Incomplete Block Designs

Hao Zhang

Introduction

In this lecture, we consider block designs where the block size is less than the number of treatments. Such an incomplete block design must be carefully designed. Otherwise, the treatments may not be comparable due to confounding with the blocks. All the designs that we discuss in this chapter are `equireplicate` ; that is, every treatment (or treatment combination) is observed r times in the experiment. These tend to be the most commonly used designs, although nonequireplicate designs are occasionally used in practice.

Design Plans and Randomization

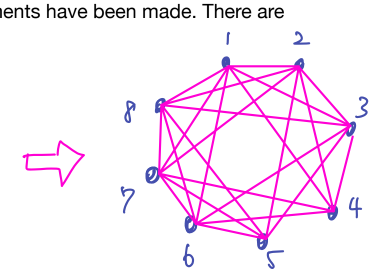
not twice

We use the symbol n_{ih} to denote the number of times that treatment i is observed in block h. When the block size is smaller than the number of treatments, each treatment should usually be observed either once or not at all in a block. Such block designs are called binary, and every n_{ih} is either 0 or 1 . For most purposes, the best binary designs are those in which pairs of treatments occur together in the same number (or nearly the same number) of blocks. These designs give rise to equal (or nearly equal) lengths of confidence intervals for pairwise comparisons of treatment effects.

The first stage of incomplete block design is to obtain as even a distribution as possible of treatment labels within the blocks. This results in an *experimental plan*. The plan in the table below shows a design with $b = 8$ blocks (labeled I, II, …, VIII) each of size $k = 3$, which can be used for an experiment with $v = 8$ treatments (labeled $1, \ldots, 8$) each observed $r = 3$ times. The treatment labels are evenly distributed in the sense that no label appears more than once per block and pairs of labels appear together in a block either once or not at all, which is "as equal as possible".

The experimental plan is often called the "design," even though it is not ready for use until the random assignments have been made. There are three steps to the randomization procedure, as follows.

- Randomly assign the block labels in the plan to the levels of the blocking factor(s).
- Randomly assign the experimental units in a block to those treatment labels allocated to that block.
- Randomly assign the treatment labels in the plan to the actual levels of the treatment factor.



connected design

Example 1.

$b=8$
$k=3$
$v=8$

$vr=bk$

therefore $r=3$
"Each treatment appears 3 times within different blocks"

Ex) treatment 3 appeared 3x I, III, and IV.

Block	(8) treatment #			Block			
I	1	3	8	V	5	7	4
II	2	4	1	VI	6	8	5
III	3	5	2	VII	7	1	6
IV	4	6	3	VIII	8	2	7

3 replicates in each block.

$\lambda = \frac{r(k-1)}{(v-1)} = \frac{3(3-1)}{8-1} = \frac{6}{7}$ not interger

Not a balancing complete block design!!!

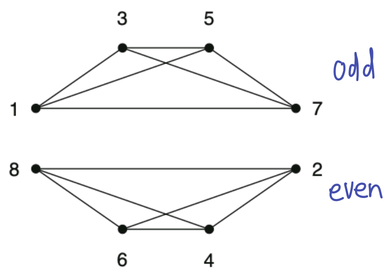
Connected Design

Consider the following experimental plan with 8 blocks, 8 treatments, and block size 3. There is a clear problem: No odd numbered treatment and even numbered treatment appeared together in one block. Therefore it is impossible to compare an odd numbered treatment with an even numbered treatment. Such a design is called disconnected.

The connectivity of a design can be illustrated by the connectivity of a graph: Draw a point for each treatment and draw a line between two points if the pair appears in a block. If for any pair of points, there is a path connecting the two points, the graph and the design are called connected. All contrasts in the treatment effects are estimable in a design if and only if the design is connected. The connectivity graph therefore provides a simple means of checking estimability.

Block				Block			
I	all odd 1	3	5	V	5	7	1
II	all even 2	4	6	VI	6	8	2

III	3	5	7	VII	7	1	3
IV	4	6	8	VIII	8	2	4



(a) Disconnected design of Table 11.3

X Incomparable to compare between odd/even

Ex) Can't compare 1 with 2.

Some Special Incomplete Block Designs

Balanced Incomplete Block Designs

A balanced incomplete block design is a design with v treatment labels, each occurring r times, and with bk experimental units grouped into b blocks of size $k < v$ in such a way that

i. Each treatment label appears either once or not at all in a block (that is, the design is binary).

ii. Each pair of labels appears together in λ blocks, where λ is a fixed integer. same # of blocks (λ)

There are three necessary conditions for the existence of a balanced incomplete block design, all of which are easy to check. These are

$$
\begin{gathered}
v r=b k \\
r(k-1)=\lambda(v-1) \\
b \geq v
\end{gathered}
$$

$r = bk/v \rightarrow$ integer ✓
if not interger, then No balanced block design exist.
$\lambda = \frac{r(k-1)}{(v-1)} \rightarrow$ interger ✓

Suggest use a computer package for the design.

Group Divisible Designs

A group divisible design is a design with $v=gl$ treatment labels (for some integers $g>1$ and $\ell>1$), each occurring r times, and bk experimental units grouped into b blocks of size $k<v$ in such a way that the units within a block are alike and units in different blocks are substantially different. The plan of the design satisfies the following conditions: (i) The $v=gl$ treatment labels are divided into g groups of ℓ labels-any two labels within a group are called first associates and any two labels in different groups are called second associates. (ii) Each treatment label appears either once or not at all in a block (that is, the design is binary). (iii) Each pair of first associates appears together in λ_1 blocks. (iv) Each pair of second associates appears together in λ_2 blocks.

Example 1 above is a group divisible design. It has the following $g=4$ groups of $\ell=2$ labels:

$$
(1,5), \quad(2,6), \quad(3,7), \quad(4,8)
$$

Labels in the same group (first associates) never appear together in a block, so $\lambda_1=0$. Labels in different groups (second associates) appear together in one block, so $\lambda_2=1$.

The following is a group divisible design with $v=12, r=3, b=6, k=6, \lambda_1=3, \lambda_2=1$.

k

Block	Treatments					
I	1	2	3	4	5	6
II	1	2	3	7	8	9
III	1	2	3	10	11	12
IV	4	5	6	7	8	9
V	4	5	6	10	11	12
VI	7	8	9	10	11	12

1,2,3 λ_1

Appears once $\lambda_2 = 1$

b

r=3 Appear 3x

v=12 types of outcomes

λ_1 and λ_2 affect the widths of confidence intervals. Therefore, it is preferable to make them as close as possible. Group divisible designs with λ_1 and λ_2 differing by one are usually regarded as the best choice of incomplete block design when no balanced incomplete block design exists.

Suggest use a computer package for the design.

Cyclic Designs

Examples

Design 1		Design 2	
Block	Treatments	Block	Treatments
1	1236	1	1236
2	2347	2	2341
3	3451	3	3452
4	4562	4	4563
5	5673	5	5614
6	6714	6	6125
7	7125		

rotating the treatment labels

A cyclic design is a design with v treatment labels, each occurring r times, and with bk experimental units grouped into $b = v$ blocks of size $k < v$ in such a way that the units within a block are alike and units in different blocks are substantially different. The experimental plan, using treatment labels $1, 2, \ldots, v$, can be obtained as follows:

i. The first block, called the initial block, consists of a selection of k distinct treatment labels.

ii. The second block is obtained from the initial block by cycling the treatment labels-that is, by replacing treatment label 1 with 2,2 with $3, \ldots, v-1$ with v, and v with 1 . The third block is obtained from the second block by cycling the treatment labels once more, and so on until the v th block is reached.

Use SAS for Incomplete Block Designs

`PROC OPTEX` in SAS searches for efficient incomplete block designs. Although one cannot specify the type of design to be generated, the software will search for the design that gives the smallest confidence region for all contrasts using the Scheffé method of multiple comparisons. If a balanced incomplete block design exists, it will usually be found by PROC OPTEX.

Here is an example of an incomplete block design with $b = 8$ blocks, each of size $k = 3$.

```
DATA CANDIDATE;
DO TREATMNT =1 to 8 ;
OUTPUT ;
END;
run;
proc print;
run;

PROC OPTEX DATA = CANDIDATE SEED =72145;
CLASS TREATMNT;
MODEL TREATMNT;
* For 8 blocks of size 3 ;
BLOCKS STRUCTURE = (8) 3 ;
EXAMINE DESIGN;
run;
```

Random Seed (Reproducible)

SAS outcome is provided below. The algorithm does 10 independent searches and `seed=xxx` where `xxx` is an integer determined where to start the search so that the algorithm yields the same results when it is re-run. The `D-Efficiency` concerns the width of confidence intervals given by Scheffe method where a higher value of D-efficiency means a shorter confidence intervals. The `A-Efficiency` concerns the width of confidence intervals for pairwise comparisons given by Tukey method where a higher value of A-efficiency means a shorter confidence intervals.

Design Number	Treatment D-Efficiency	Treatment A-Efficiency	Block Design D-Efficiency
1	75.4143	74.6667	98.0813

Higher Efficiency the Better

1	75.4143	74.6667	98.9813
2	75.4143	74.6667	98.9813
3	75.4143	74.6667	98.9813
4	75.4143	74.6667	98.9813
5	75.4143	74.6667	98.9813
6	75.4143	74.6667	98.9813
7	75.4143	74.6667	98.9813
8	75.4143	74.6667	98.9813
9	75.4143	74.6667	98.9813
10	75.4143	74.6667	98.9813

The solution of design given by the algorithm is as follows.

$b = 8$ Blocks

(size) $k = 3$ treatments per block.

$V = 7$ types of treatments

Point Number	Block Number	TREATMNT
1	1	3
2	1	2
3	1	1
4	2	5
5	2	6
6	2	7
7	3	1
8	3	6
9	3	8
10	4	3
11	4	4
12	5	4
13	5	5
14	5	4
15	5	2
16	6	7
17	6	4
18	6	1
19	7	3
20	7	8
21	7	5
22	8	8
23	8	7
24	8	2