

A Simple Procedure for Constructing Experiment Designs with Incomplete Blocks of Sizes 2 and 3

Walter T. Federer
Biometrics Unit
Cornell University
Ithaca, New York 14853

BU-1180-MA

August 1994

A SIMPLE PROCEDURE FOR CONSTRUCTING EXPERIMENT DESIGNS
WITH INCOMPLETE BLOCKS OF SIZES 2 and 3.

by

Walter T. Federer

Abstract

In order to maximize control of heterogeneity within complete blocks, an experimenter could use incomplete blocks of size $k = 2$ or 3 . In certain situations, incomplete blocks of this nature would eliminate the need for such spatial types of analyses as nearest neighbor. The intrablock efficiency factors for such designs are relatively low. However, with recovery of interblock information, Federer and Speed (1987) have presented measures of design efficiency factors which demonstrate that efficiency factors approach unity for certain ratios of the intrablock and interblock variance components. Hence with recovery of interblock information, even incomplete block designs with $k = 2$ or 3 have relatively high efficiency factors. The reduction in the intrablock error variance over the complete block error variance in many situations will provide designs with high efficiency.

A simple procedure for constructing incomplete blocks of sizes 2 and 3 is presented. It is shown how to obtain additional zero-one association confounding arrangements when $v = 4t$, t an integer, and for $v = pk$, $k < p$. It is indicated how to do the statistical analysis for these designs.

1. Introduction

Several simple construction procedures for incomplete block designs (IBDs) are available in the literature. Patterson *et al.* (1976, 1985) presented a class of IBDs denoted as alpha designs. Jarrett and Hall (1978) gave a procedure for constructing cyclic IBDs. Khare and Federer (1981) developed a procedure for constructing IBDs using "diagonalizing and variety cutting" to obtain designs for any

number of treatments. For equal block sizes, $v = pk$ with p blocks of $k \leq p$ treatments in each of the r replicates. For example, if an IBD for $v = 18$ and $k = 3$, 5 replicates can be obtained where pairs of treatments either occur together once in a block or they do not. The five first blocks of this design would be $\{1, 7, 13\}$, $\{1, 8, 15\}$, $\{1, 9, 17\}$, $\{1, 10, 14\}$ and $\{1, 11, 16\}$. When p is prime, p such replicates are possible. Khare and Federer (1981) also describe a procedure for p a prime power. Recently, Nguyen (1993) and Nguyen and Williams (1993) have discussed algorithms for constructing IBDs, row-column designs, and resolvable row-column designs. In addition, Nguyen (personal communication) has developed software, GENDEX, for constructing these designs, either in a randomized or non-randomized form. The program was found to be user-friendly.

The Khare-Federer procedure and GENDEX software have been found useful for setting up golf foursomes where players play with different individuals in each outing. Several experiment designs with block sizes of four have been constructed for agricultural researchers.

2. Even Number of Treatments

In order to construct an experiment design with incomplete blocks of size $k = 2$, for v even number of treatments, proceed as follows. First, construct the incomplete blocks for complete block 1 as

Complete block 1						
Incomplete block number						
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	...	<u>$v/2-1$</u>	<u>$v/2$</u>
1	2	3	4		$v/2-1$	$v/2$
$v/2+1$	$v/2+2$	$v/2+3$	$v/2+4$		$v-1$	v

For the incomplete blocks for complete block 2, simply move the second treatment numbers in an incomplete block one position to the left in a cyclical fashion as follows

Complete block 2						
Incomplete block number						
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	...	<u>$v/2-1$</u>	<u>$v/2$</u>
1	2	3	4		$v/2-1$	$v/2$
$v/2+2$	$v/2+3$	$v/2+4$	$v/2+5$		v	$v/2+1$

Proceeding in this fashion, the remaining of the $v/2$ incomplete block arrangements may be obtained as

Complete block 3						
Incomplete block number						
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	...	<u>$v/2-1$</u>	<u>$v/2$</u>
1	2	3	4		$v/2-1$	$v/2$
$v/2+3$	$v/2+4$	$v/2+5$	$v/2+6$		$v/2+1$	$v/2+2$
⋮						

Complete block $v/2$						
Incomplete block number						
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	...	<u>$v/2-1$</u>	<u>$v/2$</u>
1	2	3	4		$v/2-1$	$v/2$
v	$v/2+1$	$v/2+2$	$v/2+3$		$v-2$	$v-1$

The above procedure leads to incomplete block designs with binary associations in that a particular treatment either occurs with another treatment in an incomplete block or it does not, i.e., a zero-one occurrence. For example, treatment 1 occurs with treatments $v/2+1$ to v in incomplete blocks of the above $v/2$ complete blocks and does not occur with treatments 2, 3, ..., $v/2$ in any incomplete block.

3. Number of Treatments Equal to $4t$

When $v = 4t$, it is possible to construct $v/2 + v/4 = 3v/4$ confounding arrangements such that there is a binary zero-one association scheme among the treatments. $v/2$ arrangements are obtained as above, and then $v/4$ arrangements are obtained by using the above procedure on the first $v/2$ treatments and on the last $v/2$ treatments $v/2+1, \dots, v$. The procedure is illustrated with a specific example, i.e., $v = 12$ to obtain $6 + 3$ arrangements.

Complete block 1						Complete block 2						Complete block 3					
Incomplete block number						Incomplete block number						Incomplete block number					
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
7	8	9	10	11	12	8	9	10	11	12	7	9	10	11	12	7	8

Complete block 4						Complete block 5						Complete block 6					
Incomplete block number						Incomplete block number						Incomplete block number					
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
10	11	12	7	8	9	11	12	7	8	9	10	12	7	8	9	10	11

Complete block 7						Complete block 8						Complete block 9					
Incomplete block number						Incomplete block number						Incomplete block number					
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
1	2	3	7	8	9	1	2	3	7	8	9	1	2	3	7	8	9
4	5	6	10	11	12	5	6	4	11	12	10	6	4	5	12	11	10

4. Arrangements for Other Multiples of v

Extending the ideas in the above section, it is possible to obtain $v/2 + v/4 + v/8 = 7v/8$ zero-one association arrangements when $v = 8t$. When $v = 2^n$, the above procedure results in 2^{n-1} confounding arrangements leading to a balanced incomplete block design wherein every pair of treatments occurs $\lambda = 1$ times.

When $v = 3t$, for example, incomplete blocks of size $k = 3$ may be obtained as an extension of the above ideas. To illustrate, let $v = 15$. Then, five zero-one confounding arrangements may be obtained as follows:

Complete block 1					Complete block 2					Complete block 3				
Incomplete block number					Incomplete block number					Incomplete block number				
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
6	7	8	9	10	7	8	9	10	6	8	9	10	6	7
11	12	13	14	15	13	14	15	11	12	15	11	12	13	14

Complete block 4					Complete block 5				
Incomplete block number					Incomplete block number				
<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
1	2	3	4	5	1	2	3	4	5
9	10	6	7	8	10	6	7	8	9
12	13	14	15	11	14	15	11	12	13

Note that treatment 1, for example, occurs once with 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15 but does not occur with treatments 2, 3, 4, and 5 in the incomplete blocks of size $k = 3$. The same type of association occurs for every treatment. For $v = 3t$, there are t arrangements of the above form possible. Obviously the above procedure may be used for $v = pk$, $k < p$, to obtain p confounding arrangements with zero-one association schemes.

5. Efficiency Factors

An intrablock efficiency factor is defined as (e.g., Yates, 1937; Raghavarao, 1971):

$$e = (1 - 1/k) / (1 - 1/v) = v(k-1) / k(v-1) .$$

An intrablock variance of a difference between two intrablock treatment effects is $2\sigma_\epsilon^2 / re$. For interblock contrasts the variance of a difference between two treatment effects is

$$2(\sigma_\epsilon^2 + k\sigma_\beta^2) / r(1-e) ,$$

where r is the number of arrangements, σ_ϵ^2 is the intrablock error component of variance, and σ_β^2 is the interblock error component of variance. The average variance of a difference between two means recovering interblock information is

$$2\sigma_\epsilon^2 \left\{ (1 + k\gamma) / (1 + ke\gamma) \right\} / r ,$$

where $\gamma = \sigma_\beta^2 / \sigma_\epsilon^2$. One minus the ratio of the above error variance to the error variance of a randomized complete block design gives the increase in efficiency of the incomplete block design relative to a randomized complete block design.

Federer and Speed (1987) give efficiency factors for an incomplete block design as

$$e_1 = (1 + ke\gamma) / (1 + k\gamma)$$

and

$$e_2 = (1 + k\gamma) / (1 + [k + 1]\gamma) .$$

They consider that there is little to choose between the two and suggest that e_1 be the factor used.

When $\gamma = \infty$, e_1 becomes the intrablock efficiency factor e ; when $\gamma = 0$, $e_1 = 1$.

6. Statistical Analysis

Several computer software programs are available for analyzing results from IBs with recovery of interblock information. One such package is GAUSS. A program on GAUSS is given in the appendix for $v = pk$ treatments in r complete blocks (replicates) and p incomplete blocks of size k illustrated for a specific example of $v = 6$, $k = 2$, $p = 3$, and $r = 3$. If there are no missing observations, it is convenient to use $Y_{ghi} - \bar{y}_{g..}$ as the responses in the $\mathbf{Y}_{rv \times 1}$ vector of observations. This removes the mean and replicate effect from the linear model

$$Y_{ghi} = \mu + \rho_g + \beta_{gh} + \tau_i + \epsilon_{ghi} \quad (1)$$

where μ is common mean effect, ρ_g is the effect of replicate g , β_{gh} is the h th block effect in replicate g , τ_i is the i th treatment effect, and ϵ_{ghi} is a random error effect. Then

$$Y_{ghi} - \bar{y}_{g..} = \beta_{gh} + \tau_i + \epsilon_{ghi} \quad (2)$$

In the latter form, we have the simple matrix equations

$$\begin{bmatrix} k\mathbf{I}_{rp} & \mathbf{N}_{rb \times v} \\ \mathbf{N}'_{v \times rp} & r\mathbf{I}_v \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{B} \\ \mathbf{T} \end{bmatrix} \quad (3)$$

where \mathbf{I}_x is the identity matrix, \mathbf{N} is the block by treatment incidence matrix, \mathbf{B} is an $rp \times 1$ vector of block totals ($Y_{gh.} - k\bar{y}_{g..}$), \mathbf{T} is a $v \times 1$ vector of treatment totals ($Y_{. . i} - r\bar{y}_{. .}$), $\boldsymbol{\beta}$ is an $rp \times 1$ vector of block effects, and $\boldsymbol{\tau}$ is a $v \times 1$ vector of treatment effects. The intrablock solutions for block and treatment effects are:

$$\hat{\boldsymbol{\beta}} = (k\mathbf{I}_{rp} - \mathbf{N}\mathbf{N}'/r + \mathbf{J}_{rp}/r)^{-1} (\mathbf{B} - \mathbf{N}\mathbf{T}/r) \quad (4)$$

and

$$\hat{\boldsymbol{\tau}} = [r\mathbf{I}_v - \mathbf{N}'\mathbf{N}/k + \mathbf{J}_v/k]^{-1} (\mathbf{T} - \mathbf{N}'\mathbf{B}/k) \quad (5)$$

where \mathbf{J}_v is a $v \times v$ matrix of ones, and \mathbf{J}_{rp} could be of several forms, e.g. $\mathbf{I}_r * \mathbf{J}_k$ where $*$ means Kronecker product, \mathbf{I}_r is the $r \times r$ identity matrix, and \mathbf{J}_k is a $k \times k$ matrix of ones. For \mathbf{Z} a $p \times p$

matrix if zeros, \mathbf{J}_{rp} could be of the form:

$$\begin{bmatrix} \mathbf{Z} & \mathbf{J}_k & \mathbf{J}_k & \dots & \mathbf{J}_k \\ \mathbf{J}_k & \mathbf{Z} & \mathbf{J}_k & & \mathbf{J}_k \\ \vdots & & & & \\ \mathbf{J}_k & \mathbf{J}_k & \mathbf{J}_k & \dots & \mathbf{Z} \end{bmatrix} \quad (6)$$

For $p = k$, $(k\mathbf{I}_{rk} - \mathbf{N}\mathbf{N}'/r + \mathbf{J}_{rk}/r)^{-1}$ becomes a diagonal matrix. The various sums of squares in the analysis of variance (ANOVA) for the above linear model are given in table 1.

To recover interblock information simply replace $k\mathbf{I}_{rp}$ by $\mathbf{I}_{rp}(k + \sigma_\epsilon^2/\sigma_\beta^2)$ where the intrablock error mean square E/fe is taken as the estimate σ_ϵ^2 and σ_β^2 is the variance component for blocks where the expected value of $BS/r(p-1)$ is $\sigma_\epsilon^2 + \left(\frac{r-1}{r}\right) k\sigma_\beta^2$. Then, the treatment effects with recovery interblock information are:

$$\tau^* = (r\mathbf{I}_v - \mathbf{N}'\mathbf{N}/(k + \sigma_\epsilon^2/\sigma_\beta^2) + \mathbf{J}_v/k)^{-1} (\mathbf{T} - \mathbf{N}'\mathbf{B}/(k + \sigma_\epsilon^2/\sigma_\beta^2)) \quad (7)$$

Table 1. ANOVA for an IBD with v treatments in r replicates composed of p blocks of size k .

Source of variation	d.f.	sum of squares	mean square
Total	rv	$\mathbf{Y}'\mathbf{Y}$	-----
Correlation for mean		$1Y_{...}^2/rv$	
Replicates = R	$r - 1$	$\sum Y_{g...}^2/v - Y_{...}^2/rv = RS$	$RS/(r - 1)$
Treatments (ignoring blocks)	$v - 1$	$\mathbf{T}'\mathbf{T}'$	$\mathbf{T}'\mathbf{T}'/(v - 1)$
RCB error	$(r - 1)(v - 1)$	subtraction	
Blocks (eliminating treatments)	$r(p - 1)$	$\hat{\beta}'(\beta - \mathbf{N}\mathbf{T}/r) = BS$	$BS/r(p - 1)$
Intrablock error	fe	subtraction = E	E/fe
Treatments (eliminating block)	$v - 1$		

$$fe = (r - 1)(v - 1) - r(p - 1)$$

and the variance-covariance matrix is

$$\sigma_{\epsilon}^2 (\mathbf{rI}_v - \mathbf{N}'\mathbf{N}/(k + \sigma_{\epsilon}^2/\sigma_{\beta}^2) + \mathbf{J}_v /k)^{-1} \quad (8)$$

As can be seen from the GAUSS program for $v = 6$, $k = 2$, $p = 3$, and $r = 3$, the sums of squares in the ANOVA and the values for (7) and (8) are readily obtained.

7. Literature Cited

Federer, W. T. and T. P. Speed (1987). Measures of block design efficiency recovering interblock information. In *Transactions of the Fourth Army Conference on Applied Mathematics and Computing, ARO-Report 87-1*, pages 781-786.

Jarrett, R.G. and W.B. Hall (1978). Generalized cyclic incomplete block designs. *Biometrika*, 65(2): 397-401.

Khare, M. and W.T. Federer (1981). A simple construction procedure for resolvable incomplete block designs for any number of treatments. *Biometrical Journal*, 23(2): 121-132.

Nguyen, Nam-Ky (1993). An algorithm for constructing optimal resolvable incomplete block designs. *Communications In Statistics-Simula*, 22(3): 911-923.

Nguyen, Nam-Ky (1994). Computer software program GENDEX, CSIRO-IAPP Biometrics Unit, Private bag 10, Rosebank MDC, Clayton, Victoria 3169, Australia. Personal communication.

Nguyen, Nam-Ky and E.R. Williams (1993). An algorithm for constructing optimal resolvable row-column designs. *Australian Journal of Statistics*, 35(3): 363-370.

Patterson, H.D. and E.R. Williams (1976). A new class of resolvable incomplete block designs. *Biometrika*, 63(1): 83-92.

Patterson, H.D., E.R. Williams, and L. Paterson (1985). A note on resolvable incomplete block designs. *Biometrical Journal*, 27: 75-79.

Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments*. John Wiley & Sons, Inc., New York.

Yates, F. (1937). The design and analysis of factorial experiments. *Imperial Bureau Soil Science, Tech. Comm.* 35: 1-95

@

APPENDIX

The design is

1	4	1	5	1	6
2	5	2	6	2	4
3	6	3	4	3	5.

v = bk treatments in b incomplete blocks of size k with r replicates.@

v = 6; k = 2; r = 3; b = 3; format 2,1;

let Y[18,1] = -3 1 -3 1 0 4 3 3 0 0 -3 -3 0 2 -2 0 -1 1;

let N[9,6] =

```

1 0 0 1 0 0
0 1 0 0 1 0
0 0 1 0 0 1
1 0 0 0 1 0
0 1 0 0 0 1
0 0 1 1 0 0
1 0 0 0 0 1
0 1 0 1 0 0
0 0 1 0 1 0;

```

let H[4,15]=

```

1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 1 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 1 1 1 1 1 1;
1 0 0 0 0 0 0 0 0 1 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 1 0 0
0 1 0 0 0 0 0 0 0 0 1 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 1 0
0 0 1 0 0 0 0 0 0 0 0 1 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
0 0 0 1 0 0 0 0 0 1 0 0 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 1 0
0 0 0 0 1 0 0 0 0 0 1 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 1
0 0 0 0 0 1 0 0 0 0 0 1 0 0 0
0 0 0 0 0 1 0 0 0 0 0 0 1 0 0
0 0 0 0 0 0 1 0 0 1 0 0 0 0 0
0 0 0 0 0 0 1 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 1 0 0 1 0 0 0 0
0 0 0 0 0 0 0 1 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0 1 0 0 1 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0 0 1 0;

```

let X[18,15]=

"The block and treatment totals are X'*Y = tot";

tot = X'*Y;tot';G = X|H;sol = inv(G'*G)*X'*Y; sol';

Irb = eye(9);Iv = eye(6); Jv = ones(6,6);let J0[9,9] =

```

0 0 0 1 1 1 1 1 1
0 0 0 1 1 1 1 1 1
0 0 0 1 1 1 1 1 1
1 1 1 0 0 0 1 1 1
1 1 1 0 0 0 1 1 1
1 1 1 0 0 0 1 1 1
1 1 1 1 1 1 0 0 0
1 1 1 1 1 1 0 0 0
1 1 1 1 1 1 0 0 0;

```

let B[9,1] = -2 -2 4 6 0 -6 2 -2 0;

let T[6,1] = 0 -5 -4 -2 5 6;

B0 = inv(k*Irb - N*N'/r + J0/r);bl = B0*(B - N*T/r);

"Block solutions are bl =";bl';

"Intrablock treatment solutions are tr =";

```

tr = inv(r*Iv - N'*N/k + Jv/k)*(T - N'*B/k);tr';
"Block eliminating treatment sum of squares = Bss";
Bss = bl'*(B - N*T/r);Bss;
"Treatment ignoring block effects T'T/r is";T'*T/r;
"Total sum of squares Y'Y is";Y'*Y;
"The intrablock error mean square E and the blocks variance
component bvar = Bss/k(r-1)(b-1) - rE/k(r - 1) are ";E = 3;E;
bvar = Bss/8 - 3*E/4;bvar;
"Treatment effects with recovery of interblock information = trec";
trec = inv(r*Iv - N'*N/(k + E/bvar) + Jv/k)*(T - N'*B/(k + E/bvar));trec';
"Variance-covariance matrix for trec is";
varcov = E*inv(r*Iv - N'*N/(k + E/bvar) + Jv/k);varcov;

```

Ins L=69 C=1 File=D:\GAUSS\EXPMS623

Executing: D:\GAUSS\EXPMS623

The block and treatment totals are X'*Y = tot

```

-2.0 -2.0 4.0 6.0 0.0 -6.0 2.0 -2.0 0.0 0.0 -5.0 -4.0 -2.0 5.0 6.0
-1.0 -1.0 2.0 3.0 3.1E-016 -3.0 1.0 -1.0 -8.3E-017 -1.0 -1.0 -1.0 1.0 1.0 1.0

```

Block solutions are bl =

```

-1.0 -1.0 2.0 3.0 -9.2E-018 -3.0 1.0 -1.0 -4.6E-017

```

Intrablock treatment solutions are tr =

```

-1.0 -1.0 -1.0 1.0 1.0 1.0

```

Block eliminating treatment sum of squares = Bss

34.7

Treatment ignoring block effects T'T/r is

35.3

Total sum of squares Y'Y is

82.0

The intrablock error mean square E and the blocks variance

component bvar = Bss/k(r-1)(b-1) - rE/k(r - 1) are

3.0

2.1

Treatment effects with recovery of interblock information = trec

```

-0.4 -1.4 -1.2 0.0 1.4 1.6

```

Variance-covariance matrix for trec is

```

1.2 -0.2 -0.2 -0.0 -0.0 -0.0
-0.2 1.2 -0.2 -0.0 -0.0 -0.0
-0.2 -0.2 1.2 -0.0 -0.0 -0.0
-0.0 -0.0 -0.0 1.2 -0.2 -0.2
-0.0 -0.0 -0.0 -0.2 1.2 -0.2
-0.0 -0.0 -0.0 -0.2 -0.2 1.2

```