set of conditions shown below. The response variable is vibration measured as the resultant vector of three accelerometers (x, y, and z) on each test circuit board.

A	B	Treatment Combination	Replicate I	II	III	IV
−	−	(1)	18.2	18.9	12.9	14.4
+	−	a	27.2	24.0	22.4	22.5
−	+	b	15.9	14.5	15.1	14.2
+	+	ab	41.0	43.9	36.3	39.9

(a) Analyze the data from this experiment.
(b) Construct a normal probability plot of the residuals, and plot the residuals versus the predicted vibration level. Interpret these plots.
(c) Draw the AB interaction plot. Interpret this plot. What levels of bit size and speed would you recommend for routine operation?

6-6. Reconsider the experiment described in Problem 6-1. Suppose that the experimenter only performed the eight trials from replicate I. In addition, he ran four center points and obtained the following response values: 36, 40, 43, 45.
(a) Estimate the factor effects. Which effects are large?
(b) Perform an analysis of variance, including a check for pure quadratic curvature. What are your conclusions?
(c) Write down an appropriate model for predicting tool life, based on the results of this experiment. Does this model differ in any substantial way from the model in Problem 6-1, part (c)?
(d) Analyze the residuals.
(e) What conclusions would you draw about the appropriate operating conditions for this process?

6-7. An experiment was performed to improve the yield of a chemical process. Four factors were selected, and two replicates of a completely randomized experiment were run. The results are shown in the following table:

Treatment Combination	Replicate I	II	Treatment Combination	Replicate I	II
(1)	90	93	d	98	95
a	74	78	ad	72	76
b	81	85	bd	87	83
ab	83	80	abd	85	86
c	77	78	cd	99	90
ac	81	80	acd	79	75
bc	88	82	bcd	87	84
abc	73	70	$abcd$	80	80

(a) Estimate the factor effects.
(b) Prepare an analysis of variance table, and determine which factors are important in explaining yield.
(c) Write down a regression model for predicting yield, assuming that all four factors were varied over the range from −1 to +1 (in coded units).
(d) Plot the residuals versus the predicted yield and on a normal probability scale. Does the residual analysis appear satisfactory?
(e) Two three-factor interactions, ABC and ABD, apparently have large effects. Draw a cube plot in the factors A, B, and C with the average yields shown at each corner. Repeat using the factors A, B, and D. Do these two plots aid in data interpretation? Where would you recommend that the process be run with respect to the four variables?

6-8. A bacteriologist is interested in the effects of two different culture media and two different times on the growth of a particular virus. She performs six replicates of a 2^2 design, making the runs in random order. Analyze the bacterial growth data that follow and draw appropriate conclusions. Analyze the residuals and comment on the model's adequacy.

Time, h	Culture Medium 1		2	
	21	22	25	26
12	23	28	24	25
	20	26	29	27
	37	39	31	34
18	38	38	29	33
	35	36	30	35

(a) Analyze the data from this experiment. Which factors significantly affect putting performance?
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

6-20. 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 (9W, 13W), B = laser pulse frequency (4000 Hz, 12,000 Hz), C = matrix cell size (0.07 in, 0.12 in), and D = writing speed (10 in/sec, 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 UEC. The data from this experiment are shown in Table 6-26.

(a) Analyze the data from this experiment. Which factors significantly affect UEC?
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

6-21. 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?

6-22. A company markets its products by direct mail. An experiment was conducted to study the effects of three factors on the customer response rate for a particular product. The three factors are A = type of mail used (3rd class, 1st class), B = type of descriptive brochure (color, black-and-white), and C = offered price ($19.95, $24.95). The mailings are made to two groups of 8000 randomly selected customers, with 1000 customers in each group receiving each treatment combination. Each group of customers is considered as a replicate. The response variable is the number of orders placed. The experimental data are shown in the table on the next page.

Table 6-26 The 2^4 Experiment for Problem 6-20

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

Table 7-11 Analysis of Variance for Example 7-3

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Replicates	3875.0625	1	3875.0625	—	
Blocks within replicates	458.1250	2	458.1250	—	
A	41,310.5625	1	41,310.5625	16.20	0.01
B	217.5625	1	217.5625	0.08	0.78
C	374,850.5625	1	374,850.5625	146.97	<0.001
AB (rep. I only)	3528.0000	1	3528.0000	1.38	0.29
AC	94,404.5625	1	94,404.5625	37.01	<0.001
BC	18.0625	1	18.0625	0.007	0.94
ABC (rep. II only)	6.1250	1	6.1250	0.002	0.96
Error	12,752.3125	5	2550.4625		
Total	531,420.9375	15			

7-9 PROBLEMS

7-1. Consider the experiment described in Problem 6-1. Analyze this experiment assuming that each replicate represents a block of a single production shift.

7-2. Consider the experiment described in Problem 6-5. Analyze this experiment assuming that each one of the four replicates represents a block.

7-3. Consider the alloy cracking experiment described in Problem 6-15. Suppose that only 16 runs could be made on a single day, so each replicate was treated as a block. Analyze the experiment and draw conclusions.

7-4. Consider the data from the first replicate of Problem 6-1. Suppose that these observations could not all be run using the same bar stock. Set up a design to run these observations in two blocks of four observations each with ABC confounded. Analyze the data.

7-5. Consider the data from the first replicate of Problem 6-7. Construct a design with two blocks of eight observations each with $ABCD$ confounded. Analyze the data.

7-6. Repeat Problem 7-5 assuming that four blocks are required. Confound ABD and ABC (and consequently CD) with blocks.

7-7. Using the data from the 2^5 design in Problem 6-24, construct and analyze a design in two blocks with $ABCDE$ confounded with blocks.

7-8. Repeat Problem 7-7 assuming that four blocks are necessary. Suggest a reasonable confounding scheme.

7-9. Consider the data from the 2^5 design in Problem 6-24. Suppose that it was necessary to run this design in four blocks with $ACDE$ and BCD (and consequently ABE) confounded. Analyze the data from this design.

7-10. Consider the fill height deviation experiment in Problem 6-18. Suppose that each replicate was run on a separate day. Analyze the data assuming that days are blocks.

7-11. Consider the fill height deviation experiment in Problem 6-18. Suppose that only four runs could be made on each shift. Set up a design with ABC confounded in replicate 1 and AC confounded in replicate 2. Analyze the data and comment on your findings.

7-12. Consider the potting experiment in Problem 6-19. Analyze the data considering each replicate as a block.

7-13. Using the data from the 2^4 design in Problem 6-20, construct and analyze a design in two blocks with $ABCD$ confounded with blocks.

7-14. Consider the direct mail experiment in Problem 6-22. Suppose that each group of customers is in a different part of the country. Suggest an appropriate analysis for the experiment.

7-15. Design an experiment for confounding a 2^6 factorial in four blocks. Suggest an appropriate confounding scheme, different from the one shown in Table 7-8.

7-16. Consider the 2^6 design in eight blocks of eight runs each with *ABCD*, *ACE*, and *ABEF* as the independent effects chosen to be confounded with blocks. Generate the design. Find the other effects confounded with blocks.

7-17. Consider the 2^2 design in two blocks with *AB* confounded. Prove algebraically that $SS_{AB} = SS_{\text{Blocks}}$.

7-18. Consider the data in Example 7-2. Suppose that all the observations in block 2 are increased by 20. Analyze the data that would result. Estimate the block effect. Can you explain its magnitude? Do blocks now appear to be an important factor? Are any other effect estimates impacted by the change you made to the data?

7-19. Suppose that in Problem 6-1 we had confounded *ABC* in replicate I, *AB* in replicate II, and *BC* in replicate III. Calculate the factor effect estimates. Construct the analysis of variance table.

7-20. Repeat the analysis of Problem 6-1 assuming that *ABC* was confounded with blocks in each replicate.

7-21. Suppose that in Problem 6-7 *ABCD* was confounded in replicate I and *ABC* was confounded in replicate II. Perform the statistical analysis of this design.

7-22. Construct a 2^3 design with *ABC* confounded in the first two replicates and *BC* confounded in the third. Outline the analysis of variance and comment on the information obtained.

saturated design for $k = 10$ factors in $N = 6$ runs. We could have used the runs that are negative in column L equally well. This procedure will always produce a supersaturated design for $k = N - 2$ factors in $N/2$ runs. If there are fewer than $N - 2$ factors of interest, columns can be removed from the complete design.

Supersaturated designs are typically analyzed by regression model-fitting methods, such as forward selection (see Chapter 10). In this procedure, variables are selected one at a time for inclusion in the model until no other variables appear useful in explaining the response. Abraham, Chipman, and Vijayan (1999) and Holcomb, Montgomery, and Carlyle (2003) have studied analysis methods for supersaturated designs. Generally, these designs can experience large type I and type II errors, but some analysis methods can be tuned to emphasize type I errors so that the type II error rate will be moderate. In a factor screening situation, it is usually more important not to exclude an active factor than it is to conclude that inactive factors are important, so type I errors are less critical than type II errors. However, because both error rates can be large, the philosophy in using a supersaturated design should be to eliminate a large portion of the inactive factors, and not to clearly identify the few important or active factors. Holcomb, Montgomery, and Carlyle (2003) found that some types of supersaturated designs perform better than others with respect to type I and type II errors. Generally, the designs produced by search algorithms were outperformed by designs constructed from standard orthogonal designs.

Supersaturated designs have not had widespread use. However, they are an interesting and potentially useful method for experimentation with systems where there are many variables and only a very few of these are expected to produce large effects.

8-8 SUMMARY

This chapter has introduced the 2^{k-p} fractional factorial design. We have emphasized the use of these designs in screening experiments to quickly and efficiently identify the subset of factors that are active and to provide some information on interaction. The projective property of these designs makes it possible in many cases to examine the active factors in more detail. Sequential assembly of these designs via fold over is a very effective way to gain additional information about interactions that an initial experiment may identify as possibly important.

In practice, 2^{k-p} fractional factorial designs with $N = 4, 8, 16,$ and 32 runs are highly useful. Table 8-28 summarizes these designs, identifying how many factors can be used with each design to obtain various types of screening experiments. For example, the 16-run design is a full factorial for 4 factors, a one-half fraction for 5 factors, a resolution IV fraction for 6 to 8 factors, and a resolution III fraction for 9 to 15 factors. All of these designs may be constructed using the methods discussed in this chapter, and many of their alias structures are shown in Appendix Table X.

8-9 PROBLEMS

8-1. Suppose that in the chemical process development experiment described in Problem 6-7, it was only possible to run a one-half fraction of the 2^4 design. Construct the design and perform the statistical analysis, using the data from replicate I.

8-2. Suppose that in Problem 6-15, only a one-half fraction of the 2^4 design could be run. Construct the design and perform the analysis, using the data from replicate I.

8-3. Consider the plasma etch experiment described in Problem 6-18. Suppose that only a one-half fraction of the design could be run. Set up the design and analyze the data.

8-4. Problem 6-21 describes a process improvement study in the manufacturing process of an integrated circuit. Suppose that only eight runs could be made in this process. Set up an appropriate 2^{5-2} design and find the alias structure. Use the appropriate observations from Problem 6-21 as the observations in this design and estimate the factor effects. What conclusions can you draw?

8-5. ***Continuation of Problem 8-4.*** Suppose you have made the eight runs in the 2^{5-2} design in Problem 8-4. What additional runs would be required to identify the factor effects that are of interest? What are the alias relationships in the combined design?

8-6. R. D. Snee ("Experimenting with a Large Number of Variables," in *Experiments in Industry: Design, Analysis and Interpretation of Results*, by R. D. Snee, L. B. Hare, and J. B. Trout, Editors, ASQC, 1985) describes an experiment in which a 2^{5-1} design with $I = ABCDE$ was used to investigate the effects of five factors on the color of a chemical product. The factors are A = solvent/reactant, B = catalyst/reactant, C = temperature, D = reactant purity, and E = reactant pH. The results obtained were as follows:

$e = -0.63$	$d = 6.79$
$a = 2.51$	$ade = 5.47$
$b = -2.68$	$bde = 3.45$
$abe = 1.66$	$abd = 5.68$
$c = 2.06$	$cde = 5.22$
$ace = 1.22$	$acd = 4.38$
$bce = -2.09$	$bcd = 4.30$
$abc = 1.93$	$abcde = 4.05$

(a) Prepare a normal probability plot of the effects. Which effects seem active?

(b) Calculate the residuals. Construct a normal probability plot of the residuals and plot the residuals versus the fitted values. Comment on the plots.

(c) If any factors are negligible, collapse the 2^{5-1} design into a full factorial in the active factors. Comment on the resulting design, and interpret the results.

8-7. An article by J. J. Pignatiello, Jr. and J. S. Ramberg in the *Journal of Quality Technology* (Vol. 17, 1985, pp. 198–206) describes the use of a replicated fractional factorial to investigate the effect of five factors on the free height of leaf springs used in an automotive application. The factors are A = furnace temperature, B = heating time, C = transfer time, D = hold down time, and E = quench oil temperature. The data are shown below:

A	B	C	D	E	Free Height		
−	−	−	−	−	7.78	7.78	7.81
+	−	−	+	−	8.15	8.18	7.88
−	+	−	+	−	7.50	7.56	7.50
+	+	−	−	−	7.59	7.56	7.75
−	−	+	+	−	7.54	8.00	7.88
+	−	+	−	−	7.69	8.09	8.06
−	+	+	−	−	7.56	7.52	7.44
+	+	+	+	−	7.56	7.81	7.69
−	−	−	−	+	7.50	7.25	7.12
+	−	−	+	+	7.88	7.88	7.44
−	+	−	+	+	7.50	7.56	7.50
+	+	−	−	+	7.63	7.75	7.56
−	−	+	+	+	7.32	7.44	7.44
+	−	+	−	+	7.56	7.69	7.62
−	+	+	−	+	7.18	7.18	7.25
+	+	+	+	+	7.81	7.50	7.59

(a) Write out the alias structure for this design. What is the resolution of this design?

(b) Analyze the data. What factors influence the mean free height?

(c) Calculate the range and standard deviation of the free height for each run. Is there any indication that any of these factors affects variability in the free height?

(d) Analyze the residuals from this experiment, and comment on your findings.

(e) Is this the best possible design for five factors in 16 runs? Specifically, can you find a fractional design for five factors in 16 runs with a higher resolution than this one?

8-8. An article in *Industrial and Engineering Chemistry* ("More on Planning Experiments to Increase Research Efficiency," 1970, pp. 60–65) uses a 2^{5-2} design to