Randomized Complete Block Design (RCBD)

Features:
(i) Randomization procedure
(ii) Completeness- treatment applied in complete allocation to blocks
(iii) Blocking

Note:
(i) It is a two-way classification namely treatment and blocks.
(ii) The block is a replication of the treatments and blocks constitute replicates.
(iii) Treatments are allocated at random within each block.
(iv) (a) When the blocks and treatments are all the materials that are available, then the design model is fixed.
(b) When the blocks are a random sample of all plots, but the treatments are just all the available treatments, then the model is mixed.
(c) Also, if the treatments are a random sample but the blocks are not, the model is also mixed.
(d) If the blocks and treatments are respectively random samples, the model is random.
(v) The block is a collection or assembly of homogeneous plots, but blocks are not a homogeneous set.
(vi) Randomized blocks can be:
(a) Complete or Incomplete: complete because they have equal replications of the treatments.
(b) Balanced or Unbalanced: balanced when every comparison between two treatments has the same variance.

Note: A complete design is necessarily balanced while an incomplete designed may be balanced or unbalanced.

Model:
\[ y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \]
\( \mu \) is the overall mean
\( \tau_i \) is the i\textsuperscript{th} treatment effect
\( \beta_j \) is the j\textsuperscript{th} block effect
\( e_{ij} \) is the random error term
\( y_{ij} \) is the response

Assumptions:
(i) The blocks and treatment effects are additive (no interaction between the treatments and the blocks). That is \( \sum \tau_i = 0 \) and \( \sum \beta_j = 0 \).
(ii) Homoscedasticity: equal variance, expectation and same distribution of the error term. That is \( E(e_{ij}) = 0 \) and \( \text{Var}(e_{ij}) = \sigma^2 \).
(iii) Independent error. That is \( e_{ij} \sim NID(0, \sigma^2) \).

Effects of the Departure from the Assumptions of Analysis of Variance
(i) Non-additivity
(ii) Heteroscedasticity
(iii) Dependent error

**Advantages of RCBD**
(i) Grouping leads to the obtaining of more accurate result than with CRD.
(ii) Statistical analysis is straightforward.
(iii) Any number of treatments and replicates may be used even though the cost of experimentation needs to be considered when determining this.

**ANOVA Table**

<table>
<thead>
<tr>
<th>Source of Variation (SV)</th>
<th>Sum of Squares (SS)</th>
<th>Degrees of Freedom (df)</th>
<th>Mean Square (MS)</th>
<th>F ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>$SSt$</td>
<td>$t - 1$</td>
<td>$\frac{SSt}{t - 1} = A$</td>
<td>$A/C$</td>
</tr>
<tr>
<td>Blocks</td>
<td>$SSb$</td>
<td>$b - 1$</td>
<td>$\frac{SSb}{b - 1} = B$</td>
<td>$B/C$</td>
</tr>
<tr>
<td>Error</td>
<td>$SSe$</td>
<td>$(t - 1)(b - 1)$</td>
<td>$\frac{SSe}{(t - 1)(b - 1)} = C$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$SST$</td>
<td>$bt - 1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example 4:**

Four different drugs $A_1, A_2, A_3, A_4$ have been developed for the cure of a certain disease. These drugs are tried on patients in three different hospitals namely $B_1, B_2, B_3$. The result given below shows the number of cases of recovery from the disease per 100 people who have taken the drugs. The Randomized Block Design has been employed to eliminate the effect of the different hospitals. Carry out the analysis of variance at 5% level of significance.

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_1$</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>$B_2$</td>
<td>19</td>
<td>9</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>$B_3$</td>
<td>11</td>
<td>8</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>28</td>
<td>53</td>
<td>22</td>
</tr>
</tbody>
</table>

**Solution**
(i) \[ SST = 10^2 + 11^2 + \cdots + 23^2 + 5^2 - \frac{(143)^2}{12} \]
\[ = 2019 - 1704.08 = 314.92 \]

(ii) \[ SS_t = \frac{40^2 + 28^2 + 53^2 + 22^2}{3} - \frac{(143)^2}{12} \]
\[ = 1892.33 - 1704.08 = 188.25 \]

(iii) \[ SS_b = \frac{43^2 + 53^2 + 47^2}{4} - \frac{(143)^2}{12} \]
\[ = 1716.75 - 1704.08 = 12.67 \]

(iv) \[ SSE = SST - [SS_t + SS_b] \]
\[ = 314.92 - 200.92 = 114.00 \]

<table>
<thead>
<tr>
<th>SV</th>
<th>SS</th>
<th>Df</th>
<th>MS</th>
<th>Fcal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>188.25</td>
<td>3</td>
<td>62.75</td>
<td>3.30</td>
</tr>
<tr>
<td>Blocks</td>
<td>12.67</td>
<td>2</td>
<td>6.34</td>
<td>0.33</td>
</tr>
<tr>
<td>Error</td>
<td>114.00</td>
<td>6</td>
<td>19.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>314.92</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**F_{tab}**
Treatment = \( F_{0.05,3,6} = 4.76 \)
Blocks = \( F_{0.05,2,6} = 5.14 \)

**Decision:** Since \( F_{tab} > F_{cal} \) for both treatment and blocks, we fail to reject \( H_0 \) in both cases.

**Conclusion:**
(i) The drug effects (treatment) are not significant.
(ii) Hospitals effects are not significant.

**Block Efficiency**
Blocking maximizes the difference among blocks leaving the difference among plots of the same block as small as possible. Thus, the result of every RCB experiment should be examined to see how this objective has been achieved. As earlier stated, blocking reduces the magnitude of the experimental error.
The process of examination is as described below:
(i) Determine the magnitude of the reduction in experimental error due to blocking by computing the relative efficiency parameter given as:
\[ R.E = \frac{(r - 1)E_b + r(t - 1)E_e}{(rt - 1)E_e} \]

Where:
\( E_b = \) Block mean square/Replication mean square
\( E_e = \) Error mean square in the RCB analysis
\( r = \) number of replication or blocks
\( t = \) number of treatments
(ii) If the error df is less than 20, the R.E. value should be multiplied by an adjustment factor, $K$, defined as follows:

$$K = \frac{[(r - 1)(t - 1) + 1][t(r - 1) + 3]}{[(r - 1)(t - 1) + 3][t(r - 1) + 1]}$$

**Example 5:**
Considering Example 4, calculate the relative efficiency.

**Solution**

$r = 3$, $t = 4$, $E_b = 6.34$, $E_e = 19$

$$R.E = \frac{(3-1)6.34 + 3(4-1)19}{(3(4) - 1)19} = \frac{(2)6.34 + 3(3)19}{(11)19} = 12.68 + 171 = \frac{209}{209} = 0.88$$

Since the error df is less than 20, we calculate $K$.

$$K = \frac{[(3 - 1)(4 - 1) + 1][4(3 - 1) + 3]}{[(3 - 1)(4 - 1) + 3][4(3 - 1) + 1]} = \frac{[(2)(3) + 1][4(2) + 3]}{[(2)(3) + 3][4(2) + 1]} = \frac{77}{81} = 0.95$$

Adjusted R.E = $R.E \times K = 0.88 \times 0.95 = 0.84$

**Conclusion**
The result indicates that the use of the RCB design instead of a CRD decreased experimental precision by 16%.

**Note:** if for example, the adjusted R.E. is 1.65, this will be interpreted as “increased experimental precision by 65%”.

**Missing Value in the RCB Design**

Let $X$ represent the missing value. We estimate missing value in RCB design using:

$$X = \frac{rC' + tR' - G'}{(r - 1)(t - 1)}$$

Where:

$r =$ number of replicates

$t =$ number of treatments

$C' =$ column total where missing value occurs

$R' =$ row total where missing value occurs

$G' =$ overall total excluding the missing value
Example 6:

Given the data below, find the missing value.

<table>
<thead>
<tr>
<th></th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$R_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td>X</td>
<td>8.00</td>
<td>7.93</td>
</tr>
<tr>
<td>$T_2$</td>
<td>8.14</td>
<td>8.15</td>
<td>7.87</td>
</tr>
<tr>
<td>$T_3$</td>
<td>7.76</td>
<td>7.88</td>
<td>7.74</td>
</tr>
<tr>
<td>$T_4$</td>
<td>7.17</td>
<td>7.57</td>
<td>7.80</td>
</tr>
<tr>
<td>$T_5$</td>
<td>7.46</td>
<td>7.68</td>
<td>7.57</td>
</tr>
</tbody>
</table>

Solution:

$r = 3$, $t = 5$, $C' = 30.53$, $R' = 15.93$, $G' = 108.72$

$$X = \frac{rC' + tR' - G'}{(r - 1)(t - 1)}$$

$$= \frac{3(30.53) + 5(15.93) - 108.72}{(3 - 1)(5 - 1)}$$

$$= \frac{91.59 + 79.65 - 108.72}{(2)(4)}$$

$$= \frac{62.52}{8}$$

$$= 7.82$$
Latin Square (LS) Design

Features:

(i) \( V = C = R \)

Where:

\( R \) = Number of rows
\( C \) = Number of columns
\( V \) = Number of treatments

(ii) If the number of treatments is known, we have a \( V \) by \( V \) Latin Square.

(iii) Every treatment occurs once in each row and once in each column.

Advantages of Latin Square Design

(i) You can control variation in two directions.

(ii) Hopefully, you increase efficiency as compared to RCBD.

Disadvantages of Latin Square Design

(i) The number of treatments must equal the number of replicates.

(ii) The experimental error is likely to increase with the size of the square.

(iii) Small squares have very few degrees of freedom for experimental error.

(iv) You cannot evaluate interactions between:

(a) Rows and columns
(b) Rows and treatments
(c) Columns and treatments

Example of Latin Square Designs

(i) 3 × 3 Latin Square

```
A  B  C
B  C  A
C  A  B
```

(ii) 4 × 4 Latin Square

```
A  B  C  D
B  C  D  A
C  D  A  B
D  A  B  C
```
(iii) 3 × 3 Latin Square
\[ \begin{array}{ccc}
\alpha & \lambda & \beta \\
\beta & \alpha & \lambda \\
\lambda & \beta & \alpha \\
\end{array} \]

(iv) 3 × 3 Graeco-Latin Square
\[ \begin{array}{ccc}
A\alpha & B\lambda & C\beta \\
B\beta & C\alpha & A\lambda \\
C\lambda & A\beta & B\alpha \\
\end{array} \]

Note:
- (i) and (iii) are orthogonal.
- (iv) is Graeco-Latin Square (a result of super-imposing (iii) on (i))

**Analysis of Latin Square Design**

Analysis proceeds along the same lines as for randomized blocks. But, instead of the one sum of squares for blocks, systematic variation is now taken out by two sums of squares which are always called **row sum of squares** and **column sum of squares**.

**Model:**

\[ y_{ijk} = \mu + \tau_i + \eta_j + c_k + e_{ijk}; i = 1, 2, ..., v; j = 1, 2, ..., v; k = 1, 2, ..., v \]

- \( \mu \) is the overall mean
- \( \tau_i \) is the \( i \)th treatment effect
- \( \eta_j \) is the \( j \)th row effect
- \( c_k \) is the \( k \)th column effect
- \( e_{ijk} \) is the random error term
- \( y_{ijk} \) is the response

**Example 7:**
A farmer wishes to test the effects of four different fertilizers A, B, C, D on the yields of wheat. In order to eliminate sources of error due to variability in soil fertility, he uses the fertilizer in a Latin Square arrangement as shown below. Perform an analysis of variance to show if there is a significant difference at \( \alpha = 5\% \).

\[
\begin{array}{cccc}
A_{18} & C_{21} & D_{25} & B_{11} \\
D_{22} & B_{12} & A_{15} & C_{19} \\
B_{15} & A_{20} & C_{23} & D_{24} \\
C_{22} & D_{21} & B_{10} & A_{17} \\
\end{array}
\]
Solution:

<table>
<thead>
<tr>
<th></th>
<th>A_{18}</th>
<th>C_{21}</th>
<th>D_{25}</th>
<th>B_{11}</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_{15}</td>
<td>A_{15}</td>
<td>C_{19}</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_{22}</td>
<td>D_{21}</td>
<td>B_{10}</td>
<td>A_{17}</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>74</td>
<td>73</td>
<td>71</td>
<td>295</td>
<td></td>
</tr>
</tbody>
</table>

Treatments:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>48</td>
<td>85</td>
<td>92</td>
<td>295</td>
</tr>
</tbody>
</table>

(i) \[ SST = 18^2 + 21^2 + \cdots + 10^2 + 17^2 - \frac{(295)^2}{16} \]
\[ = 5769 - 5439.06 = 329.94 \]

(ii) \[ SS_{\text{row}} = \frac{75^2 + 68^2 + 82^2 + 70^2}{4} - \frac{(295)^2}{16} \]
\[ = 5468.25 - 5439.06 = 29.19 \]

(iii) \[ SS_{\text{col}} = \frac{77^2 + 74^2 + 73^2 + 71^2}{4} - \frac{(295)^2}{16} \]
\[ = 5443.75 - 5439.06 = 4.69 \]

(iv) \[ SS_t = \frac{70^2 + 48^2 + 85^2 + 92^2}{4} - \frac{(295)^2}{16} \]
\[ = 5723.25 - 5439.06 = 284.19 \]

(v) \[ SS_e = SST - [SS_{\text{row}} + SS_{\text{col}} + SS_t] \]
\[ = 329.94 - 318.07 = 11.87 \]

<table>
<thead>
<tr>
<th>SV</th>
<th>SS</th>
<th>Df</th>
<th>MS</th>
<th>Fcal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rows</td>
<td>29.19</td>
<td>3</td>
<td>9.73</td>
<td>4.91</td>
</tr>
<tr>
<td>Columns</td>
<td>4.69</td>
<td>3</td>
<td>1.56</td>
<td>0.79</td>
</tr>
<tr>
<td>Treatment</td>
<td>284.19</td>
<td>3</td>
<td>94.73</td>
<td>47.84</td>
</tr>
<tr>
<td>Error</td>
<td>11.87</td>
<td>6</td>
<td>1.98</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>329.94</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ F_{\text{tab}} = F_{0.05,3,6} = 4.76 \]

Decision and Conclusion:
- Since \( F_{\text{cal}} \) for rows > \( F_{\text{tab}} \), we reject \( H_0 \) and conclude that row effect is significant.
• Since $F_{cal}$ for columns $< F_{tab}$, we fail to reject $H_0$ and conclude that column effect is not significant.
• Since $F_{cal}$ for treatment $> F_{tab}$, we reject $H_0$ and conclude that treatment effect is significant.

Relative Efficiency for Latin Square Design
The relative efficiency of LS design as compared to a randomized complete block design (RCBD) can be computed in two ways:
(i) When rows are considered as blocks of the RCB design.
(ii) When columns are considered as blocks of the RCB design.
These two relative efficiencies are computed as follows:
\[
\text{R. E}(\text{RCB, row}) = \frac{E_r + (t-1)E_e}{t(E_e)}
\]
\[
\text{R. E}(\text{RCB, col}) = \frac{E_c + (t-1)E_e}{t(E_e)}
\]

Where:
$E_e$ is the mean square for error
$E_r$ is the mean square for row
$E_c$ is the mean square for column

Note:
When the error df in the Latin Square ANOVA is less than 20, the R.E value should be multiplied by an adjustment factor, $K$, defined as follows:
\[
K = \frac{[(t-1)(t-2)+1] + [t(r-1)^2+3]}{[(t-1)(t-2)+3] + [t(r-1)^2+1]}
\]

Example 8:
Considering Example 7, calculate the relative efficiency:
(i) When rows are blocks
(ii) When columns are blocks

Solution
$t = 4, r = 4, E_r = 9.73, E_c = 1.56, E_e = 1.98$

(i)
\[
\text{R. E}(\text{RCB, row}) = \frac{E_r + (t-1)E_e}{t(E_e)} = \frac{9.73 + (4-1)1.98}{4(1.98)} = \frac{9.73 + 5.94}{7.92} = 1.98
\]
Since the error df is less than 20, we find the adjustment factor, \( K \):

\[
K = \frac{[(t-1)(t-2) + 1] + [t(t-1)^2 + 3]}{[(t-1)(t-2) + 3] + [t(t-1)^2 + 1]} = \frac{[(4-1)(4-2) + 1] + [4(4-1)^2 + 3]}{[(4-1)(4-2) + 3] + [4(4-1)^2 + 1]}
\]

\[
= \frac{[6 + 1][36 + 3]}{6 + 3][36 + 1]} = \frac{7(39)}{9(37)} = \frac{273}{333} = 0.82
\]

(i) Adjusted \( R.E (RCB, \text{row}) = R.E \times K = 1.98 \times 0.82 = 1.62 \)

**Conclusion:**
The additional row blocking made possible by the Latin Square design increased the precision over RCBD by 62%.

(ii) Adjusted \( R.E (RCB, \text{col}) = R.E \times K = 0.95 \times 0.82 = 0.78 \)

**Conclusion:**
The additional column blocking made possible by the Latin Square design decreased the precision over RCBD by 22%.

**Missing Observation in Latin Square Design**

Let \( X \) represent the missing observation:

\[
X = \frac{P(R_{i}'' + C_{j}'' + T_{k}'' - 2G'')}{(P - 1)(P - 2)}
\]

Where:

- \( P \) = number of squares
- \( R_{i}'' \) = row total containing the missing observation
- \( C_{j}'' \) = column total containing the missing observation
- \( T_{k}'' \) = treatment total containing the missing observation
- \( G \) = Grand total
- \( G' \) = Grand total excluding missing observation