Design of Experiment

In general, Statistics deals with the methods of collection, tabulation, analysis and interpretation of data. There are some underlying assumptions to every analysis and it is for the investigator to see that the experiment is performed in a manner so that these assumptions are satisfied. The complete sequence of steps taken to ensure an objective analysis leading to valid references is called "The Design of Experiment".

The purpose of an experimental design is to obtain maximum information with minimum cost and labour.

Basic Principles of Experimental Design —

The three basic principles of experimental design are (i) Replication

(ii) Randomization

(iii) Local Control

REPLICATION: —

By replication we mean the repetition of an experiment. It help us in estimating the experimental error. We can get a more precise effect of the mean effect of any factor since $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ where $\sigma_{\bar{x}}$ is the S.D. of the means, σ^2 the true experimental error and n is number of replications.

the greater the value of n smaller the value of $\sigma_{\bar{x}}$.

RANDOMIZATION:-

However any test has some underlying principles which must be satisfied for the test to be valid. The commonest principle is that the observations i.e. errors therein are independently distributed. This can be achieved only by taking samples in a random manner or assign treatment to the experimental units in a radom manner. Thus randomization allows us to assume that the errors are independent. It may be noted that randomization does not guarantee independence but allow us to assume that independence is a fact. Although the effect of adjasence of time and space can not be estimated but randomization tends to make these effect minimum.

LOCAL CONTROL:- The local control may be defined as the balancing, blocking and grouping of experimental units employed in experimental design. While replication and randomization make a valid test of significance possible, local control makes the test more sensitive.

The Latin Square Design

A Latin square is a square arrangment of n rows and n-columns such that each symbol appears once and only once in each row and column and it is known as Latin square of order n. e.g.

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

(i)

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

(ii)

are latin square of order 5 each.

If the first row and first column has the alphabets in alphabetical order, then it is called a Standard or Reduced or Normalized Latin square design. e.g. (ii) is a standard Latin square arrangment

The number of all possible latin square arrangments are 12 of order three, 576 of order four, 161280 of order five and so on and the number of standard latin square arrangments are 1 of order three, 4 of order four and 56 of order five. In general,

The number of all possible latin square arrangment of order n $= n!(n-1)! \times$ No. of standard latin square of order n

In Latin square design, there are two restrictions (i) the number of rows and columns are equal and (ii) each treatment occurs once and only once in each row and each column.

The Latin square design is three way classification model for analysis of variance. The latin square designs are frequently used in industrial experimentation.

The mathematical model for a Latin square is

$$x_{ij} = \mu + \alpha_i + \beta_j + V_k + e_{ijk}$$

where α_i are the main effects from the first factor (say rows), β_j from second factor (say columns), V_k from third (say varieties) and e_{ijk} is the error term due to unspecified causes. Then e_{ijk} are assumed to be normal with mean 0 and variance σ^2. The mean values of α_i, β_j and V_k are adjusted to be zero.

In this case the total sum of squares S_T may be divided into four parts

$$S_T = S_a + S_b + S_c + S_d$$

where in $n \times n$ table

$$S_T = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_{ij} - \bar{x})^2 = \sum\sum x_{ij}^2 - G$$

$$S_a = n \sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2 = \sum_{i=1}^{n} \frac{S_i^2}{n} - G$$

$$S_b = n\sum_{j=1}^{n}(\bar{x}_j - \bar{x})^2 = \sum_{j=1}^{n}\frac{S_j^2}{n} - G$$

$$S_c = n\sum_{k=1}^{m}(\bar{x}_k - \bar{x})^2 = \sum_{k=1}^{n}\frac{S_k^2}{n} - G$$

$$S_d = S_T - S_a - S_b - S_c$$

with G = sum of squares due to mean:-

$$= \frac{S^2}{n^2} = \frac{1}{n^2}\left(\sum\sum x_{ij}\right)^2$$

S_i is the sum of ith row, S_j the sum of jth column and S_k sum of kth variety.

ANOVA TABLE

Source of variation	Sum of Squares	Degrees of freedom	Mean Squares	F ratios
Between rows	S_a	$n-1$	$MS_a = \frac{S_a}{n-1}$	$F = \frac{MS_a}{MS_d}$
Between column	S_b	$n-1$	$MS_b = \frac{S_b}{n-1}$	$F = \frac{MS_b}{MS_d}$
Between varieties	S_c	$n-1$	$MS_c = \frac{S_c}{n-1}$	$F = \frac{MS_c}{MS_d}$
Error	S_d	$(n-1)(n-2)$	$MS_d = \frac{S_d}{(n-1)(n-2)}$	
Total	S_T	n^2-1		