Randomised Block Design

In this session we introduce Randomised Block design and discuss its analysis. The entire topic is divided in to following subdivisions:

1. Introduction and definition
2. **The statistical Model for RBD**
3. **Least square estimates of the model parameters**
4. Statistical Analysis
5. **Analysis of variance Table(ANOVA Table)**
6. **Advantages and disadvantages**
7. **Example**
8. **Conclusion**

Randomised complete block designs(RBD)

If large number of treatments are to be compared then large number of experimental units are required and there is practical difficulties to acquire all the homogeneous experimental units.. Non- homogeneous experimental units will increase the variation among the responses and CRD may not be appropriate design to use. Alternative design in such case is to compare the p treatments is RBD

RBD: A two way layout is called RBD if there are N = pxk experimental units. Group these N experimental units in to k blocks of p units each such that with in the blocks the experiment units are relatively homogeneous in nature. With in each block the p treatments are randomly assigned to the p experimental units such that assigning the treatments to these experimental units has the same probability to appear and the assignment in different blocks are statistically independent.

The RBD utilises the principles of randomisations, replication and local control in the following way:

Randomisation: The p treatments to the p experimental units in each block are randomly applied

Replication: Since each treatment appears once and only once in each block, every treatment will appear in all the blocks. Hence each treatment replicated the number of times as the number of blocks.

Local Control:

Local control is adopted in the following way: First from the homogeneous blocks of the experimental units, then allocate each treatment randomly in each block. The error variance now will be smaller because of homogeneous blocks and some variance will be parted away from the error variance due to the difference among the blocks.

Layout: The observed data set is arranged as follows:…

	Blocks				Block Totals	Block Averages
Treatments	1	2		k		
1	y_{11}	Y_{12}	…	Y_{1n}	$y_{1.}$	$\overline{y_{1.}}$
2	y_{21}	Y_{22}.	…	Y_{2n}	$y_{2.}$	$\overline{y_{2.}}$
.	.	.	…	.	.	.
..	.	.	…	.	.	.
.	.	.	…	.	.	.
p	y_{p1}	y_{p2}	…	y_{pn}	$y_{p.}$	$\overline{y_{p.}}$
Treatment	y.	y.2	…	y.	y..	

Total	1			n		
Treatment Averages	$\overline{y_{.1}}$	$\overline{y_{.2}}$	$\overline{y_{.3}}$	$\overline{y_{.n}}$		

2. **The statistical Model for RBD**

Y_{ij} represent the j^{th} observation taken from treatment i.

We define the model

$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$, i= 1,2...p, j= 1,2 ...k

(4)

Where μ general effect

α_i – i^{th} treatment effect;

βj – j^{th}blockeffect;

ϵij- error term, independent and identically distributed random variables with mean 0 and variance σ2

There are two null hypothesis to be tested:

1. Related to the treatment effects

H_A: $\alpha_1 = \alpha_2 = \ldots \alpha_p = 0$

And the alternative hypothesis is

H_{A1}: at least one $\alpha_i = \alpha_j$ for all I,j

2. Related to the block effects

H_B: $\beta_1 = \beta_2 = \ldots \beta_p = 0$

And the alternative hypothesis is

H_{B1}: at least one$\beta_i = \beta_j$ for all I,j

3.Least square estimates of the model parameters:

The parameters μ, α_i and β_j are estimated by the method of least squares. i.e. by minimising error sum of squares. L = $\sum_{i=1}^{p}\sum_{j=1}^{k}\varepsilon_{ij}^{2} = \sum_{i=1}^{p}\sum_{j=1}^{k}(y_{ij}-\mu-\alpha_i-\beta_j)^2$

And solving the normal equations $\frac{\partial L}{\partial \mu}=0$, , $\frac{\partial L}{\partial \alpha_i}=0$ i= 1,2…p and $\frac{\partial L}{\partial \beta_j}=0$, , j=1,2…k

we obtain p+k+1 normal equations as ,

$$N\mu+\sum_{i=1}^{p}k\alpha_i+\sum_{j=1}^{k}p\beta_j=\sum_{i=1}^{p}\sum_{j=1}^{k}y_{ij}$$

$k\mu+k\alpha_i+\sum_{j=1}^{k}\beta j=\sum_{j-1}^{k}y_{ij}$, i= 1,2 …p

And

,

$p\mu+\sum_{i-1}^{p}\alpha_i+p\beta_j=\sum_{i-1}^{p}y_{ij}$, j=1.2,…k

,

These normal equations are not linearly independent, as first equation is equal to the sum of p equations corresponding to α and equal to the sum of k equations corresponding toβ. Hence no unique solution exists for μ and $\alpha_{i,\ i=}$ 1,2..p. and β_j j=1,2…k. Since we have defined the treatment effects as deviations from overall mean, hence we add independent constraint, $\sum_{i=1}^{p}\alpha i=0$, $\sum_{j=1}^{k}\beta_j=0$ and solve the simultaneous normal equations. Solving we get the solutions as

$\hat{\mu} = \bar{y}..$ and $\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}..$ i=1,2..p βj = $\hat{\beta}_i = \bar{y}._j - \bar{y}..$

The fitted model after substituting the estimates $\hat{\mu}$ and $\hat{\alpha}_i$ and $\hat{\beta}$ in the linear model we get

Yij= $\hat{\mu}$ + $\hat{\alpha}_i$ + $\hat{\beta}$ +ϵij

Or

Y_{ij}= $\bar{y}..$ + ($\bar{y}_{i.} - \bar{y}..$)+ ($\bar{y}._j - \bar{y}..$)+($y_{ij.} - \bar{y}_{i.} + \bar{y}_{.j} + \bar{y}..$)

Or

(Yij- $\bar{y}..$)= ($\bar{y}_{i.} - \bar{y}..$)+ ($\bar{y}._j - \bar{y}..$)+($y_{ij.} - \bar{y}_{i.} + \bar{y}_{.j} + \bar{y}..$), the error term is chosen that both sides of the equation are equal

Squaring both sides and summing over all the observations we get

$\sum_{i=1}^{p}\sum_{j=1}^{k}(y_{ij} - \overline{y_{..}})^2 = \sum_{i=1}^{p}k(\overline{yi.} - \overline{y_{..}})^2 + \sum_{j=1}^{k}p(\overline{y.j} - \overline{y_{..}})^2 + \sum_{i=1}^{p}\sum_{j=1}^{k}(y_{ij.} - \bar{y}_{i.} + \bar{y}_{.j} + \bar{y}..)^2$, all the cross product vanishes.

Or

SST = SSTR + + SSB+ SSE

Where

SST = $\sum_{i=1}^{p}\sum_{j=1}^{k}(y_{ij} - \overline{y_{..}})^2 = \sum_{i=1}^{p}\sum_{j=1}^{k}y_{ij}^2 - \frac{y_{..}^2}{N}$

SSTR = $\sum_{i=1}^{p}k(\overline{yi.} - \overline{y_{..}})^2 = \sum_{i=1}^{p}\frac{y_{i.}^2}{k} - \frac{y_{..}^2}{N}$

SSB= $\sum_{j=1}^{k}p(\overline{y.j} - \overline{y_{..}})^2 = \sum_{i=1}^{k}\frac{y_{.j}^2}{p} - \frac{y_{..}^2}{N}$

And SSE = $\sum_{i=1}^{p}\sum_{j=1}^{k}(y_{ij.} - \bar{y}_{i.} + \bar{y}_{.j} + \bar{y}..)^2$ SST-SSTR-SSB Computed a

4. **Statistical Analysis**: Them analysis of RBD is similar to two way ANOVA

. Under the null hypothesis H_A: $\alpha_1 = \alpha_2 = \ldots \alpha_p = 0$

the ratio

$$F_A = \frac{SSTR/(p-1)}{SSE/(p-1)(k-1)} = \frac{MSSTR}{MSSE} \qquad (1)$$

is distributed as F with p-1 and (p-1)(k-1) degrees of freedom.

. Similarly if the null hypothesis H_B: $\beta_1 = \beta_2 = \ldots \beta_p = 0$

, the ratio

$$F_B = \frac{SSB/(K-1)}{SSE/(p-1)(k-1)} = \frac{MSSB}{MSSE} \qquad (2)$$

is distributed as F with k-1 and (p-1)(k-1) degrees of freedom.

We reject H_A and conclude that there are differences in the treatment means if $F_A > F_{\alpha,,p-1,(p-1)(k-1)}$ Where F_0 is computed from equation 1 and Fα, p-1, (p-1)(k-1), is the table value referring to F table at α level significance corresponding to p-1and (p-1)(k-1) degrees freedom.

We reject H_B and conclude that there are differences in the group means if $F_B > F_{\alpha,,k-1,(p-1)(k-1)}$ Where F_B is computed from equation 2 and Fα, k-1, (p-1)(k-1), is the table value referring to F table at α level significance corresponding to k-1and (p-1)(k-1)degrees freedom.

If HB is accepted , then it indicates that the blocking is not necessary for future experimentation

5. **Analysis of variance Table(ANOVA Table)**

Sources of variation	Degree of freedom	Sum of squares	Mean sum of squares	F-Value
Treatments	p-1	SSTR	MSSTR = SSTR/(p-1)	MSSTR / MSSE

replicates	k-1	SSB	MSSB = SSB/(k-1)	MSS**B**/ MSSE
Error	(p-1) (k-1)	SSE	MSSE = SSE/(p-1) (k-1)	
total	N-1	TSS		

6. Advantages and disadvantages

Advantages of RBD ::

a) Blocking increases precision
b) Any number of blocks and any no. of treatments with in blocks can be used
c) Statistical analysis relatively simple
d) Easy to construct the design
e) When significant blocking can be achieved, differences due to error variance are eliminated from treatment contrasts.
f) RBD has greater precision than CRD

Disadvantages of the RCB designs are:

a) Missing observations within blocks complicates analysis
b) Degree of freedom for RBD smaller than for a comparable CRD
c) The design is not suitable for testing a large number of treatments, as with increase in block size, the blocks are not likely to consists of homogeneous plots and hence error sum will increase.
d) If block and treatment effects interact (that is, they are not additive). The RBD analysisis not appropriate.

7. **Example**:

Data gives the grain in yield of rice at six seeding rates(kg.Ha)

	Seeding rates(kg.ha)

	Total						
Replicates	25	50	75	100	125	150	
1	5.1	5.3	5.3	5.2	4.8	5.3	31.0
2	5.4	6.0	5.7	4.8	4.8	4.5	31.2
3	5.3	4.7	5.5	5.0	4.4	4.9	29.8
4	4.7	4.3	4.7	4.4	4.7	4.1	26.9
Total	20.5	20.3	21.2	19.4	18.7	18.8	118.9

Step 1:Calculation of correction factor(CF): $\frac{y_{..}^2}{N}$ = $118.9^2/(6x4)$ = 589.05

Step;2: Calculation of Total sum of squares: $\sum_{i=1}^{p}\sum_{j=1}^{k} y_{ij}^2 - \frac{y_{..}^2}{N}$ = $(5.1^2+5.4^2+\ldots+4.1^2)$-CF

=5.02

Step;3: Calculation of Treatment sum of squares= $\sum_{i=1}^{p} \frac{y_{i.}^2}{k} - \frac{y_{..}^2}{N}$ =

$(1/4)(20.5^2+20.3^2+21.2^2+19.4^2+18.7^2+18.8^2)$ –CF = 1.2675

Step;4: Calculation of Replicate sum of squares = $\sum_{i=1}^{k} \frac{y_{.j}^2}{p} - \frac{y_{..}^2}{N}$ =

$(1/6)(31.0^2+31.2^2+29.8^2+26.9^2)$ CF=1.965

Step;5: Calculation of Error sum of squares

ESS = 5.02-1.2675-1.965 =1.7875

Anova Table:

Sources of variation	Degree of freedom	Sum of squares	Mean sum of squares	F-Value
Treatments	p-1=5	SSTR=1.2675	0.2535	2.1267
replicates	k-1=3	SSB = 1.965	0.655	5.5788
Error	(p-1)(k-1)=15	SSE = 1.7875	0.1192	
total	N-1=23	TSS= 5.02		

Table value for Hα is F.05, 5,15 = 2. 90

Table value for Hβ is F.05, 3,15 = 3.29

Conclusion: Since F cal(2.1267) < F.05, 5,15 we do not reject Hα. Hence we conclude that all treatment means are equal

Since F cal(5.5788) > F.05, 3,15 we do reject Hβ. Hence we conclude that replicate means are significantly different ,

Calculate critical difference if necessary In case Hβ to tests pairwise comparison which is discussed in two way ANOVA.

8. **Conclusion:** In this lecture we have discussed the need for RBD and the defined RBD model. We have discussed the model under RBD and analysis of the RBD. We have derived the least square estimates of model parameters. We

summarised the complete analysis using ANOVA table. WE have discussed the advantages and disadvantages of RBD. The entire analysis we discussed solving a problem .