

Analysis of Variance and Design of Experiments-II

MODULE I

LECTURE - 2

INCOMPLETE BLOCK DESIGNS

Dr. Shalabh

Department of Mathematics & Statistics
Indian Institute of Technology Kanpur

In order 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 .

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:

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

or

$$V_j - \frac{n_{1j} B_1}{k_1} - \frac{n_{2j} B_2}{k_2} - \dots - \frac{n_{bj} B_b}{k_b} = \tau_1 \left[-\frac{n_{11} n_{1j}}{k_1} - \frac{n_{21} n_{2j}}{k_2} - \dots - \frac{n_{b1} n_{bj}}{k_b} \right] + \dots + \tau_v \left[-\frac{n_{1v} n_{1j}}{k_1} - \frac{n_{2v} n_{2j}}{k_2} - \dots - \frac{n_{bv} n_{bj}}{k_b} \right] + r_j \tau_j, \quad j = 1, \dots, v$$

or

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

or

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

where

$$Q_j = V_j - \left[\frac{n_{1j}B_1}{k_1} + \dots + \frac{n_{bj}B_b}{k_b} \right], \quad j = 1, 2, \dots, v \quad \text{are called adjusted treatment totals.}$$

[**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.]

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 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 .

Write

$$Q_j = \tau_1 \left[-\frac{n_{11}n_{1j}}{k_1} \dots - \frac{n_{b1}n_{bj}}{k_b} \right] + \dots + \tau_v \left[-\frac{n_{1v}n_{1j}}{k_1} \dots - \frac{n_{bv}n_{bj}}{k_b} \right] + r_j \tau_j, \quad j = 1, 2, \dots, v.$$

as

$$Q_j = C_{j1}\tau_1 + C_{j2}\tau_2 + \dots + C_{jv}\tau_v$$

where

$$C_{jj} = r_j - \frac{n_{1j}^2}{k_1} - \frac{n_{2j}^2}{k_2} - \dots - \frac{n_{bj}^2}{k_b}$$

$$C_{jj'} = r_j - \frac{n_{1j}n_{1j'}}{k_1} - \frac{n_{2j}n_{2j'}}{k_2} - \dots - \frac{n_{bj}n_{bj'}}{k_b}; \quad j \neq j', \quad j = 1, 2, \dots, v.$$

The $v \times v$ matrix $C = ((C_{jj'}))$, $j = 1, 2, \dots, v$; $j' = 1, 2, \dots, v$ with C_{jj} as diagonal elements and $C_{jj'}$ as off-diagonal elements is called the **C - matrix** of the incomplete block design.

C matrix is symmetric. Its row sum and column sum are zero. (proved later).

Rewrite

$$Q_j = \tau_1 \left[-\frac{n_{11}n_{1j}}{k_1} \dots - \frac{n_{b1}n_{bj}}{k_b} \right] + \dots + \tau_v \left[-\frac{n_{1v}n_{1j}}{k_1} \dots - \frac{n_{bv}n_{bj}}{k_b} \right] + r_j \tau_j, \quad j = 1, 2, \dots, v.$$

as

$$Q = C\tau.$$

This equation is called as **reduced normal equations** where

$$Q' = (Q_1, Q_2, \dots, Q_v),$$

$$\tau' = (\tau_1, \tau_2, \dots, \tau_v).$$

Equations (1) and (2) are EQUIVALENT.

Alternative presentation in matrix notations

Now let us try to represent and translate the same algebra in matrix notations.

Let $E_{mn} : m \times n$ matrix whose all elements are unity.

$N = (n_{ij})$ is $b \times v$ matrix called as incidence matrix.

$$k_i = \sum_{j=1}^v n_{ij}, \quad r_j = \sum_{i=1}^b n_{ij}, \quad n = \sum_i \sum_j n_{ij}$$

$$E_{1b}N = (r_1, r_2, \dots, r_v) = r'$$

$$NE_{v1} = (k_1, k_2, \dots, k_b)' = k.$$

For illustration, we verify one of the relationships as follows.

$$E_{1b} = (1, 1, \dots, 1)_{1 \times b}$$

$$N = \begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1v} \\ n_{21} & n_{22} & \cdots & n_{2v} \\ \vdots & \vdots & \ddots & \vdots \\ n_{b1} & n_{b2} & \cdots & n_{bv} \end{pmatrix}$$

$$E_{1b}N = (1, 1, \dots, 1)_{1 \times b} \begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1v} \\ n_{21} & n_{22} & \cdots & n_{2v} \\ \vdots & \vdots & \ddots & \vdots \\ n_{b1} & n_{b2} & \cdots & n_{bv} \end{pmatrix}_{b \times v} = \left(\sum_{i=1}^b n_{i1}, \sum_{i=1}^b n_{i2}, \dots, \sum_{i=1}^b n_{iv} \right)$$

$$= (r_1, r_2, \dots, r_v).$$

It is now clear that the treatment and blocks are not estimable as such as in the case of complete block designs. Note that

we have not made any assumption like $\sum_i \alpha_i = \sum_j \beta_j = 0$ also.

Now we introduce the general mean effect (denoted by μ) in the linear model and carry out the further analysis on the same lines as earlier.

Consider the model $y_{ijm} = \mu + \beta_i + \tau_j + \varepsilon_{ijm}$, $i = 1, 2, \dots, b$; $j = 1, 2, \dots, v$; $m = 0, 1, \dots, n_{ij}$.

The normal equations are obtained by minimizing $S = \sum_i \sum_j \sum_m \varepsilon_{ijm}^2$ with respect to the parameters μ, β_i and τ_j

and solving them, we can obtain the least squares estimators of the parameters.

Minimizing $S = \sum_i \sum_j \sum_m \varepsilon_{ijm}^2$ with respect to the parameters μ, β_i and τ_j , the normal equations are obtained as

$$n\hat{\mu} + \sum_i n_{i0}\hat{\beta}_i + \sum_j n_{0j}\hat{\tau}_j = G$$

$$n_i\hat{\mu} + n_{i0}\hat{\beta}_i + \sum_j n_{ij}\hat{\tau}_j = B_i \quad i = 1, \dots, b$$

$$n_{0j}\hat{\mu} + n_{0j}\hat{\tau}_j + \sum_i n_{ij}\hat{\beta}_i = V_j \quad j = 1, \dots, v.$$

Now we write these normal equations in matrix notations.

Denote

$$\beta = Col(\beta_1, \beta_2, \dots, \beta_b),$$

$$\tau = Col(\tau_1, \tau_2, \dots, \tau_v)$$

$$B = Col(B_1, B_2, \dots, B_b),$$

$$V = Col(V_1, V_2, \dots, V_v)$$

$$N = ((n_{ij})) : \text{incidence matrix of order } b \times v$$

where $Col(\cdot)$ denotes the column vector.

Let

$$K = diag(k_1, \dots, k_b) : b \times b \text{ diagonal matrix}$$

$$R = diag(r_1, \dots, r_v) : v \times v \text{ diagonal matrix.}$$

Then the $(b + v + 1)$ normal equations can be written as

$$\begin{pmatrix} G \\ B \\ V \end{pmatrix} = \begin{pmatrix} n & E_{1b}K & E_{1v}R \\ KE_{b1} & K & N \\ RE_{v1} & N' & R \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\tau} \end{pmatrix}. \quad (*)$$

Since we are presently interested in the testing of hypothesis related to the treatment effects, so we eliminate the block effects $\hat{\beta}$ to estimate the treatment effects. For doing so, multiply both sides on the left of equation (*) by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & I_b & -NR^{-1} \\ 0 & -N'K^{-1} & I_v \end{bmatrix}$$

where

$$R^{-1} = \text{diag}\left(\frac{1}{r_1}, \frac{1}{r_2}, \dots, \frac{1}{r_v}\right), \quad K^{-1} = \text{diag}\left(\frac{1}{k_1}, \frac{1}{k_2}, \dots, \frac{1}{k_b}\right).$$

Solving it further, we get 3 sets of equations as follows:

$$G = n\hat{\mu} + E_{1b}K\hat{\beta} + E_{1v}R\hat{\tau}$$

$$B - NR^{-1}V = [K - NR^{-1}N']\hat{\beta}$$

$$V - N'K^{-1}B = [R - N'K^{-1}N]\hat{\tau}$$

These are called as 'reduced normal equations' or 'reduced intrablock equations'.

The reduced normal equation in the treatment effects can be written as

$$Q = C\hat{\tau}$$

where

$$Q = V - N'K^{-1}B$$

$$C = R - N'K^{-1}N.$$

The vector Q is called as the vector of **adjusted treatment totals**, since it contains the treatment totals which are adjusted for the block effects and the matrix C is called as **C-matrix**.

The C matrix is symmetric and its row sums and column sums are zero.

To show that row sum is zero in C-matrix, we proceed as follows:

Row sum:

$$\begin{aligned} CE_{v1} &= RE_{v1} - N'K^{-1}NE_{v1} \\ &= (r_1, r_2, \dots, r_v)' - N'K^{-1}k^{-1} \\ &= (r_1, r_2, \dots, r_v)' - N'E_{b1} \\ &= r - r \\ &= 0 \end{aligned}$$

Similarly, the column sum can also be shown to be zero.

In order to obtain the reduced normal equation for treatment effects, we first estimated the block effects from one of the normal equation and substituted it into another normal equation related to the treatment effects. This way the adjusted treatment total vector Q (which is adjusted for block effects) is obtained.

Similarly, the reduced normal equations for the block effects can be found as follows. First we estimate the treatment effects from one of the normal equation and then it is substituted into another normal equation related to the block effects. Finally we get the **adjusted block totals** (adjusted for treatment totals).

So, similar to $Q = C\hat{\tau}$, we can obtain another equation which can be represented as

$$D\hat{\beta} = P$$

$$\text{where } D = \text{diag}(k_1, k_2, \dots, k_b) - N \text{diag}\left(\frac{1}{r_1}, \frac{1}{r_2}, \dots, \frac{1}{r_v}\right)N'$$

$$= K - NR^{-1}N'$$

$$P = B - N \text{diag}\left(\frac{1}{r_1}, \frac{1}{r_2}, \dots, \frac{1}{r_v}\right)V$$

$$= B - NR^{-1}V$$

$$\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_b)'$$

and P is the adjusted block totals which are obtained after removing the treatment effects.