Analysis of Variance and Design of Experiments-II

MODULE IV

LECTURE - 19

PARTIAL CONFOUNDING

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

The objective of confounding is to mix the less important treatment combinations with the block effect differences so that higher accuracy can be provided to the other important treatment comparisons. When such mixing of treatment contrasts and block differences is done in all the replicates, then it is termed as **total confounding.**

On the other hand, when the treatment contrast is not confounded in all the replicates but only in some of the replicates, then it is said to be partially confounded with the blocks. It is also possible that one treatment combination is confounded in some of the replicates and another treatment combination is confounded in other replicates which are different from the earlier replicates. So the treatment combinations are said to be confounded in some of the replicates and unconfounded in other replicates. So the treatment combinations are partially confounded. The partially confounded contrasts are estimated only from those replicates where they are not confounded. Since the variance of the contrast estimator is inversely proportional to the number of replicates in which it is estimable, so some factors on which information is available from all the replicates are more accurately determined.

Balanced and unbalanced partially confounded design

If all the effects of a certain order are confounded with the incomplete block differences in equal number of replicates in a design, then the design is said to be balanced partially confounded design.

If all the effects of a certain order are confounded an unequal number of times in a design, then the design is said to be unbalanced partially confounded design.

We discuss only the analysis of variance in the case of balanced partially confounded design through examples on 2^2 and 2^3 factorial experiments.

Example 1

Consider the case of 2^2 factorial as in following table in the setup of a randomized block design

Factorial effects	Treatment combinations				Divisor
	(1)	(*a*)	(*b*)	(*ab*)	
M	+	+	+	+	4
A	-	+	-	+	2
B	-	-	+	+	2
AB	+	-	-	+	2

where $y_{*i} = ((1), a, b, ab)'$ denotes the vector of total responses in the i^{th} replication and each treatment is replicated r times, $i = 1, 2, ..., r.$ If no factor is confounded then the factorial effects are estimated using all the replicates as

$$A = \frac{1}{2r}\sum_{i=1}^{r} \ell'_A y_{*i},$$

$$B = \frac{1}{2r}\sum_{i=1}^{r} \ell'_B y_{*i},$$

$$AB = \frac{1}{2r}\sum_{i=1}^{r} \ell'_{AB} y_{*i},$$

where the vectors of contrasts $\ell_A, \ell_B, \ell_{AB}$ are given by

$$\ell_A = (-1 \quad 1 \ -1 \ 1)'$$

$$\ell_B = (-1 \ -1 \quad 1 \ 1)'$$

$$\ell_{AB} = (1 \ -1 \ -1 \ 1)'.$$

We have in this case

$$\ell'_A \ell_A = \ell'_B \ell_B = \ell'_{AB} \ell_{AB} = 4.$$

The sum of squares due to *A, B* and *AB* can be accordingly modified and expressed as

$$SS_A = \frac{(\sum_{i=1}^{r} \ell'_A y_{*i})^2}{r \ell'_A \ell_A} = \frac{(ab + a - b - (1))^2}{4r}$$

$$SS_B = \frac{(\sum_{i=1}^{r} \ell'_B y_{*i})^2}{r \ell'_B \ell_B} = \frac{(ab + b - a - (1))^2}{4r}$$

and

$$SS_{AB} = \frac{(\sum_{i=1}^{r} \ell'_{AB} y_{*i})^2}{r \ell'_{AB} \ell_{AB}} = \frac{(ab + (1) - a - b)^2}{4r},$$

respectively.

Now consider a situation with 3 replicates with each consisting of 2 incomplete blocks as in the following figure :

Replicate 1 (A confounded)		Replicate 2 (B confounded)		Replicate 3 (AB confounded)	
Block 1	Block 2	Block 1	Block 2	Block 1	Block 2
ab a	b (1)	ab b	a (1)	ab (1)	a b

There are three factors *A, B* and *AB.* In case of total confounding, a factor is confounded in all the replicates. We consider here the situation of partial confounding in which a factor is not confounded in all the replicates.

Rather, the factor *A* is confounded in replicate 1, the factor *B* is confounded in replicate 2 and the interaction *AB* is confounded in replicate 3. Suppose each of the three replicate is repeated *r* times. So the observations are now available on *r* repetitions of each of the blocks in the three replicates. The partitions of replications, the blocks within replicates and plots within blocks being randomized. Now from the setup of figure,

- the factor *A* can be estimated from the replicates 2 and 3 as it is confounded in replicate 1.
- the factor *B* can be estimated from the replicates 1 and 3 as it is confounded in replicate 2 and
- the interaction *AB* can be estimated from the replicates 1 and 2 as it is confounded in replicate 3.

When A is estimated from the replicate 2 only, then its estimate is given by

$$A_{rep2} = \frac{(\sum_{i=1}^{r} \ell'_{A2} y_{*i})_{rep2}}{2r}$$

and when A is estimated from the replicate 3 only, then its estimate is given by

$$A_{rep3} = \frac{(\sum_{i=1}^{r} \ell'_{A3} y_{*i})_{rep3}}{2r}$$

where ℓ'_{A2} and ℓ'_{A3} are the suitable vectors of +1 and -1 for being the linear function to be contrasts under replicates 2 and 3, respectively. Note that ℓ'_{A2} and ℓ'_{A3} each is a (4 x 1) vector having 4 elements in it.

Now since A is estimated from both the replicates 2 and 3, so to combine them and to obtain a single estimator of A, we consider the arithmetic mean of A_{rep2} and A_{rep3} as an estimator of A which is given by

$$\begin{aligned} A_{pc} &= \frac{A_{rep2} + A_{rep3}}{2} \\ &= \frac{(\sum_{i=1}^{r} \ell'_{A2} y_{*i})_{rep2} + (\sum_{i=1}^{r} \ell'_{A3} y_{*i})_{rep3}}{4r} \\ &= \frac{\sum_{i=1}^{r} \ell^{*'}_{A} y_{*i}}{4r} \end{aligned}$$

where the (8 x 1) vector

$$\ell^{*'}_{A} = (\ell_{A2},\ \ell_{A3})$$

has 8 elements in it and subscript pc in A_{pc} denotes the estimate of A under partial confounding (pc) .

The sum of squares under partial confounding in this case is obtained as

$$SS_{Apc} = \frac{(\sum_{i=1}^{r} \ell^{*'}_{A} y_{*i})^2}{r\ell^{*'}_{A}\ell^{*}_{A}} = \frac{(\sum_{i=1}^{r} \ell^{*'}_{A} y_{*i})^2}{8r}.$$

Assuming that y_{ij} 's are independent and $Var(y_{ij}) = \sigma^2$ for all i and j, the variance of A_{pc} is given by

$$Var(A_{pc}) = \left(\frac{1}{4r}\right)^2 Var\left(\sum_{i=1}^{r} \ell^{*'}_{A} y_{*i}\right)$$

$$= \left(\frac{1}{4r}\right)^2 Var\left((\sum_{i=1}^{r} \ell^{'}_{A2} y_{*i})_{rep2} + (\sum_{i=1}^{r} \ell^{'}_{A3} y_{*i})_{rep3} \right)$$

$$= \left(\frac{1}{4r}\right)^2 \left(4r\sigma^2 + 4r\sigma^2\right)$$

$$= \frac{\sigma^2}{2r}.$$

Now suppose A is not confounded in any of the blocks in the three replicates in this example. Then A can be estimated from all the three replicates, each repeated r times. Under such a condition, an estimate of A can be obtained using the same approach as the arithmetic mean of the estimates obtained from each of the three replicates as

$$\begin{aligned} A_{pc}^{*} &= \frac{A_{rep1} + A_{rep2} + A_{rep3}}{3} \\ &= \frac{(\sum_{i=1}^{r} \ell'_{A1} y_{*i})_{rep1} + (\sum_{i=1}^{r} \ell'_{A2} y_{*i})_{rep2} + (\sum_{i=1}^{r} \ell'_{A3} y_{*i})_{rep3}}{6r} \\ &= \frac{\sum_{i=1}^{r} \ell_{A}^{**'} y_{*i}}{6r} \end{aligned}$$

where the (12 x 1) vector

$$\ell_{A}^{**'} = \left(\ell_{A1},\ \ell_{A2},\ \ell_{A3}\right)$$

has 12 elements in it. The variance of A is assuming that y_{ij}'s are independent and $Var(y_{ij}) = \sigma^{*2}$ for all i and j in this case is obtained as

$$\begin{aligned} Var(A_{pc}^{*}) &= \left(\frac{1}{6r}\right)^{2} Var\left[\left(\sum_{i=1}^{r} \ell'_{A1} y_{*i}\right)_{rep1} + \left(\sum_{i=1}^{r} \ell'_{A2} y_{*i}\right)_{rep2} + \left(\sum_{i=1}^{r} \ell'_{A3} y_{*i}\right)_{rep3}\right] \\ &= \left(\frac{1}{6r}\right)^{2} \left(4r\sigma^{*2} + 4r\sigma^{*2} + 4r\sigma^{*2}\right) \\ &= \frac{\sigma^{*2}}{3r}. \end{aligned}$$

If we compare this estimator with the earlier estimator of A for the situation where A is unconfounded in all the r replicates and estimated by

$$A = \frac{\sum_{i=1}^{r} \ell'_A y_{*i}}{2r}$$

and in the present situation of partial confounding, the corresponding estimator of A is given by

$$A^*_{pc} = \frac{A_{rep1} + A_{rep2} + A_{rep3}}{3} = \frac{\sum_{i=1}^{r} \ell^{**'}_A y_{*i}}{6r}.$$

Both the estimators, viz., A and A^*_{pc} are the same because A is based on r replications whereas A^*_{pc} is based on $3r$ replications.

If we denote $r^* = 3r$ then A^*_{pc} becomes the same as A. The expressions of variance of A and A^*_{pc} then also are same if we use $r^* = 3r$. Comparing them, we see that the information on A in the partially confounded scheme relative to that in unconfounded scheme is

$$\frac{2r/\sigma^2}{3r/\sigma^{*2}} = \frac{2}{3}\frac{\sigma^{*2}}{\sigma^2}.$$

If $\sigma^{*2} > \frac{3}{2}\sigma^2$, then the information in partially confounded design is more than the information in unconfounded design.

In total confounding case, the confounded effect is completely lost in total confounding but in the case of partial confounding, some information about the confounded effect can be recovered. For example, the two third of the total information can be recovered in this case for *A.*