F
$$TSS = \sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \cdots + \sum {X_k}^2 - \frac{T^2}{N}$$
5. Calculate the sum of squares between columns
$$SSC = \frac{(\sum X_1)^2}{n} + \frac{(\sum X_2)^2}{n} + \frac{(\sum X_3)^2}{n} + \cdots + \frac{(\sum X_k)^2}{n} - C.F$$
6. Find SSE, sum of squares with samples (Errors)
SSE = TSS- SSC
7. Find MSC, mean square between samples
MSC $= \frac{SSC}{df}, df = k - 1$
8. Find MSE, mean square within samples
MSE $= \frac{SSE}{df}, df = N - k$
9. Prepare the ANOVA table to calculate F-ratio $F_c = \frac{MSC}{MSE}$

Two way classification (Randomized Block Design)

ANOVA Table

Source of Variation	Sum of squares	Degree of freedom	Mean of square	F-ratio
Between columns(Treatments)	SSC	$v_1 = c - 1$	MSC $= \frac{SSC}{c-1}$	$F_c = \frac{MSC}{MSE}$
Between rows (Blocks)	SSR	$v_2 = r - 1$	MSR $= \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
Error or Residual	SSE	$v = (c-1)(r-1)$	MSE $= \frac{SSE}{(c-1)(r-1)}$	
Total	TSS			

Ho: There is no significant difference between treatments and between blocks.

1. Find the total number of observations N
2. Find the total value of all observations $T = \sum X_1 + \sum X_2 + \sum X_3 + \cdots + \sum X_k$
3. Find the correction factor (C.F) $= \frac{T^2}{N}$
4. Calculate the total sum of squares TSS = Sum of squares all the items – C.F
$$TSS = \sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \cdots + \sum {X_k}^2 - \frac{T^2}{N}$$
5. Calculate the sum of squares between columns
$SSC = \frac{(\sum X_1)^2}{n} + \frac{(\sum X_2)^2}{n} + \frac{(\sum X_3)^2}{n} + \cdots + \frac{(\sum X_k)^2}{n} - C.F$
(n- number of elements in each column)
6. Calculate the sum of squares between rows
$SSR = \frac{(\sum Y_1)^2}{n} + \frac{(\sum Y_2)^2}{n} + \frac{(\sum Y_3)^2}{n} + \cdots + \frac{(\sum Y_k)^2}{n} - C.F$
(n- number of elements in each row)

7. Find SSE, sum of squares within samples (Errors)
   SSE = TSS- (SSC+SSR)
8. Find MSC, mean square of columns
   MSC =$\frac{SSC}{df}, df = c - 1$
9. Find MSR, mean square of columns
   MSR =$\frac{SSR}{df}, df = r - 1$
10. Find MSE, mean square within samples
   MSE = $\frac{SSE}{df}, df = N - c - r + 1\ (or)(r-1)(c-1)$
11. Find $F_c = \frac{MSC}{MSE}$ and $F_R = \frac{MSR}{MSE}$, Prepare the ANOVA table

<u>Advantages of RBD</u>

The design of RBD is more efficient than CRD

In RBD, no restrictions are placed on number of treatments or replicants

Statistical analysis is simple and rapid.

Disadvantage

It is not suitable for large number of treatments