Chapter 9

Factorial Experiments

Introduction

Experiments are essential to the development and improvement of engineering and scientific methods. Only through experimentation can different variants of a method be compared to see which are most effective. To be useful, an experiment must be designed properly, and the data it produces must be analyzed correctly. In this chapter we will discuss the design of and the analysis of data from a class of experiments known as **factorial experiments**.

9.1 One-Factor Experiments

We begin with an example. The article "An Investigation of the $CaCO_3$-CaF_2-K_2SiO_3-SiO_2-Fe Flux System Using the Submerged Arc Welding Process on HSLA-100 and AISI-1081 Steels" (G. Fredrickson, M.S. Thesis, Colorado School of Mines, 1992) describes an experiment in which welding fluxes with differing chemical compositions were prepared. Several welds using each flux were made on AISI-1018 steel base metal. The results of hardness measurements, on the Brinell scale, of five welds using each of four fluxes are presented in Table 9.1.

TABLE 9.1 Brinell hardness of welds using four different fluxes

Flux	Sample Values					Sample Mean	Sample Standard Deviation
A	250	264	256	260	239	253.8	9.7570
B	263	254	267	265	267	263.2	5.4037
C	257	279	269	273	277	271.0	8.7178
D	253	258	262	264	273	262.0	7.4498

Figure 9.1 presents dotplots for the hardnesses using the four fluxes. Each sample mean is marked with an "X." It is clear that the sample means differ. In particular, the welds made using flux C have the largest sample mean and those using flux A have the smallest. Of course, there is uncertainty in the sample means, and the question is whether the sample means differ from each other by a greater amount than could be accounted for by uncertainty alone. Another way to phrase the question is this: Can we conclude that there are differences in the population means among the four flux types?

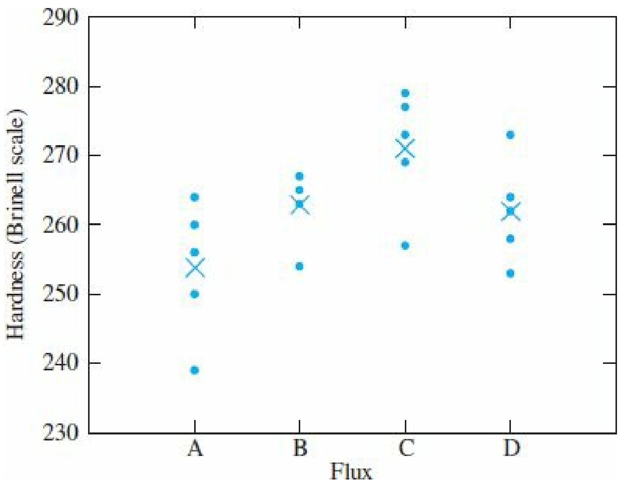

FIGURE 9.1 Dotplots for each sample in Table 9.1. Each sample mean is marked with an "X." The

sample means differ somewhat, but the sample values overlap considerably.

This is an example of a factorial experiment. In general a factorial experiment involves several variables. One variable is the **response variable**, which is sometimes called the **outcome variable** or the **dependent variable**. The other variables are called **factors**. The question addressed by a factorial experiment is whether varying the levels of the factors produces a difference in the mean of the response variable. In the experiment described in Table 9.1, the hardness is the response, and there is one factor: flux type. Since there is only one factor, this is a **one-factor experiment**. There are four different values for the flux-type factor in this experiment. These different values are called the **levels** of the factor and can also be called **treatments**. Finally, the objects upon which measurements are made are called **experimental units**. The units assigned to a given treatment are called **replicates**. In the preceding experiment, the welds are the experimental units, and there are five replicates for each treatment.

In this welding experiment, the four particular flux compositions were chosen deliberately by the experimenter, rather than at random from a larger population of fluxes. Such an experiment is said to follow a **fixed effects model**. In some experiments, treatments are chosen at random from a population of possible treatments. In this case the experiment is said to follow a **random effects model**. The methods of analysis for these two models are essentially the same, although the conclusions to be drawn from them differ. We will focus on fixed effects models. Later in this section, we will discuss some of the differences between fixed and random effects models.

Completely Randomized Experiments

In this welding experiment, a total of 20 welds were produced, five with each of the four fluxes. Each weld was produced on a different steel base plate. Therefore, to run the experiment, the experimenter had to choose, from a total of 20 base plates, a group of 5 to be welded with flux A, another group of 5 to be welded with flux B, and so on. The best way to assign the base plates to the fluxes is at random. In this way, the experimental design will not favor any one treatment over another. For example, the experimenter could number the plates from 1 to 20, and then generate a random ordering of the integers from 1 to 20. The plates whose numbers correspond to the first five numbers on the list are assigned to flux A, and so on. This is an example of a **completely randomized experiment**.

Definition

Now we describe how to estimate the parameters in the two-way ANOVA model. The fundamental idea is that the best estimate of the treatment mean is the cell mean $\overline{X}_{ij.}$, which is the average of the sample observations having that treatment. It follows that the best estimate of the quantity $\overline{\mu}_{i.}$ is the row mean $\overline{X}_{i..}$, the best estimate of the quantity $\overline{\mu}_{.j}$ is the column mean $\overline{X}_{.j.}$, and the best estimate of the population grand mean μ is the sample grand mean $\overline{X}_{...}$. We estimate the row effects α_i, the column effects β_j, and the interactions γ_{ij} by substituting these estimates into Equations (9.41) through (9.43).

$$\widehat{\alpha}_i = \overline{X}_{i..} - \overline{X}_{...} \tag{9.54}$$

$$\widehat{\beta}_j = \overline{X}_{.j.} - \overline{X}_{...} \tag{9.55}$$

$$\widehat{\gamma}_{ij} = \overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} + \overline{X}_{...} \tag{9.56}$$

The row effects, column effects, and interactions satisfy constraints given in Equation (9.44). By performing some algebra, it can be shown that their estimates satisfy the same constraints:

$$\sum_{i=1}^{I} \widehat{\alpha}_i = 0 \qquad \sum_{j=1}^{J} \widehat{\beta}_j = 0 \qquad \sum_{i=1}^{I} \widehat{\gamma}_{ij} = \sum_{j=1}^{J} \widehat{\gamma}_{ij} = 0 \tag{9.57}$$

Example 9.14

Compute the estimated row effects, column effects, and interactions for the data in Table 9.2.

Solution

Using the quantities in Table 9.4 and Equations (9.54) through (9.56), we compute

$$\widehat{\alpha}_1 = 86.42 - 79.61 = 6.81 \qquad \widehat{\alpha}_2 = 79.80 - 79.61 = 0.19$$
$$\widehat{\alpha}_3 = 75.05 - 79.61 = -4.56 \qquad \widehat{\alpha}_4 = 77.17 - 79.61 = -2.44$$

$$\widehat{\beta}_1 = 75.92 - 79.61 = -3.69 \qquad \widehat{\beta}_2 = 81.57 - 79.61 = 1.96$$
$$\widehat{\beta}_3 = 81.34 - 79.61 = 1.73$$

$$\begin{array}{lll}
\widehat{\gamma}_{11} = 2.12 & \widehat{\gamma}_{12} = 0.75 & \widehat{\gamma}_{13} = -2.87 \\
\widehat{\gamma}_{21} = -0.76 & \widehat{\gamma}_{22} = -2.36 & \widehat{\gamma}_{23} = 3.12 \\
\widehat{\gamma}_{31} = -1.06 & \widehat{\gamma}_{32} = -0.36 & \widehat{\gamma}_{33} = 1.42 \\
\widehat{\gamma}_{41} = -0.30 & \widehat{\gamma}_{42} = 1.97 & \widehat{\gamma}_{43} = -1.67
\end{array}$$

Using Two-Way ANOVA to Test Hypotheses

A two-way analysis of variance is designed to address three main questions:

1. Does the additive model hold?
2. If so, is the mean outcome the same for all levels of the row factor?
3. If so, is the mean outcome the same for all levels of the column factor?

In general, we ask questions 2 and 3 only when we believe that the additive model may hold. We will discuss this further later in this section. The three questions are addressed by performing hypothesis tests. The null hypotheses for these tests are as follows:

1. To test whether the additive model holds, we test the null hypothesis that all the interactions are equal to 0:

$$H_0: \gamma_{11} = \gamma_{12} = \cdots = \gamma_{IJ} = 0$$

If this null hypothesis is true, the additive model holds.

2. To test whether the mean outcome is the same for all levels of the row factor, we test the null hypothesis that all the row effects are equal to 0:

$$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

If this null hypothesis is true, then the mean outcome is the same for all levels of the row factor.

3. To test whether the mean outcome is the same for all levels of the column factor, we test the null hypothesis that all the column effects are equal to 0:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_J = 0$$

If this null hypothesis is true, then the mean outcome is the same for all levels of the column factor.

We now describe the standard tests for these null hypotheses. For the tests to be valid, the following conditions must hold:

Page 713

Assumptions for Two-Way ANOVA

The standard two-way ANOVA hypothesis tests are valid under the following conditions:

1. The design must be complete.
2. The design must be balanced.
3. The number of replicates per treatment, K, must be at least 2.
4. Within any treatment, the observations $X_{ij1}, \ldots, X_{ijK}$ are a simple random sample from a normal population.
5. The population variance is the same for all treatments. We denote this variance by σ^2.

Just as in one-way ANOVA, the standard tests for these null hypotheses are based on sums of squares. Specifically, they are the row sum of squares (SSA), the column sum of squares (SSB), the interaction sum of squares (SSAB), and the error sum of squares (SSE). Also of interest is the total sum of squares (SST), which is equal to the sum of the others. Formulas for these sums of squares are as follows:

$$\text{SSA} = JK\sum_{i=1}^{I}\widehat{\alpha}_i^2 = JK\sum_{i=1}^{I}(\overline{X}_{i..} - \overline{X}_{...})^2 = JK\sum_{i=1}^{I}\overline{X}_{i..}^2 - IJK\overline{X}_{...}^2 \tag{9.58}$$

$$\text{SSB} = IK\sum_{j=1}^{J}\widehat{\beta}_j^2 = IK\sum_{j=1}^{J}(\overline{X}_{.j.} - \overline{X}_{...})^2 = IK\sum_{j=1}^{J}\overline{X}_{.j.}^2 - IJK\overline{X}_{...}^2 \tag{9.59}$$

$$\begin{aligned}\text{SSAB} &= K\sum_{i=1}^{I}\sum_{j=1}^{J}\widehat{\gamma}_{ij}^2 = K\sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} + \overline{X}_{...})^2 \\ &= K\sum_{i=1}^{I}\sum_{j=1}^{J}\overline{X}_{ij.}^2 - JK\sum_{i=1}^{I}\overline{X}_{i..}^2 - IK\sum_{j=1}^{J}\overline{X}_{.j.}^2 + IJK\overline{X}_{...}^2\end{aligned} \tag{9.60}$$

$$\text{SSE} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(X_{ijk} - \overline{X}_{ij.})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}X_{ijk}^2 - K\sum_{i=1}^{I}\sum_{j=1}^{J}\overline{X}_{ij.}^2 \tag{9.61}$$

$$\text{SST} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(X_{ijk} - \overline{X}_{...})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}X_{ijk}^2 - IJK\overline{X}_{...}^2 \tag{9.62}$$

It can be seen from the rightmost expressions in Equations (9.58) through (9.62) that the total sum of squares, SST, is equal to the sum of the others. This is the analysis of variance identity for two-way ANOVA.

The Analysis of Variance Identity

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE} \tag{9.63}$$

Along with each sum of squares is a quantity known as its degrees of freedom. The sums of squares and their degrees of freedom are generally presented in an ANOVA table. Table 9.5 presents the degrees of freedom for each sum of squares, along with the computationally most convenient formula. We point out that the degrees of freedom for SST is the sum of the degrees of freedom for the other sums of squares.

TABLE 9.5 ANOVA table for two-way ANOVA

Source	Degrees of Freedom	Sum of Squares
Rows (SSA)	$I - 1$	$JK\sum_{i=1}^{I}\widehat{\alpha}_i^2 = JK\sum_{i=1}^{I}\overline{X}_{i..}^2 - IJK\overline{X}_{...}^2$
Columns (SSB)	$J - 1$	$IK\sum_{j=1}^{J}\widehat{\beta}_j^2 = IK\sum_{j=1}^{J}\overline{X}_{.j.}^2 - IJK\overline{X}_{...}^2$
Interactions (SSAB)	(I − 1)(J − 1)	$K\sum_{i=1}^{I}\sum_{j=1}^{J}\widehat{\gamma}_{ij}^2 = K\sum_{i=1}^{I}\sum_{j=1}^{J}\overline{X}_{ij.}^2 - JK\sum_{i=1}^{I}\overline{X}_{i..}^2 - IK\sum_{j=1}^{J}\overline{X}_{.j.}^2 + IJK\overline{X}_{...}^2$
Error (SSE)	$IJ(K - 1)$	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(X_{ijk} - \overline{X}_{ij.})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}X_{ijk}^2 - K\sum_{i=1}^{I}\sum_{j=1}^{J}\overline{X}_{ij.}^2$
Total (SST)	$IJK - 1$	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(X_{ijk} - \overline{X}_{...})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}X_{ijk}^2 - IJK\overline{X}_{...}^2$

Note that the magnitude of SSA depends on the magnitude of the *estimated* row effects $\widehat{\alpha}_i$. Therefore when the *true* row effects α_i are equal to 0, SSA will tend to be smaller, and when some of the true row effects are not equal to 0, SSA will tend to be larger. We will therefore reject $H_0\colon \alpha_1 = \cdots = \alpha_I = 0$ when SSA is sufficiently large. Similarly, SSB will tend to be smaller when the true column effects β_j, are all equal to 0 and larger when some column effects are not zero, and SSAB will tend to be smaller when the true interactions γ_{ij}, are all equal to 0 and larger when some interactions are not zero. We will therefore reject $H_0\colon \beta_1 = \cdots = \beta_J = 0$ when SSB is sufficiently large, and we will reject $H_0\colon \gamma_{11} = \cdots = \gamma_{IJ} = 0$ when SSAB is sufficiently large.

We can determine whether SSA, SSB, and SSAB are sufficiently large by comparing them to the error sum of squares, SSE. As in one-way ANOVA (Section 9.1), SSE depends only on the distances between the observations and their own cell means. SSE therefore measures only the random variation inherent in the process and is not affected by the values of the row effects, column effects, or interactions. To compare SSA, SSB, and SSAB with SSE, we first divide each sum of squares by its degrees of freedom, producing quantities known as **mean squares**. The mean squares, denoted MSA, MSB, MSAB, and MSE, are defined as follows:

$$\text{MSA} = \frac{\text{SSA}}{I-1} \qquad \text{MSB} = \frac{\text{SSB}}{J-1} \qquad \text{MSAB} = \frac{\text{SSAB}}{(I-1)(J-1)} \tag{9.64}$$

$$\text{MSE} = \frac{\text{SSE}}{IJ(K-1)}$$

The test statistics for the three null hypotheses are the quotients of MSA, MSB, and MSAB with MSE. The null distributions of these test statistics are *F* distributions. Specifically,

- Under H_0: $\alpha_1 = \cdots = \alpha_I = 0$, the statistic $\frac{\text{MSA}}{\text{MSE}}$ has an $F_{I-1,\, IJ(K-1)}$ distribution.
- Under H_0: $\beta_1 = \cdots = \beta_J = 0$, the statistic $\frac{\text{MSB}}{\text{MSE}}$ has an $F_{J-1,\, IJ(K-1)}$ distribution.
- Under H_0: $\gamma_{11} = \cdots = \gamma_{IJ} = 0$, the statistic $\frac{\text{MSAB}}{\text{MSE}}$ has an $F_{(I-1)(J-1),\, IJ(K-1)}$ distribution.

In practice, the sums of squares, mean squares, and test statistics are usually calculated with the use of a computer. The following output (from MINITAB) presents the ANOVA table for the data in Table 9.2.

Two-way ANOVA: Yield versus Catalyst, Reagent

Source	DF	SS	MS	F	P
Catalyst	3	877.56	292.521	9.36	0.000
Reagent	2	327.14	163.570	5.23	0.010
Interaction	6	156.98	26.164	0.84	0.550
Error	36	1125.33	31.259		
Total	47	2487.02			

S = 5.591 R-sq = 54.75% R-Sq(adj) = 40.93%

The labels DF, SS, MS, F, and P refer to degrees of freedom, sum of squares, mean square, *F* statistic, and *P*-value, respectively. As in one-way ANOVA, the mean square for error (MSE) is an estimate of the error variance σ^2 and the quantity labeled "S" is the square root of MSE and is an estimate of the error standard deviation σ. The quantities "R-sq" and "R-sq(adj)" are computed with formulas analogous to those in one-way ANOVA.

Example 9.15

Use the preceding ANOVA table to determine whether the additive model is plausible for the yield data. If the additive model is plausible, can we conclude that either the catalyst or the reagent affects the yield?

Solution

We first check to see if the additive model is plausible. The *P*-value for the interactions is 0.55, which is not small. We therefore do not reject the null hypothesis that all the interactions are equal to 0, and we conclude that the additive model is plausible. Since the additive model is plausible, we now ask whether the row or column factors affect the outcome. We see from the table that the *P*-value for the row effects (Catalyst) is approximately 0, so we conclude that the catalyst does affect the yield. Similarly, the *P*-value for the column effects (Reagent) is small (0.010), so we conclude that the reagent affects the yield as well.

Example 9.16

The article "Uncertainty in Measurements of Dermal Absorption of Pesticides" (W. Navidi and A. Bunge, *Risk Analysis*, 2002:1175–1182) describes an experiment in which a pesticide was applied to skin at various concentrations and for various lengths of time. The outcome is the amount of the pesticide that was absorbed into the skin. The following output (from MINITAB) presents the ANOVA table. Is the additive model plausible? If so, do either the concentration or the duration affect the amount absorbed?

Two-way ANOVA: Absorbed versus Concentration, Duration

Source	DF	SS	MS	F	P
Concent	2	49.991	24.996	107.99	0.000
Duration	2	19.157	9.579	41.38	0.000
Interaction	4	0.337	0.084	0.36	0.832
Error	27	6.250	0.231		
Total	35	75.735			

Solution

The *P*-value for the interaction is 0.832, so we conclude that the additive model is plausible. The *P*-values for both concentration and dose are very small. Therefore we can conclude that both concentration and duration affect the amount absorbed.

9.5 2^p Factorial Experiments

When an experimenter wants to study several factors simultaneously, the number of different treatments can become quite large. In these cases, preliminary experiments are often performed in which each factor has only two levels. One level is designated as the "high" level, and the other is designated as the "low" level. If there are p factors, there are then 2^p different treatments. Such experiments are called **2^p factorial experiments**. Often, the purpose of a 2^p experiment is to determine which factors have an important effect on the outcome. Once this is determined, more elaborate experiments can be performed, in which the factors previously found to be important are varied over several levels. We will begin by describing 2^3 factorial experiments.

Notation for 2^3 Factorial Experiments

In a 2^3 factorial design, there are three factors and $2^3 = 8$ treatments. The **main effect** of a factor is defined to be the difference between the mean response when the factor is at its high level and the mean response when the factor is at its low level. It is common to denote the main effects by *A*, *B*, and *C*. As with any factorial experiment, there can be interactions between the factors. With three factors, there are three two-way interactions, one for each pair of factors, and one three-way interaction. The two-way interactions are denoted by *AB*, *AC*, and *BC*, and the three-way interaction by *ABC*. The treatments are traditionally denoted with lowercase letters, with a letter indicating that a factor is at its high level. For example, *ab* denotes the treatment in which the first two factors are at their high level and the third factor is at its low level. The symbol "1" is used to denote the treatment in which all factors are at their low levels.

Estimating Effects in a 2^3 Factorial Experiment

Assume that there are K replicates for each treatment in a 2^3 factorial experiment. For each treatment, the cell mean is the average of the K observations for that treatment. The formulas for the effect estimates can be easily obtained from the 2^3 **sign table**, presented as Table 9.7. The signs are placed in the table as follows. For the main effects *A*, *B*, *C*, the sign is + for treatments in which the factor is at its high level, and − for treatments where the factor is at its low level. So for the main effect *A*, the sign is + for treatments *a*, *ab*, *ac*, and *abc*, and − for the rest. For the interactions, the signs are computed by taking the product of the signs in the corresponding main effects columns. For example, the signs for the two-way interaction *AB* are the products of the signs in columns *A* and *B*, and the signs for the three-way interaction *ABC* are the products of the signs in columns *A* and *B* and *C*.

TABLE 9.7 Sign table for a 2^3 factorial experiment

Treatment	Cell Mean	*A*	*B*	*C*	*AB*	*AC*	*BC*	*ABC*
1	$\overline{X}_1$	−	−	−	+	+	+	−
a	$\overline{X}_a$	+	−	−	−	−	+	+
b	$\overline{X}_b$	−	+	−	−	+	−	+
ab	$\overline{X}_{ab}$	+	+	−	+	−	−	−
c	$\overline{X}_c$	−	−	+	+	−	−	+
ac	$\overline{X}_{ac}$	+	−	+	−	+	−	−
bc	$\overline{X}_{bc}$	−	+	+	−	−	+	−
abc	$\overline{X}_{abc}$	+	+	+	+	+	+	+

Estimating main effects and interactions is done with the use of the sign table. We illustrate how to estimate the main effect of factor *A*. Factor *A* is at its high level in the rows of the table where there is a "+" sign in column *A*. Each of the cell means $\overline{X}_a$, $\overline{X}_{ab}$, $\overline{X}_{ac}$, and $\overline{X}_{abc}$ is an average response for runs made with factor *A* at its high level. We estimate the mean response for factor *A* at its high level to be the average of these cell means.

$$\text{Estimated mean response for } A \text{ at high level} = \frac{1}{4}(\overline{X}_a + \overline{X}_{ab} + \overline{X}_{ac} + \overline{X}_{abc})$$

Similarly, each row with a "—" sign in column *A* represents a treatment with factor *A* set to its low level. We estimate the mean response for factor *A* at its low level to be the average of the cell means in these rows.

$$\text{Estimated mean response for } A \text{ at low level} = \frac{1}{4}(\overline{X}_1 + \overline{X}_b + \overline{X}_c + \overline{X}_{bc})$$

The estimate of the main effect of factor *A* is the difference in the estimated mean response between its high and low levels.

$$A \text{ effect estimate} = \frac{1}{4}(-\overline{X}_1 + \overline{X}_a - \overline{X}_b + \overline{X}_{ab} - \overline{X}_c + \overline{X}_{ac} - \overline{X}_{bc} + \overline{X}_{abc})$$

The quantity inside the parentheses is called the **contrast** for factor *A*. It is computed by adding and subtracting the cell means, using the signs in the appropriate column of the sign table. Note that the number of plus signs is the same as the number of minus signs, so the sum of the coefficients is equal to 0. The effect estimate is obtained by dividing the contrast by half the number of treatments, which is $2^3/2$, or 4. Estimates for other main effects and interactions are computed in an analogous manner. To illustrate, we present the effect estimates for the main effect *C* and for the two-way interaction *AB*:

$$C \text{ effect estimate} = \frac{1}{4}(-\overline{X}_1 - \overline{X}_a - \overline{X}_b - \overline{X}_{ab} + \overline{X}_c + \overline{X}_{ac} + \overline{X}_{bc} + \overline{X}_{abc})$$

$$AB \text{ interaction estimate} = \frac{1}{4}(\overline{X}_1 - \overline{X}_a - \overline{X}_b + \overline{X}_{ab} + \overline{X}_c - \overline{X}_{ac} - \overline{X}_{bc} + \overline{X}_{abc})$$

Summary

The contrast for any main effect or interaction is obtained by adding and subtracting the cell means, using the signs in the appropriate column of the sign table.

For a 2^3 factorial experiment,

$$\text{Effect estimate} = \frac{\text{contrast}}{4} \tag{9.65}$$

Example 9.23

A 2^3 factorial experiment was conducted to estimate the effects of three factors on the yield of a chemical reaction. The factors were *A*: catalyst concentration (low or high), *B*: reagent (standard formulation or new formulation), and *C*: stirring rate (slow or fast). Three replicates were obtained for each treatment. The yields, presented in the following table, are measured as a percent of a theoretical maximum. Estimate all effects and interactions.

Treatment		Yield		Cell Mean
1	71.67	70.55	67.40	69.8733
a	78.46	75.42	81.77	78.5500
b	77.14	78.25	78.33	77.9067
ab	79.72	76.17	78.41	78.1000
c	72.65	71.03	73.54	72.4067
ac	80.10	73.91	74.81	76.2733
bc	80.20	73.49	74.86	76.1833
abc	75.58	80.28	71.64	75.8333

Solution

We use the sign table (Table 9.7) to find the appropriate sums and differences of the cell means. We present the calculations for the main effect A, the two-way interaction *BC*, and the three-way interaction *ABC*:

$$A \text{ effect estimate} = \frac{1}{4}(-69.8733 + 78.5500 - 77.9067 + 78.1000 \\ - 72.4067 + 76.2733 - 76.1833 + 75.8333) = 3.0967$$

$$BC \text{ interaction estimate} = \frac{1}{4}(69.8733 + 78.5500 - 77.9067 - 78.1000 \\ - 72.4067 - 76.2733 + 76.1833 + 75.8333) = -1.0617$$

$$ABC \text{ interaction estimate} = \frac{1}{4}(-69.8733 + 78.5500 + 77.9067 - 78.1000 \\ + 72.4067 - 76.2733 - 76.1833 + 75.8333) = 1.0667$$

We present all the estimated effects in the following table, rounded off to the same precision as the data:

Term	**Effect**
A	3.10
B	2.73
C	−0.93
AB	−3.18
AC	−1.34
BC	−1.06
ABC	1.07

For each effect, we can test the null hypothesis that the effect is equal to 0. When the null hypothesis is rejected, this provides evidence that the factors involved actually affect the outcome. To test these null hypotheses, an ANOVA table is constructed containing the appropriate sums of squares. For the tests we present to be valid, the number of replicates must be the same for each treatment and must be at least 2. In addition, the observations in each treatment must constitute a random sample from a normal population, and the populations must all have the same variance.

We compute the error sum of squares (SSE) by adding the sums of squared deviations from the sample means for all the treatments. To express this in an equation, let $s_1^2, \ldots, s_8^2$ denote the sample variances of the observations in each of the eight treatments, and let *K* be the number of replicates per treatment. Then

$$\text{SSE} = (K-1)\sum_{i=1}^{8} s_i^2 \qquad (9.66)$$

Each main effect and interaction has its own sum of squares as well. These are easy to compute.

The sum of squares for any effect or interaction is computed by squaring its contrast, multiplying by the number of replicates K, and dividing by the total number of treatments, which is $2^3 = 8$.

$$\text{Sum of squares for an effect} = \frac{K(\text{contrast})^2}{8} \tag{9.67}$$

When using Equation (9.67), it is best to keep as many digits in the effect estimates as possible, in order to obtain maximum precision in the sum of squares. For presentation in a table, effect estimates and sums of squares may be rounded to the same precision as the data.

The sums of squares for the effects and interactions have one degree of freedom each. The error sum of squares has $8(K - 1)$ degrees of freedom. The method for computing mean squares and F statistics is the same as the one presented in Section 9.3 for a two-way ANOVA table. Each mean square is equal to its sum of squares divided by its degrees of freedom. The test statistic for testing the null hypothesis that an effect or interaction is equal to 0 is computed by dividing the mean square for the effect estimate by the mean square for error. When the null hypothesis is true, the test statistic has an $F_{1,\,8(K-1)}$ distribution.

Example 9.24

Refer to Example 9.23. Construct an ANOVA table. For each effect and interaction, test the null hypothesis that it is equal to 0. Which factors, if any, seem most likely to have an effect on the outcome?

Solution

The ANOVA table follows. The sums of squares for the effects and interactions were computed by using Equation (9.67). The error sum of squares was computed by applying Equation (9.66) to the data in Example 9.23. Each F statistic is the quotient of the mean square with the mean square for error. Each F statistic has 1 and 16 degrees of freedom.

Source	Effect	Sum of Squares	df	Mean Square	*F*	*P*
A	3.10	57.54	1	57.54	7.34	0.015
B	2.73	44.72	1	44.72	5.70	0.030
C	−0.93	5.23	1	5.23	0.67	0.426
AB	−3.18	60.48	1	60.48	7.71	0.013
AC	−1.34	10.75	1	10.75	1.37	0.259
BC	−1.06	6.76	1	6.76	0.86	0.367

ABC	1.07	6.83	1	6.83	0.87	0.365
Error		125.48	16	7.84		
Total		317.78	23			

The main effects of factors *A* and *B*,aswellasthe *AB* interaction, have fairly small *P*-values. This suggests that these effects are not equal to 0 and that factors *A* and *B* do affect the outcome. There is no evidence that the main effect of factor *C*, or any of its interactions, differ from 0. Further experiments might focus on factors *A* and *B*. Perhaps a two-way ANOVA would be conducted, with each of the factors *A* and *B* evaluated at several levels, to get more detailed information about their effects on the outcome.

Interpreting Computer Output

In practice, analyses of factorial designs are usually carried out on a computer. The following output (from MINITAB) presents the results of the analysis described in Examples 9.23 and 9.24.

Page 747

Factorial Fit: Yield versus A, B, C

Estimated Effects and Coefficients for Yield (coded units)

Term Constant	Effect	Coef	SE Coef	T	P
		75.641	0.5716	132.33	0.000
A	3.097	1.548	0.5716	2.71	0.015
B	2.730	1.365	0.5716	2.39	0.030
C	−0.933	−0.467	0.5716	−0.82	0.426
A*B	−3.175	−1.587	0.5716	−2.78	0.013
A*C	−1.338	−0.669	0.5716	−1.17	0.259
B*C	−1.062	−0.531	0.5716	−0.93	0.367
A*B*C	1.067	0.533	0.5716	0.93	0.365

S = 2.80040 R-Sq = 60.51% R-Sq(adj) = 43.24%

Analysis of Variance for Yield (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	107.480	107.480	35.827	4.57	0.017
2-Way Interactions	3	77.993	77.993	25.998	3.32	0.047
3-Way Interactions	1	6.827	6.827	6.827	0.87	0.365
Residual Error	16	125.476	125.476	7.842		
Pure Error	16	125.476	125.476	7.842		
Total	23	317.776				

The table at the top of the output presents estimated effects and coefficients. The phrase "coded units" means that the values 1 and −1, rather than the actual values, are used to represent the high and low levels of each factor. The estimated effects are listed in the column labeled "Effect." In the next column are the estimated **coefficients**, each of which is equal to one-half the corresponding effect. While the effect represents the difference in the mean response between the high and low levels of a factor, the coefficient represents the difference between the mean response at the high level and the grand mean response, which is half as much. The coefficient labeled "Constant" is the mean of all the observations, that is, the sample grand mean. Every coefficient estimate has the same standard deviation, which is shown in the column labeled "SE Coef."

MINITAB uses the Student's t statistic, rather than the F statistic, to test the hypotheses that the effects are equal to zero. The column labeled "T" presents the value of the Student's t statistic, which is equal to the quotient of the coefficient estimate (Coef) and its standard deviation. Under the null hypothesis, the t statistic has a Student's t distribution with $8(K-1)$ degrees of freedom. The P-values are presented in the column labeled "P." The t test performed by MINITAB is equivalent to the F test described in Example 9.24. The $t_{8(K-1)}$ statistic can be computed by taking the square root of the $F_{1,8(K-1)}$ statistic and applying the sign of the effect estimate. The P-values are identical.

We'll discuss the analysis of variance table next. The column labeled "DF" presents degrees of freedom. The columns labeled "Seq SS" (sequential sum of squares) and "Adj SS" (adjusted sum of squares) will be identical in all the examples we will consider and will contain sums of squares. The column labeled "Adj MS" contains mean squares, or sums of squares divided by their degrees of freedom. We will now explain the rows involving error. The row labeled "Pure Error" is concerned with the error sum of squares (SSE) (Equation 9.66). There are $8(K-1) = 16$ degrees of freedom (DF) for pure error. The sum of squares for pure error, found in each of the next two columns is the error sum of squares (SSE). Under the column "Adj MS" is the mean square for error. The row above the pure error row is labeled "Residual Error." The sum of squares for residual error is equal to the sum of squares for pure error, plus the sums of squares for any main effects or interactions that are not included in the model. The degrees of freedom for the residual error sum of squares is equal to the degrees of freedom for pure error, plus the degrees of freedom (one each) for each main effect or interaction not included in the model. Since in this example all main effects and interactions are included in the model, the residual error sum of squares and its degrees of freedom are equal to the corresponding quantities for pure error. The row labeled "Total" contains the total sum of squares (SST). The total sum of squares and its degrees of freedom are equal to the sums of the corresponding quantities for all the effects, interactions, and residual error.

Going back to the top of the table, the first row is labeled "Main Effects." There are three degrees of freedom for main effects, because there are three main effects (*A*, *B*, and *C*), with one degree of freedom each. The sequential sum of squares is the sum of the sums of squares for each of the three main effects. The mean square (Adj MS) is the sum of squares divided by its degrees of freedom. The column labeled "F" presents the *F* statistic for testing the null hypothesis that all the main effects are equal to zero. The value of the *F* statistic (4.57) is equal to the quotient of the mean square for main effects (35.827) and the mean square for (pure) error (7.842). The degrees of freedom for the *F* statistic are 3 and 16, corresponding to the degrees of freedom for the two mean squares. The column labeled "P" presents the *P*-value for the *F* test. In this case the *P*-value is 0.017, which indicates that not all the main effects are zero.

The rows labeled "2-Way Interactions" and "3-Way Interactions" are analogous to the row for main effects. The *P*-value for two-way interactions is 0.047, which is reasonably strong evidence that at least some of the two-way interactions are not equal to zero. Since there is only one three-way interaction ($A * B * C$), the *P*-value in the row labeled "3-Way Interactions" is the same (0.365) as the *P*-value in the table at the top of the MINITAB output for $A * B * C$.

Recall that the hypothesis tests are performed under the assumption that all the observations have the same standard deviation σ. The quantity labeled "S" is the estimate of σ and is equal to the square root of the mean square for error (MSE). The quantities "R-sq" and "R-sq(adj)" are the coefficients of determination R^2 and the adjusted R^2, respectively, and are computed by methods analogous to those in one-way ANOVA.

Estimating Effects in a 2^p Factorial Experiment

A sign table can be used to obtain the formulas for computing effect estimates in any 2^p factorial experiment. The method is analogous to the 2^3 case. The treatments are listed in a column. The sign for any main effect is + in the rows corresponding to treatments where the factor is at its high level, and − in rows corresponding to treatments where the factor is at its low level. Signs for the interactions are found by multiplying the signs corresponding to the factors in the interaction. The estimate for any effect or interaction is found by adding and subtracting the cell means for the treatments, using the signs in the appropriate columns, to compute a contrast. The contrast is then divided by half the number of treatments, or 2^{p-1}, to obtain the effect estimate.

Summary

For a 2^p factorial experiment:

$$\text{Effect estimate} = \frac{\text{contrast}}{2^{p-1}} \tag{9.68}$$

As an example, Table 9.8 (page 750) presents a sign table for a 2^5 factorial experiment. We list the signs for the main effects and for selected interactions.

TABLE 9.8 Sign table for the main effects and selected interactions for a 2^5 factorial experiment

Treatment	*A*	*B*	*C*	*D*	*E*	*AB*	*CDE*	*ABDE*	*ABCDE*
1						+	−	+	−
a	+	−	−	−	−	−	−	−	+
b	−	+	−	−	−	−	−	−	+
ab	+	+	−	−	−	+	−	+	−
c	−	−	+	−	−	+	+	+	+
ac	+	−	+	−	−	−	+	−	−
bc	−	+	+	−	−	−	+	−	−
abc	+	+	+	−	−	+	+	+	+
d	−	−	−	+	−	+	+	−	+
ad	+	−	−	+	−	−	+	+	−
bd	−	+	−	+	−	−	+	+	−
abd	+	+	−	+	−	+	+	−	+
cd	−	−	+	+	−	+	−	−	−
acd	+	−	+	+	−	−	−	+	+
bcd	−	+	+	+	−	−	−	+	+
abcd	+	+	+	+	−	+	−	−	−
e	−	−	−	−	+	+	+	−	+
ae	+	−	−	−	+	−	+	+	−
be	−	+	−	−	+	−	+	+	−
abe	+	+	−	−	+	+	+	−	+
ce	−	−	+	−	+	+	−	−	−
ace	+	−	+	−	+	−	−	+	+
bce	−	+	+	−	+	−	−	+	+
abce	+	+	+	−	+	+	−	−	−
de	−	−	−	+	+	+	−	+	−
ade	+	−	−	+	+	−	−	−	+
bde	−	+	−	+	+	−	−	−	+

abde	+	+	−	+	+	+	−	+	−
cde	−	−	+	+	+	+	+	+	+
acde	+	−	+	+	+	−	+	−	−
bcde	−	+	+	+	+	−	+	−	−
abcde	+	+	+	+	+	+	+	+	+

Sums of squares are computed by a method analogous to that for a 2^3 experiment. To compute the error sum of squares (SSE), let $s_1, \ldots, s_{2p}$ be the sample variances of the observations in each of the 2^p treatments. Then

$$\text{SSE} = (K-1)\sum_{i=1}^{2^p} s_i^2$$

The degrees of freedom for error is $2^p(K-1)$, where K is the number of replicates per treatment. The sum of squares for each effect and interaction is equal to the square of the contrast, multiplied by the number of replicates K and divided by the number of treatments 2^p. The sums of squares for the effects and interactions have one degree of freedom each.

$$\text{Sum of squares for an effect} = \frac{K(\text{contrast})^2}{2^p} \tag{9.69}$$

F statistics for main effects and interactions are computed by dividing the sum of squares for the effect by the mean square for error. The null distribution of the F statistic is $F_{1,\,2p(K-1)}$.

Factorial Experiments without Replication

When the number of factors p is large, it is often not feasible to perform more than one replicate for each treatment. In this case, it is not possible to compute SSE, so the hypothesis tests previously described cannot be performed. If it is reasonable to assume that some of the higher-order interactions are equal to 0, then the sums of squares for those interactions can be added together and treated like an error sum of squares. Then the main effects and lower order interactions can be tested.

Example 9.25

A 2^5 factorial experiment was conducted to estimate the effects of five factors on the quality of lightbulbs manufactured by a certain process. The factors were A: plant (1 or 2), B: machine type (low or high speed), C: shift (day or evening), D: lead wire material (standard or new), and E: method of loading materials into the assembler (manual or automatic). One replicate was obtained for each treatment. Table 9.9 presents the results. Compute estimates of the main effects and interactions, and their sums of squares. Assume that the third-, fourth-, and fifth-order interactions are negligible, and add their sums of squares to use as a substitute for an error sum of squares. Use this substitute to test hypotheses concerning the main effects and second-order interactions.

TABLE 9.9

Treatment	**Outcome**
1	32.07
a	39.27
b	34.81
ab	43.07
c	31.55
ac	36.51
bc	28.80
abc	43.05
d	35.64
ad	35.91
bd	47.75
abd	51.47
cd	33.16
acd	35.32
bcd	48.26
abcd	53.28
e	25.10
ae	39.25
be	37.77
abe	46.69
ce	32.55
ace	32.56
bce	28.99
abce	48.92
de	40.60
ade	37.57
bde	47.22
abde	56.87
cde	34.51
acde	36.67
bcde	45.15

abcde	48.72

Solution

We compute the effects, using the rules for adding and subtracting observations given by the sign table, and the sums of squares, using Equation (9.69). See Table 9.10.

TABLE 9.10

Term	Effect	Sum of Squares
A	6.33	320.05
B	9.54	727.52
C	−2.07	34.16
D	6.70	358.72
E	0.58	2.66
AB	2.84	64.52
AC	0.18	0.27
AD	−3.39	91.67
AE	0.60	2.83
BC	−0.49	1.95
BD	4.13	136.54
BE	0.65	3.42
CD	−0.18	0.26
CE	−0.81	5.23
DE	0.24	0.46
ABC	1.35	14.47
ABD	−0.29	0.67
ABE	0.76	4.59
ACD	0.11	0.088
ACE	−0.69	3.75
ADE	−0.45	1.60
BCD	0.76	4.67
BCE	−0.82	5.43
BDE	−2.17	37.63
CDE	−1.25	12.48
ABCD	−2.83	63.96
ABCE	0.39	1.22
ABDE	0.22	0.37
ACDE	0.18	0.24
BCDE	−0.25	0.52

ABCDE	−1.73	23.80

Note that none of the three-, four-, or five-way interactions are among the larger effects. If some of them were, it would not be wise to combine their sums of squares. As it is, we add the sums of squares of the three-, four-, and five-way interactions. The results are presented in the following output (from MINITAB).

Factorial Fit: Response versus A, B, C, D, E

Estimated Effects and Coefficients for Response (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		39.658	0.5854	67.74	0.000
A	6.325	3.163	0.5854	5.40	0.000
B	9.536	4.768	0.5854	8.14	0.000
C	−2.066	−1.033	0.5854	−1.76	0.097
D	6.696	3.348	0.5854	5.72	0.000
E	0.576	0.288	0.5854	0.49	0.629
A*B	2.840	1.420	0.5854	2.43	0.027
A*C	0.183	0.091	0.5854	0.16	0.878
A*D	−3.385	−1.693	0.5854	−2.89	0.011
A*E	0.595	0.298	0.5854	0.51	0.618
B*C	−0.494	−0.247	0.5854	−0.42	0.679
B*D	4.131	2.066	0.5854	3.53	0.003
B*E	0.654	0.327	0.5854	0.56	0.584
C*D	−0.179	−0.089	0.5854	−0.15	0.881
C*E	−0.809	−0.404	0.5854	−0.69	0.500
D*E	0.239	0.119	0.5854	0.20	0.841

S = 3.31179 R-Sq = 90.89% R-Sq(adj) = 82.34%

Analysis of Variance for Response (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	1443.1	1443.1	288.62	26.31	0.000
2-Way Interactions	10	307.1	307.1	30.71	2.80	0.032
Residual Error	16	175.5	175.5	10.97		
Total	31	1925.7				

The estimates have not changed for the main effects or two-way interactions. The residual error sum of squares (175.5) in the analysis of variance table is found by adding the sums of squares for all the higher-order interactions that were dropped from the model. The number of degrees of freedom (16) is equal to the sum of the degrees of freedom (one each) for the 16 higher-order interactions. There is no sum of squares for pure error (SSE), because there is only one replicate per treatment. The residual error sum of squares is used as a substitute for SSE to compute all the quantities that require an error sum of squares.

We conclude from the output that factors *A*, *B*, and *D* are likely to affect the outcome. There seem to be interactions between some pairs of these factors as well. It might be appropriate to plan further experiments to focus on factors *A*, *B*, and *D*.

Using Probability Plots to Detect Large Effects

An informal method that has been suggested to help determine which effects are large is to plot the effect and interaction estimates on a normal probability plot. If in fact none of the factors affect the outcome, then the effect and interaction estimates form a simple random sample from a normal population and should lie approximately on a straight line. In many cases, most of the estimates will fall approximately on a line, and a few will plot far from the line. The main effects and interactions whose estimates plot far from the line are the ones most likely to be important. Figure 9.13 presents a normal probability plot of the main effect and interaction estimates from the data in Example 9.25. It is clear from the plot that the main effects of factors *A*, *B*, and *D*, and the *AB* and *BD* interactions, stand out from the rest.

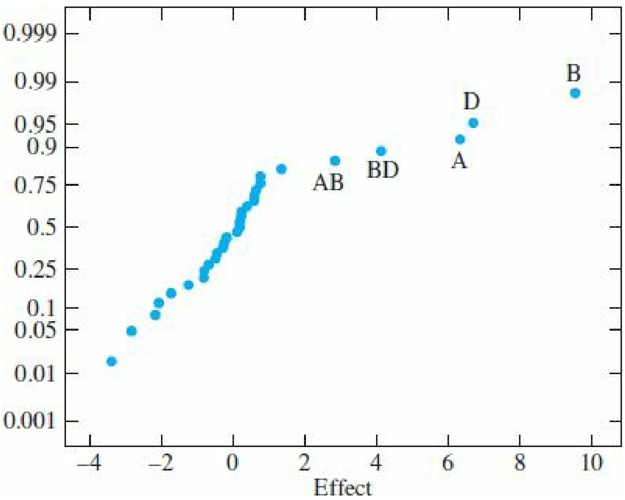

FIGURE 9.13 Normal probability plot of the effect estimates from the data in Example 9.25. The main effects of factors *A*, *B*,and *D* stand out as being larger than the rest.

Fractional Factorial Experiments

When the number of factors is large, it may not be feasible to perform even one replicate for each treatment. In these cases, observations may be taken only for some fraction of the treatments. If these treatments are chosen correctly, it is still possible to obtain information about the factors.

When each factor has two levels, the fraction must always be a power of 2, i.e., one-half, one-quarter, etc. An experiment in which half the treatments are used is called a **half-replicate;** if one-quarter of the treatments are used, it is a **quarter-replicate**, and so on. A half-replicate of a 2^p experiment is often denoted 2^{p-1}, to indicate that while there are p factors, there are only 2^{p-1} treatments being considered. Similarly, a quarter-replicate is often denoted 2^{p-2}. We will focus on half-replicate experiments.

We present a method for choosing a half-replicate of a 2^5 experiment. Such an experiment will have 16 treatments chosen from the 32 in the 2^5 experiment. To choose the 16 treatments, start with a sign table for a 2^4 design that shows the signs for the main effects and the highest-order interaction. This is presented as Table 9.11 (page 754).

TABLE 9.11 Sign table for the main effects and four-way interaction in a 2^4 factorial experiment

Treatment	*A*	*B*	*C*	*D*	*ABCD*
1	−	−	−	−	+
a	+	−	−	−	−
b	−	+	−	−	−
ab	+	+	−	−	+
c	−	−	+	−	−
ac	+	−	+	−	+
bc	−	+	+	−	+
abc	+	+	+	−	−
d	−	−	−	+	−
ad	+	−	−	+	+
bd	−	+	−	+	+
abd	+	+	−	+	−
cd	−	−	+	+	+
acd	+	−	+	+	−
bcd	−	+	+	+	−
abcd	+	+	+	+	+

Table 9.11 has the right number of treatments (16), but only four factors. To transform it into a half-replicate for a 2^5 design, we must introduce a fifth factor, E. We do this by replacing the highest-order interaction by E. This establishes the signs for the main effect of E. Then in each row where the sign for E is +, we add the letter e to the treatment, indicating that factor E is to be set to its high level for that treatment. Where the sign for E is −, factor

E is set to its low level. The resulting design is called the **principal fraction** of the 2^5 design. Table 9.12 presents the signs for the main effects and selected interactions of this design.

TABLE 9.12 Sign table for the main effects and selected interactions for the principal fraction of a 2^5 factorial experiment

Treatment	*A*	*B*	*C*	*D*	*E* = *ABCD*	*AB*	*CDE*	*ACDE*
e	−	−	−	−	+	+	+	−
a	+	−	−	−	−	−	−	−
b	−	+	−	−	−	−	−	+
abe	+	+	−	−	+	+	+	+
c	−	−	+	−	−	+	+	−
ace	+	−	+	−	+	−	−	−
bce	−	+	+	−	+	−	−	+
abc	+	+	+	−	−	+	+	+
d	−	−	−	+	−	+	+	−
ade	+	−	−	+	+	−	−	−
bde	−	+	−	+	+	−	−	+
abd	+	+	−	+	−	+	+	+
cde	−	−	+	+	+	+	+	−
acd	+	−	+	+	−	−	−	−
bcd	−	+	+	+	−	−	−	+
abcde	+	+	+	+	+	+	+	+

There is a price to be paid for using only half of the treatments. To see this, note that in Table 9.12 the ***AB*** interaction has the same signs as the ***CDE*** interaction, and the ***ACDE*** interaction has the same signs as the main effect for B. When two effects have the same signs, they are said to be **aliased**. In fact, the main effects and interactions in a half-fraction form pairs in which each member of the pair is aliased with the other. The alias pairs for this half-fraction of the 2^5 design are

{*A*, *BCDE*}	{*B*, *ACDE*}	{*C*, *ABDE*}	{*D*, *ABCE*}	{*E*, *ABCD*}
{*AB*, *CDE*}	{*AC*, *BDE*}	{*AD*, *BCE*}	{*AE*, *BCD*}	{*BC*, *ADE*}
{*BD*, *ACE*}	{*BE*, *ACD*}	{*CD*, *ABE*}	{*CE*, *ABD*}	{*DE*, *ABC*}

When two effects are aliased, their effect estimates are the same, because they involve the same signs. In fact, when the principal fraction of a design is used, the estimate of any effect actually represents the sum of that effect and its alias. Therefore for the principal fraction of a 2^5 design, each main effect estimate actually represents the sum of the main effect plus its aliased four-way interaction, and each two-way interaction estimate represents the sum of the two-way interaction and its aliased three-way interaction.

In many cases, it is reasonable to assume that the higher-order interactions are small. In the 2^5 half-replicate, if the four-way interactions are negligible, the main effect estimates will be accurate. If in addition the three-way interactions are negligible, the two-way interaction estimates will be accurate as well.

In a fractional design without replication, there is often no good way to compute an error sum of squares, and therefore no rigorous way to test the hypotheses that the effects are equal to 0. In many cases, the purpose of a fractional design is simply to identify a few factors that appear to have the greatest impact on the outcome. This information may then be used to design more elaborate experiments to investigate these factors. For this purpose, it may be enough simply to choose those factors whose effects or two-way interactions are unusually large, without performing hypothesis tests. This can be done by listing the estimates in decreasing order, and then looking to see if there are a few that are noticeably larger than the rest. Another method is to plot the effect and interaction estimates on a normal probability plot, as previously discussed.

Example 9.26

In an emulsion liquid membrane system, an emulsion (internal phase) is dispersed into an external liquid medium containing a contaminant. The contaminant is removed from the external liquid through mass transfer into the emulsion. Internal phase leakage occurs when portions of the extracted material spill into the external liquid. In the article "Leakage and Swell in Emulsion Liquid Membrane Systems: Batch Experiments" (R. Pfeiffer, W. Navidi, and A. Bunge, *Separation Science and Technology*, 2003:519–539), the effects of five factors were studied to determine the effect on leakage in a certain system. The five factors were A: surfactant concentration, *B*: internal phase lithium hydroxide concentration, *C*: membrane phase, *D*: internal phase volume fraction, and *E*: extraction vessel stirring rate. A half-fraction of a 2^5 design was used. The data are presented in the following table (in the actual experiment, each point actually represented the average of two measurements). Leakage is measured in units of percent. Assume that the third-, fourth-, and fifth-order interactions are negligible. Estimate the main effects and two-way interactions. Which, if any, stand out as being noticeably larger than the rest?

Page 756

Treatment	Leakage
e	0.61

a	0.13
b	2.23
abe	0.095
c	0.35
ace	0.075
bce	7.31
abc	0.080
d	2.03
ade	0.64
bde	11.72
abd	0.56
cde	1.45
acd	0.31
bcd	1.33
abcde	6.24

Solution

Using the sign table (Table 9.12), we compute estimates for the main effects and two-way interactions, shown in the following table.

Term	**Effect**
A	−2.36
B	3.00
C	−0.11
D	1.68
E	2.64
AB	−1.54
AC	1.43
AD	0.17
AE	−1.15
BC	0.20
BD	0.86
BE	2.65
CD	−1.30
CE	0.61
DE	1.32

Note that we do not bother to compute sums of squares for the estimates, because we have no SSE to compare them to. To determine informally which effects may be most worthy of further investigation, we rank the estimates in order of their absolute values: *B*: 3.00, *BE*: 2.65, *E*: 2.64, *A*: −2.36, *D*: 1.68, and so forth. It seems reasonable to decide that there is a fairly wide gap between the *A* and *D* effects, and therefore that factors *A*, *B*, and *E* are most likely to be important.

Exercises for Section 9.5

1. Construct a sign table for the principal fraction for a 2^4 design. Then indicate all the alias pairs.

2. Give an example of a factorial experiment in which failure to randomize can produce incorrect results.

3. A chemical reaction was run using two levels each of temperature (A), reagent concentration (B), and pH (C). For each factor, the high level is denoted 1, and the low level is denoted −1. The reaction was run twice for each combination of levels, and the yield (in percent) was recorded. The results were as follows.

A	*B*	*C*	Yields	Mean Yield
−1	−1	−1	74, 71	72.5
1	−1	−1	73, 74	73.5
−1	1	−1	78, 74	76.0
1	1	−1	86, 89	87.5
−1	−1	1	71, 68	69.5
1	−1	1	77, 77	77.0
−1	1	1	75, 85	80.0
1	1	1	92, 82	87.0

a. Compute estimates of the main effects and interactions, and the sum of squares and *P*-value for each.

b. Which main effects and interactions, if any, are important?

c. Other things being equal, will the mean yield be higher when the temperature is high or low? Explain.

4. The article "Efficient Pyruvate Production by a Multi-Vitamin Auxotroph of *Torulopsis glabrata:* Key Role and Optimization of Vitamin Levels" (Y. Li, J. Chen, et al., *Applied Microbiology and Biotechnology*, 2001:680–685) investigates the effects of the levels of several vitamins in a cell culture on the yield (in g/L) of pyruvate, a useful organic acid. The data in the following table are presented as two replicates of a 2^3 design. The factors are *A:* nicotinic acid, *B*: thiamine, and *C*: biotin. (Two statistically insignificant factors have been dropped. In the article, each factor was tested at four levels; we have collapsed these to two.)

A	*B*	*C*	Yields	Mean Yield
−1	−1	−1	0.55, 0.49	0.520

1	−1	−1	0.60, 0.42	0.510
−1	1	−1	0.37, 0.28	0.325
1	1	−1	0.30, 0.28	0.290
−1	−1	1	0.54, 0.54	0.540
1	−1	1	0.54, 0.47	0.505
−1	1	1	0.44, 0.33	0.385
1	1	1	0.36, 0.20	0.280

a. Compute estimates of the main effects and interactions, and the sum of squares and *P*-value for each.

b. Is the additive model appropriate?

c. What conclusions about the factors can be drawn from these results?

5. The article cited in Exercise 4 also investigated the effects of the factors on glucose consumption (in g/L). A single measurement is provided for each combination of factors (in the article, there was some replication). The results are presented in the following table.

A	*B*	*C*	Glucose Consumption
−1	−1	−1	68.0
1	−1	−1	77.5
−1	1	−1	98.0
1	1	−1	98.0
−1	−1	1	74.0
1	−1	1	77.0
−1	1	1	97.0
1	1	1	98.0

a. Compute estimates of the main effects and the interactions.

b. Is it possible to compute an error sum of squares? Explain.

c. Are any of the interactions among the larger effects? If so, which ones?

d. Assume that it is known from past experience that the additive model holds. Add the sums of squares for the interactions, and use that result in place of an error sum of squares to test the hypotheses that the main effects are equal to 0.

6. A metal casting process for the production of turbine blades was studied. Three factors were varied. They were *A*: the temperature of the metal, *B*: the temperature of the mold, and *C*: the pour speed. The outcome was the thickness of the blades, in mm. The results are presented in the following table.

A	*B*	*C*	Thickness
−1	−1	−1	4.61
1	−1	−1	4.51
−1	1	−1	4.60

1	1	−1	4.54
−1	−1	1	4.61
1	−1	1	4.61
−1	1	1	4.48
1	1	1	4.51

a. Compute estimates of the main effects and the interactions.

b. Is it possible to compute an error sum of squares? Explain.

c. Plot the estimates on a normal probability plot. Does the plot show that some of the factors influence the thickness? Explain.

7. The article "An Investigation into the Ball Burnishing of Aluminium Alloy 6061-T6" (M. El-Axir, *J Engineering Manufacture*, 2007:1733–1742) presents the results of study that investigated the effects of three burnishing factors on the reduction in diameter of the workpiece (in μm). The factors are A: Burnishing speed, B: Burnishing force, and C: Burnishing feed. The results presented in the following table form a 2^3 factorial design (some additional results are omitted).

A	***B***	***C***	**Reduction**
−1	−1	−1	570
1	−1	−1	353
−1	1	−1	778
1	1	−1	769
−1	−1	1	544
1	−1	1	319
−1	1	1	651
1	1	1	625

a. Compute estimates of the main effects and the interactions.

b. Is it possible to compute an error sum of squares? Explain.

c. Are any of the interactions among the larger effects? If so, which ones?

d. Someone claims that the additive model holds. Do the results tend to support this statement? Explain.

8. In a 2^p design with one replicate per treatment, it sometimes happens that the observation for one of the treatments is missing, due to experimental error or to some other cause. When this happens, one approach is to replace the missing value with the value that makes the highest-order interaction equal to 0. Refer to Exercise 7. Assume the observation for the treatment where *A*, *B*, and *C* are all at their low level (−1) is missing.

a. What value for this observation makes the three-way interaction equal to 0?

b. Using this value, compute estimates for the main effects and the interactions.

9. Safety considerations are important in the design of automobiles. The article "An Optimum Design Methodology Development Using a Statistical Technique for Vehicle Occupant Safety" (J. Hong, M. Mun, and S. Song, *Proceedings of the Institution of Mechanical*

Engineers, 2001:795–801) presents results from an occupant simulation study. The outcome variable is chest acceleration (in *g*) 3 ms after impact. Four factors were considered. They were *A*: the airbag vent characteristic, *B*: the airbag inflator trigger time, *C*: the airbag inflator mass flow rate, and *D*: the stress-strain relationship of knee foam. The results (part of a larger study) are presented in the following table. There is one replicate per treatment.

Treatment	**Outcome**
1	85.2
a	79.2
b	84.3
ab	89.0
c	66.0
ac	69.0
bc	68.5
abc	76.4
d	85.0
ad	82.0
bd	84.7
abd	82.2
cd	62.6
acd	65.4
bcd	66.3
abcd	69.0

a. Compute estimates of the main effects and the interactions.

b. If you were to design a follow-up study, which factor or factors would you focus on? Explain.

10. In a small-disc test a small, disc-shaped portion of a component is loaded until failure. The article "Optimizing the Sensitivity of the Small-Disc Creep Test to Damage and Test Conditions" (M. Evans and D. Wang, *J. Strain Analysis*, 2007:389–413) presents the results of a factorial experiment to estimate the effects of properties of the disc on the time to failure (in ms). The data in the following table are presented as a 2^5 design. The factors are A: hole diameter, B: disc diameter, C: disc thickness, D: punch head radius, and E: friction coefficient. Two other factors discussed in the article are not considered here.

Treatment	**Outcome**
1	2486.8
a	1328.1
b	2470.2
ab	1303.2
c	6817.4
ac	3845.2

bc	7045.1
abc	3992.2
d	2912.3
ad	1507.2
bd	2885.3
abd	1491.8
cd	7723.0
acd	4289.3
bcd	7952.8
abcd	4505.5
e	2508.6
ae	1319.4
be	2446.8
abe	1303.3
ce	6864.7
ace	3875.0
bce	6994.2
abce	3961.2
de	2915.0
ade	1536.7
bde	2872.8
abde	1477.9
cde	7731.6
acde	4345.1
bcde	7969.1
abcde	4494.5

a. Compute estimates of the main effects and the interactions.

b. If you were to design a follow-up experiment, which factors would you focus on? Why?

11. The article "Factorial Design for Column Flotation of Phosphate Wastes" (N. Abdel-Khalek, *Particulate Science and Technology*, 2000:57–70) describes a 2^3 factorial design to investigate the effect of superficial air velocity (A), frothier concentration (B), and superficial wash water velocity (C) on the percent recovery of P_2O_5. There were two replicates. The data are presented in the following table.

A	*B*	*C*	Percent	Recovery
−1	−1	−1	56.30	54.85
1	−1	−1	70.10	72.70
−1	1	−1	65.60	63.60
1	1	−1	80.20	78.80

−1	−1	1	50.30	48.95
1	−1	1	65.30	66.00
−1	1	1	60.53	59.50
1	1	1	70.63	69.86

a. Compute estimates of the main effects and interactions, along with their sums of squares and *P*-values.

b. Which factors seem to be most important? Do the important factors interact? Explain.

12. The article "An Application of Fractional Factorial Designs" (M. Kilgo, *Quality Engineering*, 1988:19–23) describes a2^{5-1} design (half-replicate of a 2^5 design) involving the use of carbon dioxide (CO_2) at high pressure to extract oil from peanuts. The outcomes were the solubility of the peanut oil in the CO_2 (in mg oil/liter CO_2), and the yield of peanut oil (in percent). The five factors were *A*: CO_2 pressure, *B*: CO_2 temperature, *C*: peanut moisture, *D*: CO_2 flow rate, and *E*: peanut particle size. The results are presented in the following table.

Treatment	**Solubility**	**Yield**
e	29.2	63
a	23.0	21
b	37.0	36
abe	139.7	99
c	23.3	24
ace	38.3	66
bce	42.6	71
abc	141.4	54
d	22.4	23
ade	37.2	74
bde	31.3	80
abd	48.6	33
cde	22.9	63
acd	36.2	21
bcd	33.6	44
abcde	172.6	96

a. Assuming third- and higher-order interactions to be negligible, compute estimates of the main effects and interactions for the solubility outcome.

b. Plot the estimates on a normal probability plot. Does the plot show that some of the factors influence the solubility? If so, which ones?

c. Assuming third- and higher-order interactions to be negligible, compute estimates of the main effects and interactions for the yield outcome.

d. Plot the estimates on a normal probability plot. Does the plot show that some of the factors influence the yield? If so, which ones?

13. In a 2^{5-1} design (such as the one in Exercise 12) what does the estimate of the main effect of factor *A* actually represent?

i. The main effect of *A*.

ii. The sum of the main effect of *A* and the *BCDE* interaction.

iii. The difference between the main effect of *A* and the *BCDE* interaction.

iv. The interaction between *A* and *BCDE*.

Supplementary Exercises for Chapter 9

1. The article "Gypsum Effect on the Aggregate Size and Geometry of Three Sodic Soils Under Reclamation" (I. Lebron, D. Suarez, and T. Yoshida, *Journal of the Soil Science Society of America*, 2002:92–98) reports on an experiment in which gypsum was added in various amounts to soil samples before leaching. One of the outcomes of interest was the pH of the soil. Gypsum was added in four different amounts. Three soil samples received each amount added. The pH measurements of the samples are presented in the following table.

Gypsum (g/kg)	**pH**		
0.00	7.88	7.72	7.68
0.11	7.81	7.64	7.85
0.19	7.84	7.63	7.87
0.38	7.80	7.73	8.00

Can you conclude that the pH differs with the amount of gypsum added? Provide the value of the test statistic and the *P*-value.

2. The article referred to in Exercise 1 also considered the effect of gypsum on the electric conductivity (in dS m^{-1}) of soil. Two types of soil were each treated with three different amounts of gypsum, with two replicates for each soil-gypsum combination. The data are presented in the following table.

	Soil Type			
Gypsum (g/kg)	**Las Animas**		**Madera**	
0.00	1.52	1.05	1.01	0.92
0.27	1.49	0.91	1.12	0.92
0.46	0.99	0.92	0.88	0.92

a. Is there convincing evidence of an interaction between the amount of gypsum and soil type?

b. Can you conclude that the conductivity differs among the soil types?

c. Can you conclude that the conductivity differs with the amount of gypsum added?

3. Penicillin is produced by the *Penicillium* fungus, which is grown in a broth whose sugar content must be carefully controlled. Several samples of broth were taken on each of three