Analysis of Variance and Design of Experiments

Incomplete Block Designs and Their Analysis

:::

Lecture 25

Analysis of Variance in IBD

Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur

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

Estimation of parameters:

Consider the model

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

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.

$$n\hat{\mu} + \sum_i n_{io}\hat{\beta}_i + \sum_j n_{oj}\hat{\tau}_j = G$$

$$n_i\hat{\mu} + n_{io}\hat{\beta}_i + \sum_j n_{ij}\hat{\tau}_j = B_i \qquad i = 1, ..., b$$

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

These (*b* + *v* +1) normal equations can be written as

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

Estimation of parameters:

These ($b + v$ +1) normal equations (*) can be written as

$$G = n\hat{\mu} + E_{1b}K\hat{\beta} + E_{1v}R\hat{\tau} \qquad \text{(i)}$$

$$B = KE_{b1}\hat{\mu} + K\hat{\beta} + N\hat{\tau} \qquad \text{(ii)}$$

$$V = RE_{v1}\hat{\mu} + N'\hat{\beta} + R\hat{\tau} \qquad \text{(iii)}$$

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

Premultiply equation (ii) by $N'K^{-1}$ **as**

$$N'K^{-1}B = N'K^{-1}KE_{b1}\hat{\mu} + N'K^{-1}K\hat{\beta} + N'K^{-1}N\hat{\tau}$$

and subtract it from equation (iii) as

$$V - N'K^{-1}B = (RE_{v1} - N'K^{-1}KE_{b1})\hat{\mu} + (N' - N'K^{-1}K)\hat{\beta} + (R - N'K^{-1}N)\hat{\tau}$$

or

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

Estimation of parameters:

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 adjusted for the block effects, the matrix *C* is called as *C*-matrix.

Estimation of parameters:

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, estimate the treatment effects from one of the normal equations and substitute it into another normal equation related to the block effects.

Estimation of parameters:

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

$$B = KE_{b1}\hat{\mu} + K\hat{\beta} + N\hat{\tau} \qquad \text{(ii)}$$

$$V = RE_{v1}\hat{\mu} + N'\hat{\beta} + R\hat{\tau} \qquad \text{(iii)}$$

Now we eliminate the treatment effects and obtain the block effects,

Next, pre-multiply equation (iii) by NR^{-1} **as**

$$NR^{-1}V = NR^{-1}RE_{v1}\hat{\mu} + NR^{-1}N'\hat{\beta} + NR^{-1}R\hat{\tau}$$

and subtract it from equation (ii) as

$$B - NR^{-1}V = (KE_{b1} - NR^{-1}RE_{v1})\hat{\mu} + (K - NR^{-1}N')\hat{\beta} + (N - NR^{-1}R)\hat{\tau}$$

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

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

Estimation of parameters:

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

where

$$D = diag(k_1, k_2, .., k_b) - N\, diag\left(\frac{1}{r_1}, \frac{1}{r_2}, ..., \frac{1}{r_v}\right) N' = K - NR^{-1}N'$$

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

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

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

Analysis of variance table: Block SS

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

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

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

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

Analysis of variance table: SSE

If *y* is the vector of all the observations, then error sum of squares

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

$$= \sum_i \sum_j \sum_m y_{ijm}(y_{ijm} - \hat{\mu} - \hat{\beta}_i - \hat{\tau}_j)$$

[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

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

Analysis of variance table: Partitioning SS

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

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

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

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

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

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

Analysis of variance table: Partitioning SS

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

$$\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$$

Error SS = Total SS Block SS Adjusted treatment SS

(unadjusted) (adjusted for blocks)

Analysis of variance table:

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.

Analysis of variance table:

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

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 designs, it doesn't matter at all. So the testing of hypothesis related to the block and treatment effects can be done from the same estimates.

Analysis of variance table:

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

Source	Degrees of freedom	Sum of squares	Mean squares	*F* - value
Treatment	$v-1$	$Q'\hat{\tau}$ (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	$n-1$	$y'y - \frac{G^2}{n}$		

Analysis of variance table:

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

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

  or

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

Analysis of variance table:

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.

Analysis of variance table:

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

$$\begin{aligned}
Q'E_{v1} &= V'E_{v1} - (N'K^{-1}B)'E_{v1} \\
&= (V_1,\dots,V_v)E_{v1} - B'K^{-1}NE_{v1} \\
&= (\sum_i V_i) - B'K^{-1}k' \\
&= \sum_i V_i - (B_1,\dots,B_b)diag\left(\frac{1}{k_1},\frac{1}{k_2},\dots,\frac{1}{k_b}\right)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_b \end{pmatrix} \\
&= \sum_i V_i - (B_1 \dots\dots B_b)E_{b1} \\
&= \sum_i V_i - \sum_j B_j \\
&= G - G \\
&= 0
\end{aligned}$$

so the intrablock equations are consistent

Connected designs:

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 $\ell'\tau$ **where** $l'E_{v1} = 0$ **have unique least-squares solutions. This can be easily proved.**