chap 6 Module A. mol —> utf-8

are substantial differences in costs between the three designs. The company decides to test the three designs by mailing 5000 samples of each to potential customers in four different regions of the country. Since there are known regional differences in the customer base, regions are considered as blocks. The number of responses to each mailing is as follows.

	Region			
Design	NE	NW	SE	SW
1	250	350	219	375
2	400	525	390	580
3	275	340	200	310

(a) Analyze the data from this experiment.

(b) Use the Fisher LSD method to make comparisons among the three designs to determine specifically which designs differ in the mean response rate.

(c) Analyze the residuals from this experiment.

4.14 The effect of three different lubricating oils on fuel economy in diesel truck engines is being studied. Fuel economy is measured using brake-specific fuel consumption after the engine has been running for 15 minutes. Five different truck engines are available for the study, and the experimenters conduct the following RCBD.

	Truck				
Oil	1	2	3	4	5
1	0.500	0.634	0.487	0.329	0.512
2	0.535	0.675	0.520	0.435	0.540
3	0.513	0.595	0.488	0.400	0.510

(a) Analyze the data from this experiment.

(b) Use the Fisher LSD method to make comparisons among the three lubricating oils to determine specifically which oils differ in brake-specific fuel consumption.

(c) Analyze the residuals from this experiment.

4.15 An article in the *Fire Safety Journal* ("The Effect of Nozzle Design on the Stability and Performance of Turbulent Water Jets," Vol. 4, August 1981) describes an experiment in which a shape factor was determined for several different nozzle designs at six levels of jet efflux velocity. Interest focused on potential differences between nozzle designs, with velocity considered as a nuisance variable. The data are shown below:

	Jet Efflux Velocity (m/s)					
Nozzle Design	11.73	14.37	16.59	20.43	23.46	28.74
1	0.78	0.80	0.81	0.75	0.77	0.78
2	0.85	0.85	0.92	0.86	0.81	0.83
3	0.93	0.92	0.95	0.89	0.89	0.83
4	1.14	0.97	0.98	0.88	0.86	0.83
5	0.97	0.86	0.78	0.76	0.76	0.75

(a) Does nozzle design affect the shape factor? Compare the nozzles with a scatter plot and with an analysis of variance, using $\alpha = 0.05$.

(b) Analyze the residuals from this experiment.

(c) Which nozzle designs are different with respect to shape factor? Draw a graph of the average shape factor for each nozzle type and compare this to a scaled t distribution. Compare the conclusions that you draw from this plot to those from Duncan's multiple range test.

4.16 An article in *Communications of the ACM* (Vol. 30, No. 5, 1987) studied different algorithms for estimating software development costs. Six algorithms were applied to several different software development projects and the percent error in estimating the development cost was observed. Some of the data from this experiment is shown in the table below.

(a) Do the algorithms differ in their mean cost estimation accuracy?

(b) Analyze the residuals from this experiment.

(c) Which algorithm would you recommend for use in practice?

	Project					
Algorithm	1	2	3	4	5	6
1(SLIM)	1244	21	82	2221	905	839
2(COCOMO-A)	281	129	396	1306	336	910
3(COCOMO-R)	220	84	458	543	300	794
4(COCONO-C)	225	83	425	552	291	826
5(FUNCTION POINTS)	19	11	−34	121	15	103
6(ESTIMALS)	−20	35	−53	170	104	199

4.17 An article in *Nature Genetics* (2003, Vol. 34, pp. 85–90) "Treatment-Specific Changes in Gene Expression Discriminate in vivo Drug Response in Human Leukemia Cells" studied gene expression as a function of different treatments for leukemia. Three treatment groups are as follows: mercaptopurine (MP) only; low-dose methotrexate (LDMTX) chap 6 Module A.md → utf-8

4.23 Suppose that the observation for chemical type 2 and bolt 3 is missing in Problem 4.8. Analyze the problem by estimating the missing value. Perform the exact analysis and compare the results.

4.24 Consider the hardness testing experiment in Problem 4.12. Suppose that the observation for tip 2 in coupon 3 is missing. Analyze the problem by estimating the missing value.

4.25 ***Two missing values in a randomized block.*** Suppose that in Problem 4.8 the observations for chemical type 2 and bolt 3 and chemical type 4 and bolt 4 are missing.

(a) Analyze the design by iteratively estimating the missing values, as described in Section 4.1.3.

(b) Differentiate SS_E with respect to the two missing values, equate the results to zero, and solve for estimates of the missing values. Analyze the design using these two estimates of the missing values.

(c) Derive general formulas for estimating two missing values when the observations are in *different* blocks.

(d) Derive general formulas for estimating two missing values when the observations are in the *same* block.

4.26 An industrial engineer is conducting an experiment on eye focus time. He is interested in the effect of the distance of the object from the eye on the focus time. Four different distances are of interest. He has five subjects available for the experiment. Because there may be differences among individuals, he decides to conduct the experiment in a randomized block design. The data obtained follow. Analyze the data from this experiment (use $\alpha = 0.05$) and draw appropriate conclusions.

	Subject				
Distance (ft)	**1**	**2**	**3**	**4**	**5**
4	10	6	6	6	6
6	7	6	6	1	6
8	5	3	3	2	5
10	6	4	4	2	3

4.27 The effect of five different ingredients (A, B, C, D, E) on the reaction time of a chemical process is being studied. Each batch of new material is only large enough to permit five runs to be made. Furthermore, each run requires approximately $1\frac{1}{2}$ hours, so only five runs can be made in one day. The experimenter decides to run the experiment as a Latin square so that day and batch effects may be systematically controlled. She obtains the data that follow. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

	Day				
Batch	**1**	**2**	**3**	**4**	**5**
1	$A = 8$	$B = 7$	$D = 1$	$C = 7$	$E = 3$
2	$C = 11$	$E = 2$	$A = 7$	$D = 3$	$B = 8$
3	$B = 4$	$A = 9$	$C = 10$	$E = 1$	$D = 5$
4	$D = 6$	$C = 8$	$E = 6$	$B = 6$	$A = 10$
5	$E = 4$	$D = 2$	$B = 3$	$A = 8$	$C = 8$

4.28 An industrial engineer is investigating the effect of four assembly methods (A, B, C, D) on the assembly time for a color television component. Four operators are selected for the study. Furthermore, the engineer knows that each assembly method produces such fatigue that the time required for the last assembly may be greater than the time required for the first, regardless of the method. That is, a trend develops in the required assembly time. To account for this source of variability, the engineer uses the Latin square design that follows. Analyze the data from this experiment ($\alpha = 0.05$) and draw appropriate conclusions.

	Operator			
Order of Assembly	**1**	**2**	**3**	**4**
1	$C = 10$	$D = 14$	$A = 7$	$B = 8$
2	$B = 7$	$C = 18$	$D = 11$	$A = 8$
3	$A = 5$	$B = 10$	$C = 11$	$D = 9$
4	$D = 10$	$A = 10$	$B = 12$	$C = 14$

4.29 Consider the randomized complete block design in Problem 4.9. Assume that the days are random. Estimate the block variance component.

4.30 Consider the randomized complete block design in Problem 4.12. Assume that the coupons are random. Estimate the block variance component.

4.31 Consider the randomized complete block design in Problem 4.14. Assume that the trucks are random. Estimate the block variance component.

4.32 Consider the randomized complete block design in Problem 4.16. Assume that the software projects that were used as blocks are random. Estimate the block variance component.

4.33 Consider the gene expression experiment in Problem 4.17. Assume that the subjects used in this experiment are random. Estimate the block variance component.

4.34 Suppose that in Problem 4.27 the observation from batch 3 on day 4 is missing. Estimate the missing value and perform the analysis using the value.

4.35 Consider a $p \times p$ Latin square with rows (α_i), columns (β_k), and treatments (τ_j) fixed. Obtain least squares estimates of the model parameters α_i, β_k, and τ_j.

4.36 Derive the missing value formula (Equation 4.28) for the Latin square design.

4.37 ***Designs involving several Latin squares.*** [See Cochran and Cox (1957), John (1971).] The $p \times p$ Latin square contains only p observations for each treatment. To obtain more replications, the experimenter may use several squares, say n. It is immaterial whether the squares used are the same or different. The appropriate model is

$$y_{ijkh} = \begin{array}{l} \mu + \rho_h + \alpha_{i(h)} \\ + \tau_j + \beta_{k(h)} \\ + (\tau\rho)_{jh} + \epsilon_{ijkh} \end{array} \begin{cases} i = 1, 2, \dots, p \\ j = 1, 2, \dots, p \\ k = 1, 2, \dots, p \\ h = 1, 2, \dots, n \end{cases}$$

where y_{ijkh} is the observation on treatment j in row i and column k of the hth square. Note that $\alpha_{i(h)}$ and $\beta_{k(h)}$ are the row and column effects in the hth square, ρ_h is the effect of the hth square, and $(\tau\rho)_{jh}$ is the interaction between treatments and squares.

(a) Set up the normal equations for this model, and solve for estimates of the model parameters. Assume that appropriate side conditions on the parameters are $\sum_h \hat{\rho}_h = 0$, $\sum_i \hat{\alpha}_{i(h)} = 0$, and $\sum_k \hat{\beta}_{k(h)} = 0$ for each h, $\sum_j \hat{\tau}_j = 0$, $\sum_j (\hat{\tau}\rho)_{jh} = 0$ for each h, and $\sum_h (\hat{\tau}\rho)_{jh} = 0$ for each j.

(b) Write down the analysis of variance table for this design.

4.38 Discuss how you would determine the sample size for use with the Latin square design.

4.39 Suppose that in Problem 4.27 the data taken on day 5 were incorrectly analyzed and had to be discarded. Develop an appropriate analysis for the remaining data.

4.40 The yield of a chemical process was measured using five batches of raw material, five acid concentrations, five standing times (*A, B, C, D, E*), and five catalyst concentrations $(\alpha, \beta, \gamma, \delta, \epsilon)$. The Graeco-Latin square that follows was used. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

Batch	Acid Concentration 1	2	3	4	5
1	$A\alpha = 26$	$B\beta = 16$	$C\gamma = 19$	$D\delta = 16$	$E\epsilon = 13$
2	$B\gamma = 18$	$C\delta = 21$	$D\epsilon = 18$	$E\alpha = 11$	$A\beta = 21$
3	$C\epsilon = 20$	$D\alpha = 12$	$E\beta = 16$	$A\gamma = 25$	$B\delta = 13$
4	$D\beta = 15$	$E\gamma = 15$	$A\delta = 22$	$B\epsilon = 14$	$C\alpha = 17$
5	$E\delta = 10$	$A\epsilon = 24$	$B\alpha = 17$	$C\beta = 17$	$D\gamma = 14$

4.41 Suppose that in Problem 4.28 the engineer suspects that the workplaces used by the four operators may represent an additional source of variation. A fourth factor, workplace $(\alpha, \beta, \gamma, \delta)$ may be introduced and another experiment conducted, yielding the Graeco-Latin square that follows. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

Order of Assembly	Operator 1	2	3	4
1	$C\beta = 11$	$B\gamma = 10$	$D\delta = 14$	$A\alpha = 8$
2	$B\alpha = 8$	$C\delta = 12$	$A\gamma = 10$	$D\beta = 12$
3	$A\delta = 9$	$D\alpha = 11$	$B\beta = 7$	$C\gamma = 15$
4	$D\gamma = 9$	$A\beta = 8$	$C\alpha = 18$	$B\delta = 6$

4.42 Construct a 5×5 hypersquare for studying the effects of five factors. Exhibit the analysis of variance table for this design.

4.43 Consider the data in Problems 4.28 and 4.41. Suppressing the Greek letters in problem 4.41, analyze the data using the method developed in Problem 4.37.

4.44 Consider the randomized block design with one missing value in Problem 4.24. Analyze this data by using the exact analysis of the missing value problem discussed in Section 4.1.4. Compare your results to the approximate analysis of these data given from Problem 4.24.

4.45 An engineer is studying the mileage performance characteristics of five types of gasoline additives. In the road test he wishes to use cars as blocks; however, because of a time constraint, he must use an incomplete block design. He runs the balanced design with the five blocks that follow. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.