Analysis of Variance and Design of Experiments-I

MODULE – IV

LECTURE - 24

EXPERIMENTAL DESIGNS AND THEIR ANALYSIS

Dr. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur
### Missing observations in RBD

#### One missing observation:

Suppose one observation in \((i, j)^{th}\) cell is missing and let this be \(x\).

The arrangement of observations in RBD then will look like as follows:

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>(y_{11})</td>
</tr>
<tr>
<td>2</td>
<td>(y_{12})</td>
</tr>
<tr>
<td>(j)</td>
<td>(y_{ij})</td>
</tr>
<tr>
<td>(v)</td>
<td>(y_{1v})</td>
</tr>
</tbody>
</table>

| Treatment totals | \(T_1 = y_{1o}\) | \(T_2 = y_{2o}\) | \(T_i = y_{io} + x\) | \(y_{vo}\) | Grand Total | \(G = y_{o\times o} + x\) |

where \(y_{o\times o}\) : total of known observations.
\(y_{ij}\) : total of known observations in \(j^{th}\) block
\(y_{io}\) : total of known observations in \(i^{th}\) treatment.
Correction factor \((CF) = \frac{(G')^2}{n} = \frac{(y'_{oo} + x)^2}{bv}\)

\[TSS = \sum_{i=1}^{b} \sum_{j=1}^{v} y'_{ij}^2 - CF\]

\[= (x^2 + \text{terms which are constant with respect to } x) - CF\]

\[SSBl = \frac{1}{b}[(y'_{io} + x)^2 + \text{terms which are constant with respect to } x] - CF\]

\[SSTr = \frac{1}{v}[(y'_{oj} + x)^2 + \text{terms which are constant with respect to } x] - CF\]

\[SSE = TSS - SSBl - SSTr\]

\[= x^2 - \frac{1}{b}(y'_{io} + x)^2 - \frac{1}{v}(y'_{oj} + x)^2 + \frac{(y'_{oo} + x)^2}{bv} + \text{(terms which are constant with respect to } x) - CF.\]

Find \(x\) such that \(SSE\) is minimum

\[
\frac{\partial (SSE)}{\partial x} = 0 \Rightarrow 2x - \frac{2(y'_{io} + x)}{b} - \frac{2(y'_{oj} + x)}{v} + \frac{2(y'_{oo} + x)}{bv} = 0
\]

or \(x = \frac{vy'_{io} + by'_{oj} - y'_{oo}}{(b-1)(v-1)}\).
Two missing observations

If there are two missing observation, then let they be $x$ and $y$.

- Let the corresponding row sums (block totals) are
- Column sums (treatment totals) are
- Total of known observations is $S$.

Then

$$SSE = x^2 + y^2 - \frac{1}{b}[(R_1 + x)^2 + (R_2 + y)^2] - \frac{1}{v}[(C_1 + x)^2 + (C_2 + y)^2] + \frac{1}{bv}(S + x + y)^2 + \text{terms independent of } x \text{ and } y.$$ 

Now differentiate $SSE$ with respect to $x$ and $y$, as

$$\frac{\partial (SSE)}{\partial x} = 0 \Rightarrow x - \frac{R_1 + x}{b} - \frac{C_1 + x}{b} + \frac{S + x + y}{bv} = 0$$

$$\frac{\partial (SSE)}{\partial y} = 0 \Rightarrow y - \frac{R_2 + y}{v} - \frac{C_2 + y}{v} + \frac{S + x + y}{bv} = 0.$$ 

Thus solving the following two linear equations in $x$ and $y$, we obtain the estimated missing values

$$(b - 1)(v - 1)x = bR_1 + vC_1 - S - y$$

$$(b - 1)(v - 1)y = bR_2 + vC_2 - S - x.$$
Adjustments to be done in analysis of variance

i. Obtain the within block sum of squares from incomplete data.

ii. Subtract correct error sum of squares from (i). This given the correct treatment sum of squares.

iii. Reduce the degrees of freedom of error sum of squares by the number of missing observations.

iv. No adjustments in other sum of squares are required.
**Missing observations in LSD**

Let

- $x$ be the missing observation in $(i, j, k)^{th}$ cell, i.e. $y_{ijk}, i = 1, 2, \ldots, v, j = 1, 2, \ldots, v, k = 1, 2, \ldots, v$.

- $R$: Total of known observations in $i^{th}$ row.

- $C$: Total of known observations in $j^{th}$ column.

- $T$: Total of known observation receiving the $k^{th}$ treatment.

- $S$: Total of known observations.

Now

\[
\text{Correction factor (CF)} = \frac{(S + x)^2}{v^2}
\]

\[
\text{Total sum of squares (TSS)} = x^2 + \text{term which are constant with respect to } x - CF
\]

\[
\text{Row sum of squares (SSR)} = \frac{(R + x)^2}{v} + \text{term which are constant with respect to } x - CF
\]

\[
\text{Column sum of squares (SSC)} = \frac{(C + x)^2}{v} + \text{term which are constant with respect to } x - CF
\]

\[
\text{Treatment sum of squares (SSTr)} = \frac{(T + x)^2}{v} + \text{term which are constant with respect to } x - CF
\]

\[
\text{Sum of squares due to error (SSE)} = \text{TSS} - \text{SSR} - \text{SSC} - \text{SSTr}
\]

\[
= x^2 - \frac{1}{v} \left[ (R + x)^2 + (C + x)^2 + (T + x)^2 \right] + \frac{2(S + x)^2}{v^2}.
\]
Choose $x$ such that $SSE$ is minimum.

So \[
\frac{d(SSE)}{dx} = 0
\]

\[2x - \frac{2}{v}(R + C + T + 3x) + \frac{4(S + x)}{v^2} \]

or

\[x = \frac{V(R + C + T) - 2S}{(v-1)(v-2)}.\]

**Adjustment to be done in analysis of variance**

Do all the steps as in the case of RBD.

To get the correct treatment sum of squares, proceed as follows:

- Ignore the treatment classification and consider only row and classification.
- Substitute the estimated values at the place of missing observation.
- Obtain the error sum of squares from complete data, say $SSE_1$.
- Let $SSE_2$ be the error sum of squares based on LSD obtained earlier.
- Find corrected treatment sum of squares $= SSE_2 - SSE_1$.
- Reduce of degrees of freedom of error sum of squares by the number of missing values.