Analysis of Variance and Design of Experiments

Incomplete Block Designs and Their Analysis

:::

Lecture 23

Basics and Estimation of Parameters

Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur

Slides can be downloaded from http://home.iitk.ac.in/~shalab/sp

Notations and normal equations:

Let

- **v treatments have to be compared**
- **b blocks are available**
- **k_i: Number of plots in i^{th} block (i = 1,…,b)**
- **r_j: Number of plots receiving j^{th} treatment (j = 1,…,v)**
- **n: total number of plots**

$$n = r_1 + r_2 + ... + r_v = k_1 + k_2 + ... + k_b$$

- **Each treatment may occur more than once in each block**

or

may not occur at all.

Notations and normal equations:

n_{ij} **: Number of times the *j*th treatment occurs in *i*th block.**

Similarly, $n_{ij} = 1$ **means that** j^{th} **treatment occurs in** i^{th} **block and**

$n_{ij} = 0$ **means that** j^{th} **treatment does not occurs in the** i^{th} **block.**

$$\sum_{j=1}^{v} n_{ij} = k_i \qquad i = 1, \ldots, b$$

$$\sum_{i} n_{ij} = r_j \qquad j = 1, \ldots, v$$

$$n = \sum_{i} \sum_{j} n_{ij}$$

Model:

Let y_{ijm} : denotes the response (yield) from the *m*th replicate of *j*th treatment in *i*th block and

$$y_{ijm} = \beta_i + \tau_j + \varepsilon_{ijm} \quad i = 1,...,b,\ j = 1,..,v,\ m = 1,...,n_{ij}$$

[Note: We are not considering here the general mean effect in this model for better understanding of the issues in the estimation of parameters. Later, we will consider it in the analysis.]

Model:

Following notations are used in further description.

Block totals : $B_1, B_2, ..., B_b$ **where** $B_i = \sum_j \sum_m y_{ijm}$

Treatment totals: $V_1, V_2, ..., V_v$ **where** $V_j = \sum_i \sum_m y_{ijm}$

Grand total : $Y = \sum_i \sum_j \sum_m y_{ijm}$

Generally, a design is denoted by $D(v, b, r, k, n)$ **where** ***v*, *b*, *r*, *k***

and *n* are the parameters of the design.

Model: Example

Let us consider an example to understand the meaning of these notations. Suppose there are 3 blocks (Block 1, Block 2 and Block 3) and 5 treatments (τ_1 , τ_2 , τ_3 , τ_4 , τ_5).

So b = 3 and v = 5.

These treatments are arranged in different plots in blocks as follows:

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

k_i

Number of plots in Block 1: $k_1 = 5$

Number of plots in Block 2: $k_2 = 4$

Number of plots in Block 1: $k_3 = 2$

r_1

Number of times τ_1 appears in

- Block 1 = 2
- Block 2 = 0
- Block 3 = 0.

Total number of times τ_1 appears in the entire design

is $r_1 = 2 + 0 + 0 = 2.$

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

r_2

Number of times τ_2 appears in

- **Block 1 = 2**
- **Block 2 = 1**
- **Block 3 = 2.**

Total number of times τ_2 appears in the entire design

is $r_2 = 2 + 1 + 2 = 5.$

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

r_3

Number of times τ_3 appears in

- **Block 1 = 1**
- **Block 2 = 0**
- **Block 3 = 0.**

Total number of times τ_3 appears in the entire design

is $r_3 = 1 + 0 + 0 = 1$.

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

r_4

Number of times τ_4 appears in

- **Block 1 = 0**
- **Block 2 = 1**
- **Block 3 = 0.**

Total number of times τ_4 appears in the entire design

is $r_4 = 0 + 1 + 0 = 1.$

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

r_5

Number of times τ_5 appears in

- Block 1 = 0
- Block 2 = 2
- Block 3 = 0.

Total number of times τ_5 appears in the entire design

is $r_5 = 0 + 2 + 0 = 2.$

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

n_{1j}

Total number of times τ_1 appears in Block 1: $n_{11} = 2$

Total number of times τ_2 appears in Block 1: $n_{12} = 2$

Total number of times τ_3 appears in Block 1: $n_{13} = 1$

Total number of times τ_4 appears in Block 1: $n_{14} = 0$

Total number of times τ_5 appears in Block 1: $n_{15} = 0$

Model: Example

Block 1: 5 plots	Plot 1 τ_1	Plot 2 τ_1	Plot 3 τ_2	Plot 4 τ_2	Plot 5 τ_3
Block 2: 4 plots	Plot 1 τ_2	Plot 2 τ_4	Plot 3 τ_5	Plot 4 τ_5	
Block 3: 2 plots	Plot 1 τ_2	Plot 2 τ_2			

n_{2j}

Total number of times τ_1 appears in Block 2: $n_{21} = 0$

Total number of times τ_2 appears in Block 2: $n_{22} = 1$

Total number of times τ_3 appears in Block 2: $n_{23} = 0$

Total number of times τ_4 appears in Block 2: $n_{24} = 1$

Total number of times τ_5 appears in Block 2: $n_{25} = 2$

Model: Example

	Plot 1	Plot 2	Plot 3	Plot 4	Plot 5
Block 1: 5 plots	τ_1	τ_1	τ_2	τ_2	τ_3
Block 2: 4 plots	τ_2	τ_4	τ_5	τ_5	
Block 3: 2 plots	τ_2	τ_2			

n_{3j}

Total number of times τ_1 appears in Block 3: $n_{31} = 0$

Total number of times τ_2 appears in Block 3: $n_{32} = 2$

Total number of times τ_3 appears in Block 3: $n_{33} = 0$

Total number of times τ_4 appears in Block 3: $n_{34} = 0$

Total number of times τ_5 appears in Block 3: $n_{35} = 0$

Model:

Example: y_{ijm}

y_{ijm} **: response from the** m^{th} **replicate of** j^{th} **treatment in** i^{th} **block,** i **= 1,2,3;** j **= 1,2,3,4,5;** m **= 1,2,…,** n_{ij}

Following are the notations for for y_{ijm} **different treatments in the blocks**

Block 1:	τ_1 y_{111}	τ_1 y_{112}	τ_2 y_{121}	τ_2 y_{122}	τ_3 y_{131}
Block 2:	τ_2 y_{221}	τ_4 y_{241}	τ_5 y_{251}	τ_5 y_{252}	
Block 3:	τ_2 y_{321}	τ_2 y_{322}			

Model: Example

Block 1:	τ_1 y_{111}	τ_1 y_{112}	τ_2 y_{121}	τ_2 y_{122}	τ_3 y_{131}
Block 2:	τ_2 y_{221}	τ_4 y_{241}	τ_5 y_{251}	τ_5 y_{252}	
Block 3:	τ_2 y_{321}	τ_2 y_{322}			

B_i	V_j
$B_1 = y_{111} + y_{112} + y_{121} + y_{122} + y_{131}$	$V_1 = y_{111} + y_{112}$
$B_2 = y_{221} + y_{241} + y_{251} + y_{252}$	$V_2 = y_{121} + y_{122} + y_{221} + y_{321} + y_{322}$
$B_3 = y_{321} + y_{322}$	$V_3 = y_{131}$
	$V_4 = y_{241}$
	$V_5 = y_{251} + y_{252}$

Normal equations:

Minimizing $S = \sum_i \sum_j \sum_m \varepsilon_{ijm}^2$ **with respect to** β_i **and** τ_j**, we obtain the least-squares estimators of the parameters as follows:**

$$S = \sum_i \sum_j \sum_m (y_{ijm} - \beta_i - \tau_j)^2$$

$$\frac{\partial S}{\partial \beta_i} = 0 \Rightarrow \sum_j \sum_m (y_{ijm} - \beta_i - \tau_j) = 0$$

or $$B_i - \beta_i \sum_j \sum_m 1 - \sum_j \tau_j \sum_m 1 = 0 \qquad \textbf{(1)}$$

or $$B_i = \beta_i k_i + n_{i1}\tau_1 + n_{i2}\tau_2 + ... + n_{iv}\tau_v, \quad i = 1, ..., b$$

$$B_i = \beta_i k_i + \sum_j \tau_j n_{ij} \qquad [b \textbf{ equations}]$$

Normal equations:

$$\frac{\partial S}{\partial \tau_j}=0 \Rightarrow \sum_i \sum_m (y_{ijm}-\beta_i-\tau_j)=0$$

or $$\sum_i \sum_m y_{ijm} - \sum_i \beta_i \sum_m 1 - \tau_j \sum_i \sum_m 1 = 0$$

$$V_j - \sum_i \beta_i n_{ij} - \tau_j \sum_i n_{ij} = 0 \qquad \textbf{(2)}$$

or $$V_j = n_{ij}\beta_1 + n_{2j}\beta_2 + \ldots + n_{bj}\beta_b + r_j\tau_j, \quad j=1,2,\ldots,v$$

or $$V_j = \sum_i \beta_i n_{ij} + r_j \tau_j \qquad [v \textbf{ equations}]$$

Equations (1) and (2) constitute (*b* + *v*) equations.

Normal equations:

Note that $\sum_i$ **equation** $(1) = \sum_j$ **equation** (2)

$$\sum_i B_i = \sum_j V_j$$

$$\sum_i \left(\sum_j \sum_m y_{ijm} \right) = \sum_j \left(\sum_i \sum_m y_{ijm} \right).$$

Thus there are at most (*b* + *v* - 1) degrees of freedom for estimates. So the estimates of only (*b* + *v* - 1) parameters can be obtained out of all (*b* + *v*) parameters.

[Note: We will see later that degrees of freedom may be less than or equal to (*b* + *v* - 1) in special cases. Also, note that we have not assumed any side conditions like $\sum_i \alpha_i = \sum_j \beta_j = 0$ **as in the case of complete block designs.]**

Estimation of parameters:

To obtain the estimates of the parameters, there are two options-

1. Using equation (1), eliminate β_i from equation (2) to estimate τ_j or

2. Using equation (2), eliminate τ_j from equation (1) to estimate β_i .

Estimation of parameters and treatment totals:

We consider first the approach 1., i.e., using equation (1), eliminate β_i from equation (2).

From equation (1),

$$\beta_i = \frac{1}{k_i}\left[B_i - \sum_{j=1}^{v} n_{ij}\tau_j \right].$$

Use it in (2) as follows.

$$\begin{aligned} V_j &= n_{1j}\beta_1 + \ldots + n_{bj}\beta_b + r_j\tau_j \\ &= n_{1j}\left[\frac{1}{k_1}(B_1 - n_{11}\tau_1 - \ldots - n_{1v}\tau_{1v}) \right] + n_{2j}\left[\frac{1}{k_2}(B_2 - n_{21}\tau_1 - \ldots - n_{2v}\tau_v) \right] + \ldots \\ &\quad + n_{bj}\left[\frac{1}{k_b}(B_b - n_{b1}\tau_1 - \ldots - n_{bv}\tau_v) \right] + r_j\tau_j \end{aligned}$$

Adjusted treatment totals:

$$= n_{1j}\left[\frac{1}{k_1}(B_1 - n_{11}\tau_1 - \ldots - n_{1v}\tau_{1v})\right] + n_{2j}\left[\frac{1}{k_2}(B_2 - n_{21}\tau_1 - \ldots - n_{2v}\tau_v)\right] + \ldots$$

$$+ n_{bj}\left[\frac{1}{k_b}(B_b - n_{b1}\tau_1 - \ldots - n_{bv}\tau_v)\right] + r_j\tau_j$$

or $$V_j - \frac{n_{1j}B_1}{k_1} - \frac{n_{2j}B_2}{k_2} - \ldots - \frac{n_{bj}B_b}{k_b} = \tau_1\left[-\frac{n_{11}n_{1j}}{k_1} - \frac{n_{21}n_{2j}}{k_2} - \ldots. - \frac{n_{b1}n_{bj}}{k_b}\right] + \ldots$$

$$+ \tau_v\left[-\frac{n_{1v}n_{1j}}{k_1} - \frac{n_{2v}n_{2j}}{k_2} - \ldots. - \frac{n_{bv}n_{bj}}{k_b}\right] + r_j\tau_j, \quad j = 1, \ldots, v$$

or

$$V_j - \sum_{i=1}^{b}\frac{n_{ij}B_i}{k_i} = \tau_1\left[-\frac{n_{11}n_{1j}}{k_1}\ldots - \frac{n_{b1}n_{bj}}{k_b}\right] + \ldots + \tau_v\left[-\frac{n_{1v}n_{1j}}{k_1}\ldots - \frac{n_{bv}n_{bj}}{k_b}\right] + r_j\tau_j$$

or

$$Q_j = \tau_1\left[-\frac{n_{11}n_{1j}}{k_1}\ldots - \frac{n_{b1}n_{bj}}{k_b}\right] + \ldots + \tau_v\left[-\frac{n_{1v}n_{1j}}{k_1}\ldots - \frac{n_{bv}n_{bj}}{k_b}\right] + r_j\tau_j, \quad j = 1\ldots.v$$

Adjusted treatment totals:

where $Q_j = V_j - \left[\frac{n_{1j} B_1}{k_1} + \ldots + \frac{n_{bj} B_b}{k_b} \right], \; j = 1, 2, \ldots, v$

are called adjusted treatment totals.

[Note: Compared to the earlier case, the j^{th} treatment total V_j is adjusted by a factor $\sum_{i=1}^{b} \frac{n_{ij} B_i}{k_i}$, that is why it is called "adjusted".

The adjustment is being made for the block effects because they were eliminated to estimate the treatment effects.]

Adjusted treatment totals:

Note that

k_i : **Number of plots in *i*th block.**

$\frac{B_i}{k_i}$: **is called the average (response) yield per plot from *i*th block.**

$\frac{n_{ij} B_i}{k_i}$: **is considered as an average contribution to the *j*th treatment total from the *i*th block.**

Q_j **is obtained by removing the sum of the average contributions of the *b* blocks from the *j*th treatment total** V_j **.**