Chapter 8
Experimental Design Models

The design of an experiment is extremely important. It identifies the class of linear models that can be used to analyze the results of the experiment. Perhaps more accurately, the design identifies large classes of models that are inappropriate for the analysis of the experiment. Finding appropriate models can be difficult even in well-designed experiments.

In this chapter we examine three models used to analyze results obtained from three specific experimental designs. The designs are the completely randomized design, the randomized complete block design, and the Latin square design. We also examine a particularly effective method of defining treatments for situations in which several factors are of interest.

Randomization, that is, the random application of treatments to experimental units, provides a philosophical justification for inferring that the treatments actually cause the experimental results. In addition to the models considered here, Appendix G examines estimation for completely randomized designs and randomized complete block designs under an alternative error structure that is derived from the act of randomization.

Blocking is one of the most important concepts in experimental design. Blocking is used to reduce variability for the comparison of treatments. Blocking consists of grouping experimental units that will act similarly into blocks and then (randomly) applying the treatments to the units in each block. A randomized complete block design is one in which each block has exactly the same number of units as there are treatments. Each treatment is applied once in each block. If there are more units than treatments, we can have a complete block with replications. If there are fewer units than treatments, we have an incomplete block design. A major issue in the subject of experimental design is how one should assign treatments to blocks in an incomplete block design. Latin square designs are complete block designs that incorporate two separate forms of blocking.

The analysis of balanced incomplete block designs is derived in Section 9.4. Alternative methods for blocking designs are examined in the exercises for Section 11.1 and in Section 12.11. For an excellent discussion of the concepts underlying the design of experiments, see Fisher (1935) or Cox (1958). For more detailed

R. Christensen, *Plane Answers to Complex Questions: The Theory of Linear Models*, Springer Texts in Statistics, DOI 10.1007/978-1-4419-9816-3_8,

discussion of the design and analysis of experiments, there are many good books including Hinkelmann and Kempthorne (1994), Cochran and Cox (1957), Casella (2008), or [blush, blush] Christensen (1996a).

8.1 Completely Randomized Designs

The simplest experimental design is the *completely randomized design* (CRD). It involves no blocking. The experimental technique is simply to decide how many observations are to be taken on each treatment, obtain an adequate number of experimental units, and apply the treatments to the units (randomly). The standard model for this design is

$$y_{ij} = \mu + \alpha_i + e_{ij},$$

$i = 1, \ldots, a,\ j = 1, \ldots, N_i$, $\mathrm{E}(e_{ij}) = 0$, $\mathrm{Var}(e_{ij}) = \sigma^2$, $\mathrm{Cov}(e_{ij}, e_{i'j'}) = 0$ if $(i, j) \neq (i', j')$. This is a one-way ANOVA model and is analyzed as such.

8.2 Randomized Complete Block Designs: Usual Theory

The model usually assumed for a *randomized complete block design* (RCB) is

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij},$$

$i = 1, \ldots, a,\ j = 1, \ldots, b$, $\mathrm{E}(e_{ij}) = 0$, $\mathrm{Var}(e_{ij}) = \sigma^2$, $\mathrm{Cov}(e_{ij}, e_{i'j'}) = 0$ if $(i, j) \neq (i', j')$. The β_js stand for an additive effect for each block; the α_is are an additive effect for each treatment. It is assumed that any block-treatment interaction is error so that an estimate of σ^2 is available. This randomized complete block model is just a two-way ANOVA without interaction.

In this chapter, we are presenting models that are generally used for analyzing experiments conducted with certain standard experimental designs. Many people believe that these models are useful only because they are good approximations to linear models derived from the random application of the treatments to experimental units. This randomization theory leads to the conclusion that there is no valid test for block effects. On the other hand, there is nothing in the two-way ANOVA model given above to keep one from testing block effects. It is my opinion that this contradiction arises simply because the two-way ANOVA model is not very appropriate for a randomized complete block design. For example, a basic idea of blocking is that results within a block should be more alike than results in different blocks. The two-way ANOVA, with only an additive effect for blocks, is a very simplistic model. Another key idea of blocking is that it reduces the variability of treatment comparisons. It is not clear how the existence of additive block effects reduces variability in a two-way ANOVA. The question of how blocking achieves variance reduction

is addressed in Exercise 8.1. Exercises 11.4 through 11.6 discuss alternative models for complete block designs.

Appendix G defines linear models for completely randomized designs and randomized complete block designs that are based on randomization theory. It is shown, using Theorem 10.4.5 (or Proposition 2.7.5), that least squares estimates are BLUEs for the randomization theory models.

Exercise 8.1 Using a randomized complete block design is supposed to reduce the variability of treatment comparisons. If the randomized complete block model is taken as

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad e_{ij}\text{s i.i.d. } N(0,\sigma^2),$$

$i = 1,\ldots,a,\ j = 1,\ldots,b$, argue that the corresponding variance for a completely randomized design should be $\sigma^2 + \sum_{j=1}^{b}(\mu_j - \bar{\mu}_\cdot)^2/b$, where $\mu_i = \mu + \beta_j$.

Hint: Figure out what population a completely randomized design would have to be sampled from.

8.3 Latin Square Designs

A *Latin square* is a design that allows for treatment effects and two different kinds of block effects. The number of treatments must equal the number of blocks of each kind. On occasion, the design is used with two kinds of treatments (each with the same number of levels) and one block effect.

EXAMPLE 8.3.1. A 4×4 Latin square has four treatments, say, T_1, T_2, T_3, T_4. Consider the block effects as row effects, R_1, R_2, R_3, and R_4 and column effects C_1, C_2, C_3, and C_4. We can diagram one example of a 4×4 Latin square as

	C_1	C_2	C_3	C_4
R_1	T_1	T_2	T_3	T_4
R_2	T_2	T_3	T_4	T_1
R_3	T_3	T_4	T_1	T_2
R_4	T_4	T_1	T_2	T_3

The key idea is that each treatment occurs once in each row and once in each column.

The model for an $a \times a$ Latin square design is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + e_{ijk},$$

$\mathrm{E}(e_{ijk}) = 0$, $\mathrm{Var}(e_{ijk}) = \sigma^2$, $\mathrm{Cov}(e_{ijk}, e_{i'j'k'}) = 0$ if $(i,j,k) \neq (i',j',k')$. The correspondence between the rows, columns, and treatments and the effects in the model is: α_i is the effect for row R_i, β_j is the effect for column C_j, and γ_k is the effect for treatment T_k. The subscripting in this model is unusual. The key point is that in a

Latin square if you know the row and the column, the Latin square design tells you what the treatment is, so the three subscripts do not vary freely. In particular, we can specify $i = 1,\ldots,a$, $j = 1,\ldots,a$, $k \in \{1,2,\ldots,a\}$ and $k = f(i,j)$, where for each i, $f(i,j)$ is a one-to-one function of $\{1,2,\ldots,a\}$ onto itself, and the same is true for each j. As in the case of the randomized complete block design, this model makes no distinction between treatment effects and the two sets of block effects.

To derive the analysis of the Latin square model, we need to show that after fitting μ, the spaces for the three main effects are orthogonal. Before proceeding, note again that the terms y_{ijk} are overindexed. There are a^2 terms but a^3 possible combinations of the indices. Any two of the indices serve to identify all of the observations. For example, the mean of all a^2 observations is

$$\bar{y}_{\cdots} = \frac{1}{a^2}\sum_{i=1}^{a}\sum_{j=1}^{a} y_{ijk} = \frac{1}{a^2}\sum_{i=1}^{a}\sum_{k=1}^{a} y_{ijk} = \frac{1}{a^2}\sum_{j=1}^{a}\sum_{k=1}^{a} y_{ijk}.$$

We will use triple index notation to describe the rows of the model matrix. Write the model matrix as $X = [X_0, X_1, \ldots, X_{3a}]$, where $X_0 = J$,

$$\begin{aligned} X_r &= [u_{ijk}], & u_{ijk} &= \delta_{ir}, & r &= 1,\ldots,a,\\ X_{a+s} &= [u_{ijk}], & u_{ijk} &= \delta_{js}, & s &= 1,\ldots,a,\\ X_{2a+t} &= [u_{ijk}], & u_{ijk} &= \delta_{kt}, & t &= 1,\ldots,a. \end{aligned}$$

EXAMPLE 8.3.2. The model for the 4×4 Latin square of Example 8.3.1 is

$$\begin{bmatrix} y_{111}\\ y_{122}\\ y_{133}\\ y_{144}\\ y_{212}\\ y_{223}\\ y_{234}\\ y_{241}\\ y_{313}\\ y_{324}\\ y_{331}\\ y_{342}\\ y_{414}\\ y_{421}\\ y_{432}\\ y_{443} \end{bmatrix} = \begin{bmatrix} 1&1&0&0&0&1&0&0&0&1&0&0&0\\ 1&1&0&0&0&0&1&0&0&0&1&0&0\\ 1&1&0&0&0&0&0&1&0&0&0&1&0\\ 1&1&0&0&0&0&0&0&1&0&0&0&1\\ 1&0&1&0&0&1&0&0&0&0&1&0&0\\ 1&0&1&0&0&0&1&0&0&0&0&1&0\\ 1&0&1&0&0&0&0&1&0&0&0&0&1\\ 1&0&1&0&0&0&0&0&1&1&0&0&0\\ 1&0&0&1&0&1&0&0&0&0&0&1&0\\ 1&0&0&1&0&0&1&0&0&0&0&0&1\\ 1&0&0&1&0&0&0&1&0&1&0&0&0\\ 1&0&0&1&0&0&0&0&1&0&1&0&0\\ 1&0&0&0&1&1&0&0&0&0&0&0&1\\ 1&0&0&0&1&0&1&0&0&1&0&0&0\\ 1&0&0&0&1&0&0&1&0&0&1&0&0\\ 1&0&0&0&1&0&0&0&1&0&0&1&0 \end{bmatrix} \begin{bmatrix} \mu\\ \alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \gamma_1\\ \gamma_2\\ \gamma_3\\ \gamma_4 \end{bmatrix} + e.$$

Orthogonalizing columns 1 to $3a$ of X with respect to J gives the matrix Z with columns

$$Z_0 = X_0$$
$$Z_i = X_i - \frac{X_i'J}{J'J}J = X_i - \frac{a}{a^2}J, \quad i = 1,\ldots,3a.$$

The three spaces $C(Z_1,\ldots,Z_a)$, $C(Z_{a+1},\ldots,Z_{2a})$, and $C(Z_{2a+1},\ldots,Z_{3a})$ are orthogonal. For example, with $r = 1,\ldots,a$ and $t = 1,\ldots,a$,

$$\begin{aligned} Z_r'Z_{2a+t} &= \sum_{i=1}^{a}\sum_{k=1}^{a}\left(\delta_{ir} - \frac{1}{a}\right)\left(\delta_{kt} - \frac{1}{a}\right) \\ &= \sum_{i=1}^{a}\sum_{k=1}^{a}\delta_{ir}\delta_{kt} - \sum_{i=1}^{a}\sum_{k=1}^{a}\delta_{ir}/a - \sum_{i=1}^{a}\sum_{k=1}^{a}\delta_{kt}/a + \sum_{i=1}^{a}\sum_{k=1}^{a}1/a^2 \\ &= 1-1-1+1 \\ &= 0. \end{aligned}$$

Similar computations establish the other orthogonality relationships.

Because of the orthogonality, the sum of squares for dropping, say, the α_is from the model is just the sum of squares for dropping the α_is from a one-way ANOVA model that ignores the β_j and γ_k effects. The ANOVA table is

Source	df	ANOVA SS	E(MS)
μ	1	$a^2\bar{y}_{\cdots}^2$	$\sigma^2 + a^2(\mu + \bar{\alpha}_\cdot + \bar{\beta}_\cdot + \bar{\gamma}_\cdot)^2$
α	$a-1$	$a\sum_{i=1}^{a}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2$	$\sigma^2 + \frac{a}{a-1}\sum_{i=1}^{a}(\alpha_i - \bar{\alpha}_\cdot)^2$
β	$a-1$	$a\sum_{j=1}^{a}(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2$	$\sigma^2 + \frac{a}{a-1}\sum_{j=1}^{a}(\beta_j - \bar{\beta}_\cdot)^2$
γ	$a-1$	$a\sum_{k=1}^{a}(\bar{y}_{\cdot\cdot k} - \bar{y}_{\cdots})^2$	$\sigma^2 + \frac{a}{a-1}\sum_{k=1}^{a}(\gamma_k - \bar{\gamma}_\cdot)^2$
Error	$(a-2)(a-1)$	by subtraction	σ^2
Total	a^2	$\sum_{i=1}^{a}\sum_{j=1}^{a}y_{ijk}^2$	

Estimation and testing in one of the treatment (block) spaces depends only on the appropriate projection operator. Since we have the usual one-way ANOVA projection operators, estimation and testing are performed in the usual way.

The Latin square model assumes that the $(\alpha\beta)$, $(\alpha\gamma)$, $(\beta\gamma)$, and $(\alpha\beta\gamma)$ interactions are nonexistent. This assumption is necessary in order to obtain an estimate of error. If an outside estimate of σ^2 is available, it might be hoped that the interactions could be examined. Unfortunately, it is impossible to tell from which interaction the degrees of freedom called "error" come from. For example, in the 4×4 Latin square of the example, the $(\alpha\beta)$ interaction can be broken up into 3 degrees of freedom for γ and 6 degrees of freedom for error. Since a similar result holds for each of the interactions, the 6 degrees of freedom for error involve all of the interactions.

Exercise 8.2 In the 4×4 Latin square of the examples, show that the 9 degrees of freedom for $(\alpha\beta)$ interaction are being divided into 3 degrees of freedom for γ and 6 degrees of freedom for error.

A *Graeco-Latin square* is a Latin square in which a second group of a treatments has been applied so that each treatment in the second group occurs once in each row, once in each column, and once with each of the treatments from the first group.

Exercise 8.3 Derive the analysis for the Graeco-Latin square given below. Use the model $y_{hijk} = \mu + \alpha_h + \beta_i + \gamma_j + \eta_k + e_{hijk}$.

	C_1	C_2	C_3	C_4	C_5
R_1	$T_1\tau_1$	$T_2\tau_3$	$T_3\tau_5$	$T_4\tau_2$	$T_5\tau_4$
R_2	$T_2\tau_2$	$T_3\tau_4$	$T_4\tau_1$	$T_5\tau_3$	$T_1\tau_5$
R_3	$T_3\tau_3$	$T_4\tau_5$	$T_5\tau_2$	$T_1\tau_4$	$T_2\tau_1$
R_4	$T_4\tau_4$	$T_5\tau_1$	$T_1\tau_3$	$T_2\tau_5$	$T_3\tau_2$
R_5	$T_5\tau_5$	$T_1\tau_2$	$T_2\tau_4$	$T_3\tau_1$	$T_4\tau_3$

Extend the analysis to an arbitrary $a \times a$ Graeco-Latin square.

8.4 Factorial Treatment Structures

For each experimental design considered in this chapter, we have assumed the existence of "a" treatments. Sometimes the treatments are chosen in such a way that the treatment space can be conveniently broken into orthogonal subspaces. One of the most common methods of doing this is to choose treatments with factorial structure.

Suppose that two or more different kinds of treatments are of interest. Each kind of treatment is called a *factor*. Each factor is of interest at some number of different levels. A very efficient way of gaining information on all of the levels of all of the factors is to use what is called a *factorial design*. In a factorial design the treatments are taken to be all possible combinations of the levels of the different factors. Since a factorial design refers only to the treatment structure, factorial designs can be used with all of the designs considered in this chapter as well as with balanced incomplete block designs (cf. Section 9.4) and split plot designs (cf. Section 11.3).

EXAMPLE 8.4.1. An experiment is to be conducted examining the effects of fertilizer on potato yields. Of interest are two kinds of fertilizer, a nitrogen-based fertilizer and a phosphate-based fertilizer. The two types of fertilizer are factors. The nitrogen fertilizer is to be examined at two levels: no nitrogen fertilizer (n_0) and a single dose of nitrogen fertilizer (n_1). The phosphate fertilizer has three levels: no phosphate fertilizer (p_0), a single dose of phosphate (p_1), and a double dose of phosphate (p_2). The treatments are taken to be all six of the possible combinations:

$$n_0p_0 \quad n_0p_1 \quad n_0p_2 \quad n_1p_0 \quad n_1p_1 \quad n_1p_2.$$

The use of treatments with factorial structure has a number of advantages. One is that it allows study of the interrelationships (interactions) between the factors. In

Example 8.4.1, it is possible to examine whether the levels of phosphate have different effects depending on whether or not nitrogen was applied. Another advantage is that if there are no interactions, the experimental material is used very efficiently. Suppose that there is no interaction between nitrogen and phosphate. The effect of nitrogen is the difference in yields between experimental units that have the same level of phosphate but different levels of nitrogen, that is,

$$n_0p_0 - n_1p_0,$$

$$n_0p_1 - n_1p_1,$$

and

$$n_0p_2 - n_1p_2.$$

In each case, the only difference in the pair of treatments is the difference in nitrogen. At the same time, the difference in, say, the effects of p_1 and p_2 can be examined by looking at the differences

$$n_0p_1 - n_0p_2$$

and

$$n_1p_1 - n_1p_2.$$

The same treatments are used to obtain estimates of both the nitrogen effect and the phosphate effect.

We now examine the analysis of designs having factorial treatment structure. Consider a randomized complete block design with c blocks, where the treatments are all combinations of two factors, say A and B. Suppose factor A has a levels and factor B has b levels. The total number of treatments is ab. Rewriting the model of Section 2 with τ_h denoting treatment effects and ξ_k denoting block effects, we get

$$y_{hk} = \mu + \tau_h + \xi_k + e_{hk}, \tag{1}$$

$h = 1, \ldots, ab$, $k = 1, \ldots, c$. In this model, each treatment is indicated by an index h. Since the treatments consist of all combinations of two factors, it makes sense to use two indices to identify treatments: one index for each factor. With this idea we can rewrite model (1) as

$$y_{ijk} = \mu + \tau_{ij} + \xi_k + e_{ijk}, \tag{2}$$

$i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, c$.

Exercise 8.4 For model (1), $C(M_\tau)$ is given by Proposition 4.2.3. What is $C(M_\tau)$ in the notation of model (2)?

The orthogonal breakdown of the treatment space follows from a reparameterization of model (2). Model (2) can be rewritten as

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \xi_k + e_{ijk}. \tag{3}$$

The new parameterization is simply $\tau_{ij} = \alpha_i + \beta_j + (\alpha\beta)_{ij}$. Using the ideas of Sections 7.1 and 7.2, the treatment space $C(M_\tau)$ can be broken up into three orthogonal subspaces: one for factor A, $C(M_\alpha)$, one for factor B, $C(M_\beta)$, and one for interaction, $C(M_{\alpha\beta})$. The analysis of model (3) follows along the lines of Chapter 7. Model (3) is just a balanced three-way ANOVA in which some of the interactions (namely, any interactions that involve blocks) have been thrown into the error.

In practice, it is particularly important to be able to relate contrasts in the treatments (τ_{ij}s) to contrasts in the main effects (α_is and β_js) and contrasts in the interactions ($(\alpha\beta)_{ij}$s). The relationship is demonstrated in Exercise 8.5. The relationship is illustrated in the following example.

EXAMPLE 8.4.2. Consider the treatments of Example 8.4.1. There is one contrast in nitrogen,

	n_0	n_1
N	1	−1

The two orthogonal polynomial contrasts in phosphate are:

	p_0	p_1	p_2
P linear	−1	0	1
P quadratic	1	−2	1

The main effect contrasts define two orthogonal interaction contrasts: $N-P$ linear

	p_0	p_1	p_2
n_0	−1	0	1
n_1	1	0	−1

and $N-P$ quadratic

	p_0	p_1	p_2
n_0	1	−2	1
n_1	−1	2	−1

The corresponding contrasts in the six treatments are:

	n_0p_0	n_0p_1	n_0p_2	n_1p_0	n_1p_1	n_1p_2
N	1	1	1	−1	−1	−1
P linear	−1	0	1	−1	0	1
P quadradic	1	−2	1	1	−2	1
$N-P$ linear	−1	0	1	1	0	−1
$N-P$ quadradic	1	−2	1	−1	2	−1

It is easily verified that these five contrasts are orthogonal.

Exercise 8.5 Show that the contrasts in the τ_{ij}s corresponding to the contrasts $\sum\lambda_i\alpha_i$ and $\sum\sum\lambda_i\eta_j(\alpha\beta)_{ij}$ are $\sum\sum\lambda_i\tau_{ij}$ and $\sum\sum\lambda_i\eta_j\tau_{ij}$, respectively.

Hint: Any contrast in the τ_{ij}s corresponds to a vector in $C(M_\tau)$, just as any contrast in the α_is corresponds to a vector in $C(M_\alpha) \subset C(M_\tau)$. Recall that contrasts are only defined up to constant multiples; and that contrasts in the α_is also involve the interactions when interaction exists.

8.5 More on Factorial Treatment Structures

We now present a more theoretical discussion of factorial treatment structures and some interesting new models.

For two factors, say α at s levels and η at t levels, there are a total of $p \equiv st$ treatment groups. Start by considering a linear model for the data $y_{ijk} = (\alpha\eta)_{ij} + e_{ijk}$. In matrix form, the linear model $Y = X\beta + e$ has X as the indicator matrix used for a one-way ANOVA model and $\beta = [(\alpha\eta)_{11}, \ldots, (\alpha\eta)_{st}]'$ containing a separate effect for every combination of the two factors.

More generally, if we have r factors α_i with s_i levels respectively, the factorial structure defines $p = \prod_{i=1}^{r} s_i$ groups. We again take X to be the one-way ANOVA indicator matrix and define $\beta = [\alpha_{i_1,\ldots,i_r}]$ as a vector providing a separate effect for each combination of the factors.

The idea is that factorial models are defined by specifying a linear structure for the group parameters. This consists of putting a linear constraint on β, say $\beta = U_{p\times q}\gamma$ for some known matrix U, cf. Section 3.3. For example, in the two-factor model, one such linear constraint is to force an additive main effects model for the data, $y_{ijk} = \mu + \alpha_i + \eta_j + e_{ijk}$. In matrix terms using Kronecker products, this amounts to specifying that

$$\beta = \Big([J_s \otimes J_t], [I_s \otimes J_t], [J_s \otimes I_t]\Big) \begin{bmatrix} \mu \\ \alpha \\ \eta \end{bmatrix} \equiv U\gamma, \tag{1}$$

where $\alpha = (\alpha_1, \ldots, \alpha_s)'$ and $\eta = (\eta_1, \ldots, \eta_t)'$. As in Section 3.3, the linear model for the data has been transformed to $Y = XU\gamma + e$, where the model matrix is now XU, which is a reduced model relative to the original, i.e., $C(XU) \subset C(X)$. Note that if we define a linear structure for the parameters, $\beta = U\gamma$ and a reduced structure, i.e., $\beta = U_0\gamma_0$, where $C(U_0) \subset C(U)$, then this also defines a reduced linear model for the data in that $C(XU_0) \subset C(XU)$.

I have no difficulty considering all such models to be factorial models, but McCullagh (2000) proposes a more stringent definition involving "selection invariance." The idea is that if you drop various indices from various factors, the model should somehow remain invariant. This idea can be executed in the following way: Begin with a linear structure $\beta = U\gamma$ but partition U into sets of columns $U = [U_0, U_1, \ldots, U_m]$ and then specify that

$$U_i = [V_{i1} \otimes V_{i2} \otimes \cdots \otimes V_{ir}]. \tag{2}$$

Such structures should be sufficient to satisfy the basic idea of selection invariance. However, as will be seen later, other interesting selection invariant factorial models require us to consider matrices U_i that are linear combinations of matrices with the structure (2).

An interesting aspect of factorial structures is dealing with factors that have the same levels. Such factors are called *homologous*. Example 7.5.1 involves genotypes of mothers and genotypes of litters, but the genotypes are identical for the mothers and the litters, so it provides an example of homologous factors. In the additive two-factor model $y_{ijk} = \mu + \alpha_i + \eta_j + e_{ijk}$ with $s = t$ and homologous factors, we might consider situations such as *symmetric additive effects* where $y_{ijk} = \mu + \alpha_i + \alpha_j + e_{ijk}$ or *alternating (skew symmetric) additive effects* where $y_{ijk} = \mu + \alpha_i - \alpha_j + e_{ijk}$. As illustrated in (1), we can write the additive model for the parameters in matrix form as

$$\beta = U\gamma = [J_{st}, U_1, U_2] \begin{bmatrix} \mu \\ \alpha \\ \eta \end{bmatrix}.$$

We can now specify symmetric additive effects by specifying $\alpha = \eta$ to get

$$\beta = [J_{st}, U_1, U_2] \begin{bmatrix} \mu \\ \alpha \\ \alpha \end{bmatrix} = [J_{st}, (U_1 + U_2)] \begin{bmatrix} \mu \\ \alpha \end{bmatrix},$$

thus defining the linear model $Y = X[J_{st}, (U_1 + U_2)] \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + e$. Similarly, specifying alternating additive effects $\alpha = -\eta$ leads to $Y = X[J_{st}, (U_1 - U_2)] \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + e$. In a 3×3 example with $\beta = [(\alpha\eta)_{11}, (\alpha\eta)_{12}, \ldots, (\alpha\eta)_{33}]'$, the additive main effects model has

$$U = [J_{st}, U_1, U_2] = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

The symmetric additive effects and alternating additive effects models have

$$[J_{st}, U_1 + U_2] = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 2 \end{bmatrix}, \quad [J_{st}, (U_1 - U_2)] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 0 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

respectively. Given the simple structure of the original one-way ANOVA matrix X, the reduced model matrices $X[J_{st}, (U_1 + U_2)]$ and $X[J_{st}, (U_1 - U_2)]$ have structures very similar to $[J_{st}, (U_1 + U_2)]$ and $[J_{st}, (U_1 - U_2)]$. However, these linear structures for the parameters are not of the form (2), hence the need to consider linear combinations of terms like those in (2).

Models specifying such things as simple symmetry $(\alpha\eta)_{ij} = (\alpha\eta)_{ji}$ can also be specified quite easily by defining an appropriate U matrix, e.g.,

$$\begin{bmatrix} (\alpha\eta)_{11} \\ (\alpha\eta)_{12} \\ (\alpha\eta)_{13} \\ (\alpha\eta)_{21} \\ (\alpha\eta)_{22} \\ (\alpha\eta)_{23} \\ (\alpha\eta)_{31} \\ (\alpha\eta)_{32} \\ (\alpha\eta)_{33} \end{bmatrix} = \beta = U\phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{12} \\ \phi_{13} \\ \phi_{22} \\ \phi_{23} \\ \phi_{33} \end{bmatrix}.$$

These also fit into the class of linear combinations of matrices with the pattern (2), e.g., the column of U associated with ϕ_{23} can be written as

$$\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) + \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right).$$

These ideas also apply to generalized linear models. For example, in log-linear models, symmetry is sometimes an interesting model, cf. Christensen (1997, Exercise 2.7.10), and the symmetric additive effects model is the model of marginal homogeneity, cf. Christensen (1997, Exercise 10.8.6). See McCullagh (2000) for a more extensive and theoretical treatment of these ideas.

8.6 Additional Exercises

Exercise 8.6.1 A study was performed to examine the effect of two factors on increasing muscle mass in weight lifters. The first factor was dietary protein level. The levels were use of a relatively low protein diet (L) and use of a high protein diet (H). The second factor was the use of anabolic steroids. The first level consisted of no steroid use (N) and the second level involved the use of steroids (S). Subjects were chosen so that one subject was in each combination of four height groups (i) and four weight groups (j). Treatments are identified as LN, LS, HN, and HS. The dependent variable is a measure of increase in muscle mass during the treatment period. The study was replicated in two different years. The height groups and weight groups changed from the first year to the second year. The design and data are listed below. Heights are the columns of the squares and weights are the rows. Analyze the data.

Year 1

		Weight			
	Trt(y_{ijk})	1	2	3	4
	1	$LN(2.7)$	$LS(4.6)$	$HS(9.3)$	$HN(0.0)$
Height	2	$LS(2.0)$	$LN(5.0)$	$HN(9.1)$	$HS(4.5)$
	3	$HN(6.4)$	$HS(10.2)$	$LS(6.1)$	$LN(2.9)$
	4	$HS(8.3)$	$HN(6.3)$	$LN(6.3)$	$LS(0.9)$

Year 2

		Weight			
	Trt(y_{ijk})	1	2	3	4
	1	$LN(8.6)$	$LS(3.3)$	$HN(7.6)$	$HS(9.0)$
Height	2	$LS(8.9)$	$HN(4.7)$	$HS(12.2)$	$LN(5.2)$
	3	$HN(10.0)$	$HS(7.6)$	$LN(12.3)$	$LS(5.4)$
	4	$HS(10.0)$	$LN(0.5)$	$LS(5.0)$	$HN(3.7)$

Exercise 8.6.2 Show that the set of indices $i = 1,\ldots,a$, $j = 1,\ldots,a$, and $k = (i+j+a-1) \bmod (a)$ determines a Latin square design.

Hint: Recall that $t \bmod (a)$ means t modulo a and is defined as the remainder when t is divided by a.