

Analysis of Variance and Design of Experiments

Incomplete Block Designs and Their Analysis

...

Lecture 28

Interblock Analysis of Variance



Shalabh

**Department of Mathematics and Statistics
Indian Institute of Technology Kanpur**



Slides can be downloaded from <http://home.iitk.ac.in/~shalab/sp1>

Interblock analysis of incomplete block design:

The purpose of block designs is to reduce the variability of response by removing the part of the variability as block numbers.

If in fact, this removal is illusory, the block effects being all equal, then the estimates are less accurate than those obtained by ignoring the block effects and using the estimates of treatment effects.

On the other hand, if the block effect is very marked, the reduction in the basic variability may be sufficient to ensure a reduction of the actual variances for the block analysis.

Interblock analysis of incomplete block design:

In the intrablock analysis related to treatments, the treatment effects are estimated after eliminating the block effects.

If the block effects are marked, then the block comparisons may also provide information about the treatment comparison.

So a question arises how to utilize the block information additionally to develop an analysis of variance to test the hypothesis about the significance of treatment effects.

Such an analysis can be derived regarding the block effects as random variables.

Interblock analysis of incomplete block design:

This assumption involves the random allocation of different blocks of the design to be the blocks of the material selected (at random from the population of possible blocks) in addition to the random allocation of treatments occurring in a block to the units of the block selected to contain them.

Now the two responses from the same block are correlated because the error associated with each contains the block number in common.

Such an analysis of incomplete block design is termed as interblock analysis.

Interblock analysis of incomplete block design:

To illustrate the idea behind the interblock analysis and how the block comparisons also contain information about the treatment comparisons, consider an allocation of four selected treatments in two blocks each. The outputs are recorded as follows:

Block 1: y_{14} y_{16} y_{17} y_{19}

Block 2: y_{23} y_{25} y_{26} y_{27} .

The block totals are

$$B_1 = y_{14} + y_{16} + y_{17} + y_{19},$$

$$B_2 = y_{23} + y_{25} + y_{26} + y_{27}.$$

Interblock analysis of incomplete block design:

Following the model $y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$, $i = 1, 2$, $j = 1, 2, \dots, 9$,

we have for block 1, $y_{14} = \mu + \beta_1 + \tau_4 + \varepsilon_{14}$,

$$y_{16} = \mu + \beta_1 + \tau_6 + \varepsilon_{16},$$

$$y_{17} = \mu + \beta_1 + \tau_7 + \varepsilon_{17},$$

$$y_{19} = \mu + \beta_1 + \tau_9 + \varepsilon_{19},$$

and for block 2, $y_{23} = \mu + \beta_2 + \tau_3 + \varepsilon_{23}$,

$$y_{25} = \mu + \beta_2 + \tau_5 + \varepsilon_{25},$$

$$y_{26} = \mu + \beta_2 + \tau_6 + \varepsilon_{26},$$

$$y_{27} = \mu + \beta_2 + \tau_7 + \varepsilon_{27},$$

and thus $B_1 - B_2 = 4(\beta_1 - \beta_2) + (\tau_4 + \tau_6 + \tau_7 + \tau_9) - (\tau_3 + \tau_5 + \tau_6 + \tau_7)$
 $+ (\varepsilon_{14} + \varepsilon_{16} + \varepsilon_{17} + \varepsilon_{19}) - (\varepsilon_{23} + \varepsilon_{25} + \varepsilon_{26} + \varepsilon_{27}).$

Interblock analysis of incomplete block design:

$$B_1 - B_2 = 4(\beta_1 - \beta_2) + (\tau_4 + \tau_6 + \tau_7 + \tau_9) - (\tau_3 + \tau_5 + \tau_6 + \tau_7) \\ + (\varepsilon_{14} + \varepsilon_{16} + \varepsilon_{17} + \varepsilon_{19}) - (\varepsilon_{23} + \varepsilon_{25} + \varepsilon_{26} + \varepsilon_{27}).$$

If we assume additionally that the block effects β_1 and β_2 are random with mean zero, then

$$E(B_1 - B_2) = (\tau_4 + \tau_9) - (\tau_3 + \tau_5)$$

which reflects that the block comparisons can also provide information about the treatment comparisons.

Interblock analysis of incomplete block design:

The intrablock analysis of an incomplete block design is based on estimating the treatment effects (or their contrasts) by eliminating the block effects.

Since different treatments occur in different blocks, so one may expect that the block totals may also provide some information on the treatments.

The interblock analysis utilizes the information on block totals to estimate the treatment differences.

Interblock analysis of incomplete block design:

The block effects are assumed to be random and so we consider the set up of mixed effect model in which the treatment effects are fixed but the block effects are random.

This approach is applicable only when the number of blocks is more than the number of treatments.

We consider here the interblock analysis of binary proper designs for which $n_{ij} = 0$ or 1 and $k_1 = k_2 = \dots = k_b = k$ in connection with the intrablock analysis.

Model and Normal Equations:

Let y_{ij} denotes the response from j^{th} treatment in i^{th} block from the model.

$$y_{ij} = \mu^* + \beta_i^* + \tau_j + \varepsilon_{ij}, i = 1, 2, \dots, b; j = 1, 2, \dots, v$$

where

μ^* : is the general mean effect;

β_i^* : is the random additive i^{th} block effect;

τ_j : is the fixed additive j^{th} treatment effect; and

ε_{ij} : is the i.i.d. random error with $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Model and Normal Equations:

Since the block effect is now considered to be random, so we additionally assume that β_i^* ($i = 1, 2, \dots, b$) are independently distributed following $N(0, \sigma_\beta^2)$ and are uncorrelated with ε_{ij} .

One may note that we can not assume here $\sum_i \beta_i^* = 0$ as in other cases of fixed-effect models. In place of this, we take $E(\beta_i^*) = 0$.

Also, y_{ij} 's are no longer independently distributed but

$$\text{Var}(y_{ij}) = \sigma_\beta^2 + \sigma^2,$$

$$\text{Cov}(y_{ij}, y_{i'j'}) = \begin{cases} \sigma_\beta^2 & \text{if } i = i', j \neq j' \\ 0 & \text{otherwise.} \end{cases}$$

Model and Normal Equations:

In the case of interblock analysis, we work with the block totals

B_i in place of y_{ij} where

$$\begin{aligned} B_i &= \sum_{j=1}^v n_{ij} y_{ij} \\ &= \sum_{j=1}^v n_{ij} (\mu^* + \beta_i^* + \tau_j + \varepsilon_{ij}) \\ &= k\mu^* + \sum_j n_{ij} \tau_j + f_i \end{aligned}$$

where $f_i = \beta_i^* k + \sum_j n_{ij} \varepsilon_{ij}$, $(i = 1, 2, \dots, b)$ are independent and normally distributed with mean 0 and

$$\text{Var}(f_i) = k^2 \sigma_\beta^2 + k\sigma^2 = \sigma_f^2$$

Model and Normal Equations:

Thus $E(B_i) = k\mu^* + \sum_j n_{ij}\tau_j,$

$$\text{Var}(B_i) = \sigma_f^2; \quad i = 1, 2, \dots, b,$$

$$\text{Cov}(B_i, B_{i'}) = 0; \quad i \neq i'; i, i' = 1, 2, \dots, b.$$

In matrix notations, the model under consideration can be written as

$$B = k\mu^* E_{b1} + N\tau + f$$

where $f = (f_1, f_2, \dots, f_b)'$.

In order to obtain the estimates of μ^* and τ , we minimise the sum of squares due to error $f = (f_1, f_2, \dots, f_b)'$, i.e., minimize

$$(B - k\mu^* E_{b1} - N\tau)'(B - k\mu^* E_{b1} - N\tau)$$

with respect to μ^* and τ .

Model and Normal Equations:

The estimates of μ^* and τ are the solutions of following normal equations:

$$\begin{pmatrix} kE'_{b1} \\ N' \end{pmatrix} \begin{pmatrix} kE_{b1} & N \end{pmatrix} \begin{pmatrix} \tilde{\mu} \\ \tilde{\tau} \end{pmatrix} = \begin{pmatrix} kE'_{b1} \\ N' \end{pmatrix} B$$

or
$$\begin{pmatrix} k^2 E'_{b1} E_{b1} & kE'_{b1} N \\ kN' E_{b1} & N' N \end{pmatrix} \begin{pmatrix} \tilde{\mu} \\ \tilde{\tau} \end{pmatrix} = \begin{pmatrix} kG \\ N' B \end{pmatrix}$$

or
$$\begin{pmatrix} k^2 b & kE'_{v1} R \\ kRE_{v1} & N' N \end{pmatrix} \begin{pmatrix} \tilde{\mu} \\ \tilde{\tau} \end{pmatrix} = \begin{pmatrix} kG \\ N' B \end{pmatrix} \quad (\text{using } N' E_{b1} = r = RE_{v1})$$

or
$$k^2 b \tilde{\mu} + kE'_{v1} R \tilde{\tau} = kG \quad (1)$$

$$kRE_{v1} \tilde{\mu} + N' N \tilde{\tau} = N' B. \quad (2)$$

Estimates of μ^* and τ in interblock analysis:

Using the side condition $E_{v1}' R \tilde{\tau} = 0$ (or equivalently $\sum_j r_j \tilde{\tau}_j = 0$)

and assuming $N'N$ to be nonsingular, the equation (1) gives

$$\tilde{\mu} = \frac{G}{bk}.$$

Substituting $\tilde{\mu}$ in the equation (2) gives $\tilde{\tau}$.

$$\begin{aligned}\tilde{\tau} &= (N'N)^{-1} \left(N'B - \frac{RE_{v1}G}{b} \right) \\ &= (N'N)^{-1} \left(N'B - \frac{kGN'E_{v1}}{bk} \right) \quad (\text{using } RE_{v1} = r = N'E_{v1}) \\ &= (N'N)^{-1} \left(N'B - \frac{G}{bk} N'NE_{v1} \right) \\ &= (N'N)^{-1} N'B - \frac{GE_{v1}}{bk}.\end{aligned}$$

Estimates of μ^* and τ in interblock analysis:

Generally, we are not interested merely in the interblock analysis of variance but we want to utilize the information from interblock analysis along with the intrablock information to improve upon the statistical inferences.

After obtaining the interblock estimate of treatment effects, the next question that arises is how to use this information for improved estimation of treatment effects and use it further for the testing of significance of treatment effects.

Such an estimate will be based on the use of more information, so it is expected to provide better statistical inferences.