# CHAPTER 6

# Reserved Problems

**6.1R** In Problem 6.13, the engineer was also interested in potential fatigue differences resulting from the two types of bottles. As a measure of the amount of effort required, he measured the elevation of the heart rate (pulse) induced by the task. The results follow. Analyze the data and draw conclusions. Analyze the residuals and comment on the model's adequacy.

| | Worker | | | |
|---|---|---|---|---|
| **Bottle Type** | **1** | | **2** | |
| Glass | 39 | 45 | 20 | 13 |
| | 58 | 35 | 16 | 11 |
| Plastic | 44 | 35 | 13 | 10 |
| | 42 | 21 | 16 | 15 |

**6.2R** Calculate approximate 95 percent confidence limits for the factor effects in Problem 6.1R. Do the results of this analysis agree with the analysis of variance performed in Problem 6.1R?

**6.3R** A nickel–titanium alloy is used to make components for jet turbine aircraft engines. Cracking is a potentially serious problem in the final part because it can lead to nonrecoverable failure. A test is run at the parts producer to determine the effect of four factors on cracks. The four factors are pouring temperature ($A$), titanium content ($B$), heat treatment method ($C$), and amount of grain refiner used ($D$). Two replicates of a $2^4$ design are run, and the length of crack (in mm $\times 10^{-2}$) induced in a sample coupon subjected to a standard test is measured. The data are shown in Table P6.1R.

**■ TABLE P6.1R**
**The Experiment for problem 6.3R**

| | | | | | Replicate | |
|---|---|---|---|---|---|---|
| ***A*** | ***B*** | ***C*** | ***D*** | **Treatment Combination** | **I** | **II** |
| − | − | − | − | (1) | 7.037 | 6.376 |
| + | − | − | − | *a* | 14.707 | 15.219 |
| − | + | − | − | *b* | 11.635 | 12.089 |
| + | + | − | − | *ab* | 17.273 | 17.815 |
| − | − | + | − | *c* | 10.403 | 10.151 |
| + | − | + | − | *ac* | 4.368 | 4.098 |
| − | + | + | − | *bc* | 9.360 | 9.253 |
| + | + | + | − | *abc* | 13.440 | 12.923 |
| − | − | − | + | *d* | 8.561 | 8.951 |
| + | − | − | + | *ad* | 16.867 | 17.052 |
| − | + | − | + | *bd* | 13.876 | 13.658 |
| + | + | − | + | *abd* | 19.824 | 19.639 |
| − | − | + | + | *cd* | 11.846 | 12.337 |
| + | − | + | + | *acd* | 6.125 | 5.904 |
| − | + | + | + | *bcd* | 11.190 | 10.935 |
| + | + | + | + | *abcd* | 15.653 | 15.053 |

**(a)** Estimate the factor effects. Which factor effects appear to be large?

**(b)** Conduct an analysis of variance. Do any of the factors affect cracking? Use $\alpha = 0.05$.

**(c)** Write down a regression model that can be used to predict crack length as a function of the significant main effects and interactions you have identified in part (b).

**(d)** Analyze the residuals from this experiment.

**(e)** Is there an indication that any of the factors affect the variability in cracking?

**(f)** What recommendations would you make regarding process operations? Use interaction and/or main effect plots to assist in drawing conclusions.

**6.4R** ***Continuation of Problem 6.2R***. One of the variables in the experiment described in Problem 6.2R, heat treatment method ($C$), is a categorical variable. Assume that the remaining factors are continuous.

**(a)** Write two regression models for predicting crack length, one for each level of the heat treatment method variable. What differences, if any, do you notice in these two equations?

**(b)** Generate appropriate response surface contour plots for the two regression models in part (a).

**(c)** What set of conditions would you recommend for the factors $A$, $B$, and $D$ if you use heat treatment method $C = +$?

**(d)** Repeat part (c) assuming that you wish to use heat treatment method $C = -$.

**6.5R** Semiconductor manufacturing processes have long and complex assembly flows, so matrix marks and automated 2d-matrix readers are used at several process steps throughout factories. Unreadable matrix marks negatively affect factory run rates because manual entry of part data is required before manufacturing can resume. A $2^4$ factorial experiment was conducted to develop a 2d-matrix laser mark on a metal cover that protects a substrate-mounted die. The design factors are $A$ = laser power (9 and 13 W), $B$ = laser pulse frequency (4000 and 12,000 Hz), $C$ = matrix cell size (0.07 and 0.12 in.), and $D$ = writing speed (10 and 20 in./sec), and the response variable is the unused error correction (UEC). This is a measure of the unused portion of the redundant information embedded in the 2d-matrix. A UEC of 0 represents the lowest reading that still results in a decodable matrix, while a value of 1 is the highest reading. A DMX Verifier was used to measure the UEC. The data from this experiment are shown in Table P6.2R.

**(a)** Analyze the data from this experiment. Which factors significantly affect the UEC?

**(b)** Analyze the residuals from this experiment. Are there any indications of model inadequacy?

**6.6R** Reconsider the experiment described in Problem 6.20. Suppose that four center points are available and that the UEC response at these four runs is 0.98, 0.95, 0.93, and 0.96, respectively. Reanalyze the experiment incorporating a test for curvature into the analysis. What conclusions can you draw? What recommendations would you make to the experimenters?

■ **TABLE P6.2R**
**The $2^4$ Experiment for Problem 6.5R**

| Standard Order | Run Order | Laser Power | Pulse Frequency | Cell Size | Writing Speed | UEC |
|---|---|---|---|---|---|---|
| 8 | 1 | 1.00 | 1.00 | 1.00 | −1.00 | 0.8 |
| 10 | 2 | 1.00 | −1.00 | −1.00 | 1.00 | 0.81 |
| 12 | 3 | 1.00 | 1.00 | −1.00 | 1.00 | 0.79 |
| 9 | 4 | −1.00 | −1.00 | −1.00 | 1.00 | 0.6 |
| 7 | 5 | −1.00 | 1.00 | 1.00 | −1.00 | 0.65 |
| 15 | 6 | −1.00 | 1.00 | 1.00 | 1.00 | 0.55 |
| 2 | 7 | 1.00 | −1.00 | −1.00 | −1.00 | 0.98 |
| 6 | 8 | 1.00 | −1.00 | 1.00 | −1.00 | 0.67 |
| 16 | 9 | 1.00 | 1.00 | 1.00 | 1.00 | 0.69 |
| 13 | 10 | −1.00 | −1.00 | 1.00 | 1.00 | 0.56 |
| 5 | 11 | −1.00 | −1.00 | 1.00 | −1.00 | 0.63 |
| 14 | 12 | 1.00 | −1.00 | 1.00 | 1.00 | 0.65 |
| 1 | 13 | −1.00 | −1.00 | −1.00 | −1.00 | 0.75 |
| 3 | 14 | −1.00 | 1.00 | −1.00 | −1.00 | 0.72 |
| 4 | 15 | 1.00 | 1.00 | −1.00 | −1.00 | 0.98 |
| 11 | 16 | −1.00 | 1.00 | −1.00 | 1.00 | 0.63 |

**6.7R** ***A missing value in a $2^k$ factorial***. It is not unusual to find that one of the observations in a $2^k$ design is missing due to faulty measuring equipment, a spoiled test, or some other reason. If the design is replicated $n$ times ($n > 1$), some of the techniques discussed in Chapter 5 can be employed. However, for an unreplicated factorial ($n = 1$) some other method must be used. One logical approach is to estimate the missing value with a number that makes the highest order interaction contrast zero. Apply this technique to the experiment in Example 6.2 assuming that run *ab* is missing. Compare the results with the results of Example 6.2.

**6.8R** The book by Davies (*Design and Analysis of Industrial Experiments*) describes an experiment to study the yield of isatin. The factors studied and their levels are as follows:

| Factor | Low (−) | High (+) |
|---|---|---|
| A: Acid strength (%) | 87 | 93 |
| B: Reaction time (min) | 15 | 30 |
| C: Amount of acid (mL) | 35 | 45 |
| D: Reaction temperature (°C) | 60 | 70 |

The data from the $2^4$ factorial are shown in Table P6.3R.

**(a)** Fit a main-effects-only model to the data from this experiment. Are any of the main effects significant?

■ **TABLE P6.3R**
**The $2^4$ Factorial Experiment in Problem 6.8R**

| A | B | C | D | Yield |
|---|---|---|---|---|
| −1 | −1 | −1 | −1 | 6.08 |
| 1 | −1 | −1 | −1 | 6.04 |
| −1 | 1 | −1 | −1 | 6.53 |
| 1 | 1 | −1 | −1 | 6.43 |
| −1 | −1 | 1 | −1 | 6.31 |
| 1 | −1 | 1 | −1 | 6.09 |
| −1 | 1 | 1 | −1 | 6.12 |
| 1 | 1 | 1 | −1 | 6.36 |
| −1 | −1 | −1 | 1 | 6.79 |
| 1 | −1 | −1 | 1 | 6.68 |
| −1 | 1 | −1 | 1 | 6.73 |
| 1 | 1 | −1 | 1 | 6.08 |
| −1 | −1 | 1 | 1 | 6.77 |
| 1 | −1 | 1 | 1 | 6.38 |
| −1 | 1 | 1 | 1 | 6.49 |
| 1 | 1 | 1 | 1 | 6.23 |

**(b)** Analyze the residuals. Are there any indications of model inadequacy or violation of the assumptions?

**(c)** Find an equation for predicting the yield of isatin over the design space. Express the equation in both coded and engineering units.

**(d)** Is there any indication that adding interactions to the model would improve the results that you have obtained?

**6.9R** ***Hierarchical models***. Several times we have used the hierarchy principle in selecting a model; that is, we have included nonsignificant lower order terms in a model because they were factors involved in significant higher order terms. Hierarchy is certainly not an absolute principle that must be followed in all cases. To illustrate, consider the model resulting from Problem 6.5, which required that a nonsignificant main effect be included to achieve hierarchy. Using the data from Problem 6.5.

**(a)** Fit both the hierarchical and the nonhierarchical models.

**(b)** Calculate the PRESS statistic, the adjusted $R^2$, and the mean square error for both models.

**(c)** Find a 95 percent confidence interval on the estimate of the mean response at a cube corner ($x_1 = x_2 = x_3 = \pm 1$). *Hint:* Use the results of Problem 6.33.

**(d)** Based on the analyses you have conducted, which model do you prefer?

**6.10R** An engineer has conducted an experiment using a single replicate of a $2^3$ factorial design. After analyzing the data, the following results are available:

| Term | Effect Estimate | Sum of Squares | % Contribution |
|---|---|---|---|
| A | −3.75 | 28.125 | _________ |
| B | _____ | 6.125 | 1.73574 |
| C | 0.25 | 0.125 | 0.0354233 |
| AB | −2.75 | 15.125 | 4.28622 |
| AC | 3.25 | _____ | 5.98654 |
| BC | −1.75 | 6.125 | 1.73574 |
| ABC | −11.75 | 276.125 | 78.2501 |

Answer the following questions.

**(a)** What is the effect estimate for factor B? ________ __________

**(b)** What is the sum of squares for the AC interaction? ___________

**(c)** What is the percent contribution for factor A? _______________

**(d)** this an I-optimal design (circle the **correct** answer)? **Yes** **No**

**(e)** Suppose that the experimenter also ran 4 center points (the response data is not included in the calculations above) and observed the following results: 20, 22, 23, and 21. The mean square for pure error is

**a.** 2.00
**b.** 1.85
**c.** 1.67
**d.** 1.50
**e.** None of the above

**(f)** Suppose that the average of the 8 runs in the factorial portion of the experiment is 24. The value of the $F$-statistic for pure quadratic curvature is

**a.** 10.51
**b.** 9.98
**c.** 8.75
**d.** 9.50
**e.** None of the above

**(g)** The table below shows the coded and natural (engineering) units for the factors. Find the regression model for predicting the response in terms of the natural units for the inorganic catalyst assuming that the model contains only the main effects of A and C and the AC interaction.

| Factor | Natural levels | Coded levels (x's) |
|---|---|---|
| **A – time** | **80, 120 (minutes)** | **−1, 1** |
| **B - temperature** | **150, 200 (deg C)** | **−1, 1** |
| **C –catalyst type** | **organic, inorganic** | **−1, 1** |

**6.11R** Lenth's method is a technique for determining which factors are potentially active in an unreplicated $2^k$ design. **True** **False**

**6.12R** All $2^k$ designs are D-optimal designs. **True** **False**

**6.13R** A design that is G-optimal is one for which the average variance of the predicted response over the design space is minimized. **True** **False**

**6.14R** In a $2^k$ design if center points are added, the observations at the center do not have any impact of the estimates of the main effects or interaction terms in the model. **True** **False**

**6.15R** A $2^k$ design with additional center points is a D-optimal design. **True** **False**

**6.16R** An engineer has conducted an experiment using a single replicate of a $2^4$ factorial design. After analyzing the data, the following results are available:

| Model Term | Effect Estimate | Sum of Squares | % Contribution |
|---|---|---|---|
| Intercept | X | X | X |
| A-A | −9.75 | 380.25 | 40.4952 |
| B-B | −0.25 | | 0.0266241 |
| C-C | −0.25 | 0.25 | 0.0266241 |
| D-D | | 90.25 | 9.61129 |
| AB | 4.5 | 81 | X |
| AC | 0.5 | 1 | 0.106496 |
| AD | −3.5 | 49 | 5.21832 |
| BC | −1 | 4 | 0.425985 |
| BD | −2 | 16 | 1.70394 |
| CD | 1 | 4 | 0.425985 |
| ABC | −6.25 | 156.25 | 16.64 |
| ABD | 4.75 | 90.25 | 9.61129 |
| ACD | −0.25 | 0.25 | 0.0266241 |
| BCD | −2.75 | 30.25 | 3.22151 |
| ABCD | 3 | X | 3.83387 |

In addition, you know the following: the grand average of all 16 observations is 83.75 and the total sum of squares is 939.00.

Answer the following questions.

**(a)** Fill in the blanks in the table above,

**(b)** Is this an I-optimal design (circle the **correct** answer)? **Yes** **No**

**(c)** Suppose that the experimenter also ran 4 center points (the response data is not included in the calculations above) and observed the following results: 88 89, 92, and 87. The mean square for pure error is

**a.** 2.00
**b.** 3.85
**c.** 4.67
**d.** 5.25
**e.** None of the above

**(d)** The value of the $F$-statistic for pure quadratic curvature is

**a.** 10.51
**b.** 9.98
**c.** 8.75
**d.** 9.50
**e.** None of the above

**(e)** The table below shows the coded and natural (engineering) units for the factors. Find the regression model(s) for predicting the response in terms of the natural units assuming that the model contains only the main effects of A and D and the AD interaction.

| Factor | Natural levels | Coded levels (x's) |
|---|---|---|
| **A – time** | **80, 120 (minutes)** | **−1, 1** |
| **B - temperature** | **150, 200 (deg C)** | **−1, 1** |
| **C - pressure** | **(100, 150) psi** | **−1, 1** |
| **D - catalyst type** | **Organic, inorganic** | **−1, 1** |

**6.17R** Center points can be used to determine:

(**I**) an estimate of pure error

(**II**) an estimate of the sum of the quadratic effects

(**III**) if a process is running "as usual"

**(a)** I and II

**(b)** II and III

**(c)** I and III

**(d)** all of the above

**(e)** none of the above

**6.18R** The residuals from the model fit to data from a designed experiment

**(a)** are independent random variables.

**(b)** are always normally distributed.

**(c)** always have constant variance.

**(d)** none of the above (a-c) are true

**6.19R** If an observation is missing in an unreplicated 2k design, a reasonable strategy is to replace the observation with the average of the actual observations. **True** **False**