Chapter 8. Supplemental Text Material

S8-1. Yates's Method for the Analysis of Fractional Factorials

Computer programs are almost always used for the analysis of fractional factorial. However, we may use Yates's algorithm for the analysis of a 2^{k-1} fractional factorial design by initially considering the data as having been obtained from a full factorial in k - 1 variables. The treatment combinations for this full factorial are listed in standard order, and then an additional letter (or letters) is added in parentheses to these treatment combinations to produce the actual treatment combinations run. Yates's algorithm then proceeds as usual. The actual effects estimated are identified by multiplying the effects associated with the treatment combinations in the full 2^{k-1} design by the defining relation of the 2^{k-1} fractional factorial.

The procedure is demonstrated in Table 1 below using the data from Example 8-1. This is a 2^{4-1} fractional. The data are arranged as a full 2^3 design in the factors A, B, and C. Then the letter d is added in parentheses to yield the actual treatment combinations that were performed. The effect estimated by, say, the second row in this table is $A + BCD$ since A and BCD are aliases.

Table 1. Yates's Algorithm for the 2_{IV}^{4-1} Fractional Factorial in Example 8-1

Treatment Combination	Response	(1)	(2)	(3)	Effect	Effect Estimate $2 \times (3) / N$
(1)	45	145	255	566	-	-
$a(d)$	100	110	311	76	$A+BCD$	19.00
$b(d)$	45	135	75	6	$B+ACD$	1.5
ab	65	176	1	-4	$AB+CD$	-1.00
$c(d)$	75	55	-35	56	$C+ABD$	14.00
ac	60	20	41	-74	$AC+BD$	-18.50
bc	80	-15	-15	76	$BC+AD$	19.00
$abc(d)$	96	16	16	66	$ABC+D$	16.50

S8-2 Alias Structures in Fractional Factorials and Other Designs

In this chapter we show how to find the alias relationships in a 2^{k-p} fractional factorial design by use of the complete defining relation. This method works well in simple designs, such as the regular fractions we use most frequently, but it does not work as well in more complex settings, such as some of the irregular fractions and partial fold-over designs. Furthermore, there are some fractional factorials that do not have defining relations, such as Plackett-Burman designs, so the defining relation method will not work for these types of designs at all.

Fortunately, there is a general method available that works satisfactorily in many situations. The method uses the polynomial or regression model representation of the model, say

$$\mathbf{y} = \mathbf{X}_1\beta_1 + \varepsilon$$

where $\mathbf{y}$ is an $n \times 1$ vector of the responses, $\mathbf{X}_1$ is an $n \times p_1$ matrix containing the design matrix expanded to the form of the model that the experimenter is fitting, β_1 is an $p_1 \times 1$ vector of the model parameters, and ε is an $n \times 1$ vector of errors. The least squares estimate of β_1 is

$$\hat{\beta}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}$$

Suppose that the **true** model is

$$\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon$$

where $\mathbf{X}_2$ is an $n \times p_2$ matrix containing additional variables that are not in the fitted model and β_2 is a $p_2 \times 1$ vector of the parameters associated with these variables. It can be easily shown that

$$\begin{aligned} E(\hat{\beta}_1) &= \beta_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\beta_2 \\ &= \beta_1 + \mathbf{A}\beta_2 \end{aligned}$$

where $\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$ is called the **alias matrix**. The elements of this matrix operating on β_2 identify the alias relationships for the parameters in the vector β_1.

We illustrate the application of this procedure with a familiar example. Suppose that we have conducted a 2^{3-1} design with defining relation $I = ABC$ or $I = x_1x_2x_3$. The model that the experimenter plans to fit is the main-effects-only model

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \varepsilon$$

In the notation defined above,

$$\beta_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \text{ and } \mathbf{X}_1 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Suppose that the true model contains all the two-factor interactions, so that

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \varepsilon$$

and

$$\beta_2 = \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix}, \text{ and } \mathbf{X}_2 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Now

$$(\mathbf{X}_1'\mathbf{X}_1)^{-1} = \frac{1}{4}\mathbf{I}_4 \text{ and } \mathbf{X}_1'\mathbf{X}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix}$$

Therefore

$$E(\hat{\beta}_1) = \beta_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\beta_2$$

$$E\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix}\begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix}$$

$$= \begin{bmatrix} \beta_0 \\ \beta_1 + \beta_{23} \\ \beta_2 + \beta_{13} \\ \beta_3 + \beta_{12} \end{bmatrix}$$

The interpretation of this, of course, is that each of the main effects is aliased with one of the two-factor interactions, which we know to be the case for this design. While this is a very simple example, the method is very general and can be applied to much more complex designs.

S8-3. More About Fold-Over and Partial Fold-Over of Fractional Factorials

In the textbook, we illustrate how a fractional factorial design can be augmented with additional runs to separate effects that are aliased. A fold-over is another design that is the same size as the original fraction. So if the original experiment has 16 runs, the fold-over will require another 16 runs.

Sometimes it is possible to augment a 2^{k-p} fractional factorial with fewer than an additional 2^{k-p} runs. This technique is generally referred to as a partial fold over of the original design.

For example, consider the 2^{5-2} design shown in Table 2. The alias structure for this design is shown below the table.

Table 2. A 2^{5-2} Design

Std ord	Run ord	Block	Factor A:A	Factor B:B	Factor C:C	Factor D:D	Factor E:E
2	1	Block 1	1	-1	-1	-1	-1
6	2	Block 1	1	-1	1	-1	1
3	3	Block 1	-1	1	-1	-1	1
1	4	Block 1	-1	-1	-1	1	1
8	5	Block 1	1	1	1	1	1
5	6	Block 1	-1	-1	1	1	-1
4	7	Block 1	1	1	-1	1	-1
7	8	Block 1	-1	1	1	-1	-1

[A] = A + BD + CE
[B] = B + AD + CDE
[C] = C + AE + BDE
[D] = D + AB + BCE
[E] = E + AC + BCD
[BC] = BC + DE + ABE + ACD
[BE] = BE + CD + ABC + ADE

Now suppose that after running the eight trials in Table 2, the largest effects are the main effects *A*, *B*, and *D*, and the *BC* + *DE* interaction. The experimenter believes that all other effects are negligible. Now this is a situation where fold-over of the original design is not an attractive alternative. Recall that when a resolution III design is folded over by reversing all the signs in the test matrix, the combined design is resolution IV. Consequently, the *BC* and *DE* interactions will still be aliased in the combined design. One could alternatively consider reversing signs in individual columns, but these approaches will essentially require that another eight runs be performed.

The experimenter wants to fit the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_4 x_4 + \beta_{23} x_2 x_3 + \beta_{45} x_4 x_5 + \varepsilon$$

where $x_1 = A, x_2 = B, x_3 = C, x_4 = D,$ and $x_5 = E$. Recall that a **partial fold-over** is a design containing fewer than eight runs that can be used to augment the original design and will allow the experimenter to fit this model. One way to select the runs for the partial fold-over is to select points from the remaining unused portion of the 2^5 such that the variances of the model coefficients in the above regression equation are minimized. This augmentation strategy is based on the idea of a **D-optimal design**, discussed in Chapter 11.

Design-Expert can utilize this strategy to find a partial fold-over. The design produced by the computer program is shown in Table 3. This design completely dealiases the *BC* and *DE* interactions.

Table 3. The Partially-Folded Fractional Design

Std ord	Run ord	Block	Factor A:A	Factor B:B	Factor C:C	Factor D:D	Factor E:E
2	1	Block 1	1	-1	-1	-1	-1
6	2	Block 1	1	-1	1	-1	1
3	3	Block 1	-1	1	-1	-1	1
1	4	Block 1	-1	-1	-1	1	1
8	5	Block 1	1	1	1	1	1
5	6	Block 1	-1	-1	1	1	-1
4	7	Block 1	1	1	-1	1	-1
7	8	Block 1	-1	1	1	-1	-1
9	9	Block 2	-1	-1	-1	-1	1
10	10	Block 2	1	1	1	1	-1
11	11	Block 2	-1	-1	1	-1	-1
12	12	Block 2	1	1	-1	1	1

Notice that the D-optimal partial fold-over design requires four additional trials. Furthermore, these trials are arranged in a second block that is orthogonal to the first block of eight trials.

This strategy is very useful in 16-run resolution IV designs, situations in which a full fold-over would require another 16 trials. Often a partial fold-over with four or eight runs can be used as an alternative. In many cases, a partial fold with only four runs over can be constricted using the D-optimal approach.

As a second example, consider the 2^{6-2} resolution IV design shown in Table 4. The alias structure for the design is shown below the table.

Table 4. A 2^{6-2} Resolution IV Design

Std ord	Run ord	Block	Factor A:A	Factor B:B	Factor C:C	Factor D:D	Factor E:E	Factor F:F
10	1	Block 1	1	-1	-1	1	1	1
11	2	Block 1	-1	1	-1	1	1	-1
2	3	Block 1	1	-1	-1	-1	1	-1
12	4	Block 1	1	1	-1	1	-1	-1
16	5	Block 1	1	1	1	1	1	1
15	6	Block 1	-1	1	1	1	-1	1
8	7	Block 1	1	1	1	-1	1	-1
7	8	Block 1	-1	1	1	-1	-1	-1
5	9	Block 1	-1	-1	1	-1	1	1
1	10	Block 1	-1	-1	-1	-1	-1	-1
6	11	Block 1	1	-1	1	-1	-1	1
4	12	Block 1	1	1	-1	-1	-1	1
14	13	Block 1	1	-1	1	1	-1	-1
13	14	Block 1	-1	-1	1	1	1	-1
9	15	Block 1	-1	-1	-1	1	-1	1
3	16	Block 1	-1	1	-1	-1	1	1

[A] = A + BCE + DEF
[B] = B + ACE + CDF
[C] = C + ABE + BDF
[D] = D + AEF + BCF
[E] = E + ABC + ADF
[F] = F + ADE + BCD
[AB] = AB + CE
[AC] = AC + BE
[AD] = AD + EF
[AE] = AE + BC + DF
[AF] = AF + DE
[BD] = BD + CF
[BF] = BF + CD
[ABD] = ABD + ACF + BEF + CDE
[ABF] = ABF + ACD + BDE + CEF

Suppose that the main effects of factors *A, B, C,* and *E* are large, along with the *AB* + *CE* interaction chain. A full fold-over of this design would involve reversing the signs in columns *B, C, D, E,* and *F*. This would, of course, require another 16 trials. A standard partial fold using the method described in the textbook would require 8 additional runs. The D-optimal partial fold-over approach requires only four additional runs. The augmented design, obtained from Design-Expert, is shown in Table 5. These four runs form a second block that is orthogonal to the first block of 16 runs, and allows the interactions of interest in the original alias chain to be separately estimated.

Remember that partial fold over designs are irregular fractions. They are not orthogonal and as a result, the effect estimates are correlated. This correlation between effect estimates causes inflation in the standard errors of the effects; that is, the effects are not estimated as precisely as they would have been in an orthogonal design. However, this disadvantage may be offset by the decrease in the number of runs that the partial fold over requires.

Table 5. The Partial Fold-Over

Std	Run	Block	Factor A:A	Factor B:B	Factor C:C	Factor D:D	Factor E:E	Factor F:F
12	1	Block 1	1	1	-1	1	-1	-1
15	2	Block 1	-1	1	1	1	-1	1
2	3	Block 1	1	-1	-1	-1	1	-1
9	4	Block 1	-1	-1	-1	1	-1	1
5	5	Block 1	-1	-1	1	-1	1	1
8	6	Block 1	1	1	1	-1	1	-1
11	7	Block 1	-1	1	-1	1	1	-1
14	8	Block 1	1	-1	1	1	-1	-1
13	9	Block 1	-1	-1	1	1	1	-1
4	10	Block 1	1	1	-1	-1	-1	1
10	11	Block 1	1	-1	-1	1	1	1
6	12	Block 1	1	-1	1	-1	-1	1
7	13	Block 1	-1	1	1	-1	-1	-1
16	14	Block 1	1	1	1	1	1	1
3	15	Block 1	-1	1	-1	-1	1	1
1	16	Block 1	-1	-1	-1	-1	-1	-1
17	17	Block 2	1	-1	1	-1	-1	-1
18	18	Block 2	-1	1	-1	-1	-1	-1
19	19	Block 2	-1	-1	1	1	1	1
20	20	Block 2	1	1	-1	1	1	1