Analysis of Variance and Design of Experiments-II

MODULE I

LECTURE - 3

INCOMPLETE BLOCK DESIGNS

Dr. Shalabh
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur

Analysis of variance table

Under the null hypothesis $H_0 : \tau = 0,$ the design is one way analysis of variance set up with blocks as classifications. In this setup, we have the following:

$$\text{Sum of squares due to blocks} = \sum_{i=1}^{b} \frac{B_i^2}{k_i} - \frac{G^2}{n}$$

$$= (B_1, B_2, ..., B_b)' \, diag\left(\frac{1}{k_1}, \frac{1}{k_2}, ..., \frac{1}{k_b}\right) \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_b \end{pmatrix} - \frac{G^2}{n}$$

$$= B'K^{-1}B - \frac{G^2}{n}$$

If y is the vector of all the observations, then

$$\text{Error sum of squares} (S_e) = \sum_i \sum_j \sum_m (y_{ijm} - \hat{\mu} - \hat{\beta}_i - \hat{\tau}_j)^2$$

$$= \sum_i \sum_j \sum_m y_{ijm} (y_{ijm} - \hat{\mu} - \hat{\beta}_i - \hat{\tau}_j) \quad \text{[Using normal equations, other terms will be zero]}$$

$$= \sum_i \sum_j \sum_m y_{ijm}^2 - \hat{\mu} G - \sum_j \hat{\tau}_j V_j - \sum_i \hat{\beta}_i B_i$$

$$= y'y - \hat{\mu} G - V'\hat{\tau} - B'\hat{\beta}.$$

Using original normal equations given by

$$B = KE_{b1}\hat{\mu} + K\hat{\beta} + N\hat{\tau},$$

we have

$$\hat{\beta} = K^{-1}B - E_{b1}\hat{\mu} - K^{-1}N\hat{\tau}.$$

Since $G = V'E_{1b} = B'E_{b1},$ substituting $\hat{\beta}$ in S_e gives

$$S_e = y'y - G\hat{\mu} - B'[K^{-1}B - E_{b1}\hat{\mu} - K^{-1}N\hat{\tau}] - V'\hat{\tau}$$

$$= y'y - G\hat{\mu} - B'[K^{-1}B - E_{b1}\hat{\mu} - K^{-1}N\hat{\tau}] - V'\hat{\tau}$$

$$= y'y - G\hat{\mu} - B'K^{-1}B + G\hat{\mu} + B'K^{-1}N\hat{\tau} - V'\hat{\tau}$$

$$= y'y - B'K^{-1}B + (B'K^{-1}N - V')\hat{\tau}$$

$$= \left(y'y - \frac{G^2}{n}\right) - \left(B'K^{-1}B - \frac{G^2}{n}\right) - (V - N'K^{-1}B)'\hat{\tau}$$

$$\underset{\substack{\downarrow \\ \text{Error SS}}}{S_e} \;=\; \underset{\substack{\downarrow \\ \text{Total SS}}}{\left(y'y - \frac{G^2}{n}\right)} - \underset{\substack{\downarrow \\ \text{Block SS} \\ \text{(unadjusted)}}}{\left(B'K^{-1}B - \frac{G^2}{n}\right)} - \underset{\substack{\downarrow \\ \text{Adjusted treatment SS} \\ \text{(adjusted for blocks)}}}{Q'\hat{\tau}}$$

The degrees of freedom associated with the different sum of squares are as follows:

Block SS (unadjusted)	: $b - 1$
Treatment SS (adjusted)	: $v - 1$
Error SS	: $n - b - v + 1$
Total SS	: $n - 1$

The adjusted treatment sum of squares and the sum of squares due to error are independently distributed and follow a chi-square distribution with $(v - 1)$ and $(n - b - v + 1)$ degrees of freedom, respectively.

Under H_0, $$\frac{Q'\hat{\tau}/(v-1)}{S_e/(n-b-v+1)} \sim F(v-1, n-b-v+1).$$

The analysis of variance table for $H_0 : \tau = 0$ is as follows:

Analysis of variance table

Source	Degrees of freedom	Sum of squares	Mean squares	*F* - value
Treatments	$v - 1$	$Q'\hat{\tau}$ (Adjusted)	$\frac{Q'\hat{\tau}}{v-1}$	$F = \frac{Q'\hat{\tau}/(v-1)}{S_e/(n-b-v+1)}$
Blocks	$b - 1$	$B'K^{-1}B - \frac{G^2}{n}$ (Unadjusted)		
Error	$(n - b - v + 1)$	S_e	$\frac{S_e}{n-b-v+1}$	
Total	$n - 1$	$y'y - \frac{G^2}{n}$		

Thus in an incomplete block design, it matters whether we are estimating the block effects first and then the treatment effects are estimated

or

first estimate the treatment effects and then the block effects are estimated.

In complete block design, it doesn't matter at all. So the test of hypothesis related to the block and treatment effects can be carried out at the same time.

A reason for this is as follows: In an incomplete block design, either the

- Adjusted sum of squares due to treatments, unadjusted sum of squares due to blocks and corresponding sum of squares due to errors are orthogonal

or

- Adjusted sum of squares due to blocks, unadjusted sum of squares due to treatments and corresponding sum of squares due to errors are orthogonal .

Note that the adjusted sum of squares due to treatment and the adjusted sum of squares due to blocks are not orthogonal. So

either

Error S.S = *Total SS*– *SS block* (Unadjusted) – *SS treat* (Adjusted)

holds true

or

Error S.S = *Total SS* – *SS block* (Adjusted) – *SS treat* (Unadjusted)

holds true due to **Fisher Cochran theorem**.

Since $CE_{v1} = 0,$ so *C* is a nonsingular matrix. Also, since

$$Q'E_{v1} = V'E_{v1} - (N'K^{-1}B)'E_{v1}$$

$$= (V_1, ..., V_v)E_{v1} - B'K^{-1}NE_{v1}$$

$$= (\sum_i V_i) - B'K^{-1}k'$$

$$= \sum_i V_i - (B_1, ..., B_b)diag\left(\frac{1}{k_1}, \frac{1}{k_2}, ..., \frac{1}{k_b}\right)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_b \end{pmatrix}$$

$$= \sum_i V_i - (B_1.....B_b)E_{b1}$$

$$= \sum_i V_i - \sum_j B_j$$

$$= G - G$$

$$= 0$$

so the intrablock equations are consistent.

We will confine our attention to those designs for which *rank*(*C*) = *v* - 1. These are called **connected designs** and for which all contrasts in the treatments, i.e., all linear combinations $l'\tau$ where $l'E_{v1} = 0$ have unique least squares solutions.

This we prove now as follows.

Let G^* and H^* be any two generalized inverses of C by which we mean that they are square matrix of order v such that G^*Q and H^*Q are both the solution vectors to the intrablock equation, i.e., $\hat{\tau} = G^*Q$ and $\hat{\tau} = H^*Q$.

Then

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

$$\Rightarrow Q = CG^*Q$$

and $Q = CH^*Q$ for all Q

so that $C(G^* - H^*)Q = 0$.

It follows that $(G^* - H^*)Q$ can be written as a^*E_{v1} where a is any scalar which may be zero.

Let ℓ be a vector such that $\ell' E_{v1} = 0$. The two estimates of $\ell'\tau$ are $l'G^*Q$ and $l'H^*Q$ but

$$\begin{aligned}
\ell' G^*Q - \ell' H^*Q &= \ell'(G^* - H^*)Q \\
&= \ell' a^* E_{v1} \\
&= a^* \ell' E_{v1} \\
&= 0
\end{aligned}$$

$\Rightarrow$ $l'\tau$ is unique.

Theorem: The adjusted treatment totals are orthogonal to the block totals.

Proof: It is enough to prove that

$$Cov(B_i, Q_j) = 0 \text{ for all } i, j.$$

Now
$$Cov(B_i, Q_j) = Cov\left[B_i, V_j - \sum_i \left(\frac{n_{ij}}{k_i}\right) B_i\right]$$
$$= Cov(B_i, V_j) - \frac{n_{ij}}{k_i} Var(B_i)$$

because the block totals are mutually orthogonal, see how:

For $y_{11}, y_{12}, ..., y_{1v},$ the block total $B_1 = \sum_{j=1}^{v} y_{1j}.$

For $y_{21}, y_{22}, ..., y_{2v},$ the block total $B_2 = \sum_{j=1}^{v} y_{2j}.$

$$Var(B_1) = \sum_{j=1}^{v} Var(y_{1j}) = v\sigma^2 \quad \text{as } Cov(y_{1j}, y_{1k}) = 0 \text{ for } j \neq k$$

$$Var(B_2) = \sum_{j=1}^{v} Var(y_{2j}) = v\sigma^2 \quad \text{as } Cov(y_{2j}, y_{2k}) = 0 \text{ for } j \neq k$$

$$Var(B_1) + Var(B_2) = 2v\sigma^2 \quad \text{as } Cov(y_{1j}, y_{2k}) = 0 \text{ for } j \neq k$$

$\Rightarrow$ B_1 and B_2 are mutually orthogonal as all y_{ij}'s are independent.

As B_i and V_j have n_{ij} observations in common and the observations are mutually independent, so

$$Cov(B_i, V_j) = n_{ij}\sigma^2$$

$$Var(B_i) = k_i\sigma^2$$

Thus $Cov(B_i, Q_j) = n_{ij}\sigma^2 - \frac{n_{ij}}{k_i} k_i\sigma^2 = 0.$

Hence proved.