10.3 THE RANDOMIZED BLOCK DESIGN

In section 10.1, the one-way ANOVA F test to evaluate differences in the means of c groups was developed. The one-way ANOVA F test is used in situations in which n homogeneous items or individuals (i.e., *experimental units*) are randomly assigned to the c levels of a factor of interest (i.e., the *treatment groups*). Such designed one-factor experiments are referred to as *one-way* or *completely randomized* designs.

Alternatively, in section 9.2 the t test for the mean difference was used in situations involving repeated measurements or matched samples in order to evaluate differences between two treatment conditions. Suppose, that there are more than two treatment groups or levels of a factor of interest. In such cases, the heterogeneous sets of items or individuals that have been matched (or on whom repeated measurements have been taken) are called **blocks**. The numerical response or outcome of each treatment group and block combination is then obtained.

Experimental situations where blocks are used are referred to as **randomized block designs**. Although treatments and blocks are both used in a randomized block design, the focus of the analysis is on the differences among the c different treatment groups. The purpose of blocking is to remove as much variability as possible from the experimental error so that the differences among the c treatment groups are more evident. A randomized block design is often more efficient statistically than a completely randomized design and therefore obtains more precise results (see references 1, 4, 7, and 8).

To compare a completely randomized design with a randomized block design, return to the Using Statistics scenario concerning the Perfect Parachute Company. Suppose that a completely randomized design is used with 12 observations, one parachute woven during each of 12 different shifts. Any variability among the shifts becomes part of the experimental error, and therefore differences among the four suppliers might be hard to detect. To reduce the experimental error, a randomized block experiment is designed where three shifts are used and four parachutes are woven during each shift (one parachute using fibers from Supplier 1, one parachute using fibers from Supplier 2, and so forth). The three shifts are considered the blocks while the treatment factor is still the four suppliers. The advantage of the randomized block design is that the variability among the three shifts is removed from the experimental error. Therefore, this design often provides more precise results concerning differences among the four suppliers.

Tests for the Treatment and Block Effects

Recall from Figure 10.1 that, in the completely randomized design, the total variation (SST) is subdivided into variation attributable to differences *among* the c groups (SSA) and variation due to chance or attributable to inherent variation *within* the c groups (SSW). Within-group variation is considered experimental error, and among-group variation is attributable to treatment effects.

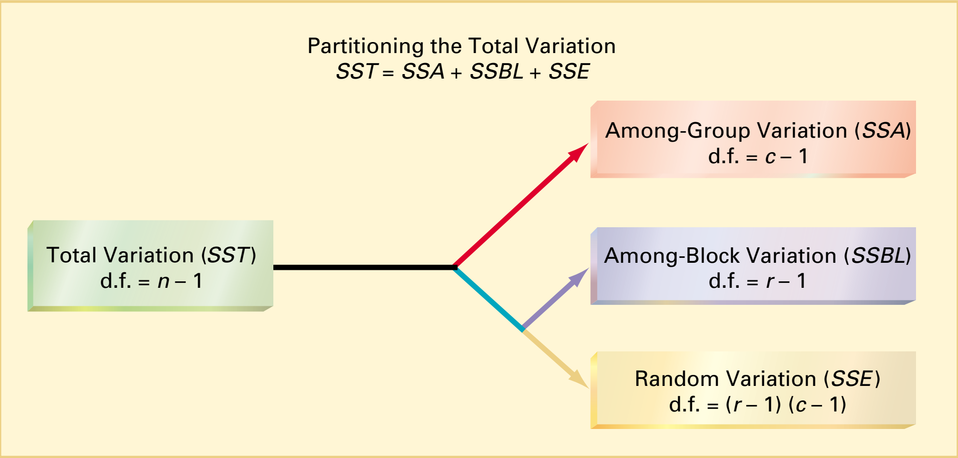
To filter out the effects of the blocking from the experimental error in the randomized block design, the within-group variation (SSW) needs to be further subdivided into variation attributable to differences among the blocks ($SSBL$) and variation attributable to inherent random error (SSE). Therefore, as presented in Figure 10.18, in a randomized block design the total variation in the outcome measurements is the sum of among-group variation (SSA), among-block variation ($SSBL$), and inherent random error (SSE).

To develop the ANOVA procedure for the randomized block design, the following terms need to be defined:

$$\begin{aligned}
r &= \text{the number of blocks} \\
c &= \text{the number of groups or treatment levels} \\
n &= \text{the total number of observations (where } n = rc) \\
X_{ij} &= \text{the value in the } i\text{th block for the } j\text{th group} \\
\bar{X}_{i.} &= \text{the mean of all the values in block } i \\
\bar{X}_{.j} &= \text{the mean of all the values for group } j \\
\sum_{j=1}^{c}\sum_{i=1}^{r} X_{ij} &= \text{the grand total}
\end{aligned}$$

FIGURE 10.18

Partitioning the total variation in a randomized block design model



The total variation, also called sum of squares total (SST), is a measure of the variation among all the observations. SST is obtained by summing the squared differences between each individual observation and the overall or grand mean $\bar{\bar{X}}$ that is based on all n observations. SST is computed using Equation (10.19).

TOTAL VARIATION

$$SST = \sum_{j=1}^{c}\sum_{i=1}^{r}(X_{ij} - \bar{\bar{X}})^2 \qquad \textbf{(10.19)}$$

where

$$\bar{\bar{X}} = \frac{\sum_{j=1}^{c}\sum_{i=1}^{r} X_{ij}}{rc} \text{ (i.e., the overall or grand mean)}$$

The among-group variation, also called the sum of squares among groups (SSA), is measured by the sum of the squared differences between the sample mean of each group $\bar{X}_{.j}$ and the overall or grand mean $\bar{\bar{X}}$, weighted by the number of blocks r. The among-group variation is computed using Equation (10.20).

AMONG-GROUP VARIATION

$$SSA = r\sum_{j=1}^{c}(\bar{X}_{.j} - \bar{\bar{X}})^2 \qquad \textbf{(10.20)}$$

where

$$\bar{X}_{.j} = \frac{\sum_{i=1}^{r} X_{ij}}{r} \text{ (i.e., the group means)}$$

The **among-block variation**, also called the **sum of squares among blocks** (***SSBL***), is measured by the sum of the squared differences between the mean of each block $\bar{X}_{i.}$ and the overall or grand mean $\bar{\bar{X}}$, weighted by the number of groups c. The among-block variation is computed using Equation (10.21).

AMONG-BLOCK VARIATION

$$SSBL = c\sum_{i=1}^{r}(\bar{X}_{i.} - \bar{\bar{X}})^2 \qquad \textbf{(10.21)}$$

where

$$\bar{X}_{i.} = \frac{\sum_{j=1}^{c} X_{ij}}{c} \text{ (i.e., the block means)}$$

The inherent random variation or error, also called the **sum of squares error** (***SSE***), is measured by the sum of the squared differences among all the observations after the effect of the particular treatments and blocks have been accounted for. *SSE* is computed using Equation (10.22).

RANDOM ERROR

$$SSE = \sum_{j=1}^{c}\sum_{i=1}^{r}(X_{ij} - \bar{X}_{.j} - \bar{X}_{i.} + \bar{\bar{X}})^2 \qquad \textbf{(10.22)}$$

Since there are c treatment levels of the factor being compared, there are $c - 1$ degrees of freedom associated with the sum of squares among groups (*SSA*). Similarly, since there are r blocks, there are $r - 1$ degrees of freedom associated with the sum of squares among blocks (*SSBL*). Moreover, there are $n - 1$ degrees of freedom associated with the sum of squares total (*SST*) because each observation X_{ij} is being compared to the overall or grand mean $\bar{\bar{X}}$ based on all n observations. Therefore, since the degrees of freedom for each of the sources of variation must add to the degrees of freedom for the total variation, the degrees of freedom for the sum of squares error (*SSE*) component is obtained by subtraction and algebraic manipulation. The degrees of freedom is given by $(r - 1)(c - 1)$.

If each of the component sums of squares is divided by its associated degrees of freedom, you obtain the three *variances* or mean square terms (***MSA***, ***MSBL***, and ***MSE***) needed for ANOVA depicted in Equation (10.23a–c).

OBTAINING THE MEAN SQUARES

$$MSA = \frac{SSA}{c - 1} \qquad \textbf{(10.23a)}$$

$$MSBL = \frac{SSBL}{r - 1} \qquad \textbf{(10.23b)}$$

$$MSE = \frac{SSE}{(r - 1)(c - 1)} \qquad \textbf{(10.23c)}$$

If the assumptions pertaining to the analysis of variance are met, the null hypothesis of no differences in the c population means (i.e., no treatment effects):

$$H_0: \mu_{.1} = \mu_{.2} = \cdots = \mu_{.c}$$

is tested against the alternative that not all the c population means are equal:

$$H_1: \text{Not all } \mu_{.j} \text{ are equal (where } j = 1, 2, \ldots, c)$$

by computing the test statistic F given in Equation (10.24).

RANDOMIZED BLOCK *F*-TEST STATISTIC FOR DIFFERENCES IN *c* MEANS

$$F = \frac{MSA}{MSE} \tag{10.24}$$

The F-test statistic follows an F distribution with $c - 1$ degrees of freedom for the MSA term and $(r - 1)(c - 1)$ degrees of freedom for the MSE term. For a given level of significance α, the null hypothesis is rejected if the computed F-test statistic is greater than the upper-tail critical value F_U from the F distribution with $c - 1$ and $(r - 1)(c - 1)$ degrees of freedom, respectively, in the numerator and denominator (see Table E.5). That is, the decision rule is:

$$\text{Reject } H_0 \text{ if } F > F_U;$$
$$\text{otherwise do not reject } H_0.$$

To examine whether it is advantageous to use a randomized block design, some statisticians suggest that the ***F* test for blocking effects** be performed. The null hypothesis of no block effects:

$$H_0: \mu_{1.} = \mu_{2.} = \cdots = \mu_{r.}$$

is tested against the alternative:

$$H_1: \text{Not all } \mu_{i} \text{ are equal (where } i = 1, 2, \ldots, r)$$

The F statistic is given in Equation (10.25).

F-TEST STATISTIC FOR BLOCK EFFECTS

$$F = \frac{MSBL}{MSE} \tag{10.25}$$

The null hypothesis is rejected at the α level of significance if the F-test statistic is greater than the upper-tail critical value F_U from the F distribution with $r - 1$ and $(r - 1)(c - 1)$ degrees of freedom, respectively, in the numerator and denominator (see Table E.5). That is, the decision rule is:

$$\text{Reject } H_0 \text{ if } F > F_U;$$
$$\text{otherwise do not reject } H_0.$$

Some statisticians believe that this test is unnecessary since the sole purpose of establishing the blocks is to provide a more efficient method of testing for treatment effects by reducing the experimental error.

As in section 10.1, the results of an analysis-of-variance procedure are usually displayed in an ANOVA summary table, the format for which is presented in Table 10.8.

TABLE 10.8

Analysis-of-variance table for the randomized block design

Source	Degrees of Freedom	Sum of Squares	Mean Square (Variance)	F
Among treatments	$c - 1$	SSA	$MSA = \dfrac{SSA}{c-1}$	$F = \dfrac{MSA}{MSE}$
Among blocks	$r - 1$	$SSBL$	$MSBL = \dfrac{SSBL}{r-1}$	$F = \dfrac{MSBL}{MSE}$
Error	$(r-1)(c-1)$	SSE	$MSE = \dfrac{SSE}{(r-1)(c-1)}$	
Total	$rc - 1$	SST		

To illustrate the randomized block design, suppose that a fast-food chain having four branches in a certain geographical area wants to evaluate the service at these restaurants. The customer services director for the chain hires six investigators with varied experiences in food-service evaluations to act as raters. To reduce the effect of the variability from rater-to-rater, a randomized block design is used with raters serving as the blocks. The four restaurants are the treatment groups of interest.

The six raters are randomly assigned an order to evaluate the service at each of the four restaurants. A rating scale from 0 (low) to 100 (high) is used. The results are summarized in Table 10.9, along with the group totals, group means, block totals, block means, grand total, and grand mean.

TABLE 10.9

Restaurant ratings for four branches of a fast-food chain

FFCHAIN.xls

	Restaurants					
Raters	**A**	**B**	**C**	**D**	**Totals**	**Means**
1	70	61	82	74	287	71.75
2	77	75	88	76	316	79.00
3	76	67	90	80	313	78.25
4	80	63	96	76	315	78.75
5	84	66	92	84	326	81.50
6	78	68	98	86	330	82.50
Totals	465	400	546	476	1,887	
Means	77.50	66.67	91.00	79.33	78.625	

In addition, from Table 10.9,

$$r = 6 \qquad c = 4 \qquad n = rc = 24$$

and,

$$\bar{\bar{X}} = \frac{\sum_{j=1}^{c}\sum_{i=1}^{r} X_{ij}}{rc} = \frac{1{,}887}{24} = 78.625$$

To analyze the results of this randomized block design, Figure 10.19 illustrates output from Microsoft Excel.

FIGURE 10.19

Microsoft Excel output for the fast-food-chain study

	A	B	C	D	E	F	G
1	Anova: Two-Factor Without Replication						
2							
3	*SUMMARY*	*Count*	*Sum*	*Average*	*Variance*		
4	Rater 1	4	287	71.75	76.25		
5	Rater 2	4	316	79	36.66667		
6	Rater 3	4	313	78.25	90.91667		
7	Rater 4	4	315	78.75	184.9167		
8	Rater 5	4	326	81.5	121		
9	Rater 6	4	330	82.5	161		
10							
11	Restaurant A	6	465	77.5	21.5		
12	Restaurant B	6	400	66.66667	23.46667		
13	Restaurant C	6	546	91	33.2		
14	Restaurant D	6	476	79.33333	23.46667		
15							
16							
17	ANOVA						
18	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
19	Rows Raters	283.375	5	56.675	3.781835	0.020456	2.901295
20	Columns	1787.458	3	595.8194	39.75811	2.23E-07	3.287383
21	Error	224.7917	15	14.98611			
22	Restaurants						
23	Total	2295.625	23				

When using the 0.05 level of significance to test for differences among the restaurant branches, the decision rule is to reject the null hypothesis (H_0: $\mu_{.1} = \mu_{.2} = \mu_{.3} = \mu_{.4}$) if the calculated F value is greater than 3.29, the upper-tail critical value from the F distribution with 3 and 15 degrees of freedom in the numerator and denominator, respectively (see Figure 10.20). Since $F = 39.758 > F_U = 3.29$, or since the p-value $= 0.000 < 0.05$, you can reject H_0 and conclude that there is evidence of a difference in the average rating among the different restaurants. The extremely small p-value indicates that if the means from the four restaurant branches are equal, there is virtually no chance to observe differences this large or larger among the sample means as that observed in this study. Thus, there is little degree of belief in the null hypothesis. You may therefore conclude that the alternative hypothesis is correct.

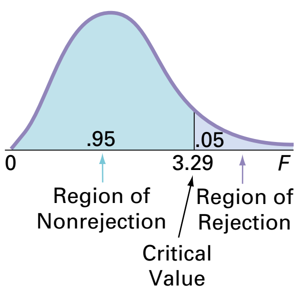


FIGURE 10.20

Regions of rejection and nonrejection for the fast-food-chain study at the 0.05 level of significance with 3 and 15 degrees of freedom

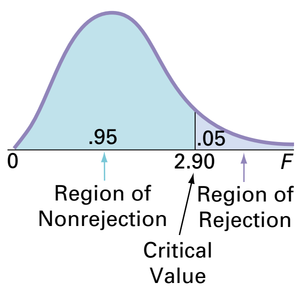


FIGURE 10.21

Regions of rejection and nonrejection for the fast-food-chain study at the 0.05 level of significance with 5 and 15 degrees of freedom

As a check on the effectiveness of blocking, you can test for a difference among the raters. The decision rule, using the 0.05 level of significance, is to reject the null hypothesis (H_0: $\mu_{1.} = \mu_{2.} = \cdots = \mu_{6.}$) if the calculated F value is greater than 2.90, the upper-tail critical value from the F distribution with 5 and 15 degrees of freedom in the numerator and denominator, respectively (see Figure 10.21). Since $F = 3.782 > F_U = 2.90$, or the p-value $= 0.02 < 0.05$, you reject H_0 and conclude that there is evidence of a difference among the raters. Thus, you conclude that the blocking has been advantageous in reducing the experimental error.

Excel Strategy Performing Analysis of Variance Using Randomized Block Designs

You perform the analysis of variance for a randomized block design by using the Data Analysis Anova: Two Factor Without Replication procedure. PHStat2 does not perform this test and because the calculations to perform this test are complex, you cannot easily manually implement a worksheet design to perform the test.

For example, to perform analysis of variance for the randomized block design of the restaurant rating data of Table 10.9, open to the **Ratings Data** worksheet of the **Chapter 10.xls** workbook and:

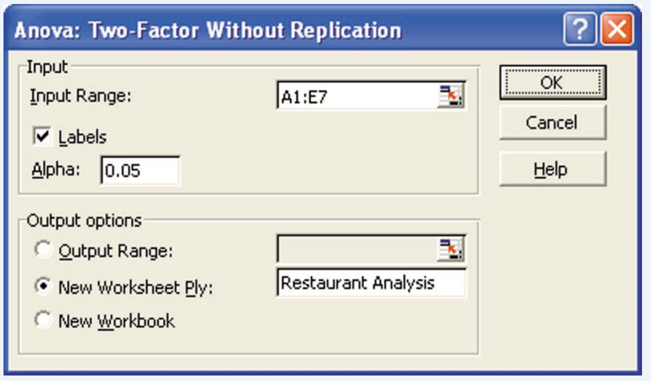
CHAPTER 10.xls

*Retrieve the Figure 10.19 worksheet of the **Chapter 10.xls** workbook to view the analysis of variance for the randomized block design.*

Select **Tools | Data Analysis**.

Select **Anova: Two Factor Without Replication** from the list box in the Data Analysis dialog box and click the **OK** button.

In the Anova: Two Factor Without Replication dialog box (shown at right):



Enter **A1:E7** in the **Input Range** edit box.

Select the **Labels** check box.

Enter **0.05** in the **Alpha** edit box.

Select the **New Worksheet Ply** option button and enter a name for the new sheet.

Click the **OK** button.

This procedure generates a worksheet that does **not** dynamically change. Therefore, you will need to rerun the procedure if you make any changes to the underlying data in order to update the results of the test.

This procedure uses unstacked *data in which the data values for each group are found in separate columns. If your data are stacked, appearing in a single column adjoining a column that identifies the group to which individual data values belong, rearrange your data into separate columns using one of the techniques discussed in section EH9.2.*

In addition to the assumptions of the one-way analysis of variance previously mentioned in section 10.1, you need to assume that there is no *interacting effect* between the treatments and the blocks. In other words, you need to assume that any differences between the treatments (the restaurants) are consistent across the entire set of blocks (the raters). The concept of *interaction* is discussed in section 10.2.

Now that the randomized block model has been developed and used in the fast-food-chain study, the question arises as to what effect the blocking had on the analysis. That is, did the blocking result in an increase in precision in comparing the different treatment groups? To answer this question, the **estimated relative efficiency** (*RE*) of the randomized block design as compared with the completely randomized design can be calculated using Equation (10.26).

ESTIMATED RELATIVE EFFICIENCY

$$RE = \frac{(r-1)MSBL + r(c-1)MSE}{(rc-1)MSE} \tag{10.26}$$

Using Figure 10.19,

$$RE = \frac{(5)(56.675) + (6)(3)(14.986)}{(23)(14.986)} = 1.60$$

This value for relative efficiency means that 1.6 times as many observations in each treatment group would be needed in a one-way ANOVA design as compared to the randomized block design in order to obtain the same precision for comparison of treatment group means.

Multiple Comparisons: The Tukey Procedure

As in the case of the completely randomized design, once the null hypothesis of no differences between the treatment groups has been rejected, you can determine which of the treatment groups are significantly different from the others. For the randomized block design, a procedure developed by John Tukey (see references 7, 8, and 9) is used. The critical range for the **Tukey procedure** is given in Equation (10.27).

OBTAINING THE CRITICAL RANGE

$$\text{critical range} = Q_U\sqrt{\frac{MSE}{r}} \tag{10.27}$$

where the statistic Q_U = the upper-tail critical value from a Studentized range distribution having c degrees of freedom in the numerator and $(r-1)(c-1)$ degrees of freedom in the denominator. Values for the Studentized range distribution are found in Table E.9.

Each of the $c(c-1)/2$ pairs of means is compared against the one critical range. A specific pair, say group j versus group j' is declared significantly different if the absolute difference in the sample means $|\bar{X}_{.j} - \bar{X}_{.j'}|$ is greater than this critical range.

To apply the Tukey procedure, return to the fast-food-chain study. Since there are four restaurants, there are $4(4-1)/2 = 6$ possible pairwise comparisons to be made. From Figure 10.19, the absolute mean differences are

1. $|\bar{X}_{.1} - \bar{X}_{.2}| = |77.50 - 66.67| = 10.83$
2. $|\bar{X}_{.1} - \bar{X}_{.3}| = |77.50 - 91.00| = 13.50$
3. $|\bar{X}_{.1} - \bar{X}_{.4}| = |77.50 - 79.33| = 1.83$
4. $|\bar{X}_{.2} - \bar{X}_{.3}| = |66.67 - 91.00| = 24.33$
5. $|\bar{X}_{.2} - \bar{X}_{.4}| = |66.67 - 79.33| = 12.66$
6. $|\bar{X}_{.3} - \bar{X}_{.4}| = |91.00 - 79.33| = 11.67$

To determine the critical range, use Figure 10.19 to obtain $MSE = 14.986$ and $r = 6$. From Table E.9 [for $\alpha = .05$, $c = 4$, and $(r-1)(c-1) = 15$], Q_U, the upper-tail critical value of the test statistic with

4 degrees of freedom in the numerator and 15 degrees of freedom in the denominator is 4.08. Using Equation (10.27),

$$\text{Critical range} = 4.08\sqrt{\frac{14.986}{6}} = 6.448$$

Note that all pairwise comparisons except $|\bar{X}_{.1} - \bar{X}_{.4}|$ are greater than the critical range. Therefore, you conclude that there is evidence of a significant difference in the average rating between all pairs of restaurant branches except for branches A and D. In addition, branch C has the highest ratings (i.e., is most preferred) and branch B has the lowest (i.e., is least preferred).

PROBLEMS FOR SECTION 10.3

Learning the Basics

10.42 Given a randomized block experiment having a single factor of interest with five treatment levels and seven blocks, answer the following.

a. How many degrees of freedom are there in determining the among-group variation?
b. How many degrees of freedom are there in determining the among-block variation?
c. How many degrees of freedom are there in determining the inherent random variation or error?
d. How many degrees of freedom are there in determining the total variation?

10.43 From problem 10.42:

a. If $SSA = 60$, $SSBL = 75$, and $SST = 210$, what is SSE?
b. What is MSA?
c. What is $MSBL$?
d. What is MSE?
e. What is the value of the test statistic F for the difference in the five means?
f. What is the value of the test statistic F for the block effects?

10.44 From problems 10.42 and 10.43:

a. Form the ANOVA summary table and fill in all values in the body of the table.
b. At the 0.05 level of significance, when testing for differences in the five means, what is the upper-tail critical value from the F distribution?
c. State the decision rule for testing the null hypothesis that all five groups have equal population means.
d. What is your statistical decision?
e. At the 0.05 level of significance, when testing for block effects, what is the upper-tail critical value from the F distribution?
f. State the decision rule for testing the null hypothesis that there are no block effects.
g. What is your statistical decision?

10.45 From problems 10.42, 10.43, and 10.44:

a. To perform the Tukey procedure, how many degrees of freedom are there in the numerator and how many degrees of freedom are there in the denominator of the Studentized range distribution?
b. At the 0.05 level of significance, what is the upper-tail critical value from the Studentized range distribution?
c. To perform the Tukey procedure, what is the critical range?

10.46 Given a randomized block experiment having one factor containing three treatment levels and seven blocks:

a. How many degrees of freedom are there in determining the among-group variation?
b. How many degrees of freedom are there in determining the among-block variation?
c. How many degrees of freedom are there in determining the inherent random variation or error?
d. How many degrees of freedom are there in determining the total variation?

10.47 From problem 10.46, if $SSA = 36$ and the randomized block F-test statistic is 6.0:

a. What is MSE?
b. What is SSE?
c. What is $SSBL$ if the F-test statistic for block effects is 4.0?
d. What is SST?
e. At the 0.01 level of significance, is there evidence of a treatment effect?
f. At the 0.01 level of significance, is there evidence of a block effect?

10.48 Given a randomized block experiment having one factor containing four treatment levels and eight blocks, from the ANOVA summary table below, fill in all the missing results.

Source	Degrees of Freedom	Sum of Squares	Mean Square (Variance)	F
Among treatments	$c - 1 = ?$	$SSA = ?$	$MSA = 80$	$F = ?$
Among blocks	$r - 1 = ?$	$SSBL = 540$	$MSBL = ?$	$F = 5.0$
Error	$(r - 1)(c - 1) = ?$	$SSE = ?$	$MSE = ?$	
Total	$rc - 1 = ?$	$SST = ?$		

10.49 From problem 10.46:

a. At the 0.05 level of significance, what is the upper-tail critical value from the F distribution when testing for differences among the four treatment level means?

b. State the decision rule for testing the null hypothesis that all four groups have equal population means.

c. What is your statistical decision?

d. At the 0.05 level of significance, what is the upper-tail critical value from the F distribution when testing for block effects?

e. State the decision rule for testing the null hypothesis of no block effects.

f. What is your statistical decision?

Applying the Concepts

Note: We recommend that Microsoft Excel be used to solve Problems 10.50–10.54.

10.50 A taste-testing experiment has been designed so that four brands of Colombian coffee are to be rated by nine experts. To avoid any carryover effects, the tasting sequence for the four brews is randomly determined for each of the nine expert tasters until a rating on a 7-point scale (1 = extremely unpleasing, 7 = extremely pleasing) is given for each of four characteristics: taste, aroma, richness, and acidity. The following table displays the summated ratings—accumulated over all four characteristics. **COFFEE.xls**

	Brand			
Expert	***A***	***B***	***C***	***D***
C.C.	24	26	25	22
S.E.	27	27	26	24
E.G.	19	22	20	16
B.L.	24	27	25	23
C.M.	22	25	22	21
C.N.	26	27	24	24
G.N.	27	26	22	23
R.M.	25	27	24	21
P.V.	22	23	20	19

At the 0.05 level of significance, completely analyze the data to determine whether there is evidence of a difference in the summated ratings of the four brands of Colombian coffee and, if so, which of the brands are rated highest (i.e., best). What can you conclude?

10.51 A student team in a business statistics course designed an experiment to investigate whether the brand of bubblegum used affected the size of bubbles they could blow. The students believed that Kyle was an expert at blowing bubbles, and his expertise might negatively affect the results of a completely randomized design. Thus, to reduce the person-to-person variability, they decided to use a randomized block design using themselves as blocks. Four brands of bubblegum were tested. A student chewed two pieces of a brand of gum and then blew two bubbles, attempting to make them as big as possible. Another student measured the diameters of the bubbles at their biggest point. The following table gives the combined diameters of the bubbles (in inches) for the 16 observations. **BUBBLEGUM.xls**

	Brand of Bubblegum			
Student	**Bazooka**	**Bubbletape**	**Bubbleyum**	**Bubblicious**
Kyle	8.75	9.50	8.50	11.50
Sarah	9.50	4.00	8.50	11.00
Leigh	9.25	5.50	7.50	7.50
Isaac	9.50	8.50	7.50	7.50

a. Using a 0.05 level of significance, is there evidence of a difference in the mean diameter of the bubbles produced by the different brands?

b. If appropriate, use the Tukey procedure to determine which brands of bubblegum differ. (Use a = 0.05.)

c. Do you think that there was a significant block effect in this experiment? Explain.

d. Do you think Kyle is the best at blowing big bubbles?

10.52 The manager of a nationally known real estate agency has just completed a training session on appraisals for three newly hired agents. To evaluate the effectiveness of his training, the manager wishes to determine whether there is any difference in the appraised values placed on houses by these three different individuals. A sample of 12 houses is selected by the manager and each agent is assigned the task of placing an appraised value (in thousands of dollars) on the 12 houses. The results are summarized in the data file: **REAPPR3.xls**

a. At the 0.05 level of significance, use the randomized block design to determine if there is evidence of a difference in the mean appraised values given by the three agents.

b. What assumptions are necessary to perform this test?

c. What conclusions about the effectiveness of the training in terms of the ability of different agents to rate the properties in a similar way can the manager reach? Explain.

10.53 Philips Semiconductors is a leading European manufacturer of integrated circuits. Integrated circuits are produced on silicon wafers, which are ground to target thickness early in the production process. The wafers are positioned in different locations on a grinder and kept in place through vacuum decompression. One of the goals of process improvement is to reduce the variability in the thickness of the wafers in different positions and in different batches. Data were collected from a sample of 30 batches. In each batch the thickness of the wafers on positions 1 and 2 (outer circle), 18 and 19 (middle circle), and 28 (inner circle) was measured. The results are given in the **CIRCUITS.xls** file. At the 0.01 level of significance, completely analyze the data to determine whether there is evidence of a difference in the mean thickness of the wafers for the five positions and, if so, which of the positions are different. What can you conclude?

Source: K. C. B. Roes and R. J. M. M. Does, "Shewhart-type Charts in Nonstandard Situations," Technometrics, *37, 1995, 15–24.*

10.54 The data in the **CONCRETE2.xls** file represent the compressive strength in thousands of pounds per square inch (psi) of 40 samples of concrete taken 2, 7, and 28 days after pouring.

Source: O. Carrillo-Gamboa and R. F. Gunst, "Measurement-Error-Model Collinearities," Technometrics, *34, 1992, 454–464.*

a. At the 0.05 level of significance, is there evidence of a difference in the average compressive strength after 2, 7, and 28 days?
b. If appropriate, use the Tukey procedure to determine the days that differ in average compressive strength. (Use $\alpha = 0.05$.)
c. Determine the relative efficiency of the randomized block design as compared with the one-way ANOVA (completely randomized) design.
d. Obtain box-and-whisker plots of the compressive strength for the different time periods.
e. Based on the results of (a), (b), and (d), is there a pattern in the compressive strength over the three time periods?