Chapter 4 Supplemental Text Material

S4-1. Relative Efficiency of the RCBD

In Example 4-1 we illustrated the noise-reducing property of the randomized complete block design (RCBD). If we look at the portion of the total sum of squares not accounted for by treatments (302.14; see Table 4-4), about 63 percent (192.25) is the result of differences between blocks. Thus, if we had run a completely randomized design, the mean square for error MS_E would have been much larger, and the resulting design would not have been as sensitive as the randomized block design.

It is often helpful to estimate the relative efficiency of the RCBD compared to a completely randomized design (CRD). One way to define this relative efficiency is

$$R = \frac{(df_b + 1)(df_r + 3)}{(df_b + 3)(df_r + 1)} \cdot \frac{\sigma_r^2}{\sigma_b^2}$$

where σ_r^2 and σ_b^2 are the experimental error variances of the completely randomized and randomized block designs, respectively, and df_r and df_b are the corresponding error degrees of freedom. This statistic may be viewed as the increase in replications that is required if a CRD is used as compared to a RCBD if the two designs are to have the same sensitivity. The ratio of degrees of freedom in R is an adjustment to reflect the different number of error degrees of freedom in the two designs.

To compute the relative efficiency, we must have estimates of σ_r^2 and σ_b^2**.** We can use the mean square for error MS_E from the RCBD to estimate σ_b^2, and it may be shown [see Cochran and Cox (1957), pp. 112-114] that

$$\hat{\sigma}_r^2 = \frac{(b-1)\,MS_{Blocks} + b(a-1)\,MS_E}{ab-1}$$

is an unbiased estimator of the error variance of a the CRD. To illustrate the procedure, consider the data in Example 4-1. Since MS_E = 7.33, we have

$$\hat{\sigma}_b^2 = 7.33$$

and

$$\begin{aligned}\hat{\sigma}_r^2 &= \frac{(b-1)MS_{Blocks} + b(a-1)MS_E}{ab-1} \\ &= \frac{(5)38.45 + 6(3)7.33}{4(6)-1} \\ &= 14.10\end{aligned}$$

Therefore our estimate of the relative efficiency of the RCBD in this example is

$$
\begin{aligned}
R &= \frac{(df_b+1)(df_r+3)}{(df_b+3)(df_r+1)} \cdot \frac{\sigma_r^2}{\sigma_b^2} \\
&= \frac{(15+1)(20+3)}{(15+3)(20+1)} \cdot \frac{14.10}{7.33} \\
&= 1.87
\end{aligned}
$$

This implies that we would have to use approximately twice times as many replicates with a completely randomized design to obtain the same sensitivity as is obtained by blocking on the metal coupons.

Clearly, blocking has paid off handsomely in this experiment. However, suppose that blocking was not really necessary. In such cases, if experimenters choose to block, what do they stand to lose? In general, the randomized complete block design has $(a-1)(b-1)$ error degrees of freedom. If blocking was unnecessary and the experiment was run as a completely randomized design with b replicates we would have had $a(b-1)$ degrees of freedom for error. Thus, incorrectly blocking has cost $a(b-1)-(a-1)(b-1)=b-1$ degrees of freedom for error, and the test on treatment means has been made less sensitive needlessly. However, if block effects really are large, then the experimental error may be so inflated that significant differences in treatment means could possibly remain undetected. (Remember the incorrect analysis of Example 4-1.) As a general rule, when the importance of block effects is in doubt, the experimenter should block and gamble that the block means are different. If the experimenter is wrong, the slight loss in error degrees of freedom will have little effect on the outcome as long as a moderate number of degrees of freedom for error are available.

S4-2. Partially Balanced Incomplete Block Designs

Although we have concentrated on the balanced case, there are several other types of incomplete block designs that occasionally prove useful. BIBDs do not exist for all combinations of parameters that we might wish to employ because the constraint that λ be an integer can force the number of blocks or the block size to be excessively large. For example, if there are eight treatments and each block can accommodate three treatments, then for λ to be an integer the smallest number of replications is $r = 21$. This leads to a design of 56 blocks, which is clearly too large for most practical problems. To reduce the number of blocks required in cases such as this, the experimenter can employ **partially balanced incomplete block** designs, or PBIDs, in which some pairs of treatments appear together λ_1 times, some pairs appear together λ_2 times, . . ., and the remaining pairs appear together λ_m times. Pairs of treatments that appear together λ_i times are called *i*th *associates*. The design is then said to have m **associate classes**.

An example of a PBID is shown in Table 1. Some treatments appear together $\lambda_1 = 2$ times (such as treatments 1 and 2), whereas others appear together only $\lambda_2 = 1$ times (such as treatments 4 and 5). Thus, the design has two associate classes. We now describe the **intrablock analysis** for these designs.

A partially balanced incomplete block design with two associate classes is described by the following parameters:

1 . There are a treatments arranged in b blocks. Each block contains k runs and each treatment appears in r blocks.

2. Two treatments which are ith associates appear together in λ_i blocks, i = 1, 2.

3. Each treatment has exactly n_i ith associates, i = 1,2. The number n_i is independent of the treatment chosen.

4. If two treatments are ith associates, then the number of treatments that are jth associates of one treatment and kth associates of the other treatment is p^i_{jk}, (i , j ,k = 1, 2). It is convenient to write the p^i_{jk} as (2 x 2) matrices with p^i_{jk} the jkth element of the ith matrix. 4

For the design in Table 1 we may readily verify that $a = 6$, $b = 6$, $k = 3$, $r = 3$, $\lambda_1 = 2$, $\lambda_2 = 1$, $n_1 = 1$, $n_2 = 4$,

$$\{p^1{}_{jk}\} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \qquad \text{and} \qquad \{p^2{}_{jk}\} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

Table 1. A Partially Balanced incomplete Block Design with Two Associate Classes

Block	Treatment Combinations		
1	1	2	3
2	3	4	5
3	2	5	6
4	1	2	4
5	3	4	6
6	1	5	6

We now show how to determine the p^i_{jk} . Consider any two treatments that are first associates, say 1 and 2. For treatment 1, the only first associate is 2 and the second associates are 3, 4, 5, and 6. For treatment 2, the only first associate is 1and the second associates are 3, 4, 5, and 6. Combining this information produces Table 2. Counting the number of treatments in the cells of this table, have the $\{p^1_{jk}\}$ given above. The elements $\{p^2_{jk}\}$ are determined similarly.

The linear statistical model for the partially balanced incomplete block design with two associate classes is

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

where μ is the overall mean, τ_i is the ith treatment effect, β_j is the jth block effect, and ε_{ij} is the NID(0, σ^2) random error component. We compute a total sum of squares, a block sum of squares (unadjusted), and a treatment sum of squares (adjusted). As before, we call

$$Q_i = y_{i.} - \frac{1}{k}\sum_{j=1}^{b} n_{ij} y_{.j}$$

the adjusted total for the ith treatment. We also define

$$S_1(Q_i) = \sum_s Q_s \qquad \text{s and i are first associates}$$

$$\Delta = k^{-2}\{(rk - r + \lambda_1)(rk - r + \lambda_2) + (\lambda_1 + \lambda_2)$$

$$c_1 = (k\Delta)^{-1}[\lambda_1(rk - r + \lambda_2) + (\lambda_1 - \lambda_2)(\lambda_2 p^1_{12} - \lambda_1 p^2_{12})]$$

$$c_2 = (k\Delta)^{-1}[\lambda_2(rk - r + \lambda_1) + (\lambda_1 - \lambda_2)(\lambda_2 p^1_{12} - \lambda_1 p^2_{12})]$$

The estimate of the ith treatment effect is

$$\hat{\tau}_i = \frac{1}{r(k-1)}[(k - c_2)Q_i + (c_1 - c_2)S_1(Q_i)]$$

and the adjusted treatment sum of squares is

$$SS_{Treatments(adjusted)} = \sum_{i=1}^{a} \hat{\tau}_i Q_i$$

The analysis of variance is summarized in Table 3. To test H_0: $\tau_i = 0$, we use $F_0 = MS_{Treatments(adjusted)} / MS_E$.

Table 2. Relationship of Treatments to 1 and 2

	Treatment 2	
Treatment 1	1st Associate	2nd Associate
1st associate		
2nd associate		3,4,5,6

Table 3. Analysis of Variance for the Partially Balanced Incomplete Block Design with Two Associate Classes

Source of Variation	Sum of Squares	Degrees of Freedom
Treatments (adjusted)	$\sum_{i=1}^{a} \hat{\tau}_i Q_i$	a-1
Blocks	$\frac{1}{k}\sum_{j=1}^{b} y^2_{.j} - \frac{y^2..}{bk}$	b-1
Error	Subtraction	bk-b-a+1
Total	$\sum_i \sum_j y^2_{ij} - \frac{y^2..}{bk}$	bk-1

We may show that the variance of any contrast of the form $\hat{\tau}_u - \hat{\tau}_v$ is

$$V(\tau_u - \tau_v) = \frac{2(k - c_i)\sigma^2}{r(k-1)}$$

where treatments *u* and *v* are *i*th associates (i = 1, 2). This indicates that comparisons between treatments are not all estimated with the same precision. This is a consequence of the partial balance of the design.

We have given only the intrablock analysis. For details of the interblock analysis, refer to Bose and Shimamoto (1952) or John (1971). The second reference contains a good discussion of the general theory of incomplete block designs. An extensive table of partially balanced incomplete block designs with two associate classes has been given by Bose, Clatworthy, and Shrikhande (1954).

S4-3. Youden Squares

Youden squares are "incomplete" Latin square designs in which the number of columns does not equal the number of rows and treatments. For example, consider the design shown in Table 4. Notice that if we append the column *(E, A, B, C, D*) to this design, the result is a 5 × 5 Latin square. Most of these designs were developed by W. J. Youden, hence their name.

Although a Youden square is always a Latin square from which at least one column (or row or diagonal) is missing, it is not necessarily true that every Latin square with more than one column (or row or diagonal) missing is a Youden square. The arbitrary removal of more than one column, say, for a Latin square may destroy its balance. In general, a

Youden square is a symmetric balanced incomplete block design in which rows correspond to blocks and each treatment occurs exactly once in each column or "position" of the block. Thus, it is possible to construct Youden squares from all symmetric balanced incomplete block designs, as shown by Smith and Hartley (l948). A table of Youden squares is given in Davies (1956), and other types of incomplete Latin squares are discussed by Cochran and Cox (1957, Chapter 13).

Table 4. A Youden Square for Five Treatments (A, *B*, *C*, *D*, *E*)

	Column			
Row	1	2	3	4
1	A	B	C	D
2	B	C	D	E
3	C	D	E	A
4	D	E	A	B
5	E	A	B	C

The linear model for a Youden square is

$$Y_{ijh} = \mu + \alpha_i + \tau_j + \beta_h + \varepsilon_{ijh}$$

where, μ is the overall mean, α_i is the *i*th block effect τ_j is the *j*th treatment effect, β_h is the *h*th position effect, and ε_{ijh} is the usual NID(0, σ^2) error term. Since positions occur exactly once in each block and once with each treatment, positions are orthogonal to blocks and treatments. The analysis of the Youden square is similar to the analysis of a balanced incomplete block design, except that a sum of squares between the position totals may also be calculated.

Example of a Youden Square

An industrial engineer is studying the effect of five illumination levels on the occurrence of defects in an assembly operation. Because time may be a factor in the experiment, she has decided to run the experiment in five blocks, where each block is a day of the week**.** However, the department in which the experiment is conducted has four work stations and these stations represent a potential source of variability. The engineer decided to run a Youden square with five rows (days or blocks), four columns (work stations), and five treatments (the illumination levels). The coded data are shown in Table 5.

Table 5. The Youden Square Design used in the Example

Day (Block)	Work Station					Treatment totals
	1	2	3	4	$y_{i..}$	
1	A=3	B=1	C=-2	D=0	2	$y_{.1.}$=12 (A)
2	B=0	C=0	D=-1	E=7	6	$y_{.2.}$=2 (B)
3	C=-1	D=0	E=5	A=3	7	$y_{.3.}$=-4 (C)
4	D=-1	E=6	A=4	B=0	9	$y_{.4.}$=-2 (D)
5	E=5	A=2	B=1	C=-1	7	$y_{.5.}$=23 (E)
$y_{..h}$	6	9	7	9	$y_{...}$=31	

Considering this design as a balanced incomplete block, we find a = *b* = 5, *r*= *k* = 4, and k = 3. Also,

$$SS_T = \sum_i \sum_j \sum_h y^2{}_{ijh} - \frac{y^2 \cdots}{N} = 183.00 - \frac{(31)^2}{20} = 134.95$$

$$Q_1 = 12 - \frac{1}{4}(2 + 7 + 9 + 7) = 23/4$$

$$Q_2 = 2 - \frac{1}{4}(2 + 6 + 9 + 7) = -16/4$$

$$Q_3 = -4 - \frac{1}{4}(2 + 6 + 7 + 7) = -38/4$$

$$Q_4 = -2 - \frac{1}{4}(2 + 6 + 7 + 9) = -32/4$$

$$Q_5 = 23 - \frac{1}{4}(6 + 7 + 9 + 7) = 63/4$$

$$SS_{Treatments(adjusted)} = \frac{k \sum_{i=1}^{a} Q_i^2}{\lambda a}$$

$$= \frac{4[(23/4)^2 + (-16/4)^2 + (-38/4)^2 + (-32/4)^2 + (63/4)^2}{(3)(5)} = 120.37$$

Also,

$$SS_{Days} = \sum_{i=1}^{b} \frac{y^2{}_{i..}}{k} - \frac{y^2{}_{...}}{N} = \frac{(2)^2 + (6)^2 + (7)^2 + (9)^2 + (7)^2}{4} - \frac{(31)^2}{20} = 6.70$$

$$SS_{Stations} = \sum_{h=1}^{k} \frac{y^2{}_{..h}}{b} - \frac{y^2{}_{...}}{N} = \frac{(6)^2 + (9)^2 + (7)^2 + (9)^2}{5} - \frac{(31)^2}{20} = 1.35$$

and

$SS_E = SS_T - SS_{Treatments\ (adjusted)} - SS_{Days} - SS_{Stations}$
$= 134.95 - 120.37 - 6.70 - 1.35 = 6.53$

Block or day effects may be assessed by computing the adjusted sum of squares for blocks. This yields

$Q_1' = 2 - \frac{1}{4}(12 + 2 - 4 - 2) = 0/4$

$Q_2' = 6 - \frac{1}{4}(2 - 3 - 2 + 23) = 5/4$

$Q_3' = 7 - \frac{1}{4}(12 - 4 - 2 + 23) = -1/4$

$Q_4' = 9 - \frac{1}{4}(12 + 2 - 2 + 23) = 1/4$

$Q_5' = 7 - \frac{1}{4}(12 + 2 - 4 + 23) = -5/4$

$$SS_{Days(adjusted)} = \frac{r\sum_{j=1}^{b} Q_j'^2}{\lambda b}$$

$$= \frac{4[(0/4)^2 + (5/4)^2 + (-1/4)^2 + (1/4)^2 + (-5/4)^2]}{(3)(5)} = 0.87$$

The complete analysis of variance is shown in Table 6. Illumination levels are significantly different at 1 percent.

Table 6 Analysis of Variance for the Youden Square Example

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Illumination level, adjusted	120.37	4	30.09	36.87[a]
Days, unadjusted	6.70	4	-	
Days, adjusted	(0.87)	(4)	0.22	
Work Station	1.35	3	0.45	
Error	6.53	8	0.82	
Total	134.95	19		

[a] Significant at 1 percent.

S4-4. Lattice Designs

Consider a balanced incomplete block design with k^2 treatments arranged in $b = k(k + 1)$ blocks with k runs per block and $r = k + 1$ replicates. Such a design is called a **balanced lattice**. An example is shown in Table 7 for k^2=9 treatments in 12 blocks of 3 runs each. Notice that the blocks can be grouped into sets such that each set contains a complete replicate. The analysis of variance for the balanced lattice design proceeds like that for a balanced incomplete block design, except that a sum of squares for replicates is computed and removed from the sum of squares for blocks. Replicates will have k degrees of freedom and blocks will have k^2-1 degrees of freedom.

Lattice designs are frequently used in situations where there are a large number of treatment combinations. In order to reduce the size of the design, the experimenter may resort to **partially** balanced lattices. We very briefly describe some of these designs. Two replicates of a design for k^2 treatments in 2k blocks of k runs are called a **simple lattice**. For example, consider the first two replicates of the design in Table 7. The partial balance is easily seen, since, for example, treatment 2 appears in the same block with treatments 1, 3, 5, and 8, but does not appear at all with treatments 4, 6, 7, and 9. A lattice design with k^2 treatments in 3k blocks grouped into three replicates is called a **triple lattice**. An example would be the first three replicates in Table 7. A lattice design for k^2 treatments in 4k blocks arranged in four replicates is called a **quadruple lattice**.

Table 7. A 3 x 3 Balanced Lattice Design

Block	**Replicate 1**			**Block**	**Replicate 3**		
1	1	2	3	7	1	5	9
2	4	5	6	8	7	2	6
3	7	8	9	9	4	8	3
Block	**Replicate 2**			**Block**	**Replicate 4**		
1	1	4	7	10	1	8	6
2	2	5	8	11	4	2	9
3	3	6	9	12	7	5	3

There are other types of lattice designs that occasionally prove useful. For example, the **cubic lattice** design can be used for k^3 treatments in k^2 blocks of k runs. A lattice design for $k(k + 1)$ treatments in $k + 1$ blocks of size k is called a **rectangular lattice**. Details of the analysis of lattice designs and tables of plans are given in Cochran and Cox (1957).

Supplemental References

Bose, R. C. and T. Shimamoto (1952). "Classification and Analysis of Partially Balanced Incomplete Block Designs with Two Associate Classes". *Journal of the American Statistical Association*, Vol. 47, pp. 151-184.

Bose, R. C. W. H. Clatworthy, and S. S. Shrikhande (1954). *Tables of Partially Balanced Designs with Two Associate Classes*. Technical Bulletin No. 107, North Carolina Agricultural Experiment Station.

Smith, C. A. B. and H. O. Hartley (1948). "Construction of Youden Squares". *Journal of the Royal Statistical Society Series B*, Vol. 10, pp. 262-264.