Lecture-9 Completely Randomized Design

Complete and incomplete block designs:

In most of the experiments, the available experimental units are grouped into blocks having more or less identical characteristics to remove the blocking effect from the experimental error. Such design is termed as block designs. The number of experimental units in a block is called the **block size**.

If size of block = number of treatments and each treatment in each block is randomly allocated, then it is a full replication and the design is called a complete block design.

In case, the number of treatments is so large that a full replication in each block makes it too heterogeneous with respect to the characteristic under study, then smaller but homogeneous blocks can be used. In such a case, the blocks do not contain a full replicate of the treatments. Experimental designs with blocks containing an incomplete replication of the treatments are called incomplete block designs.

***Completely randomized design (CRD)**:

The CRD is the simplest design. Suppose there are v treatments to be compared.

- All experimental units are considered the same and no division or grouping among them exist.
- In CRD, the v treatments are allocated randomly to the whole set of experimental units, without making any effort to group the experimental units in any way for more homogeneity.

- Design is entirely flexible in the sense that any number of treatments or replications may be used.

- The number of replications for different treatments need not be equal and may vary from treatment to treatment depending on the knowledge (if any) on the variability of the observations on individual treatments as well as on the accuracy required for the estimate of individual treatment effect. Example: Suppose there are 4 treatments and 20 experimental units, then - the treatment 1 is replicated, say 3 times and is given to 3 experimental units, - the treatment 2 is replicated, say 5 times and is given to 5 experimental units, - the treatment 3 is replicated, say 6 times and is given to 6 experimental units and - finally, the treatment 4 is replicated [20-(6+5+3)=]6 times and is given to the remaining 6 experimental units.

All the variability among the experimental units goes into experimented error.

- CRD is used when the experimental material is homogeneous.

- CRD is often inefficient.

- CRD is more useful when the experiments are conducted inside the lab.

- CRD is well suited for the small number of treatments and for the homogeneous experimental material.

Layout of CRD :

Following steps are needed to design a CRD

- Divide the entire experimental material or area into a number of experimental units, say n.
- Fix the number of replications for different treatments in advance (for given total number of available experimental units).
- No local control measure is provided as such except that the error variance can be reduced by choosing a homogeneous set of experimental units.

Procedure:

Suppose that we have v treatments under comparison and ith treatment is to be replicated ri times; i=1, 2… v, then the total number of experimental units necessary for these experiment is $n = \sum_{i=1}^{v} ri$. In CRD, we allocate v treatments completely at random to the n units subject to the condition that the ith treatment appears in ri units; i=1, 2,…, v. in particular case if r_i=r; i=1, 2,…, v viz. if each treatment is repeated an equal number of times say r times then n=rv and randomization gives every group an equal chance of receiving the treatments.

ADVANTAGES OF CRD:-

1. It is easy to layout the design.
2. It results in the maximum use of the experimental units since all the experimental materials can be used.
3. CRD provides 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.
4. CRD allows complete flexibility in the number of treatments and the number of their replications which may vary from treatment to treatment.
5. The relative loss of information due to missing data is smaller in comparison with any other design.

DIS-ADVANTAGES OF CRD:-

1. The main objection against CRD is that the principle of local control has not been used.
2. CRD is rarely used in field experimentation because the plots are not homogeneous. In certain circumstances the design suffers from the disadvantage of being inherently less informative than other more sophisticated layouts. This usually happens, if the experimental material is not homogeneous.
3. Since randomization is not restricted in any direction to ensure that the units receiving one treatment are similar to those receiving the other treatments, the whole variation

among the experimental units is included in the residual variance. This makes the design less efficient and results in less sensitivity in detecting significant effects. Thus, CRD is rarely used.

APPLICATIONS OF CRD:-

1. CRD is most useful in laboratory technique and methodological studies. For e.g.:- it is used in physics, chemistry and biological experiments and also in some green house studies where either the experimental material is homogeneous or the variability between the units can be reduced.
2. CRD is also recommended in situations where an appreciable fraction of units is likely to be destroyed or fail to respond.

STATISTICAL ANALYSIS OF CRD:-

This design provides one-way classified data according to the level of single factor treatments. Assuming various effects are additive in nature the linear mathematical model for CRD is given by,

$x_{ij}=\mu+t_i+e_{ij}$; i=1,2,…,v ; j=1,2,…,ri -------(1)

where x_{ij} = the yield from the jth unit receiving the ith treatment.

μ = general mean effect

t_i = effect due to ith treatment.

e_{ij} = error effect due to chance.

Along with usual assumptions:-

i. All sample observations are independent.
ii. Various effects are additive in nature.
iii. $e_{ij} \sim N(0, {\sigma_e}^2)$
iv. $\sum_{i=1}^{v} riti = 0$

NULL-HYPOTHESIS:-

We set up the null-hypothesis as:-

H_0: $\mu_{1.} = \mu_{2.} = \ldots\ldots = \mu_{v.} = \mu$ or

H_t: $t_1 = t_2 = \ldots\ldots = t_v = 0$ or

Treatments are homogeneous or Treatments are equally effective.

ESTIMATES OF THE PARAMETERS IN THE MODEL:-

Let $\bar{x}_{i.} = \frac{1}{r_i}\sum_{j=1}^{ri} x_{ij}$

=mean of the i[th] treatment

$\bar{x}_{..} = \frac{1}{n}\sum_{i=1}^{v}\sum_{j=1}^{ri} x_{ij}$

= overall mean

The parameters involved in the mathematical model given in (1) can be estimated using the principles of least squares on minimizing the error sum of squares.

Viz. we minimize; $E = \sum_i \sum_j {e_{ij}}^2$

$=\sum_i \sum_j (x_{ij} - \mu - t_i)^2$

The normal equations for estimating μ and t_i are $\frac{\partial E}{\partial \mu} = 0$ *and* $\frac{\partial E}{\partial ti} = 0$ respectively.

Thus, $\frac{\partial E}{\partial \mu} = 0 \Rightarrow -2\sum_i \sum_j (x_{ij} - \mu - t_i) = 0$

$\Rightarrow \sum_i \sum_j x_{ij} - n\mu - \sum_i r_i t_i = 0$

$\Rightarrow \sum_i \sum_j x_{ij} - n\mu = 0 \quad (\because \sum_i r_i t_i = 0)$

$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_i \sum_j x_{ij}$

$\Rightarrow \hat{\mu} = \bar{x}_{..}$

Similarly, $\frac{\partial E}{\partial ti} = 0 \Rightarrow -2\sum_{j=1}^{ri}(x_{ij} - \mu - t_i) = 0$

$\Rightarrow \sum_{j=1}^{ri} x_{ij} - r_i\mu - r_i t_i = 0$

$\Rightarrow \hat{t}_i = \frac{1}{r_i}\sum_{j=1}^{ri} x_{ij} - \hat{\mu}$

$\Rightarrow \hat{t}_i = \bar{x}_{i.} - \bar{x}_{..}$

Now, $\sum_{i=1}^{v}\sum_{j=1}^{ri}(x_{ij} - \bar{x}_{..})^2$

$= \sum_i \sum_j (x_{ij} - \bar{x}_{i.} + \bar{x}_{i.} - \bar{x}_{..})^2$

$= \sum_i \sum_j (\bar{x}_{ij} - \bar{x}_{i.})^2 + \sum_i r_i(\bar{x}_{i.} - \bar{x}_{..})^2 + 0.$

Thus, total S.S = S.S.E. + S.S.t.

Where, total S.S. = total sum of squares = $\sum_{i=1}^{v}\sum_{j=1}^{ri}(x_{ij} - \bar{x}_{..})^2$

S.S.E. = Sum of squares due to error = $\sum_i \sum_j (\bar{x}_{ij} - \bar{x}_{i.})^2$

S.S.t. = Sum of squares due to treatments = $\sum_i r_i(\bar{x}_{i.} - \bar{x}_{..})^2$

ANOVA TABLE FOR CRD:-

Sources of variation	**Degrees of freedom**	**Sum of squares**	**Mean sum of squares**	**F-ratio**
Treatments	v-1	$S_t{}^2 = \sum_i r_i(\bar{x}_{i.} - \bar{x}_{..})^2$	$s_t{}^2 = \frac{S_t{}^2}{v-1}$	$F = \frac{s_t{}^2}{s_E{}^2}$
Error	n-v	$S_E{}^2 = \sum_i \sum_j (\bar{x}_{ij} - \bar{x}_{i.})^2$	$s_E{}^2 = \frac{S_E{}^2}{n-v}$	
Total	n-1	$\sum_{i=1}^{v} \sum_{j=1}^{ri} (x_{ij} - \bar{x}_{..})^2$		

Where, $F = \frac{s_t{}^2}{s_E{}^2} \sim F - distribution\ with\ (v-1, n-v) degrees\ of\ freedom.$

If $F_{cal} < F_{tab}$ at α% level of significance, we may accept H_t: $t_1 = t_2 = \ldots = t_v = 0$ otherwise it is rejected.