Chapter 9. Supplemental Text Material

S9-1. Yates's Algorithm for the 3^k Design

Computer methods are used almost exclusively for the analysis of factorial and fractional designs. However, Yates's algorithm can be modified for use in the 3^k factorial design. We will illustrate the procedure using the data in Example 5-1. The data for this example are originally given in Table 5-1. This is a 3^2 design used to investigate the effect of material type (*A*) and temperature (*B*) on the life of a battery. There are $n = 4$ replicates.

The Yates' procedure is displayed in Table 1 below. The treatment combinations are written down in standard order; that is, the factors are introduced one at a time, each level being combined successively with every set of factor levels above it in the table. (The standard order for a 3^3 design would be 000, 100, 200, 010, 110, 210, 020, 120, 220, 001, . . .). The Response column contains the total of all observations taken under the corresponding treatment combination. The entries in column (1) are computed as follows. The first third of the column consists of the sums of each of the three sets of three values in the Response column. The second third of the column is the third minus the first observation in the same set of three. This operation computes the linear component of the effect. The last third of the column is obtained by taking the sum of the first and third value minus twice the second in each set of three observations. This computes the quadratic component. For example, in column (1), the second, fifth, and eighth entries are 229 + 479 + 583 = 1291, -229 + 583 = 354, and 229 - (2)(479) + 583 = -146, respectively. Column (2) is obtained similarly from column (1). In general, *k* columns must be constructed.

The Effects column is determined by converting the treatment combinations at the left of the row into corresponding effects. That is, 10 represents the linear effect of *A*, A_L, and 11 represents the AB_{LXL} component of the *AB* interaction. The entries in the Divisor column are found from

$$2^r 3^t n$$

where *r* is the number of factors in the effect considered, *t* is the number of factors in the experiment minus the number of linear terms in this effect, and *n* is the number of replicates. For example, B_L has the divisor $2^1 \times 3^1 \times 4 = 24$.

The sums of squares are obtained by squaring the element in column (2) and dividing by the corresponding entry in the Divisor column. The Sum of Squares column now contains all of the required quantities to construct an analysis of variance table if both of the design factors *A* and *B* are quantitative. However, in this example, factor *A* (material type) is qualitative; thus, the linear and quadratic partitioning of *A* is not appropriate. Individual observations are used to compute the total sum of squares, and the error sum of squares is obtained by subtraction.

Table 1. Yates's Algorithm for the 3^2 Design in Example 5-1

Treatment Combination	Response	(1)	(2)	Effects	Divisor	Sum of Squares
00	539	1738	3799			
10	623	1291	503	A_L	$2^1 \times 3^1 \times 4$	10, 542.04
20	576	770	-101	A_Q	$2^1 \times 3^2 \times 4$	141.68
01	229	37	-968	B_L	$2^1 \times 3^1 \times 4$	39,042.66
11	479	354	75	AB_{LXL}	$2^2 \times 3^0 \times 4$	351.56
21	583	112	307	AB_{QXL}	$2^2 \times 3^1 \times 4$	1,963.52
02	230	-131	-74	B_Q	$2^1 \times 3^2 \times 4$	76.96
12	198	-146	-559	AB_{LXQ}	$2^2 \times 3^1 \times 4$	6,510.02
22	342	176	337	AB_{QXQ}	$2^2 \times 3^2 \times 4$	788.67

The analysis of variance is summarized in Table 2. This is essentially the same results that were obtained by conventional analysis of variance methods in Example 5-1.

Table 2. Analysis of Variance for the 3^2 Design in Example 5-1

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-value
$A = A_L \times A_Q$	10, 683.72	2	5,341.86	7.91	0.0020
B, Temperature	39,118.72	2	19,558.36	28.97	<0.0001
B_L	(39, 042.67)	(1)	39,042.67	57.82	<0.0001
B_Q	(76.05)	(1)	76.05	0.12	0.7314
AB	9,613.78	4	2,403.44	3.576	0.0186
$A \times B_L = AB_{LXL} + AB_{QXL}$	(2,315.08)	(2)	1,157.54	1.71	0.1999
$A \times B_Q = AB_{LXQ} + AB_{QXQ}$	(7,298.70)	(2)	3,649.75	5.41	0.0106
Error	18,230.75	27	675.21		
Total	77,646.97	35			

S9-2. Aliasing in Three-Level and Mixed-Level Designs

In the supplemental text material for Chapter 8 (Section 8-2) we gave a general method for finding the alias relationships for a fractional factorial design. 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,

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

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

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

The **true** model is assumed to be

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

where $\mathbf{X}_2$ is an $n \times p_2$ matrix containing additional variables not in the fitted model and β_2 is a $p_2 \times 1$ vector of the parameters associated with these additional variables. The parameter estimates in the fitted model are not unbiased, since

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

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

This procedure can be used to find the alias relationships in three-level and mixed-level designs. We now present two examples.

<u>Example 1</u>

Suppose that we have conducted an experiment using a 3^2 design, and that we are interested in fitting the following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11}(x_1^2 - \bar{x}_1^2) + \beta_{22}(x_2^2 - \bar{x}_2^2) + \varepsilon$$

This is a complete quadratic polynomial. The pure second-order terms are written in a form that orthogonalizes these terms with the intercept. We will find the aliases in the parameter estimates if the true model is a reduced cubic, say

$$\begin{aligned} y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11}(x_1^2 - \bar{x}_1^2) + \beta_{22}(x_2^2 - \bar{x}_2^2) \\ & + \beta_{111} x_1^3 + \beta_{222} x_2^3 + \beta_{122} x_1 x_2^2 + \varepsilon \end{aligned}$$

Now in the notation used above, the vector β_1 and the matrix $\mathbf{X}_1$ are defined as follows:

$$\beta_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \\ \beta_{11} \\ \beta_{22} \end{bmatrix}, \quad \text{and} \quad \mathbf{X}_1 = \begin{bmatrix} 1 & -1 & -1 & 1 & 1/3 & 1/3 \\ 1 & -1 & 0 & 0 & 1/3 & -2/3 \\ 1 & -1 & 1 & -1 & 1/3 & 1/3 \\ 1 & 0 & -1 & 0 & -2/3 & 1/3 \\ 1 & 0 & 0 & 0 & -2/3 & -2/3 \\ 1 & 0 & 1 & 0 & -2/3 & 1/3 \\ 1 & 1 & -1 & -1 & 1/3 & 1/3 \\ 1 & 1 & 0 & 0 & 1/3 & -2/3 \\ 1 & 1 & 1 & 1 & 1/3 & 1/3 \end{bmatrix}$$

Now

$$\mathbf{X}_1'\mathbf{X}_1 = \begin{bmatrix} 9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

and the other quantities we require are

$$\mathbf{X}_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad \beta_2 = \begin{bmatrix} \beta_{111} \\ \beta_{222} \\ \beta_{122} \end{bmatrix}, \quad \text{and} \quad \mathbf{X}_1'\mathbf{X}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 6 & 0 & 4 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The expected value of the fitted model parameters is

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

or

$$E\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_{12}\\ \hat{\beta}_{11}\\ \hat{\beta}_{22}\end{bmatrix}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_{12}\\ \beta_{11}\\ \beta_{22}\end{bmatrix}+\begin{bmatrix}9&0&0&0&0&0\\0&6&0&0&0&0\\0&0&6&0&0&0\\0&0&0&4&0&0\\0&0&0&0&2&0\\0&0&0&0&0&2\end{bmatrix}^{-1}\begin{bmatrix}0&0&0\\6&0&4\\0&6&0\\0&0&0\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}\beta_{111}\\ \beta_{222}\\ \beta_{122}\end{bmatrix}$$

The alias matrix turns out to be

$$\mathbf{A}=\begin{bmatrix}0&0&0\\1&0&2/3\\0&1&0\\0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

This leads to the following alias relationships:

$$E(\hat{\beta}_0)=\beta_0$$
$$E(\hat{\beta}_1)=\beta_1+\beta_{111}+(2/3)\beta_{122}$$
$$E(\hat{\beta}_2)=\beta_2+\beta_{222}$$
$$E(\hat{\beta}_{12})=\beta_{12}$$
$$E(\hat{\beta}_{11})=\beta_{11}$$
$$E(\hat{\beta}_{22})=\beta_{22}$$

Example 2

This procedure is very useful when the design is a mixed-level fractional factorial. For example, consider the mixed-level design in Table 9-10 of the textbook. This design can accommodate four two-level factors and a single three-level factor. The resulting resolution III fractional factorial is shown in Table 3.

Since the design is resolution III, the appropriate model contains the main effects

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{55}(x_5^2-\bar{x}_5^2)+\varepsilon,$$

where the model terms

$$\beta_5x_5 \text{ and } \beta_{55}(x_5^2-\bar{x}_5^2)$$

represent the linear and quadratic effects of the three-level factor x_5. The quadratic effect of x_5 is defined so that it will be orthogonal to the intercept term in the model.

Table 3. A Mixed-Level Resolution III Fractional Factorial

x_1	x_2	x_3	x_4	x_5
-1	1	1	-1	-1
1	-1	-1	1	-1
-1	-1	1	1	0
1	1	-1	-1	0
-1	1	-1	1	0
1	-1	1	-1	0
-1	-1	-1	-1	1
1	1	1	1	1

Now suppose that the **true** model has interaction:

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_{55}(x_5^2 - \bar{x}_5^2) \\ &+ \beta_{12} x_1 x_2 + \beta_{15} x_1 x_5 + \beta_{155} x_1 (x_5^2 - \bar{x}_5^2) + \varepsilon \end{aligned}$$

So in the true model the two-level factors x_1 and x_2 interact, and x_1 interacts with both the linear and quadratic effects of the three-level factor x_5. Straightforward, but tedious application of the procedure described above leads to the alias matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and the alias relationships are computed from

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

This results in

$$E(\hat{\beta}_0) = \beta_0$$
$$E(\hat{\beta}_1) = \beta_1$$
$$E(\hat{\beta}_2) = \beta_2 + (1/2)\beta_{15}$$
$$E(\hat{\beta}_3) = \beta_3 + (1/2)\beta_{15}$$
$$E(\hat{\beta}_4) = \beta_4 + (1/2)\beta_{155}$$
$$E(\hat{\beta}_5) = \beta_5 + \beta_{12}$$
$$E(\hat{\beta}_{55}) = \beta_{55}$$

The linear and quadratic components of the interaction between x_1 and x_5 are aliased with the main effects of $x_2, x_3,$ and x_4, and the x_1x_2 interaction aliases the linear component of the main effect of x_5.