UNIT 9 COMPLETELY RANDOMISED DESIGN

Structure

9.1 INTRODUCTION

The modern concepts of experimental designs were primarily given by Ronald A. Fisher in the 1920s and 1930s at "Rothamasted Experimental Station", an agricultural research station of London. In Fisher's first book on design of experiments, he showed how valid conclusions could be drawn efficiently from experiments with natural fluctuation such as temperature, soil conditions and rainfall, that is, in the presence of nuisance variables. The known nuisance variables usually cause systematic biases in groups of results (e.g. batch-to-batch variables). The unknown nuisance variables usually cause random variability in the results and are called inherent variability or noise.

The experimental design was first used in an agricultural context, the method has been applied successfully in the military and in industry since the 1940s. Besse Day, working at U. S. Naval Experimentation Laboratory, used experimental designs to solve problems such as finding the cause of bad welds at the naval shipyards during World War II. George Box, employed by Imperial Chemical Industries before coming to the United States, is a leading developer of experimental design produced for optimizing chemical process. W. Edwards Deming taught statistical methods, including experimental designs, to Japanese scientist and engineers in the early 1950's at a time when "Made in Japan" meant poor quality. Genichi Taguchi, the most well known of this group of Japanese scientists is famous for his quality improvement methods. One of the companies where Taguchi first applied his methods was Toyota. Since the late 1970's, U.S. industry has become interested again in

quality improvement initiatives, now known as "Total Quality" & "Six-sigma" programs. Design of experiments is considered an advanced method in the six sigma programs, which were pioneered at Motorola & GE.

According to Bernad Ostle, "The design of experiment is, the complete sequence of steps taken ahead of time to ensure that the appropriate data will be obtained in a way which permits an objective analysis to valid inferences with respect to stated problem".

In any field of study either in life sciences or some other, it is essential to plan an experiment, i.e. what is the object and which type of data is required. In order to make use of time and energy spent on experiment, it should be planned with a careful designing. Once a design of experiment is decided, the observations are obtained from it and with the technique of analysis of variance, the data is analysed.

Basic definitions of experimental design are described in Section 9.2. The principles of design of experiments i.e. randomisation, replication and local control are explained in Section 9.3 whereas size and shape of plots are described in Section 9.4. In Section 9.5, layout and statistical analysis of completely randomised design are explained. The least square estimates of effects, variance of the estimates and expectation of sum of squares are also given in Section 9.5. Suitability of CRD is descrived in Section 9.6.

Objectives

After studying this unit, you would be able to

- describe the experimental design;
- explain the planning and classification of experimental designs;
- describe the principles of design of experiments;
- explain the completely randomised design;
- describe the layout of CRD;
- explain the statistical analysis of CRD; and
- explain the advantages and disadvantages as well as the suitability of CRD.

9.1.1 Planning of an Experiment

There are some basic points regarding the planning of an experiment, which should be under consideration. These are as follows:

1. **The Experiment should be Free from Bias**

   An experiment must be planned so that it gives an unbiased estimate of the values we wish to measure. It is a matter of the design being such that no bias on the part of the experimenter can possibly enter into the results. This is achieved mainly by randomisation.

2. **There must be a Measure of Error**

   The true experiment is one that is strictly objective. It should furnish a measure of error and this error alone should be the measuring stick of significance.

3. **There must be a Clearly Defined Objective**
   For an experiment it is essential to specify the objects perfectly. In other words the objective of the experiment should be clearly defined.
4. **The Experiment should have Sufficient Accuracy**
   The accuracy of an experiment can be brought by the elimination of technical errors and by increasing replications. The number of replications should be decided to produce a given degree of accuracy.

9.1.2 Classification of Experimental Designs

Statisticians by themselves do not design experiments, but they have developed a number of structured schedules called "experimental designs", which they recommend for the taking of measurements. These designs have certain rational relationships to the purposes, needs and physical limitations of experiments. Designs also offer certain advantages in economy of experimentation and provide straightforward estimates of experimental effects and valid estimates of variance. There are a number of ways in which experiment designs might be classified, for example, the following:

1. By the number of experimental factors to be investigated (e.g., single-factor versus multifactor designs)
2. By the structure of the experimental design (e.g., blocked, factorial, nested, or response-surface design)
3. By the kind of information which the experiment is primarily intended to provide (e.g. estimates of effects, estimates of variance, or empirical mappings).

9.2 BASIC DEFINITIONS OF EXPERIMENTAL DESIGN

Several fundamental terms are widely used throughout this section. They may be defined as follows:

1. **Treatment**
   In an experiment, there are some variants under study, the effects of which are measured and tested (compared). These variants will be referred to as treatments. For example, to test the effects of three fertilizers, i.e., Nitrogen, Phosphorus and Potash on the yield of a certain crop. Then Nitrogen, Phosphorus and Potash are called treatments.
2. **Yield**
   The response of the treatment is measured by some indicator such as crop production, milk production, body temperature, mileage of engine set, etc. Such an indicator is called yield. The treatments are applied to some units such as field plots, sample of cows, sample of patients, sample of engine, sets, etc. and the effect on the yield is observed.
3. **Experimental Units**
   A unit to which one treatment applied is called experimental unit. It is the smallest division of an experimental material to which the treatment applied and on which the variable under study is measured. In carrying out an experiment, we should clear as to what constitute the experimental unit.

It can be understood that in a field of agriculture it is called plot, in the field of animal husbandry it may be a cow (cattle), in the field of medicine it may be a patient and in the field of automobile industry it may be engine set and so on.

4. **Experimental Material**
We have already explained the concept of experimental unit. The experimental material is nothing but a set of experimental units. For example, a piece of land, a group of cows, a number of patients and a group of engine sets, etc. Actually, an experimental material is that material on which some set of treatments are applied and tested.

5. **Blocks**
The experimental material is divided into a number of groups or strata which are so formed that they are within homogeneous and between heterogeneous. These groups or strata are called blocks.

6. **Experimental Error**
There is always a variation between the yields of the different plots even when they get the same treatment. This variation exists due to non-assignable causes, which cannot be detected and explained. These are taken to be of random type. This unexplained random part of variation is termed as experimental error. This include all types of extraneous variation due to, (i) inherent variability in the experimental units, (ii) error associated with the measurement made and (iii) lack of representativeness of the sample of the population understudy.

7. **Precision**
The precision of an experiment is measured by the reciprocal of the variance of a mean, i.e.

$$\frac{1}{v(\bar{x})} = \frac{1}{\sigma^2_{\bar{x}}} = \frac{n}{\sigma^2}$$

As n, the replication number increases, precision also increases.

8. **Uniformity Trial**
We know that to increase the efficiency of a design, the plots should be arranged into homogeneous blocks. It can be done only if we have a correct idea about the fertility variation of the field. This is achieved through uniformity trial. It is known that fertility of soil does not increase or decrease uniformly in any direction but it is distributed over the entire field in an erratic manner. By a uniformity trial, we mean a trial in which the field (experimental material) is divided into small units (plots) and the same treatment is applied on each of the units and their yields are recorded. From these yields we can draw a fertility control map which gives us a graphic picture of the variation of the soil fertility and enables us to form a good idea about the nature of the soil fertility variation. This fertility control map is obtained by joining the points of equal fertility through lines.
A uniformity trial gives us an idea about the

1. Fertility gradient of the field,
2. Determination of the shape of the plots to be used,
3. Optimum size of plots,
4. Estimation of number of replications required for achieving certain degree of accuracy.

9.3 PRINCIPLES OF DESIGN OF EXPERIMENTS

Good experimentation is an art and depends heavily upon the prior knowledge and abilities of the experimenter. Designing an experiment means deciding how the observations or measurements should be taken to answer a particular question in a valid, efficient and economical way. If a design is properly designed, then there will exists an appropriate way of analsing the data. From an ill-designed experiment no conclusion can be drawn.

The fundamental principles in design of experiments are the solutions to the problems in experimentation posed by the two types of nuisance factors and serve to improve the efficiency of experiments. For the validity of the design Prof. R.A. Fisher gave three principles of design of experiments, those fundamental principles are:

- Randomisation
- Replication
- Local Control

9.3.1 Randomisation

The principle of randomisation is essential for a valid estimate of the experimental error and to minimize the bias in the results. In the words of Cochran and Cox, "Randomisation is analogous to insurance in that it is a precaution against disturbances that may or may not occur and they may or may not be serious if they do occur". Thus, randomisation is so done that each treatment should get an equal chance. We mean that the treatments should be allocated randomly, i.e., by the help of random numbers. The following are the advantages of randomisations:

1. It provides a basis for the test of significance because randomisation ensures the independence of the observations which is one of the assumptions for the analysis of variance.
2. It is also a device for eliminating bias. Bias creeps in experiment, when the treatments are not assigned randomly to the units. This bias may be personal or subjective. The randomisation ensures the validity of the results.

9.3.2 Replication

"Replication" is the repetition, the rerunning of an experiment or measurement in order to increase precision or to provide the means for measuring precision. A single replicate consists of a single observation or experimental run. Replication provides an opportunity for the effects of uncontrolled factors or factors unknown to the experimenter to balance out and thus, through randomisation, acts as a bias-decreasing tool. Suppose a pain relieving drug A is applied to 4 patients, we say that drug A is replicated four times. By repeating a treatment it is possible to obtain a more reliable estimate because it reduces the experimental error. Further by repeating a treatment number of times we can judge the average performance of a treatment and the situation becomes clearer. Basically there are following uses of replication:

1. It enables us to obtain a more precise estimate of the treatments effects.
2. The next important purpose of replication is to provide an estimate of the experimental error without which we cannot test the significance of the difference between any two treatments. The estimate of experimental error is obtained by considering the difference in the plots receiving the same treatment in different replications and there is no other alternative of obtaining this estimate.
3. For a desired amount of precision, the minimum number of replications can be obtained.

9.3.3 Local Control

This method is used to attain the accuracy or to reduce the experimental error without increasing unduly the number of replications. Local control is a technique that handles the experimental material in such a way that the effects of variability are reduced. In local control, experimental units are divided into a number of homogeneous groups called blocks. These blocks are so formed that they are homogeneous within and heterogeneous between. This blocking of experiment may be row-wise, column-wise or both according to the number of factors responsible for heterogeneity. Different types of blocking constitute different types of experimental designs. The following are the advantages of local control:

1. By means of local control, the experimental error is reduced considerably and the efficiency of the design is increased.
2. By means of local control the test procedure becomes more sensitive or powerful.

Besides the above three principles, there are some other general principles in designing an experiment. Familiarity with the treatments and experimental material is an asset. Selection of experimental site is an asset. Selection of experimental site should be carefully done. Within block variability should be reduced.

9.4 SIZE AND SHAPE OF THE PLOTS

In field experiments, the size and shape of plots as well as of blocks influence the experimental error. The total available experimental area remaining fixed, an increase in size of plots will automatically decrease the number of plots and indirectly increases the block size. In order to reduce the flow of experimental material from one plot to another, it is customary to leave out strips of land between consecutive plots and also between blocks. These non-experimental areas are known as guard area. The size and shape of the plot should be such that we make a compromise between statistical and practical requirements i.e. if plot size is x and the variance of the plot is V(x), then V(x) is minimum (statistical consideration) and there should be no disturbance for agricultural operations (practical requirements).

The size and shape of block will ordinarily be determined by the size and shape of plots and the number of plots in a block. It is desirable from the point of view of error control to have small variations among the plots within a block and large variation among the blocks i.e. in general the division of experimental material into blocks is made in such a way that plots within blocks are as homogeneous as possible.

Based on these fundamental principles, we have certain designs. The analysis of those designs is based on the theory of least squares which gives the best estimates of the treatments effects and was initiated by Fisher (1926) followed by Yates (1936), Bose & Nair (1939) and Rao (1976). The following three designs are frequently used:

1. Completely Randomised Design
2. Randomised Block Design
3. Latin Square Design

9.5 COMPLETELY RANDOMISED DESIGN

The simplest of all the design is completely randomised design (CRD) which is applied in the case when the experimental materials are homogeneous. CRD is based on two principles i.e. randomisation and replication. The third principle, i.e. local control is not used because it is assumed that experimental materials are homogeneous. In this, the treatments are allocated randomly to the experimental units and each treatment is assigned to different experimental units completely at random (can be repeated any number of times) that is why it is called completely randomised design.

Suppose we have k treatments under comparison and the i^{th} treatment is to be replicated n_i times for i =1, 2, …, k, then the total number of units required for the design are $n = \sum_{i=1}^{k} n_i$. We allocate the k treatments completely at random to n units such that i^{th} treatment appears n_i times in the experiment.

9.5.1 Layout of Completely Randomised Design

The term layout refers to the allocation of different treatment to the experimental units. We have already said that treatments are allocated completely at random to the different experimental units. Every experimental unit has the same chance of receiving a particular treatment.

Suppose we want to test the effect of three pain relieving drugs A, B and C on twelve patients. Then we first number all the patients (units) from 1 to 12. Then from a random number table of one digit we pick up 12 numbers which are less than 4. Suppose the numbers are 1, 3, 2, 1, 3, 2, 1, 3, 2, 2, 3, 1. Thus the drug A is allotted to patient 1, drug C is allotted to patient number 2 and so on. It can be shown below:

(1) A	(2) C	(3) B	(4) A	(5) C	(6) B
(7) A	(8) C	(9) B	(10) B	(11) C	(12) A

It is clear from the above layout that the replications of A, B and C are equal. If the number of replications for each treatment is 5, 4 and 3 respectively, we number the experimental units in a convenient way from 1 to 12. We then get a random permutation of the experimental units. To the first 5 of the units in the random permutation we assign treatment A, to the next 4 units treatment B is assigned and the treatment C is assigned to the remaining 3 units.

9.5.2 Statistical Analysis of Completely Randomised Design

Statistical analysis of a CRD is analogous to the ANOVA for one-way classified data for fixed effect model, the linear model (assuming various effect to be additive) becomes

$$y_{ij} = \mu + \alpha_i + e_{ij}\ ,\ i = 1, 2, 3, \ldots, k;\ \ j = 1, 2, 3, ..., n_i \qquad \ldots(1)$$

where y_{ij} is the yield or response from the j^{th} unit receiving the i^{th} treatment, μ is the general mean effect, α_i is the effect due to the i^{th} treatment, where μ and α_i are constants so that $\sum_{i=1}^{k} n_i \alpha_i = 0$ and e_{ij} is identically and independently distributed (i.i.d.) $N(0, \sigma_e^2)$. Then, $n = \sum_{i=1}^{k} n_i$ is the total number of experimental units.

The analysis of model given in equation (1) is as same as that of fixed effect model of one-way classified data, discussed in Unit 6 of MST-005. If we write

$\sum_i \sum_j y_{ij} = y_{..} = G$ = Grand total of the n observations, and

$\sum_{j=1}^{n_i} y_{ij} = y_{i.} = T_{i.}$ = Total response in the units receiving the i^{th} treatment,

Then, as in ANOVA (one-way classified data),

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} + \bar{y}_{i.}\right)^2 + \sum_{i=1}^{k} n_i \left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

i.e. $\qquad TSS = SSE + SST$

where, TSS, SST and SSE are the Total Sum of Squares, Sum of Squares due to Treatments (between treatments SS) and Sum of Square due to Error (within treatment SS) given respectively by

$$TSS = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{..}\right)^2$$

$$SST = \sum_{i=1}^{k} n_i \left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 = S_T^2$$

and $\qquad SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{i.}\right)^2 = S_E^2$

ANOVA Table for CRD

Source of Variation	DF	SS	MSS	Variance Ratio (F)
Treatments	k-1	$SST = S_T^2$	$MSST = \frac{S_T^2}{(k-1)}$	$F_T = \frac{MSST}{MSSE}$
Error	n-k	$SSE = S_E^2$	$MSSE = \frac{S_E^2}{(n-k)}$	
Total	n-1	$S_T^2 + S_E^2$		

Under the null hypothesis, H_0: $\alpha_1 = \alpha_2 = \ldots. = \alpha_k$ against the alternative that all α's are not equal, the test statistic

$$F_T = \frac{MSST}{MSSE} \sim F(k-1, n-k)$$

i.e., F_T follows F distribution with $(k-1, n-k)$ df.

If $F_T > F_{(k-1, n-k)}$ (α) then H_0 is rejected at α level of significance and we conclude that treatments differ significantly. If $F_T < F_{(k-1, n-k)}$ (α) then H_0 may be accepted i.e. the data do not provide any evidence to prefer one treatment to the other and as such all of them can be considered alike.

If the treatments show significant effect then we would be interested to find out which pair of treatments differs significantly. For this instead of calculating Student's t-test for different pairs of treatment means we calculate the least significant difference at the given level of significance. This least difference is called as critical different (CD) and CD at α level of significance is given by

CD = Standard error of difference between two treatment means × $t_{\alpha/2}$ for error degrees of freedom.

We have

$$\text{Var}\,(\bar{y}_{i.} - \bar{y}_{.j}) = \frac{\sigma_e^2}{n_i} + \frac{\sigma_e^2}{n_j} = \sigma_e^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)$$

$$\text{Standard Error}\,(\bar{y}_{i.} - \bar{y}_{.j}) = \sigma_e\left(\frac{1}{n_i} + \frac{1}{n_j}\right)^{1/2}$$

Hence, the critical difference (CD) for $(\bar{y}_{i.} - \bar{y}_{.j})$

$$= t_{\alpha/2}\,(\text{for error df}) \times \left[\text{MSSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)\right]^{1/2}$$

Since MSSE provides an unbiased estimate of σ_e^2.

If each treatment is replicated n times, that is $n_i = n$ for i=1, 2, ..., k then

$$\text{CD for difference of mean} = (t_{\alpha/2} \text{ for error df}) \times \left[\text{MSSE} \times \left(\frac{2}{n}\right)\right]^{1/2}$$

9.5.3 Least Square Estimates of Effects

The completely randomised model in equation (1) in Sub-section 9.5.2 is a fixed effect model. Proceeding exactly as in Section 6.4 of Unit 6, we shall get

$$\hat{\mu} = \frac{y_{..}}{n} = \bar{y}_{..} \text{ and } \hat{\alpha}_i = \bar{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \qquad \ldots (2)$$

9.5.4 Variance of the Estimates

Proceeding exactly as in Section 6.7 of Unit 6, we shall get

$$\text{Var}\,(\hat{\mu}) = \frac{\sigma_e^2}{n}; \text{ where } n = \sum_{i=1}^{k} n_i \qquad \ldots (3)$$

and $\text{Var}(\alpha_i) = \text{Var}(\alpha_i) = \sigma_e^2 \left(\frac{1}{n_i} - \frac{1}{\sum_{i=1}^{k} n_i} \right)$ … (4)

If we assume that each treatment is replicated an equal number of times i.e., if $n_i = n$, (say), i = 1, 2, …, k; then $n = \sum_{i=1}^{k} n_i = nk$

Hence, from equations (3) and (4), we get

$$\text{Var}(\hat{\mu}) = \frac{\sigma_e^2}{nk} \text{ and } \text{Var}(\hat{\alpha}_i) = \text{Var}(\alpha_i) = \sigma_e^2 \left(\frac{k-1}{nk} \right) \quad \text{… (5)}$$

9.5.5 Expectation of Sum of Squares

Proceeding exactly as in Section 6.7 of Unit 6 [fixed effect model for one-way classified data], we get

$$E(SST) = E\left[\sum_{i=1}^{k} n_i (\bar{y}_{i.} - \bar{y}_{..})^2\right] = (k-1)\ \sigma_e^2 + \sum_{i=1}^{k} n_i \alpha_i^2$$

$$E(MSST) = E\left[\frac{S_T^2}{(k-1)}\right] = \sigma_e^2 + \frac{1}{k-1}\sum_{i=1}^{k} n_i \alpha_i^2 \quad \text{… (13)}$$

$$E(SSE) = (n-k)\,\sigma_e^2$$

$$E(MSSE) = E\left[\frac{S_E^2}{(n-k)}\right] = \sigma_e^2 \quad \text{… (14)}$$

The method of analysis of completely randomised design would be similar to one-way ANOVA, which has been illustrated below with the following example:

Example 1: A person wanted to purchase a lot of electric drills. He got quotations from five manufacturers. For the selection, he wanted to conduct an experiment to estimate the time taken by each making a hole in a metallic sheet. As the sheet might not be uniform all over in respect of thickness and hardness, he marked 20 places on the sheet and applied four drills from each concern in 4 randomly selected places to make holes. The time for making each hole was recorded and these formed the observations. The observations in seconds are shown below in brackets along with marks of the drills denoted by D_1, D_2, D_3, D_4 and D_5.

$D_1(19)\ D_3(22)\ D_4(20)\ D_1(20)$

$D_5(29)\ D_2(24)\ D_5(30)\ D_3(24)$

$D_2(26)\ D_4(25)\ D_1(16)\ D_2(22)$

$D_5(28)\ D_3(25)\ D_5(31)\ D_4(28)$

$D_4(27)\ D_1(16)\ D_2(27)\ D_3(20)$

Conduct the experiment by adopting a completely randomised design.

Solution: The analysis of the given design is done by one-way analysis of variance method. The data is analysed and computation results are given as below:

The totals of time records for 4 holes by each of the different makes are denoted by T_1, T_2, T_3, T_4 and T_5 are shown below.

$$T_1 = 71,\ T_2 = 99,\ T_3 = 91,\ T_4 = 100,\ T_5 = 118$$

Grand Total (G) = 479

$$\text{Correction Factor (CF)} = \frac{G^2}{N} = \frac{(479)^2}{20} = 11472.05$$

$$\text{Total Sum of Squares (TSS)} = 21^2 + 18^2 + 22^2 + \ldots + 31^2 + 20^2 - 11472.05$$

$$= 11847 - 11472.05 = 374.95$$

Sum of Squares due to Makes (SSM)

$$= \frac{(71)^2 + (99)^2 + (91)^2 + (100)^2 + (118)^2}{4} - 11472.05$$

$$= 11761.75 - 11472.05 = 289.70$$

$$\text{Sum of Squares due to Error (SSE)} = \text{TSS} - \text{SSM}$$

$$= 374.95 - 289.70 = 85.25$$

Analysis of Variance Table

Sources of Variation	DF	SS	MSS	F
Makes	4	289.7	72.425	12.75
Error	15	85.25	5.68	
Total	19	374.95		

The tabulated value of F at 1 per cent level of significance for 4 and 15 df is 4.89. Thus, the calculated value of F viz. 12.75 shows that Make to Make variation is highly significant thereby indicating that the hypothesis that the time periods taken by the different Makes in boring a hole are, on an average, the same, is rejected. So multiple camparison test will be applied for different Makes.

Mean for Different Makes

Makes				
$\bar{D}_1$	$\bar{D}_2$	$\bar{D}_3$	$\bar{D}_4$	$\bar{D}_5$
17.74	24.75	22.75	25.00	29.50

$$\text{SE} = \sqrt{\frac{2\text{MSSE}}{n}} = \sqrt{\frac{2 \times 5.68}{4}} = 1.69$$

Critical difference at 1 % level of significance

$$\text{CD} = t_{\alpha/2} \text{ (for error df)} \times \text{SE} = 3.055 \times 1.69 = 5.16$$

The initial difference indicates that the Make D_5 is significantly better than all the other Makes.

Pair of Treatments	Difference	CD	Inference
D_1, D_2	$\lvert\overline{D_1} - \overline{D_2}\rvert = 7.01$	5.16	Significant
D_1, D_3	$\lvert\overline{D_1} - \overline{D_3}\rvert = 5.01$	5.16	Insignificant
D_1, D_4	$\lvert\overline{D_1} - \overline{D_4}\rvert = 7.26$	5.16	Significant
D_1, D_5	$\lvert\overline{D_1} - \overline{D_5}\rvert = 11.26$	5.16	Significant
D_2, D_3	$\lvert\overline{D_2} - \overline{D_3}\rvert = 2.00$	5.16	Insignificant
D_2, D_4	$\lvert\overline{D_2} - \overline{D_4}\rvert = 0.25$	5.16	Insignificant
D_2, D_5	$\lvert\overline{D_2} - \overline{D_5}\rvert = 4.75$	5.16	Insignificant
D_3, D_4	$\lvert\overline{D_3} - \overline{D_4}\rvert = 2.25$	5.16	Insignificant
D_3,D_5	$\lvert\overline{D_3} - \overline{D_5}\rvert = 6.75$	5.16	Significant
D_4, D_5	$\lvert\overline{D_4} - \overline{D_5}\rvert = 4.5$	5.16	Insignificant

E1) Carryout the ANOVA for the given following data of yields of 5 varieties, 7 observations on each variety:

Variety	Observations						
	1	**2**	**3**	**4**	**5**	**6**	**7**
1	13	15	14	14	17	15	16
2	11	11	10	10	15	9	12
3	10	13	12	15	14	13	13
4	16	18	13	17	19	14	15
5	12	12	11	10	12	10	10

9.6 SUITABILITY OF CRD

The following are some situations, in which one can apply the complete randomised design:

1. The CRD is used in the situations where experimental materials are homogeneous. That is why, CRD is mostly used in chemical, biological and banking experiments, where the experimental material is thoroughly mixed powder, liquid or chemical.
2. The CRD is used in the situations where the observations on some units are missing or destroyed. This feature of missing observation does not disturb the analysis of the design.

3. In agricultural experiments, this design is not used because experimental material is not homogeneous.

9.6.1 Advantages and Disadvantages of CRD

Advantages of CRD

1. In this design any number of treatments and replications can be used. There may be different number of replications for different treatments.
2. Analysis is simple and easy even if the number of replication is unequal for each treatment. In such case experimental error will differ from treatment to treatment.
3. If some of the observations are missing or destroyed or not available due to some reasons, the analysis can be done without any problem.
4. It provides large degree of freedom for error sum of squares. This increases the sensitivity of the experiment.
5. In CRD there is no condition on the number of replication of the treatments, they can be increased or decreased according to the need of the experimenter. Thus, the design is flexible.

Disadvantages of CRD

1. The main disadvantage of CRD is that the principle of local control has not been used in this design. Due to this fact, the experimental error is inflated. This is the main reason for the criticism of CRD.
2. In agricultural experiments, the design is seldom used because the experimental material is not homogenous.

9.7 SUMMARY

In this unit, we have discussed:

1. The experimental design;
2. The planning and classification of experimental designs;
3. Principles of design of experiments;
4. Completely randomised design;
5. Layout of CRD;
6. The statistical analysis of CRD; and
7. Advantages and disadvantages as well as suitability of CRD.

9.8 SOLUTIONS / ANSWERS

E1) The analysis of the given design is done by one-way analysis of variance method. The data is analysed and computation results are given as below:

Correction Factor (CF) = 6072.03

Raw Sum of Squares (RSS) = 6293

Total Sum of Squares (TSS) = 220.97

Sum of Squares due to Variety (SSV) = 138.40

Sum of Squares due to Error (SSE) = TSS – SSV

= 220.97 - 138.40 = 82.57

ANOVA Table

Source of Variation	DF	SS	MSS	Variance Ratio	
				Calculated	Tabulated
Variety	4	138.40	34.60	12.58	2.66
Error	30	82.57	2.75		
Total	34	220.97			

Null Hypothesis H_0: $\mu_1 = \mu_2 = \ldots = \mu_5$

Since, calculated value of F is greater than the tabulated value of F, we reject the null hypothesis and conclude that variety effects are significantly different.

Mean for Different Varieties

Varieties				
$\bar{D}_1$	$\bar{D}_2$	$\bar{D}_3$	$\bar{D}_4$	$\bar{D}_5$
14.86	11.14	12.86	16.00	11.00

$$SE = \sqrt{\frac{2MSSE}{n}} = \sqrt{\frac{2 \times 13.34}{7}} = 1.95$$

Critical difference at 1 % level of significance

$= t_{\alpha/2}$ (for error df) $\times$ SE $= 3.055 \times 1.95 = 5.96$

The initial difference indicates that the Variety D_4 is significantly better than all the other Varieties.

Pair of Treatments	Difference	CD	Inference
D_1, D_2	$\lvert\bar{D}_1 - \bar{D}_2\rvert = 3.72$	5.96	Insignificant
D_1, D_3	$\lvert\bar{D}_1 - \bar{D}_3\rvert = 2.00$	5.96	Insignificant
D_1, D_4	$\lvert\bar{D}_1 - \bar{D}_4\rvert = 1.14$	5.96	Insignificant
D_1, D_5	$\lvert\bar{D}_1 - \bar{D}_5\rvert = 3.86$	5.96	Insignificant
D_2, D_3	$\lvert\bar{D}_2 - \bar{D}_3\rvert = 1.72$	5.96	Insignificant
D_2, D_4	$\lvert\bar{D}_2 - \bar{D}_4\rvert = 4.86$	5.96	Insignificant
D_2, D_5	$\lvert\bar{D}_2 - \bar{D}_5\rvert = 0.14$	5.96	Insignificant
D_3, D_4	$\lvert\bar{D}_3 - \bar{D}_4\rvert = 3.14$	5.96	Insignificant
D_3, D_5	$\lvert\bar{D}_3 - \bar{D}_5\rvert = 1.86$	5.96	Insignificant
D_4, D_5	$\lvert\bar{D}_4 - \bar{D}_5\rvert = 5.00$	5.96	Insignificant