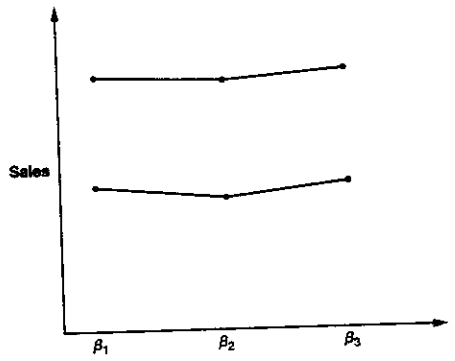

FIGURE 9.24 Example of Treatment Mean Plot Showing No Interaction Effects and No Factor *B* Main Effects.

EXERCISES

3. **Satisfaction.** A company designs applications software for a large number of firms. A study is conducted to assess the level of satisfaction of these firms with the software supplied. The effects of two factors on the level of satisfaction are to be investigated:

   1. the industry of the firm that receives the software (INDUSTRY)
   2. the contact person from whom the software was purchased (CONTACT)

   Four industries are served and there are three contact people. For each combination of industries and contacts, two firms are randomly selected and surveyed. A satisfaction score (SATIS) is obtained from the questionnaire administered. The data are available in a file named SATIS9 on the CD. The ANOVA for this two-factor study is shown in Figure 9.25. Use the results to determine whether industry or contact person affects satisfaction. Use a 5% level of significance. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

9.4 ANALYSIS OF COVARIANCE

Analysis of covariance (ANCOVA) is a procedure sometimes used with models containing some quantitative and some qualitative independent variables. In ANCOVA, however, the main interest is in the qualitative variables. In this respect, ANCOVA might be viewed as a modification of ANOVA procedures rather than as a special case of regression analysis (it is, in fact, both). The term ANCOVA is used when the quantitative independent variables are added to the ANOVA model to reduce the variance of the error terms and thus provide more precise measurement of the treatment effects. These quantitative variables should be constructed in such a way that they are not influenced by the treatments.

Although some statistical programs do provide specific routines for performing ANCOVA, these are not discussed in this text. ANCOVA can be performed using regression routines with appropriately constructed indicator variables.

FIGURE 9.25 ANOVA Output for Satisfaction Exercise.

Source	DF	Sum of Squares	Mean Square	F Stat	P Value
INDUSTRY	3	716.125	238.708	49.82	0.000
CONTACT	2	0.250	0.125	0.03	0.974
Interaction	6	5.750	0.958	0.20	0.970
Error	12	57.500	4.792		
Total	23	779.625			

ADDITIONAL EXERCISES

4. **ANOVA Tables**

   a. Assume that one-way analysis of variance is to be performed. Complete the ANOVA table for this analysis:

Source	DF	Sum of Squares	Mean Square	F Stat
Treatment	3			
Error		20		
Total	23	170		

   b. Assume that a randomized block experiment is to be performed. Complete the ANOVA table for this analysis:

Source	DF	Sum of Squares	Mean Square	F Stat
Blocks	4			
Treatment		80	40	
Error	39		2	
Total	45	198		

5. **Assembly Line.** Three assembly lines are used to produce a certain component for a computer. To examine the production rate of the assembly lines, a random sample of six hourly periods is chosen for each assembly line, and the number of components produced during these periods for each line is recorded. These data are available in a file named ASSEMBL9 on the CD. In the MINITAB file, the number of components produced will be in one column. A second column will indicate the production line (coded as 1, 2, or 3). In the Excel spreadsheet, the production numbers will be in a separate column for each production line.

   Is there a difference in the average production rates for the three assembly lines? Use a 5% level of significance in answering this question. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

6. **Salaries.** A large state university is interested in comparing salaries of its graduates (BA or BS) in the following areas: business, education, engineering, and liberal arts. Five graduates in each major are randomly selected and their starting salaries recorded. These data are available in a file named SALMAJ9 on the CD. In the MINITAB file, the variable SALARY will be one column. The variable MAJOR is a second column coded as 1 = business, 2 = education, 3 = engineering, and 4 = liberal arts. The salary data are in four separate columns in the Excel spreadsheet.

   a. The university wants to know if there is a difference in the population average salaries for the four majors. Use a 5% level of significance in making the decision. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

   b. Find a 95% interval estimate for the difference between the business and engineering mean salaries.

   c. Use the Tukey or Bonferroni method to compare all possible pairs of means with a family-wise confidence level of 90%. Use the familywise comparisons to determine if there is a significant difference in the population mean salaries for

      (1) business and education majors
      (2) business and engineering majors
      (3) education and liberal arts majors

      If so, which major has a higher mean salary in each comparison?

7. **Test Scores.** A large financial planning firm wants to compare the results of three training programs

for its staff. Twenty-seven employees are selected for the study. These employees are grouped into three blocks of nine for the comparison. The blocks are based on number of years since college graduation, with block 1 being the most recent graduates and block 3 the most distant graduates. After the training programs are completed, the employees are tested and the test scores recorded. The data are in a file named TESTSCR9 on the CD. In the MINITAB file there are three columns: SCORE contains the test scores, PROGRAM coded as 1, 2, or 3, and BLOCK coded as 1, 2, or 3. The Excel spreadsheet is set up as it should be for an Excel randomized block analysis. See Using the Computer for more information.

a. Determine whether there is a difference in average test scores because of the training programs. Use a 5% level of significance. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

b. Construct a 95% confidence interval estimate of the difference between the means for program 1 and program 2.

8. **Employee Productivity.** A study of employee productivity is to be conducted with employees who enter data at computers. The amount of data entered is the dependent variable. Two factors that may influence the dependent variable are examined. One factor is the type of keyboard used (three types are available in the company). The second factor is the time of day (morning or afternoon). Four employees are randomly assigned to each type of keyboard. Two employees' production levels are recorded for a period of 1 hour in the morning. The other two are recorded in the afternoon. The production level is the number of forms completely entered by each employee. These data are available in a file named PRODRAT9 on the CD. In the MINITAB file, there are three columns: NUMBER is the number of forms processed by each employee, KEYBOARD is the keyboard type (coded as 1, 2, or 3), and TIME is coded as 1 for morning and 2 for afternoon. In the Excel spreadsheet, the data will be set up for a two-way analysis of variance. See Using the Computer for more information.

Do the two factors, KEYBOARD and TIME, appear to influence production rate (NUMBER)? Use a 5% level of significance for any hypothesis tests used. State any hypotheses to be tested, the decision rule, the test statistic, and your decision. How would you describe the influence you observed in words?

9. **Bill's Sales.** Bill's is a popular restaurant/bar in southwest Fort Worth, Texas. Daily sales data for Bill's for the period from October 14 through December 8 are available in a file named BILLS9 on the CD. In the MINITAB file, SALES will be in a single column. The column labeled DAY represents the day of the week, coded as 1 = Monday, 2 = Tuesday, 3 = Wednesday, 4 = Thursday, 5 = Friday, 6 = Saturday, and 7 = Sunday. In the Excel spreadsheet sales for each day will be in a separate column. Determine whether there is a difference in Bill's average sales on different days of the week. Use a 5% level of significance for any hypothesis tests used. State any hypotheses to be tested, the decision rule, the test statistic, and your decision. If there is a difference, on which day (or days) do sales appear to be highest? What does this suggest to the owner of Bill's about staffing?

10. **Executives' Salaries.** The file EXECSAL9 on the CD contains the company name, name of the CEO, the 2002 salary, and the 2002 salary plus bonus for samples of executives from 11 different industries. The data are from the *Wall Street Journal* 2002 CEO Compensation Survey (Monday April 14, 2003). The industries are coded as follows:

1 = Basic Materials
2 = Energy
3 = Industrial
4 = Cyclical
5 = Noncyclical
6 = Technology
7 = Financial
8 = Utilities
9 = Health Care
10 = Telecommunications

SALARY and SALBONUS will each be in a single column. The column labeled INDUSTRY will contain the industry classification, coded as indicated.

a. Determine whether there is a difference in average salaries for CEOs in the different industry classifications. Use a 5% level of significance for any hypothesis tests used. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

b. Determine whether there is a difference in the combined average salary and bonus for CEOs in the different industry classifications. Use a 5% level of significance for any hypothesis tests used. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

11. **Comparing Suppliers.** A company purchases diodes from four different suppliers. The company's engineers would like to determine if the average lifetimes of the diodes are essentially the same. They randomly select three diodes from each supplier's shipment and test the lifetimes. The resulting data are available in a file named SUPPLY9 on the CD. In the MINITAB file, the variable LIFETIME will be one column. The variable SUPPLIER is a second column coded as 1, 2, 3, or 4 to indicate the supplier. The lifetime data are in four separate columns in the Excel spreadsheet.

a. The engineers want to know if there is a difference in the population average lifetimes of the diodes for the four suppliers. Use a 5% level of significance in making the decision. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

b. Find a 95% interval estimate for the difference between the mean lifetimes for suppliers 1 and 2.

c. Use the Tukey or Bonferroni method to compare all possible pairs of means with a family-wise confidence level of 90%.

12. **Comparing Inspectors.** A manufacturer employs five persons who visually inspect circuit boards for flaws in the printed circuitry. A circuit board that is rejected at visual inspection but does not have the flaw claimed by the inspector is referred to as a "false reject." Since false rejects add to manufacturing costs, the manufacturer wants to determine whether the false reject averages are the same for the five visual inspectors. To do so, the boards rejected by each inspector are checked for false rejects over six 1-week periods and the numbers of false rejects (FALSE) recorded. The resulting data are available in a file named INSPECT9 on the CD. In the MINITAB file, the variable FALSE will be one column. The variable INSPECTOR is a second column coded as 1, 2, 3, 4, or 5 to indicate the inspector. The data are in five separate columns in the Excel spreadsheet. The manufacturer wants to know if there is a difference in the population average number of false rejects for the five inspectors. Use a 5% level of significance in making the decision. State any hypotheses to be tested, the decision rule, the test statistic, and your decision.

(*Source*: This example is from *Engineer Statistics: The Industrial Experience*, pages 476–477.[2])

USING THE COMPUTER

The Using the Computer section in each chapter describes how to perform the computer analyses in the chapter using Excel, MINITAB, and SAS. For further detail on Excel, MINITAB, and SAS, see Appendix C.

EXCEL

One-Way Analysis of Variance

Figure 9.26 shows the data arrangement for a one-way ANOVA in Excel. The data for each factor level must be in a separate column. In the example shown, there are nine factor levels, so there are nine columns of data. The columns do not have to be of equal length. For a one-way ANOVA, choose Anova: Single Factor from the Data

[2]See References for complete publication information.