CHAPTER 4

Randomized Blocks, Latin Squares, and Related Designs

CHAPTER OUTLINE

4.1 THE RANDOMIZED COMPLETE BLOCK DESIGN
- 4.1.1 Statistical Analysis of the RCBD
- 4.1.2 Model Adequacy Checking
- 4.1.3 Some Other Aspects of the Randomized Complete Block Design
- 4.1.4 Estimating Model Parameters and the General Regression Significance Test

4.2 THE LATIN SQUARE DESIGN

4.3 THE GRAECO-LATIN SQUARE DESIGN

4.4 BALANCED INCOMPLETE BLOCK DESIGNS
- 4.4.1 Statistical Analysis of the BIBD
- 4.4.2 Least Squares Estimation of the Parameters
- 4.4.3 Recovery of Interblock Information in the BIBD

SUPPLEMENTAL MATERIAL FOR CHAPTER 4

S4.1 Relative Efficiency of the RCBD

S4.2 Partially Balanced Incomplete Block Designs

S4.3 Youden Squares

S4.4 Lattice Designs

The supplemental material is on the textbook website www.wiley.com/college/montgomery.

4.1 The Randomized Complete Block Design

In any experiment, variability arising from a nuisance factor can affect the results. Generally, we define a **nuisance factor** as a design factor that probably has an effect on the response, but we are not interested in that effect. Sometimes a nuisance factor is **unknown and uncontrolled**; that is, we don't know that the factor exists, and it may even be changing levels while we are conducting the experiment. **Randomization** is the design technique used to guard against such a "lurking" nuisance factor. In other cases, the nuisance factor is **known but uncontrollable**. If we can at least observe the value that the nuisance factor takes on at each run of the experiment, we can compensate for it in the statistical analysis by using the **analysis of covariance**, a technique we will discuss in Chapter 14. When the nuisance source of variability is **known and controllable**, a design technique called **blocking** can be used to systematically eliminate its effect on the statistical comparisons among treatments. Blocking is an extremely important design technique used extensively in industrial experimentation and is the subject of this chapter.

To illustrate the general idea, reconsider the hardness testing experiment first described in Section 2.5.1. Suppose now that we wish to determine whether or not four different tips produce different readings on a hardness testing machine. An experiment such as this might be part of a

■ **TABLE 4.1**
Randomized Complete Block Design for the Hardness Testing Experiment

Test Coupon (Block)			
1	**2**	**3**	**4**
Tip 3	Tip 3	Tip 2	Tip 1
Tip 1	Tip 4	Tip 1	Tip 4
Tip 4	Tip 2	Tip 3	Tip 2
Tip 2	Tip 1	Tip 4	Tip 3

gauge capability study. The machine operates by pressing the tip into a metal test coupon, and from the depth of the resulting depression, the hardness of the coupon can be determined. The experimenter has decided to obtain four observations on Rockwell C-scale hardness for each tip. There is only one factor—tip type—and a completely randomized single-factor design would consist of randomly assigning each one of the $4 \times 4 = 16$ runs to an **experimental unit**, that is, a metal coupon, and observing the hardness reading that results. Thus, 16 different metal test coupons would be required in this experiment, one for each run in the design.

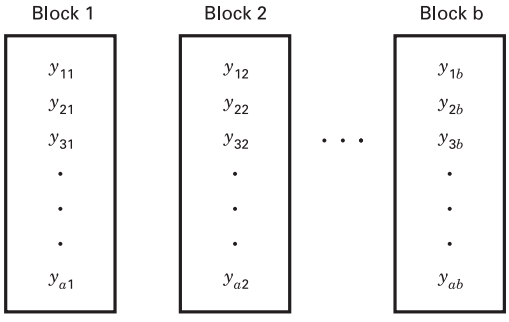
There is a potentially serious problem with a completely randomized experiment in this design situation. If the metal coupons differ slightly in their hardness, as might happen if they are taken from ingots that are produced in different heats, the experimental units (the coupons) will contribute to the variability observed in the hardness data. As a result, the experimental error will reflect *both* random error *and* variability between coupons.

We would like to make the experimental error as small as possible; that is, we would like to remove the variability between coupons from the experimental error. A design that would accomplish this requires the experimenter to test each tip once on each of four coupons. This design, shown in Table 4.1, is called a **randomized complete block design** (**RCBD**). The word "complete" indicates that each block (coupon) contains all the treatments (tips). By using this design, the blocks, or coupons, form a more homogeneous experimental unit on which to compare the tips. Effectively, this design strategy improves the accuracy of the comparisons among tips by eliminating the variability among the coupons. Within a block, the order in which the four tips are tested is randomly determined. Notice the similarity of this design problem to the paired *t*-test of Section 2.5.1. The randomized complete block design is a generalization of that concept.

The RCBD is one of the most widely used experimental designs. Situations for which the RCBD is appropriate are numerous. Units of test equipment or machinery are often different in their operating characteristics and would be a typical blocking factor. Batches of raw material, people, and time are also common nuisance sources of variability in an experiment that can be systematically controlled through blocking.[1]

Blocking may also be useful in situations that do not necessarily involve nuisance factors. For example, suppose that a chemical engineer is interested in the effect of catalyst feed rate on the viscosity of a polymer. She knows that there are several factors, such as raw material source, temperature, operator, and raw material purity that are very difficult to control in the full-scale process. Therefore she decides to test the catalyst feed rate factor in blocks, where each block consists of some combination of these uncontrollable factors. In effect, she is using the blocks to test the **robustness** of her process variable (feed rate) to conditions she cannot easily control. For more discussion of this, see Coleman and Montgomery (1993).

[1] A special case of blocking occurs where the blocks are experimental units such as people, and each block receives the treatments own time or the treatment effects are measured at different times. These are called **repeated measures** designs. They are discussed in chapter 15.



■ **FIGURE 4.1** **The randomized complete block design**

4.1.1 Statistical Analysis of the RCBD

Suppose we have, in general, a treatments that are to be compared and b blocks. The randomized complete block design is shown in Figure 4.1. There is one observation per treatment in each block, and the order in which the treatments are run within each block is determined randomly. Because the only randomization of treatments is within the blocks, we often say that the blocks represent a **restriction on randomization**.

The **statistical model** for the RCBD can be written in several ways. The traditional model is an **effects model**:

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{4.1}$$

where μ is an overall mean, τ_i is the effect of the ith treatment, β_j is the effect of the jth block, and ϵ_{ij} is the usual NID $(0, \sigma^2)$ random error term. We will initially consider treatments and blocks to be fixed factors. The case of random blocks, which is very important, is considerd in Section 4.1.3. Just as in the single-factor experimental design model in Chapter 3, the effects model for the RCBD is an overspecified model. Consequently, we usually think of the treatment and block effects as deviations from the overall mean so that

$$\sum_{i=1}^{a} \tau_i = 0 \quad \text{and} \quad \sum_{j=1}^{b} \beta_j = 0$$

It is also possible to use a **means model** for the RCBD, say

$$y_{ij} = \mu_{ij} + \epsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

where $\mu_{ij} = \mu + \tau_i + \beta_j$. However, we will use the effects model in Equation 4.1 throughout this chapter.

In an experiment involving the RCBD, we are interested in testing the equality of the treatment means. Thus, the hypotheses of interest are

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_a$$
$$H_1: \text{at least one } \mu_i \neq \mu_j$$

Because the ith treatment mean $\mu_i = (1/b)\Sigma_{j=1}^{b} (\mu + \tau_i + \beta_j) = \mu + \tau_i$, an equivalent way to write the above hypotheses is in terms of the treatment effects, say

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_a = 0$$
$$H_1: \tau_i \neq 0 \text{ at least one } i$$

The analysis of variance can be easily extended to the RCBD. Let $y_{i.}$ be the total of all observations taken under treatment i, $y_{.j}$ be the total of all observations in block j, $y_{..}$ be the

grand total of all observations, and $N = ab$ be the total number of observations. Expressed mathematically,

$$y_{i.} = \sum_{j=1}^{b} y_{ij} \qquad i = 1, 2, \ldots, a \tag{4.2}$$

$$y_{.j} = \sum_{i=1}^{a} y_{ij} \qquad j = 1, 2, \ldots, b \tag{4.3}$$

and

$$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij} = \sum_{i=1}^{a} y_{i.} = \sum_{j=1}^{b} y_{.j} \tag{4.4}$$

Similarly, $\bar{y}_{i.}$ is the average of the observations taken under treatment i, $\bar{y}_{.j}$ is the average of the observations in block j, and $\bar{y}_{..}$ is the grand average of all observations. That is,

$$\bar{y}_{i.} = y_{i.}/b \quad \bar{y}_{.j} = y_{.j}/a \quad \bar{y}_{..} = y_{..}/N \tag{4.5}$$

We may express the total corrected sum of squares as

$$\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} [(\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}]^2 \tag{4.6}$$

By expanding the right-hand side of Equation 4.6, we obtain

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 = {} & b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})^2 \\ & + \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 + 2\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i.} - \bar{y}_{..})(\bar{y}_{.j} - \bar{y}_{..}) \\ & + 2\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \\ & + 2\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})\end{aligned}$$

Simple but tedious algebra proves that the three cross products are zero. Therefore,

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 = b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..})^2 \tag{4.7}$$

represents a partition of the total sum of squares. This is the fundamental ANOVA equation for the RCBD. Expressing the sums of squares in Equation 4.7 symbolically, we have

$$SS_T = SS_{\text{Treatments}} + SS_{\text{Blocks}} + SS_E \tag{4.8}$$

Because there are N observations, SS_T has $N - 1$ degrees of freedom. There are a treatments and b blocks, so $SS_{\text{Treatments}}$ and SS_{Blocks} have $a - 1$ and $b - 1$ degrees of freedom, respectively. The error sum of squares is just a sum of squares between cells minus the sum of squares for treatments and blocks. There are ab cells with $ab - 1$ degrees of freedom between them, so SS_E has $ab - 1 - (a - 1) - (b - 1) = (a - 1)(b - 1)$ degrees of freedom. Furthermore, the degrees of freedom on the right-hand side of Equation 4.8 add to the total on the left; therefore, making the usual normality assumptions on the errors, one may use Theorem 3-1 to show

that $SS_{\text{Treatments}}/\sigma^2$, $SS_{\text{Blocks}}/\sigma^2$, and SS_E/σ^2 are independently distributed chi-square random variables. Each sum of squares divided by its degrees of freedom is a mean square. The expected value of the mean squares, if treatments and blocks are fixed, can be shown to be

$$E(MS_{\text{Treatments}}) = \sigma^2 + \frac{b\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E(MS_{\text{Blocks}}) = \sigma^2 + \frac{a\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$E(MS_E) = \sigma^2$$

Therefore, to test the equality of treatment means, we would use the test statistic

$$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$$

which is distributed as $F_{a-1,(a-1)(b-1)}$ if the null hypothesis is true. The critical region is the upper tail of the F distribution, and we would reject H_0 if $F_0 > F_{\alpha,a-1,(a-1)(b-1)}$. A P-value approach can also be used.

We may also be interested in comparing block means because, if these means do not differ greatly, blocking may not be necessary in future experiments. From the expected mean squares, it seems that the hypothesis $H_0{:}\beta_j = 0$ may be tested by comparing the statistic $F_0 = MS_{\text{Blocks}}/MS_E$ to $F_{\alpha,b-1,(a-1)(b-1)}$. However, recall that randomization has been applied only to treatments *within* blocks; that is, the blocks represent a **restriction on randomization**. What effect does this have on the statistic $F_0 = MS_{\text{Blocks}}/MS_E$? Some differences in treatment of this question exist. For example, Box, Hunter, and Hunter (2005) point out that the usual analysis of variance F test can be justified on the basis of randomization only,[2] without direct use of the normality assumption. They further observe that the test to compare block means cannot appeal to such a justification because of the randomization restriction; but if the errors are NID(0, σ^2), the statistic $F_0 = MS_{\text{Blocks}}/MS_E$ can be used to compare block means. On the other hand, Anderson and McLean (1974) argue that the randomization restriction prevents this statistic from being a meaningful test for comparing block means and that this F ratio really is a test for the equality of the block means plus the randomization restriction [which they call a restriction error; see Anderson and McLean (1974) for further details].

In practice, then, what do we do? Because the normality assumption is often questionable, to view $F_0 = MS_{\text{Blocks}}/MS_E$ as an exact F test on the equality of block means is not a good general practice. For that reason, we exclude this F test from the analysis of variance table. However, as an approximate procedure to investigate the effect of the blocking variable, examining the ratio of MS_{Blocks} to MS_E is certainly reasonable. If this ratio is large, it implies that the blocking factor has a large effect and that the noise reduction obtained by blocking was probably helpful in improving the precision of the comparison of treatment means.

The procedure is usually summarized in an ANOVA table, such as the one shown in Table 4.2. The computing would usually be done with a statistical software package. However, computing formulas for the sums of squares may be obtained for the elements in Equation 4.7 by working directly with the identity

$$y_{ij} - \bar{y}_{..} = (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})$$

[2] Actually, the normal-theory F distribution is an approximation to the randomization distribution generated by calculating F_0 from every possible assignment of the responses to the treatments.

■ **TABLE 4.2**
Analysis of Variance for a Randomized Complete Block Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}}$	$a - 1$	$\frac{SS_{\text{Treatments}}}{a - 1}$	$\frac{MS_{\text{Treatments}}}{MS_E}$
Blocks	SS_{Blocks}	$b - 1$	$\frac{SS_{\text{Blocks}}}{b - 1}$	
Error	SS_E	$(a - 1)(b - 1)$	$\frac{SS_E}{(a - 1)(b - 1)}$	
Total	SS_T	$N - 1$		

These quantities can be computed in the columns of a spreadsheet (Excel). Then each column can be squared and summed to produce the sum of squares. Alternatively, computing formulas can be expressed in terms of treatment and block totals. These formulas are

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{4.9}$$

$$SS_{\text{Treatments}} = \frac{1}{b}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N} \tag{4.10}$$

$$SS_{\text{Blocks}} = \frac{1}{a}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N} \tag{4.11}$$

and the error sum of squares is obtained by subtraction as

$$SS_E = SS_T - SS_{\text{Treatments}} - SS_{\text{Blocks}} \tag{4.12}$$

EXAMPLE 4.1

A medical device manufacturer produces vascular grafts (artificial veins). These grafts are produced by extruding billets of polytetrafluoroethylene (PTFE) resin combined with a lubricant into tubes. Frequently, some of the tubes in a production run contain small, hard protrusions on the external surface. These defects are known as "flicks." The defect is cause for rejection of the unit.

The product developer responsible for the vascular grafts suspects that the extrusion pressure affects the occurrence of flicks and therefore intends to conduct an experiment to investigate this hypothesis. However, the resin is manufactured by an external supplier and is delivered to the medical device manufacturer in batches. The engineer also suspects that there may be significant batch-to-batch variation, because while the material should be consistent with respect to parameters such as molecular weight, mean particle size, retention, and peak height ratio, it probably isn't due to manufacturing variation at the resin supplier and natural variation in the material. Therefore, the product developer decides to investigate the effect of four different levels of extrusion pressure on flicks using a randomized complete block design considering batches of resin as blocks. The RCBD is shown in Table 4.3. Note that there are four levels of extrusion pressure (treatments) and six batches of resin (blocks). Remember that the order in which the extrusion pressures are tested within each block is random. The response variable is yield, or the percentage of tubes in the production run that did not contain any flicks.

■ TABLE 4.3
Randomized Complete Block Design for the Vascular Graft Experiment

Extrusion Pressure (PSI)	Batch of Resin (Block) 1	2	3	4	5	6	Treatment Total
8500	90.3	89.2	98.2	93.9	87.4	97.9	556.9
8700	92.5	89.5	90.6	94.7	87.0	95.8	550.1
8900	85.5	90.8	89.6	86.2	88.0	93.4	533.5
9100	82.5	89.5	85.6	87.4	78.9	90.7	514.6
Block Totals	350.8	359.0	364.0	362.2	341.3	377.8	$y_{..} = 2155.1$

To perform the analysis of variance, we need the following sums of squares:

$$SS_T = \sum_{i=1}^{4}\sum_{j=1}^{6} y_{ij}^2 - \frac{y_{..}^2}{N}$$
$$= 193{,}999.31 - \frac{(2155.1)^2}{24} = 480.31$$
$$SS_{\text{Treatments}} = \frac{1}{b}\sum_{i=1}^{4} y_{i.}^2 - \frac{y_{..}^2}{N}$$
$$= \frac{1}{6}[(556.9)^2 + (550.1)^2 + (533.5)^2 + (514.6)^2] - \frac{(2155.1)^2}{24} = 178.17$$
$$SS_{\text{Blocks}} = \frac{1}{a}\sum_{j=1}^{6} y_{j}^2 - \frac{y_{..}^2}{N}$$
$$= \frac{1}{4}[(350.8)^2 + (359.0)^2 + \cdots + (377.8)^2] - \frac{(2155.1)^2}{24} = 192.25$$
$$SS_E = SS_T - SS_{\text{Treatments}} - SS_{\text{Blocks}}$$
$$= 480.31 - 178.17 - 192.25 = 109.89$$

The ANOVA is shown in Table 4.4. Using $\alpha = 0.05$, the critical value of F is $F_{0.05,3,15} = 3.29$. Because $8.11 > 3.29$, we conclude that extrusion pressure affects the mean yield. The P-value for the test is also quite small. Also, the resin batches (blocks) seem to differ significantly, because the mean square for blocks is large relative to error.

■ TABLE 4.4
Analysis of Variance for the Vascular Graft Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Treatments (extrusion pressure)	178.17	3	59.39	8.11	0.0019
Blocks (batches)	192.25	5	38.45		
Error	109.89	15	7.33		
Total	480.31	23			

It is interesting to observe the results we would have obtained from this experiment had we not been aware of randomized block designs. Suppose that this experiment had been run as a completely randomized design, and (by chance) the same design resulted as in Table 4.3. The incorrect analysis of these data as a completely randomized single-factor design is shown in Table 4.5.

Because the P-value is less than 0.05, we would still reject the null hypothesis and conclude that extrusion pressure significantly affects the mean yield. However, note that the mean

■ **TABLE 4.5**
Incorrect Analysis of the Vascular Graft Experiment as a Completely Randomized Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Extrusion pressure	178.17	3	59.39	3.95	0.0235
Error	302.14	20	15.11		
Total	480.31	23			

square for error has more than doubled, increasing from 7.33 in the RCBD to 15.11. All of the variability due to blocks is now in the error term. This makes it easy to see why we sometimes call the RCBD a noise-reducing design technique; it effectively increases the signal-to-noise ratio in the data, or it improves the precision with which treatment means are compared. This example also illustrates an important point. If an experimenter fails to block when he or she should have, the effect may be to inflate the experimental error, and it would be possible to inflate the error so much that important differences among the treatment means could not be identified.

Sample Computer Output. Condensed computer output for the vascular graft experiment in Example 4.1, obtained from Design-Expert and JMP is shown in Figure 4.2. The Design-Expert output is in Figure 4.2*a* and the JMP output is in Figure 4.2*b*. Both outputs are very similar, and match the manual computation given earlier. Note that JMP computes an *F*-statistic for blocks (the batches). The sample means for each treatment are shown in the output. At 8500 psi, the mean yield is $\bar{y}_{1.} = 92.82$, at 8700 psi the mean yield is $\bar{y}_{2.} = 91.68$, at 8900 psi the mean yield is $\bar{y}_{3.} = 88.92$, and at 9100 psi the mean yield is $\bar{y}_{4.} = 85.77$. Remember that these sample mean yields estimate the treatment means μ_1, μ_2, μ_3, and μ_4. The model residuals are shown at the bottom of the Design-Expert output. The residuals are calculated from

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$

and, as we will later show, the fitted values are $\hat{y}_{ij} = \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}$, so

$$e_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..} \tag{4.13}$$

In the next section, we will show how the residuals are used in **model adequacy checking**.

Multiple Comparisons. If the treatments in an RCBD are fixed, and the analysis indicates a significant difference in treatment means, the experimenter is usually interested in multiple comparisons to discover *which* treatment means differ. Any of the multiple comparison procedures discussed in Section 3.5 may be used for this purpose. In the formulas of Section 3.5, simply replace the number of replicates in the single-factor completely randomized design (n) by the number of blocks (b). Also, remember to use the number of error degrees of freedom for the randomized block $[(a - 1)(b - 1)]$ instead of those for the completely randomized design $[a(n - 1)]$.

The Design-Expert output in Figure 4.2 illustrates the Fisher LSD procedure. Notice that we would conclude that $\mu_1 = \mu_2$, because the *P*-value is very large. Furthermore, μ_1 differs from all other means. Now the *P*-value for H_0:$\mu_2 = \mu_3$ is 0.097, so there is some evidence to conclude that $\mu_2 \neq \mu_3$, and $\mu_2 \neq \mu_4$ because the *P*-value is 0.0018. Overall, we would conclude that lower extrusion pressures (8500 psi and 8700 psi) lead to fewer defects.

Response: Yield
ANOVA for Selected Factorial Model
Analysis of Variance Table [Partial Sum of Squares]

Source	Sum of Squares	DF	Mean Square	*F* Value	Prob > *F*
Block	192.25	5	38.45		
Model	178.17	3	59.39	8.11	0.0019
A	*178.17*	*3*	*59.39*	*8.11*	*0.0019*
Residual	109.89	15	7.33		
Cor Total	480.31	23			

Std. Dev.	2.71	R-Squared	0.6185
Mean	89.80	Adj R-Squared	0.5422
C.V.	3.01	Pred R-Squared	0.0234
PRESS	281.31	Adeq Precision	9.759

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1–8500	92.82	1.10
2-8700	91.68	1.10
3-8900	88.92	1.10
4-9100	85.77	1.10

Treatment	Mean Difference	DF	Standard Error	*t* for H_0 Coeff=0	Prob > |*t*|
1 vs. 2	1.13	1	1.56	0.73	0.4795
1 vs. 3	3.90	1	1.56	2.50	0.0247
1 vs. 4	7.05	1	1.56	4.51	0.0004
2 vs. 3	2.77	1	1.56	1.77	0.0970
2 vs. 4	5.92	1	1.56	3.79	0.0018
3 vs. 4	3.15	1	1.56	2.02	0.0621

Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier *t*	Run Order
1	90.30	90.72	–0.42	0.375	–0.197	0.003	–0.190	1
2	89.20	92.77	–3.57	0.375	–1.669	0.186	–1.787	6
3	98.20	94.02	4.18	0.375	1.953	0.254	2.185	9
4	93.90	93.57	0.33	0.375	0.154	0.002	0.149	13
5	87.40	88.35	–0.95	0.375	–0.442	0.013	–0.430	19
6	97.90	97.47	0.43	0.375	0.201	0.003	0.194	23
7	92.50	89.59	2.91	0.375	1.361	0.124	1.405	4
8	89.50	91.64	–2.14	0.375	–0.999	0.067	–0.999	5
9	90.60	92.89	–2.29	0.375	–1.069	0.076	–1.075	10
10	94.70	92.44	2.26	0.375	1.057	0.075	1.062	16
11	87.00	87.21	–0.21	0.375	–0.099	0.001	–0.096	20
12	95.80	96.34	–0.54	0.375	–0.251	0.004	–0.243	21
13	85.50	86.82	–1.32	0.375	–0.617	0.025	–0.604	3
14	90.80	88.87	1.93	0.375	0.902	0.054	0.896	8
15	89.60	90.12	–0.52	0.375	–0.243	0.004	–0.236	12
16	86.20	89.67	–3.47	0.375	–1.622	0.175	–1.726	15
17	88.00	84.45	3.55	0.375	1.661	0.184	1.776	17
18	93.40	93.57	–0.17	0.375	–0.080	0.000	–0.077	22
19	82.50	83.67	–1.17	0.375	–0.547	0.020	–0.534	2
20	89.50	85.72	3.78	0.375	1.766	0.208	1.917	7
21	85.60	86.97	–1.37	0.375	–0.641	0.027	–0.628	11
22	87.40	86.52	0.88	0.375	0.411	0.011	0.399	14
23	78.90	81.30	–2.40	0.375	–1.120	0.084	–1.130	18
24	90.70	90.42	0.28	0.375	0.130	0.001	0.126	24

Note: Predicted values include block corrections.

(a)

■ **FIGURE 4.2** Computer output for Example 4.1. (*a*) Design-Expert; (*b*) JMP

Oneway Analysis of Yield By Pressure
Block
Batch

Oneway Anova
Summary of Fit

Rsquare	0.771218
Adj Rsquare	0.649201
Root Mean Square Error	2.706612
Mean of Response	89.79583
Observations (or Sum Wgts)	24

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	*F* Ratio	Prob > F
Pressure	3	178.17125	59.3904	8.1071	0.0019
Batch	5	192.25208	38.4504	5.2487	0.0055
Error	15	109.88625	7.3257		
C.Total	23	480.30958			

Means for Oneway Anova

Level	Number	Mean	Std. Error	Lower 95%	Upper 95%
8500	6	92.8167	1.1050	90.461	95.172
8700	6	91.6833	1.1050	89.328	94.039
8900	6	88.9167	1.1050	86.561	91.272
9100	6	85.7667	1.1050	83.411	88.122

Std. Error uses a pooled estimate of error variance

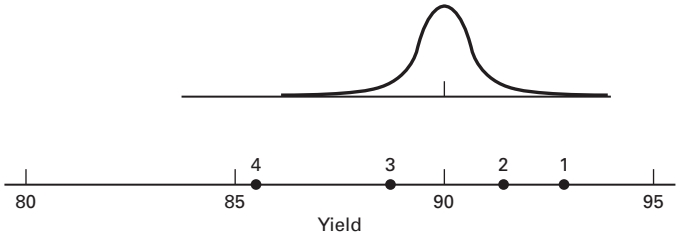
Block Means

Batch	Mean	Number
1	87.7000	4
2	89.7500	4
3	91.0000	4
4	90.5500	4
5	85.3250	4
6	94.4500	4

(b)

■ **FIGURE 4.2** (*Continued*)

We can also use the graphical procedure of Section 3.5.1 to compare mean yield at the four extrusion pressures. Figure 4.3 plots the four means from Example 4.1 relative to a scaled t distribution with a scale factor $\sqrt{MS_E/b} = \sqrt{7.33/6} = 1.10$. This plot indicates that the two lowest pressures result in the same mean yield, but that the mean yields for 8700 psi and



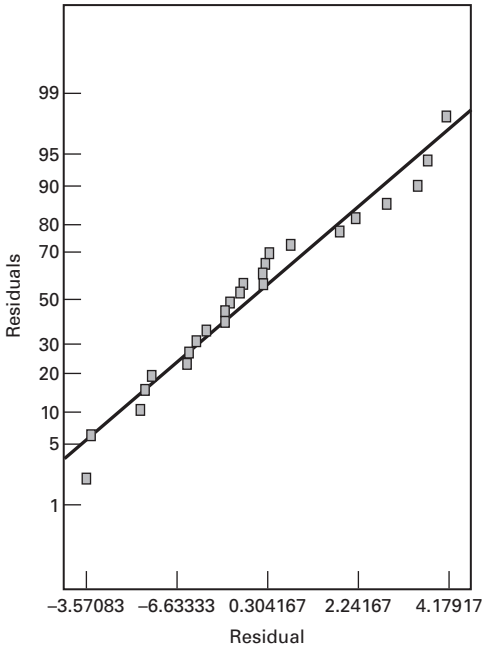
■ **FIGURE 4.3** **Mean yields for the four extrusion pressures relative to a scaled t distribution with a scale factor $\sqrt{MS_E/b} = \sqrt{7.33/6} = 1.10$**

8900 psi (μ_2 and μ_3) are also similar. The highest pressure (9100 psi) results in a mean yield that is much lower than all other means. This figure is a useful aid in interpreting the results of the experiment and the Fisher LSD calculations in the Design-Expert output in Figure 4.2.

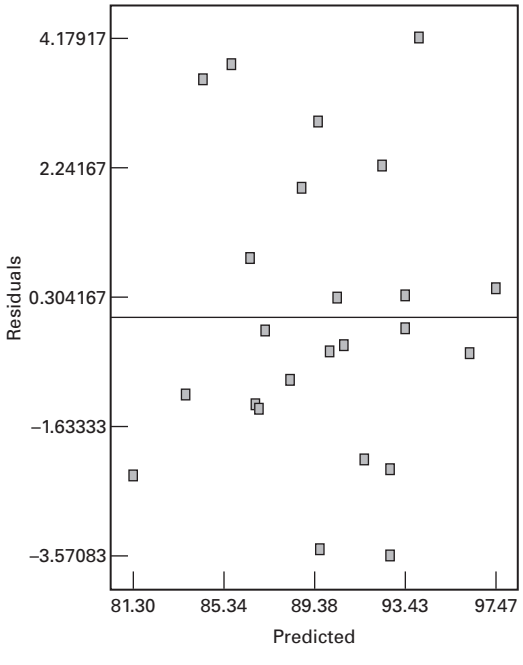
4.1.2 Model Adequacy Checking

We have previously discussed the importance of checking the adequacy of the assumed model. Generally, we should be alert for potential problems with the normality assumption, unequal error variance by treatment or block, and block–treatment interaction. As in the completely randomized design, residual analysis is the major tool used in this diagnostic checking. The residuals for the randomized block design in Example 4.1 are listed at the bottom of the Design-Expert output in Figure 4.2.

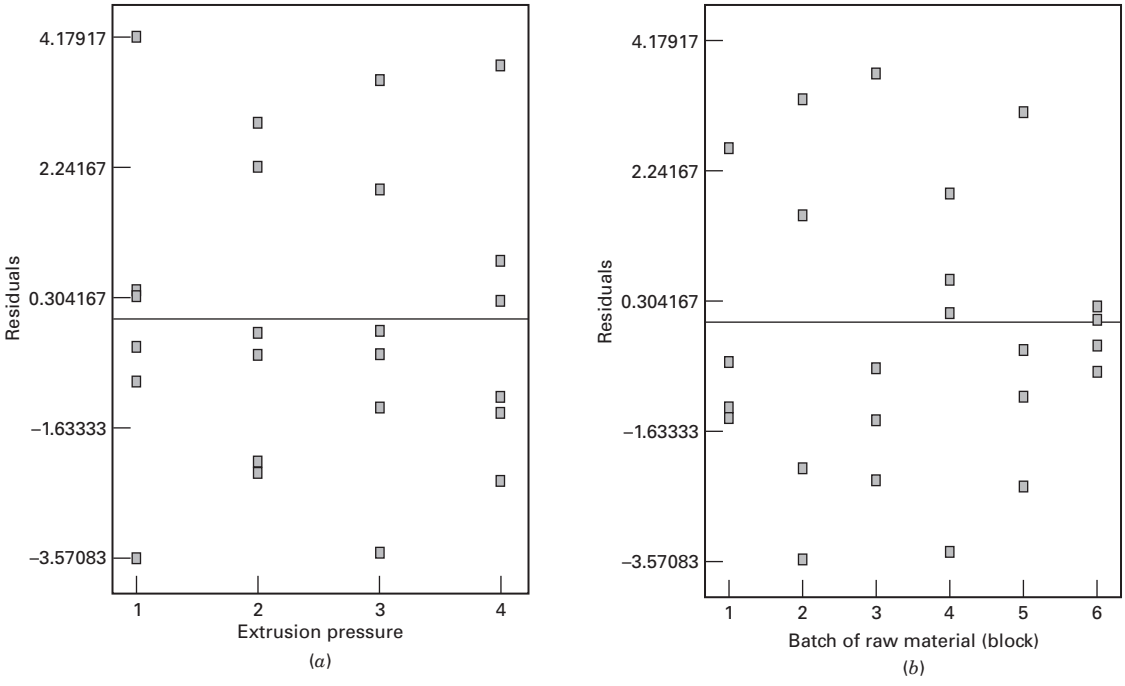
A normal probability plot of these residuals is shown in Figure 4.4. There is no severe indication of nonnormality, nor is there any evidence pointing to possible outliers. Figure 4.5 plots the residuals versus the fitted values $\hat{y}_{ij}$. There should be no relationship between the size of the residuals and the fitted values $\hat{y}_{ij}$. This plot reveals nothing of unusual interest. Figure 4.6 shows plots of the residuals by treatment (extrusion pressure) and by batch of resin or block. These plots are potentially very informative. If there is more scatter in the residuals for a particular treatment, that could indicate that this treatment produces more erratic response readings than the others. More scatter in the residuals for a particular block could indicate that the block is not homogeneous. However, in our example, Figure 4.6 gives no indication of inequality of variance by treatment but there is an indication that there is less variability in the yield for batch 6. However, since all of the other residual plots are satisfactory, we will ignore this.



■ **FIGURE 4.4** Normal probability plot of residuals for Example 4.1



■ **FIGURE 4.5** Plot of residuals versus $\hat{y}_{ij}$ for Example 4.1



■ **FIGURE 4.6** **Plot of residuals by extrusion pressure (treatment) and by batches of resin (block) for Example 4.1**

Sometimes the plot of residuals versus $\hat{y}_{ij}$ has a curvilinear shape; for example, there may be a tendency for negative residuals to occur with low $\hat{y}_{ij}$ values, positive residuals with intermediate $\hat{y}_{ij}$ values, and negative residuals with high $\hat{y}_{ij}$ values. This type of pattern is suggestive of **interaction** between blocks and treatments. If this pattern occurs, a transformation should be used in an effort to eliminate or minimize the interaction. In Section 5.3.7, we describe a statistical test that can be used to detect the presence of interaction in a randomized block design.

4.1.3 Some Other Aspects of the Randomized Complete Block Design

Additivity of the Randomized Block Model. The linear statistical model that we have used for the randomized block design

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

is completely **additive**. This says that, for example, if the first treatment causes the expected response to increase by five units ($\tau_1 = 5$) and if the first block increases the expected response by 2 units ($\beta_1 = 2$), the expected increase in response of *both* treatment 1 *and* block 1 together is $E(y_{11}) = \mu + \tau_1 + \beta_1 = \mu + 5 + 2 = \mu + 7$. In general, treatment 1 *always* increases the expected response by 5 units over the sum of the overall mean and the block effect.

Although this simple additive model is often useful, in some situations it is inadequate. Suppose, for example, that we are comparing four formulations of a chemical product using six batches of raw material; the raw material batches are considered blocks. If an impurity in batch 2 affects formulation 2 adversely, resulting in an unusually low yield, but does not affect the other formulations, an **interaction** between formulations (or treatments) and batches (or

blocks) has occurred. Similarly, interactions between treatments and blocks can occur when the response is measured on the wrong scale. Thus, a relationship that is multiplicative in the original units, say

$$E(y_{ij}) = \mu\tau_i\beta_j$$

is linear or additive in a log scale since, for example,

$$\ln E(y_{ij}) = \ln \mu + \ln \tau_i + \ln \beta_j$$

or

$$E(y_{ij}^*) = \mu^* + \tau_i^* + \beta_j^*$$

Although this type of interaction can be eliminated by a transformation, not all interactions are so easily treated. For example, transformations do not eliminate the formulation–batch interaction discussed previously. Residual analysis and other diagnostic checking procedures can be helpful in detecting nonadditivity.

If interaction is present, it can seriously affect and possibly invalidate the analysis of variance. In general, the presence of interaction inflates the error mean square and may adversely affect the comparison of treatment means. In situations where both factors, as well as their possible interaction, are of interest, **factorial designs** must be used. These designs are discussed extensively in Chapters 5 through 9.

Random Treatments and Blocks. Our presentation of the randomized complete block design thus far has focused on the case when both the treatments and blocks were considered as fixed factors. There are many situations where either treatments or blocks (or both) are random factors. It is very common to find that the blocks are random. This is usually what the experimenter would like to do, because we would like for the conclusions from the experiment to be valid across the population of blocks that the ones selected for the experiments were sampled from. First, we consider the case where the treatments are fixed and the blocks are random. Equation 4.1 is still the appropriate statistical model, but now the block effects are random, that is, we assume that the β_j, $j = 1, 2,..., b$ are $NID(0, \sigma^2_\beta)$ random variables. This is a special case of a mixed model (because it contains both fixed and random factors). In Chapters 13 and 14 we will discuss mixed models in more detail and provide several examples of situations where they occur. Our discussion here is limited to the RCBD.

Assuming that the RCBD model Equation 4.1 is appropriate, if the blocks are random and the treatments are fixed we can show that:

$$\begin{aligned} E(y_{ij}) &= \mu + \tau_i, \qquad i = 1, 2,..., a \\ V(y_{ij}) &= \sigma^2_\beta + \sigma^2 \\ Cov(y_{ij}, y_{i'j'}) &= 0, \quad j \neq j' \\ Cov(y_{ij}, y_{i'j}) &= \sigma^2_\beta \quad i \neq i' \end{aligned} \tag{4.14}$$

Thus, the variance of the observations is constant, the covariance between any two observations in different blocks is zero, but the covariance between two observations from the same block is σ^2_β. The expected mean squares from the usual ANOVA partitioning of the total sum of squares are

$$\begin{aligned} E(MS_{\text{Treatments}}) &= \sigma^2 + \frac{b\sum_{i=1}^{a}\tau_i^2}{a-1} \\ E(MS_{\text{Blocks}}) &= \sigma^2 + a\sigma^2_\beta \\ E(MS_E) &= \sigma^2 \end{aligned} \tag{4.15}$$

The appropriate statistic for testing the null hypothesis of no treatment effects (all $\tau_i = 0$) is

$$F_0 = \frac{MS_{\text{Treatment}}}{MS_E}$$

which is exactly the same test statistic we used in the case where the blocks were fixed. Based on the expected mean squares, we can obtain an ANOVA-type estimator of the variance component for blocks as

$$\hat{\sigma}_\beta^2 = \frac{MS_{\text{Blocks}} - MS_E}{a} \tag{4.16}$$

For example, for the vascular graft experiment in Example 4.1 the estimate of σ_β^2 is

$$\hat{\sigma}_\beta^2 = \frac{MS_{\text{Blocks}} - MS_E}{a} = \frac{38.45 - 7.33}{4} = 7.78$$

This is a method-of-moments estimate and there is no simple way to find a confidence interval on the block variance component σ_β^2. The REML method would be preferred here. Table 4.6 is the JMP output for Example 4.1 assuming that blocks are random. The REML estimate of σ_β^2 is exactly the same as the ANOVA estimate, but REML automatically produces the standard error of the estimate (6.116215) and the approximate 95 percent confidence interval. JMP gives the test for the fixed effect (pressure), and the results are in agreement with those originally reported in Example 4.1. REML also produces the point estimate and CI for the error variance σ^2. The ease with which confidence intervals can be constructed is a major reason why REML has been so widely adopted.

Now consider a situation where there is an interaction between treatments and blocks. This could be accounted for by adding an interaction term to the original statistical model Equation 4.1. Let $(\tau\beta)_{ij}$ be the interaction effect of treatment I in block j. Then the model is

$$y_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{4.17}$$

The interaction effect is assumed to be random because it involves the random block effects. If $\sigma_{\tau\beta}^2$ is the variance component for the block treatment interaction, then we can show that the expected mean squares are

$$\begin{aligned} E(MS_{\text{Treatments}}) &= \sigma^2 + \sigma_{\tau\beta}^2 + \frac{b\sum_{i=1}^{a}\tau_i^2}{a-1} \\ E(MS_{\text{Blocks}}) &= \sigma^2 + a\sigma_\beta^2 \\ E(MS_E) &= \sigma^2 + \sigma_{\tau\beta}^2 \end{aligned} \tag{4.18}$$

From the expected mean squares, we see that the usual F-statistic $F = MS_{\text{Treatments}}/MS_E$ would be used to test for no treatment effects. So another advantage of the random block model is that the assumption of no interaction in the RCBD is not important. However, if blocks are fixed and there is interaction, then the interaction effect is not in the expected mean square for treatments but it is in the error expected mean square, so there would not be a statistical test for the treatment effects.

TABLE 4.6
JMP Output for Example 4.1 with Blocks Assumed Random

Response Y

Summary of Fit

RSquare	0.756688
RSquare Adj	0.720192
Root Mean Square Error	2.706612
Mean of Response	89.79583
Observations (or Sum Wgts)	24

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Block	1.0621666	7.7811667	6.116215	−4.206394	19.768728	51.507
Residual		7.32575	2.6749857	3.9975509	17.547721	48.493
Total		15.106917				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	Block	Residual
Block	37.408085	−1.788887
Residual	−1.788887	7.1555484

Fixed Effect Tests

Source	Nparm	DF	DFDen	F Ratio	Prob > F
Pressure	3	3	15	8.1071	0.0019*

Choice of Sample Size. Choosing the **sample size**, or the **number of blocks** to run, is an important decision when using an RCBD. Increasing the number of blocks increases the number of replicates and the number of error degrees of freedom, making design more sensitive. Any of the techniques discussed in Section 3.7 for selecting the number of replicates to run in a completely randomized single-factor experiment may be applied directly to the RCBD. For the case of a fixed factor, the operating characteristic curves in Appendix Chart V may be used with

$$\Phi^2 = \frac{b\sum_{i=1}^{a}\tau_i^2}{a\sigma^2} \tag{4.19}$$

where there are $a - 1$ numerator degrees of freedom and $(a - 1)(b - 1)$ denominator degrees of freedom.

EXAMPLE 4.2

Consider the RCBD for the vascular grafts described in Example 4.1. Suppose that we wish to determine the appropriate number of blocks to run if we are interested in detecting a true maximum difference in yield of 6 with a reasonably high probability and an estimate of the standard deviation of the errors is $\sigma = 3$. From Equation 3.45, the minimum value of Φ^2 is (writing b, the number of blocks, for n)

$$\Phi^2 = \frac{bD^2}{2a\sigma^2}$$

where D is the maximum difference we wish to detect. Thus,

$$\Phi^2 = \frac{b(6)^2}{2(4)(3)^2} = 0.5b$$

If we use $b = 5$ blocks, $\Phi = \sqrt{0.5b} = \sqrt{0.5(5)} = 1.58$, and there are $(a - 1)(b - 1) = 3(4) = 12$ error degrees of freedom. Appendix Chart V with $\nu_1 = a - 1 = 3$ and $\alpha = 0.05$ indicates that the β risk for this design is approximately 0.55 (power $= 1 - \beta = 0.45$). If we use $b = 6$ blocks, $\Phi = \sqrt{0.5b} = \sqrt{0.5(6)} = 1.73$, with $(a - 1)(b - 1) = 3(5) = 15$ error degrees of freedom, and the corresponding β risk is approximately 0.4 (power $= 1 - \beta = 0.6$). Because the batches of resin are expensive and the cost of experimentation is high, the experimenter decides to use six blocks, even though the power is only about 0.6 (actually many experiments work very well with power values of only 0.5 or higher).

Estimating Missing Values. When using the RCBD, sometimes an observation in one of the blocks is missing. This may happen because of carelessness or error or for reasons beyond our control, such as unavoidable damage to an experimental unit. A missing observation introduces a new problem into the analysis because treatments are no longer **orthogonal to blocks**; that is, every treatment does not occur in every block. There are two general approaches to the missing value problem. The first is an **approximate analysis** in which the missing observation is estimated and the usual analysis of variance is performed just as if the estimated observation were real data, with the error degrees of freedom reduced by 1. This approximate analysis is the subject of this section. The second is an **exact analysis**, which is discussed in Section 4.1.4.

Suppose the observation y_{ij} for treatment i in block j is missing. Denote the missing observation by x. As an illustration, suppose that in the vascular graft experiment of Example 4.1 there was a problem with the extrusion machine when the 8700 psi run was conducted in the fourth batch of material, and the observation y_{24} could not be obtained. The data might appear as in Table 4.7.

In general, we will let y'_{ij} represent the grand total with one missing observation, $y'_{i.}$ represent the total for the treatment with one missing observation, and $y'_{.j}$ be the total for the block with one missing observation. Suppose we wish to estimate the missing observation x

■ **TABLE 4.7**
Randomized Complete Block Design for the Vascular Graft Experiment with One Missing Value

Extrusion Pressures (PSI)	Batch of Resin (Block) 1	2	3	4	5	6	
8500	90.3	89.2	98.2	93.9	87.4	97.9	556.9
8700	92.5	89.5	90.6	x	87.0	95.8	455.4
8900	85.5	90.8	89.6	86.2	88.0	93.4	533.5
9100	82.5	89.5	85.6	87.4	78.9	90.7	514.6
Block totals	350.8	359.0	364.0	267.5	341.3	377.8	$y'_{..} = 2060.4$

■ **TABLE 4.8**
Approximate Analysis of Variance for Example 4.1 with One Missing Value

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Extrusion pressure	166.14	3	55.38	7.63	0.0029
Batches of raw material	189.52	5	37.90		
Error	101.70	14	7.26		
Total	457.36	23			

so that x will have a minimum contribution to the error sum of squares. Because $SS_E = \Sigma_{i=1}^{a}\Sigma_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$, this is equivalent to choosing x to minimize

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{1}{b}\sum_{i=1}^{a}\left(\sum_{j=1}^{b} y_{ij}\right)^2 - \frac{1}{a}\sum_{j=1}^{b}\left(\sum_{i=1}^{a} y_{ij}\right)^2 + \frac{1}{ab}\left(\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}\right)^2$$

or

$$SS_E = x^2 - \frac{1}{b}(y'_{i.} + x)^2 - \frac{1}{a}(y'_{.j} + x)^2 + \frac{1}{ab}(y'_{..} + x)^2 + R \tag{4.20}$$

where R includes all terms not involving x. From $dSS_E/dx = 0$, we obtain

$$x = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} \tag{4.21}$$

as the estimate of the missing observation.

For the data in Table 4.7, we find that $y'_{2.} = 455.4$, $y'_{.4} = 267.5$, and $y'_{..} = 2060.4$. Therefore, from Equation 4.16,

$$x \equiv y_{24} = \frac{4(455.4) + 6(267.5) - 2060.4}{(3)(5)} = 91.08$$

The usual analysis of variance may now be performed using $y_{24} = 91.08$ and reducing the error degrees of freedom by 1. The analysis of variance is shown in Table 4.8. Compare the results of this approximate analysis with the results obtained for the full data set (Table 4.4).

If several observations are missing, they may be estimated by writing the error sum of squares as a function of the missing values, differentiating with respect to each missing value, equating the results to zero, and solving the resulting equations. Alternatively, we may use Equation 4.21 iteratively to estimate the missing values. To illustrate the iterative approach, suppose that two values are missing. Arbitrarily estimate the first missing value, and then use this value along with the real data and Equation 4.21 to estimate the second. Now Equation 4.21 can be used to reestimate the first missing value, and following this, the second can be reestimated. This process is continued until convergence is obtained. In any missing value problem, the error degrees of freedom are reduced by one for each missing observation.

4.1.4 Estimating Model Parameters and the General Regression Significance Test

If both treatments and blocks are fixed, we may estimate the parameters in the RCBD model by least squares. Recall that the linear statistical model is

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{4.22}$$

Applying the rules in Section 3.9.2 for finding the normal equations for an experimental design model, we obtain

$$
\begin{aligned}
\mu\colon\ & ab\hat{\mu} + b\hat{\tau}_1 + b\hat{\tau}_2 + \cdots + b\hat{\tau}_a + a\hat{\beta}_1 + a\hat{\beta}_2 + \cdots + a\hat{\beta}_b = y_{..} \\
\tau_1\colon\ & b\hat{\mu} + b\hat{\tau}_1 + \hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_b = y_{1.} \\
\tau_2\colon\ & b\hat{\mu} + b\hat{\tau}_2 + \hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_b = y_{2.} \\
& \vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\
\tau_a\colon\ & b\hat{\mu} + b\hat{\tau}_a + \hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_b = y_{a.} \\
\beta_1\colon\ & a\hat{\mu} + \hat{\tau}_1 + \hat{\tau}_2 + \cdots + \hat{\tau}_a + a\hat{\beta}_1 = y_{.1} \\
\beta_2\colon\ & a\hat{\mu} + \hat{\tau}_1 + \hat{\tau}_2 + \cdots + \hat{\tau}_a + a\hat{\beta}_2 = y_{.2} \\
& \vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\
\beta_b\colon\ & a\hat{\mu} + \hat{\tau}_1 + \hat{\tau}_2 + \cdots + \hat{\tau}_a + a\hat{\beta}_b = y_{.b}
\end{aligned}
\tag{4.23}
$$

Notice that the second through the $(a + 1)$st equations in Equation 4.23 sum to the first normal equation, as do the last b equations. Thus, there are two linear dependencies in the normal equations, implying that two constraints must be imposed to solve Equation 4.23. The usual constraints are

$$
\sum_{i=1}^{a} \hat{\tau}_i = 0 \qquad \sum_{j=1}^{b} \hat{\beta}_j = 0 \tag{4.24}
$$

Using these constraints helps simplify the normal equations considerably. In fact, they become

$$
\begin{aligned}
ab\,\hat{\mu} &= y_{..} \\
b\hat{\mu} + b\hat{\tau}_i &= y_{i.} \qquad i = 1, 2, \ldots, a \\
a\hat{\mu} + a\hat{\beta}_j &= y_{.j} \qquad j = 1, 2, \ldots, b
\end{aligned}
\tag{4.25}
$$

whose solution is

$$
\begin{aligned}
\hat{\mu} &= \bar{y}_{..} \\
\hat{\tau}_i &= \bar{y}_{i.} - \bar{y}_{..} \qquad i = 1, 2, \ldots, a \\
\hat{\beta}_j &= \bar{y}_{.j} - \bar{y}_{..} \qquad j = 1, 2, \ldots, b
\end{aligned}
\tag{4.26}
$$

Using the solution to the normal equation in Equation 4.26, we may find the estimated or fitted values of y_{ij} as

$$
\begin{aligned}
\hat{y}_{ij} &= \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j \\
&= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) \\
&= \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}
\end{aligned}
$$

This result was used previously in Equation 4.13 for computing the residuals from a randomized block design.

The general regression significance test can be used to develop the analysis of variance for the randomized complete block design. Using the solution to the normal equations given by Equation 4.26, the reduction in the sum of squares for fitting the **full model** is

$$\begin{aligned} R(\mu, \tau, \beta) &= \hat{\mu} y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} + \sum_{j=1}^{b} \hat{\beta}_j y_{.j} \\ &= \bar{y}_{..} y_{..} + \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..}) y_{i.} + \sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..}) y_{.j} \\ &= \frac{y_{..}^2}{ab} + \sum_{i=1}^{a} \bar{y}_{i.} y_{i.} - \frac{y_{..}^2}{ab} + \sum_{j=1}^{b} \bar{y}_{.j} y_{.j} - \frac{y_{..}^2}{ab} \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} + \sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab} \end{aligned}$$

with $a + b - 1$ degrees of freedom, and the error sum of squares is

$$\begin{aligned} SS_E &= \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2 - R(\mu, \tau, \beta) \\ &= \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2 - \sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \sum_{j=1}^{b} \frac{y_{.j}^2}{a} + \frac{y_{..}^2}{ab} \\ &= \sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \end{aligned}$$

with $(a - 1)(b - 1)$ degrees of freedom. Compare this last equation with SS_E in Equation 4.7.

To test the hypothesis H_0: $\tau_i = 0$, the **reduced model** is

$$y_{ij} = \mu + \beta_j + \epsilon_{ij}$$

which is just a single-factor analysis of variance. By analogy with Equation 3.5, the reduction in the sum of squares for fitting the reduced model is

$$R(\mu, \beta) = \sum_{j=1}^{b} \frac{y_{.j}^2}{a}$$

which has b degrees of freedom. Therefore, the sum of squares due to $\{\tau_i\}$ after fitting μ and $\{\beta_j\}$ is

$$\begin{aligned} R(\tau|\mu, \beta) &= R(\mu, \tau, \beta) - R(\mu, \beta) \\ &= R(\text{full model}) - R(\text{reduced model}) \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} + \sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab} - \sum_{j=1}^{b} \frac{y_{.j}^2}{a} \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{ab} \end{aligned}$$

which we recognize as the treatment sum of squares with $a - 1$ degrees of freedom (Equation 4.10).

The block sum of squares is obtained by fitting the **reduced model**

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

which is also a single-factor analysis. Again, by analogy with Equation 3.5, the reduction in the sum of squares for fitting this model is

$$R(\mu, \tau) = \sum_{i=1}^{a} \frac{y_{i.}^2}{b}$$

with a degrees of freedom. The sum of squares for blocks $\{\beta_j\}$ after fitting μ and $\{\tau_i\}$ is

$$\begin{aligned} R(\beta|\mu, \tau) &= R(\mu, \tau, \beta) - R(\mu, \tau) \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} + \sum_{j=1}^{b} \frac{y_{j}^2}{a} - \frac{y_{..}^2}{ab} - \sum_{i=1}^{a} \frac{y_{i.}^2}{b} \\ &= \sum_{j=1}^{b} \frac{y_{j}^2}{a} - \frac{y_{..}^2}{ab} \end{aligned}$$

with $b - 1$ degrees of freedom, which we have given previously as Equation 4.11.

We have developed the sums of squares for treatments, blocks, and error in the randomized complete block design using the general regression significance test. Although we would not ordinarily use the general regression significance test to actually analyze data in a randomized complete block, the procedure occasionally proves useful in more general randomized block designs, such as those discussed in Section 4.4.

Exact Analysis of the Missing Value Problem. In Section 4.1.3 an approximate procedure for dealing with missing observations in the RCBD was presented. This approximate analysis consists of estimating the missing value so that the error mean square is minimized. It can be shown that the approximate analysis produces a biased mean square for treatments in the sense that $E(MS_{\text{Treatments}})$ is larger than $E(MS_E)$ if the null hypothesis is true. Consequently, too many significant results are reported.

The missing value problem may be analyzed exactly by using the general regression significance test. The missing value causes the design to be **unbalanced**, and because all the treatments do not occur in all blocks, we say that the treatments and blocks are not **orthogonal**. This method of analysis is also used in more general types of randomized block designs; it is discussed further in Section 4.4. Many computer packages will perform this analysis.

4.2 The Latin Square Design

In Section 4.1 we introduced the randomized complete block design as a design to reduce the residual error in an experiment by removing variability due to a known and controllable nuisance variable. There are several other types of designs that utilize the blocking principle. For example, suppose that an experimenter is studying the effects of five different formulations of a rocket propellant used in aircrew escape systems on the observed burning rate. Each formulation is mixed from a batch of raw material that is only large enough for five formulations to be tested. Furthermore, the formulations are prepared by several operators, and there may be substantial differences in the skills and experience of the operators. Thus, it would seem that there are two nuisance factors to be "averaged out" in the design: batches of raw material and operators. The appropriate design for this problem consists of testing each formulation exactly once in each batch of raw material and for each formulation to be prepared exactly once by each of five operators. The resulting design, shown in Table 4.9, is called a **Latin square design**. Notice that the design is a square arrangement and that the five formulations (or treatments) are denoted by the Latin letters *A, B, C, D*, and *E*; hence the name Latin square.

■ **TABLE 4.9**
Latin Square Design for the Rocket Propellant Problem

Batches of Raw Material	Operators				
	1	**2**	**3**	**4**	**5**
1	$A = 24$	$B = 20$	$C = 19$	$D = 24$	$E = 24$
2	$B = 17$	$C = 24$	$D = 30$	$E = 27$	$A = 36$
3	$C = 18$	$D = 38$	$E = 26$	$A = 27$	$B = 21$
4	$D = 26$	$E = 31$	$A = 26$	$B = 23$	$C = 22$
5	$E = 22$	$A = 30$	$B = 20$	$C = 29$	$D = 31$

We see that both batches of raw material (rows) and operators (columns) are orthogonal to treatments.

The Latin square design is used to eliminate two nuisance sources of variability; that is, it systematically allows blocking in two directions. Thus, the rows and columns actually represent **two restrictions on randomization**. In general, a Latin square for p factors, or a $p \times p$ Latin square, is a square containing p rows and p columns. Each of the resulting p^2 cells contains one of the p letters that corresponds to the treatments, and each letter occurs once and only once in each row and column. Some examples of Latin squares are

4×4	5×5	6×6
A B D C	*A D B E C*	*A D C E B F*
B C A D	*D A C B E*	*B A E C F D*
C D B A	*C B E D A*	*C E D F A B*
D A C B	*B E A C D*	*D C F B E A*
	E C D A B	*F B A D C E*
		E F B A D C

Latin squares are closely related to a popular puzzle called a sudoku puzzle that originated in Japan (sudoku means "single number" in Japanese). The puzzle typically consists of a 9×9 grid, with nine additional 3×3 blocks contained within. A few of the spaces contain numbers and the others are blank. The goal is to fill the blanks with the integers from 1 to 9 so that each row, each column, and each of the nine 3×3 blocks making up the grid contains just one of each of the nine integers. The additional constraint that a standard 9×9 sudoku puzzle have 3×3 blocks that also contain each of the nine integers reduces the large number of possible 9×9 Latin squares to a smaller but still quite large number, approximately 6×10^{21}.

Depending on the number of clues and the size of the grid, sudoku puzzles can be extremely difficult to solve. Solving an $n \times n$ sudoku puzzle belongs to a class of computational problems called *NP*-complete (the *NP* refers to non-polynomial computing time). An *NP*-complete problem is one for which it's relatively easy to check whether a particular answer is correct but may require an impossibly long time to solve by any simple algorithm as n gets larger.

Solving a sudoku puzzle is also equivalent to "coloring" a graph—an array of points (vertices) and lines (edges) in a particular way. In this case, the graph has 81 vertices, one for each cell of the grid. Depending on the puzzle, only certain pairs of vertices are joined by an edge. Given that some vertices have already been assigned a "color" (chosen from the nine number possibilities), the problem is to "color" the remaining vertices so that any two vertices joined by an edge don't have the same "color."

The **statistical model** for a Latin square is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \end{cases} \tag{4.27}$$

where y_{ijk} is the observation in the ith row and kth column for the jth treatment, μ is the overall mean, α_i is the ith row effect, τ_j is the jth treatment effect, β_k is the kth column effect, and ϵ_{ijk} is the random error. Note that this is an **effects model**. The model is completely **additive**; that is, there is no interaction between rows, columns, and treatments. Because there is only one observation in each cell, only two of the three subscripts i, j, and k are needed to denote a particular observation. For example, referring to the rocket propellant problem in Table 4.8, if $i = 2$ and $k = 3$, we automatically find $j = 4$ (formulation D), and if $i = 1$ and $j = 3$ (formulation C), we find $k = 3$. This is a consequence of each treatment appearing exactly once in each row and column.

The analysis of variance consists of partitioning the total sum of squares of the $N = p^2$ observations into components for rows, columns, treatments, and error, for example,

$$SS_T = SS_{\text{Rows}} + SS_{\text{Columns}} + SS_{\text{Treatments}} + SS_E \tag{4.28}$$

with respective degrees of freedom

$$p^2 - 1 = p - 1 + p - 1 + p - 1 + (p - 2)(p - 1)$$

Under the usual assumption that ϵ_{ijk} is NID $(0, \sigma^2)$, each sum of squares on the right-hand side of Equation 4.28 is, upon division by σ^2, an independently distributed chi-square random variable. The appropriate statistic for testing for no differences in treatment means is

$$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$$

which is distributed as $F_{p-1,(p-2)(p-1)}$ under the null hypothesis. We may also test for no row effect and no column effect by forming the ratio of MS_{Rows} or MS_{Columns} to MS_E. However, because the rows and columns represent restrictions on randomization, these tests may not be appropriate.

The computational procedure for the ANOVA in terms of treatment, row, and column totals is shown in Table 4.10. From the computational formulas for the sums of squares, we see that the analysis is a simple extension of the RCBD, with the sum of squares resulting from rows obtained from the row totals.

■ **TABLE 4.10**
Analysis of Variance for the Latin Square Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}} = \frac{1}{p}\sum_{j=1}^{p} y_{\cdot j \cdot}^2 - \frac{y_{\cdots}^2}{N}$	$p - 1$	$\frac{SS_{\text{Treatments}}}{p-1}$	$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$
Rows	$SS_{\text{Rows}} = \frac{1}{p}\sum_{i=1}^{p} y_{i\cdots}^2 - \frac{y_{\cdots}^2}{N}$	$p - 1$	$\frac{SS_{\text{Rows}}}{p-1}$	
Columns	$SS_{\text{Columns}} = \frac{1}{p}\sum_{k=1}^{p} y_{\cdots k}^2 - \frac{y_{\cdots}^2}{N}$	$p - 1$	$\frac{SS_{\text{Columns}}}{p-1}$	
Error	SS_E (by subtraction)	$(p - 2)(p - 1)$	$\frac{SS_E}{(p-2)(p-1)}$	
Total	$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{\cdots}^2}{N}$	$p^2 - 1$		

EXAMPLE 4.3

Consider the rocket propellant problem previously described, where both batches of raw material and operators represent randomization restrictions. The design for this experiment, shown in Table 4.8, is a 5×5 Latin square. After coding by subtracting 25 from each observation, we have the data in Table 4.11. The sums of squares for the total, batches (rows), and operators (columns) are computed as follows:

$$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$$
$$= 680 - \frac{(10)^2}{25} = 676.00$$

$$SS_{\text{Batches}} = \frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N}$$
$$= \frac{1}{5}[(-14)^2 + 9^2 + 5^2 + 3^2 + 7^2] - \frac{(10)^2}{25} = 68.00$$

$$SS_{\text{Operators}} = \frac{1}{p}\sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N}$$
$$= \frac{1}{5}[(-18)^2 + 18^2 + (-4)^2 + 5^2 + 9^2] - \frac{(10)^2}{25} = 150.00$$

The totals for the treatments (Latin letters) are

Latin Letter	Treatment Total
A	$y_{.1.} = 18$
B	$y_{.2.} = -24$
C	$y_{.3.} = -13$
D	$y_{.4.} = 24$
E	$y_{.5.} = 5$

The sum of squares resulting from the formulations is computed from these totals as

$$SS_{\text{Formulations}} = \frac{1}{p}\sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$$
$$= \frac{18^2 + (-24)^2 + (-13)^2 + 24^2 + 5^2}{5} - \frac{(10)^2}{25} = 330.00$$

The error sum of squares is found by subtraction

$$SS_E = SS_T - SS_{\text{Batches}} - SS_{\text{Operators}} - SS_{\text{Formulations}}$$
$$= 676.00 - 68.00 - 150.00 - 330.00 = 128.00$$

The analysis of variance is summarized in Table 4.12. We conclude that there is a significant difference in the mean burning rate generated by the different rocket propellant formulations. There is also an indication that differences between operators exist, so blocking on this factor was a good precaution. There is no strong evidence of a difference between batches of raw material, so it seems that in this particular experiment we were unnecessarily concerned about this source of variability. However, blocking on batches of raw material is usually a good idea.

■ **TABLE 4.11**
Coded Data for the Rocket Propellant Problem

Batches of Raw Material	Operators 1	2	3	4	5	$y_{i..}$
1	$A = -1$	$B = -5$	$C = -6$	$D = -1$	$E = -1$	-14
2	$B = -8$	$C = -1$	$D = 5$	$E = 2$	$A = 11$	9
3	$C = -7$	$D = 13$	$E = 1$	$A = 2$	$B = -4$	5
4	$D = 1$	$E = 6$	$A = 1$	$B = -2$	$C = -3$	3
5	$E = -3$	$A = 5$	$B = -5$	$C = 4$	$D = 6$	7
$y_{..k}$	-18	18	-4	5	9	$10 = y_{...}$

■ TABLE 4.12
Analysis of Variance for the Rocket Propellant Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Formulations	330.00	4	82.50	7.73	0.0025
Batches of raw material	68.00	4	17.00		
Operators	150.00	4	37.50		
Error	128.00	12	10.67		
Total	676.00	24			

As in any design problem, the experimenter should investigate the adequacy of the model by inspecting and plotting the residuals. For a Latin square, the residuals are given by

$$\begin{aligned} e_{ijk} &= y_{ijk} - \hat{y}_{ijk} \\ &= y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..k} + 2\bar{y}_{...} \end{aligned}$$

The reader should find the residuals for Example 4.3 and construct appropriate plots.

A Latin square in which the first row and column consists of the letters written in alphabetical order is called a **standard Latin square**, which is the design shown in Example 4.4. A standard Latin square can always be obtained by writing the first row in alphabetical order and then writing each successive row as the row of letters just above shifted one place to the left. Table 4.13 summarizes several important facts about Latin squares and standard Latin squares.

As with any experimental design, the observations in the Latin square should be taken in random order. The proper randomization procedure is to select the particular square employed at random. As we see in Table 4.13, there are a large number of Latin squares of a particular size, so it is impossible to enumerate all the squares and select one randomly. The usual procedure is

■ TABLE 4.13
Standard Latin Squares and Number of Latin Squares of Various Sizes[a]

Size	3 × 3	4 × 4	5 × 5	6 × 6	7 × 7	$p \times p$
Examples of standard squares	*A B C* *B C A* *C A B*	*A B C D* *B C D A* *C D A B* *D A B C*	*A B C D E* *B A E C D* *C D A E B* *D E B A C* *E C D B A*	*A B C D E F* *B C F A D E* *C F B E A D* *D E A B F C* *E A D F C B* *F D E C B A*	*A B C D E F G* *B C D E F G A* *C D E F G A B* *D E F G A B C* *E F G A B C D* *F G A B C D E* *G A B C D E F*	*ABC . . . P* *BCD . . . A* *CDE . . . B* ⋮ *PAB . . . (P − 1)*
Number of standard squares	1	4	56	9408	16,942,080	—
Total number of Latin squares	12	576	161,280	818,851,200	61,479,419,904,000	$p!(p-1)! \times$ (number of standard squares)

[a]Some of the information in this table is found in Fisher and Yates (1953). Little is known about the properties of Latin squares larger than 7 × 7.

to select an arbitrary Latin square from a table of such designs, as in Fisher and Yates (1953), or start with a standard square, and then arrange the order of the rows, columns, and letters at random. This is discussed more completely in Fisher and Yates (1953).

Occasionally, one observation in a Latin square is missing. For a $p \times p$ Latin square, the missing value may be estimated by

$$y_{ijk} = \frac{p(y'_{i..} + y'_{.j.} + y'_{..k}) - 2y'_{...}}{(p-2)(p-1)} \tag{4.29}$$

where the primes indicate totals for the row, column, and treatment with the missing value, and $y'_{...}$ is the grand total with the missing value.

Latin squares can be useful in situations where the rows and columns represent factors the experimenter actually wishes to study and where there are no randomization restrictions. Thus, three factors (rows, columns, and letters), each at p levels, can be investigated in only p^2 runs. This design assumes that there is no interaction between the factors. More will be said later on the subject of interaction.

Replication of Latin Squares. A disadvantage of small Latin squares is that they provide a relatively small number of error degrees of freedom. For example, a 3×3 Latin square has only two error degrees of freedom, a 4×4 Latin square has only six error degrees of freedom, and so forth. When small Latin squares are used, it is frequently desirable to replicate them to increase the error degrees of freedom.

A Latin square may be replicated in several ways. To illustrate, suppose that the 5×5 Latin square used in Example 4.4 is replicated n times. This could have been done as follows:

1. Use the same batches and operators in each replicate.
2. Use the same batches but different operators in each replicate (or, equivalently, use the same operators but different batches).
3. Use different batches and different operators.

The analysis of variance depends on the method of replication.

Consider case 1, where the same levels of the row and column blocking factors are used in each replicate. Let y_{ijkl} be the observation in row i, treatment j, column k, and replicate l. There are $N = np^2$ total observations. The ANOVA is summarized in Table 4.14.

■ **TABLE 4.14**
Analysis of Variance for a Replicated Latin Square, Case 1

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Treatments}}}{p-1}$	$\frac{MS_{\text{Treatments}}}{MS_E}$
Rows	$\frac{1}{np}\sum_{i=1}^{p} y_{i...}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Rows}}}{p-1}$	
Columns	$\frac{1}{np}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Columns}}}{p-1}$	
Replicates	$\frac{1}{p^2}\sum_{l=1}^{n} y_{...l}^2 - \frac{y_{....}^2}{N}$	$n-1$	$\frac{SS_{\text{Replicates}}}{n-1}$	
Error	Subtraction	$(p-1)[n(p+1)-3]$	$\frac{SS_E}{(p-1)[n(p+1)-3]}$	
Total	$\sum\sum\sum\sum y_{ijkl}^2 - \frac{y_{....}^2}{N}$	np^2-1		

■ **TABLE 4.15**
Analysis of Variance for a Replicated Latin Square, Case 2

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$\frac{1}{np}\sum_{j=1}^{p} y_{\cdot j\cdot\cdot}^2 - \frac{y_{\cdots\cdot}^2}{N}$	$p-1$	$\frac{SS_{\text{Treatments}}}{p-1}$	$\frac{MS_{\text{Treatments}}}{MS_E}$
Rows	$\frac{1}{p}\sum_{l=1}^{n}\sum_{i=1}^{p} y_{i\cdot\cdot l}^2 - \sum_{l=1}^{n}\frac{y_{\cdots l}^2}{p^2}$	$n(p-1)$	$\frac{SS_{\text{Rows}}}{n(p-1)}$	
Columns	$\frac{1}{np}\sum_{k=1}^{p} y_{\cdot\cdot k\cdot}^2 - \frac{y_{\cdots\cdot}^2}{N}$	$p-1$	$\frac{SS_{\text{Columns}}}{p-1}$	
Replicates	$\frac{1}{p^2}\sum_{l=1}^{n} y_{\cdots l}^2 - \frac{y_{\cdots\cdot}^2}{N}$	$n-1$	$\frac{SS_{\text{Replicates}}}{n-1}$	
Error	Subtraction	$(p-1)(np-1)$	$\frac{SS_E}{(p-1)(np-1)}$	
Total	$\sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{\cdots\cdot}^2}{N}$	np^2-1		

Now consider case 2 and assume that new batches of raw material but the same operators are used in each replicate. Thus, there are now five new rows (in general, p new rows) within each replicate. The ANOVA is summarized in Table 4.15. Note that the source of variation for the rows really measures the variation between rows within the n replicates.

Finally, consider case 3, where new batches of raw material and new operators are used in each replicate. Now the variation that results from both the rows and columns measures the variation resulting from these factors within the replicates. The ANOVA is summarized in Table 4.16.

There are other approaches to analyzing replicated Latin squares that allow some interactions between treatments and squares (refer to Problem 4.30).

Crossover Designs and Designs Balanced for Residual Effects. Occasionally, one encounters a problem in which time periods are a factor in the experiment. In general, there are p treatments to be tested in p time periods using np experimental units. For example, a human performance analyst is studying the effect of two replacement fluids on dehydration

■ **TABLE 4.16**
Analysis of Variance for a Replicated Latin Square, Case 3

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$\frac{1}{np}\sum_{j=1}^{p} y_{\cdot j\cdot\cdot}^2 - \frac{y_{\cdots\cdot}^2}{N}$	$p-1$	$\frac{SS_{\text{Treatments}}}{p-1}$	$\frac{MS_{\text{Treatments}}}{MS_E}$
Rows	$\frac{1}{p}\sum_{l=1}^{n}\sum_{i=1}^{p} y_{i\cdot\cdot l}^2 - \sum_{l=1}^{n}\frac{y_{\cdots l}^2}{p^2}$	$n(p-1)$	$\frac{SS_{\text{Rows}}}{n(p-1)}$	
Columns	$\frac{1}{p}\sum_{l=1}^{n}\sum_{k=1}^{p} y_{\cdot\cdot kl}^2 - \sum_{l=1}^{n}\frac{y_{\cdots l}^2}{p^2}$	$n(p-1)$	$\frac{SS_{\text{Columns}}}{n(p-1)}$	
Replicates	$\frac{1}{p^2}\sum_{l=1}^{n} y_{\cdots l}^2 - \frac{y_{\cdots\cdot}^2}{N}$	$n-1$	$\frac{SS_{\text{Replicates}}}{n-1}$	
Error	Subtraction	$(p-1)[n(p-1)-1]$	$\frac{SS_E}{(p-1)[n(p-1)-1]}$	
Total	$\sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{\cdots\cdot}^2}{N}$	np^2-1		

	Latin Squares																			
	I		II		III		IV		V		VI		VII		VIII		IX		X	
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Period 1	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*A*	*B*	*B*	*A*	*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*
Period 2	*B*	*A*	*A*	*B*	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*A*

■ **FIGURE 4.7** **A crossover design**

■ **TABLE 4.17**
Analysis of Variance for the Crossover Design in Figure 4.7

Source of Variation	Degrees of Freedom
Subjects (columns)	19
Periods (rows)	1
Fluids (letters)	1
Error	18
Total	39

in 20 subjects. In the first period, half of the subjects (chosen at random) are given fluid *A* and the other half fluid *B*. At the end of the period, the response is measured and a period of time is allowed to pass in which any physiological effect of the fluids is eliminated. Then the experimenter has the subjects who took fluid *A* take fluid *B* and those who took fluid *B* take fluid *A*. This design is called a **crossover design**. It is analyzed as a set of 10 Latin squares with two rows (time periods) and two treatments (fluid types). The two columns in each of the 10 squares correspond to subjects.

The layout of this design is shown in Figure 4.7. Notice that the rows in the Latin square represent the time periods and the columns represent the subjects. The 10 subjects who received fluid *A* first (1, 4, 6, 7, 9, 12, 13, 15, 17, and 19) are randomly determined.

An abbreviated analysis of variance is summarized in Table 4.17. The subject sum of squares is computed as the corrected sum of squares among the 20 subject totals, the period sum of squares is the corrected sum of squares among the rows, and the fluid sum of squares is computed as the corrected sum of squares among the letter totals. For further details of the statistical analysis of these designs see Cochran and Cox (1957), John (1971), and Anderson and McLean (1974).

It is also possible to employ Latin square type designs for experiments in which the treatments have a **residual effect**—that is, for example, if the data for fluid *B* in period 2 still reflected some effect of fluid *A* taken in period 1. Designs balanced for residual effects are discussed in detail by Cochran and Cox (1957) and John (1971).

4.3 The Graeco-Latin Square Design

Consider a $p \times p$ Latin square, and superimpose on it a second $p \times p$ Latin square in which the treatments are denoted by Greek letters. If the two squares when superimposed have the property that each Greek letter appears once and only once with each Latin letter, the two Latin squares are said to be **orthogonal**, and the design obtained is called a **Graeco-Latin square**. An example of a 4×4 Graeco-Latin square is shown in Table 4.18.

■ **TABLE 4.18**
4 × 4 Graeco-Latin Square Design

	Column			
Row	**1**	**2**	**3**	**4**
1	$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
2	$B\delta$	$A\gamma$	$D\beta$	$C\alpha$
3	$C\beta$	$D\alpha$	$A\delta$	$B\gamma$
4	$D\gamma$	$C\delta$	$B\alpha$	$A\beta$

The Graeco-Latin square design can be used to control systematically three sources of extraneous variability, that is, to block in *three* directions. The design allows investigation of four factors (rows, columns, Latin letters, and Greek letters), each at p levels in only p^2 runs. Graeco-Latin squares exist for all $p \geq 3$ except $p = 6$.

The statistical model for the Graeco-Latin square design is

$$y_{ijkl} = \mu + \theta_i + \tau_j + \omega_k + \Psi_l + \epsilon_{ijkl} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \\ l = 1, 2, \ldots, p \end{cases} \tag{4.30}$$

where y_{ijkl} is the observation in row i and column l for Latin letter j and Greek letter k, θ_i is the effect of the ith row, τ_j is the effect of Latin letter treatment j, ω_k is the effect of Greek letter treatment k, Ψ_l is the effect of column l, and ϵ_{ijkl} is an NID $(0, \sigma^2)$ random error component. Only two of the four subscripts are necessary to completely identify an observation.

The analysis of variance is very similar to that of a Latin square. Because the Greek letters appear exactly once in each row and column and exactly once with each Latin letter, the factor represented by the Greek letters is orthogonal to rows, columns, and Latin letter treatments. Therefore, a sum of squares due to the Greek letter factor may be computed from the Greek letter totals, and the experimental error is further reduced by this amount. The computational details are illustrated in Table 4.19. The null hypotheses of equal row, column, Latin letter, and Greek letter treatments would be tested by dividing the corresponding mean square by mean square error. The rejection region is the upper tail point of the $F_{p-1,(p-3)(p-1)}$ distribution.

■ **TABLE 4.19**
Analysis of Variance for a Graeco-Latin Square Design

Source of Variation	**Sum of Squares**	**Degrees of Freedom**
Latin letter treatments	$SS_L = \frac{1}{p}\sum_{j=1}^{p} y_{j..}^2 - \frac{y_{....}^2}{N}$	$p - 1$
Greek letter treatments	$SS_G = \frac{1}{p}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	$p - 1$
Rows	$SS_{\text{Rows}} = \frac{1}{p}\sum_{i=1}^{p} y_{i...}^2 - \frac{y_{....}^2}{N}$	$p - 1$
Columns	$SS_{\text{Columns}} = \frac{1}{p}\sum_{l=1}^{p} y_{...l}^2 - \frac{y_{....}^2}{N}$	$p - 1$
Error	SS_E (by subtraction)	$(p - 3)(p - 1)$
Total	$SS_T = \sum_i \sum_j \sum_k \sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	$p^2 - 1$