**Chapter 6**

**The 2*k* Factorial Design**

**Solutions**

**6.1.** In a 24 factorial design, the number of degrees of freedom for the model, assuming the complete factorial model, is

(a) **7**

(b)5

(c)6

(d)11

(e)12

(f)None of the above

**6.2.** A 23 factorial is replicated twice. The number of pure error or residual degrees of freedom are

(a)4

(b)12

(c)15

(d)2

(e) **8**

(f)None of the above

**6.3.** A 23 factorial is replicated twice. The ANOVA indicates that all main effects are significant but the interactions are not significant. The interaction terms are dropped from the model. The number of residual degrees of freedom for the reduced model are

(a) **12**

(b)8

(c)16

(d)14

(e)10

(f)None of the above

**6.4.** A 23 factorial is replicated three times. The ANOVA indicates that all main effects are significant, but two of the interactions are not significant. The interaction terms are dropped from the model. The number of residual degrees of freedom for the reduced model are

(a)12

(b)14

(c)6

(d)10

(e)8

(f) **None of the above**

**6.5S.** An engineer is interested in the effects of cutting speed (*A*), tool geometry (*B*), and cutting angle on the life (in hours) of a machine tool. Two levels of each factor are chosen, and three replicates of a 23 factorial design are run. The results are as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Treatment |  | Replicate |  |
| *A* | *B* | *C* | Combination | I | II | III |
| - | - | - | (1) | 22 | 31 | 25 |
| + | - | - | *a* | 32 | 43 | 29 |
| - | + | - | *b* | 35 | 34 | 50 |
| + | + | - | *ab* | 55 | 47 | 46 |
| - | - | + | *c* | 44 | 45 | 38 |
| + | - | + | *ac* | 40 | 37 | 36 |
| - | + | + | *bc* | 60 | 50 | 54 |
| + | + | + | *abc* | 39 | 41 | 47 |

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

From the normal probability plot of effects below, factors *B*, *C*, and the *AC* interaction appear to be significant.



(b) Use the analysis of variance to confirm your conclusions for part (a).

The analysis of variance confirms the significance of factors *B*, *C*, and the *AC* interaction.

Design Expert Output

**Response:** **Life** **in hours**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1612.67 7 230.38 7.64 0.0004 significant

A 0.67 1 0.67 0.022 0.8837

*B 770.67 1 770.67 25.55 0.0001*

*C 280.17 1 280.17 9.29 0.0077*

*AB 16.67 1 16.67 0.55 0.4681*

*AC 468.17 1 468.17 15.52 0.0012*

*BC 48.17 1 48.17 1.60 0.2245*

*ABC 28.17 1 28.17 0.93 0.3483*

### Pure Error 482.67 16 30.17

Cor Total 2095.33 23

The Model F-value of 7.64 implies the model is significant. There is only

a 0.04% chance that a "Model F-Value" this large could occur due to noise.

The reduced model ANOVA is shown below. Factor *A* was included to maintain hierarchy.

Design Expert Output

**Response:** **Life** **in hours**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1519.67 4 379.92 12.54 < 0.0001 significant

*A* *0.67* *1* *0.67* *0.022* *0.8836*

*B* *770.67* *1* *770.67* *25.44* *< 0.0001*

*C* *280.17* *1* *280.17* *9.25* *0.0067*

*AC* *468.17* *1* *468.17* *15.45* *0.0009*

Residual 575.67 19 30.30

*Lack of Fit* *93.00* *3* *31.00* *1.03* *0.4067* *not significant*

*Pure Error* *482.67* *16* *30.17*

Cor Total 2095.33 23

The Model F-value of 12.54 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Effects *B, C* and *AC* are significant at 1%.

(c) Write down a regression model for predicting tool life (in hours) based on the results of this experiment.



Design Expert Output

**Coefficient** **Standard** **95% CI** **95% CI**

**Factor** **Estimate** **DF** **Error** **Low** **High** **VIF**

Intercept 40.83 1 1.12 38.48 43.19

A-Cutting Speed 0.17 1 1.12 -2.19 2.52 1.00

B-Tool Geometry 5.67 1 1.12 3.31 8.02 1.00

C-Cutting Angle 3.42 1 1.12 1.06 5.77 1.00

AC -4.42 1 1.12 -6.77 -2.06 1.00

**Final Equation in Terms of Coded Factors:**

Life =

+40.83

+0.17 \* A

+5.67 \* B

+3.42 \* C

-4.42 \* A \* C

**Final Equation in Terms of Actual Factors:**

Life =

+40.83333

+0.16667 \* Cutting Speed

+5.66667 \* Tool Geometry

+3.41667 \* Cutting Angle

-4.41667 \* Cutting Speed \* Cutting Angle

The equation in part (c) and in the given in the computer output form a “hierarchical” model, that is, if an interaction is included in the model, then all of the main effects referenced in the interaction are also included in the model.

(d) Analyze the residuals. Are there any obvious problems?



There is nothing unusual about the residual plots.

(e) Based on the analysis of main effects and interaction plots, what levels of *A*, *B*, and *C* would you recommend using?

Since *B* has a positive effect, set *B* at the high level to increase life. The *AC* interaction plot reveals that life would be maximized with *C* at the high level and *A* at the low level.



**6.6S.** Reconsider part (c) of Problem 6.5S. Use the regression model to generate response surface and contour plots of the tool life response. Interpret these plots. Do they provide insight regarding the desirable operating conditions for this process?

The response surface plot and the contour plot in terms of factors *A* and *C* with *B* at the high level are shown below. They show the curvature due to the *AC* interaction. These plots make it easy to see the region of greatest tool life.



**6.7.** Find the standard error of the factor effects and approximate 95 percent confidence limits for the factor effects in Problem 6.5. Do the results of this analysis agree with the conclusions from the analysis of variance?



|  |  |  |
| --- | --- | --- |
| Variable | Effect |  |
| *A* | 0.333 |  |
| *B* | 11.333 | \* |
| *AB* | -1.667 |  |
| *C* | 6.833 | \* |
| *AC* | -8.833 | \* |
| *BC* | -2.833 |  |
| *ABC* | -2.167 |  |

The 95% confidence intervals for factors *B, C* and *AC* do not contain zero. This agrees with the analysis of variance approach.

**6.8.** Plot the factor effects from Problem 6.5 on a graph relative to an appropriately scaled *t* distribution. Does this graphical display adequately identify the important factors? Compare the conclusions from this plot with the results from the analysis of variance.





This method identifies the same factors as the analysis of variance.

**6.9S.** A router is used to cut locating notches on a printed circuit board. The vibration level at the surface of the board as it is cut is considered to be a major source of dimensional variation in the notches. Two factors are thought to influence vibration: bit size (*A*) and cutting speed (*B*). Two bit sizes (1/16 and 1/8 inch) and two speeds (40 and 90 rpm) are selected, and four boards are cut at each set of conditions shown below. The response variable is vibration measured as a resultant vector of three accelerometers (*x*, *y*, and *z*) on each test circuit board.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Treatment |  | Replicate |  |  |
| *A* | *B* | Combination | 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.

Design Expert Output

**Response:** **Vibration**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1638.11 3 546.04 91.36 < 0.0001 significant

*A* *1107.23* *1* *1107.23* *185.25* *< 0.0001*

*B* *227.26* *1* *227.26* *38.02* *< 0.0001*

*AB* *303.63* *1* *303.63* *50.80* *< 0.0001*

Residual 71.72 12 5.98

*Lack of Fit* *0.000* *0*

*Pure Error* *71.72* *12* *5.98*

Cor Total 1709.83 15

The Model F-value of 91.36 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

(b) Construct a normal probability plot of the residuals, and plot the residuals versus the predicted vibration level. Interpret these plots.



There is nothing unusual about the residual plots.

(c) Draw the *AB* interaction plot. Interpret this plot. What levels of bit size and speed would you recommend for routine operation?

To reduce the vibration, use the smaller bit. Once the small bit is specified, either speed will work equally well, because the slope of the curve relating vibration to speed for the small tip is approximately zero. The process is robust to speed changes if the small bit is used.



**6.10S.** Reconsider the experiment described in Problem 6.5S. 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?



Effects *B, C,* and *AC* appear to be large.

(b) Perform an analysis of variance, including a check for pure quadratic curvature. What are your conclusions?



Design Expert Output

**Response:** **Life** **in hours**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1048.88 7 149.84 9.77 0.0439 significant

*A 3.13 1 3.13 0.20 0.6823*

*B 325.13 1 325.13 21.20 0.0193*

*C 190.12 1 190.12 12.40 0.0389*

*AB 6.13 1 6.13 0.40 0.5722*

*AC 378.12 1 378.12 24.66 0.0157*

*BC 55.12 1 55.12 3.60 0.1542*

*ABC 91.12 1 91.12 5.94 0.0927*

Curvature 0.042 1 0.042 2.717E-003 0.9617 not significant

Pure Error 46.00 3 15.33

Cor Total 1094.92 11

The Model F-value of 9.77 implies the model is significant. There is only

a 4.39% chance that a "Model F-Value" this large could occur due to noise.

The "Curvature F-value" of 0.00 implies the curvature (as measured by difference between the

average of the center points and the average of the factorial points) in the design space is not

significant relative to the noise. There is a 96.17% chance that a "Curvature F-value"

this large could occur due to noise.

Design Expert Output

**Response:** **Life** **in hours**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 896.50 4 224.13 7.91 0.0098 significant

*A 3.13 1 3.13 0.11 0.7496*

*B 325.12 1 325.12 11.47 0.0117*

*C 190.12 1 190.12 6.71 0.0360*

*AC 378.12 1 378.12 13.34 0.0082*

Residual 198.42 7 28.35

*Lack of Fit 152.42 4 38.10 2.49 0.2402 not significant*

*Pure Error 46.00 3 15.33*

Cor Total 1094.92 11

The Model F-value of 7.91 implies the model is significant. There is only

a 0.98% chance that a "Model F-Value" this large could occur due to noise.

Effects *B, C* and *AC* are significant at 5%. There is no effect of curvature.

(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)?

The model shown in the *Design Expert* output below does not differ substantially from the model in Problem 6.5, part (c).

Design Expert Output

**Final Equation in Terms of Coded Factors:**

Life =

+40.88

+0.62 \* A

+6.37 \* B

+4.87 \* C

-6.88 \* A \* C

(d) Analyze the residuals.



(e) What conclusions would you draw about the appropriate operating conditions for this process?

To maximize life run with *B* at the high level, *A* at the low level and *C* at the high level



**6.11.** 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 | Replicate | Replicate | Treatment | Replicate | Replicate |
| Combination | I | II | Combination | 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.

Design Expert Output

Term Effect SumSqr % Contribtn

Model Intercept

Error A -9.0625 657.031 40.3714

Error B -1.3125 13.7812 0.84679

Error C -2.6875 57.7813 3.55038

Error D 3.9375 124.031 7.62111

Error AB 4.0625 132.031 8.11267

Error AC 0.6875 3.78125 0.232339

Error AD -2.1875 38.2813 2.3522

Error BC -0.5625 2.53125 0.155533

Error BD -0.1875 0.28125 0.0172814

Error CD 1.6875 22.7812 1.3998

Error ABC -5.1875 215.281 13.228

Error ABD 4.6875 175.781 10.8009

Error ACD -0.9375 7.03125 0.432036

Error BCD -0.9375 7.03125 0.432036

Error ABCD 2.4375 47.5313 2.92056

(b) Prepare an analysis of variance table, and determine which factors are important in explaining yield.

Design Expert Output

**Response:** **yield**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1504.97 15 100.33 13.10 < 0.0001 significant

*A* *657.03* *1* *657.03* *85.82* *< 0.0001*

*B* *13.78* *1* *13.78* *1.80* *0.1984*

*C* *57.78* *1* *57.78* *7.55* *0.0143*

*D* *124.03* *1* *124.03* *16.20* *0.0010*

*AB* *132.03* *1* *132.03* *17.24* *0.0007*

*AC* *3.78* *1* *3.78* *0.49* *0.4923*

*AD* *38.28* *1* *38.28* *5.00* *0.0399*

*BC* *2.53* *1* *2.53* *0.33* *0.5733*

*BD* *0.28* *1* *0.28* *0.037* *0.8504*

*CD* *22.78* *1* *22.78* *2.98* *0.1038*

*ABC* *215.28* *1* *215.28* *28.12* *< 0.0001*

*ABD* *175.78* *1* *175.78* *22.96* *0.0002*

*ACD* *7.03* *1* *7.03* *0.92* *0.3522*

*BCD* *7.03* *1* *7.03* *0.92* *0.3522*

*ABCD* *47.53* *1* *47.53* *6.21* *0.0241*

Residual 122.50 16 7.66

*Lack of Fit* *0.000* *0*

*Pure Error* *122.50* *16* *7.66*

Cor Total 1627.47 31

The Model F-value of 13.10 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, C, D, AB, AD, ABC, ABD, ABCD are significant model terms.

, and therefore, factors *A* and *D* and interactions *AB, ABC*, and *ABD* are significant at 1%. Factor *C* and interactions *AD* and *ABCD* are significant at 5%.

(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).

Model with hierarchy maintained:

Design Expert Output

**Final Equation in Terms of Coded Factors:**

yield =

+82.78

-4.53 \* A

-0.66 \* B

-1.34 \* C

+1.97 \* D

+2.03 \* A \* B

+0.34 \* A \* C

-1.09 \* A \* D

-0.28 \* B \* C

-0.094 \* B \* D

+0.84 \* C \* D

-2.59 \* A \* B \* C

+2.34 \* A \* B \* D

-0.47 \* A \* C \* D

-0.47 \* B \* C \* D

+1.22 \* A \* B \* C \* D

Model without hierarchy terms:

Design Expert Output

**Final Equation in Terms of Coded Factors:**

yield =

+82.78

-4.53 \* A

-1.34 \* C

+1.97 \* D

+2.03 \* A \* B

-1.09 \* A \* D

-2.59 \* A \* B \* C

+2.34 \* A \* B \* D

+1.22 \* A \* B \* C \* D

Confirmation runs might be run to see if the simpler model without hierarchy is satisfactory.

(d) Does the residual analysis appear satisfactory?

There appears to be one large residual both in the normal probability plot and in the plot of residuals versus predicted.



(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?



Run the process at *A* low *B* low, *C* low and *D* high.

**6.12.** 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 22 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.

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

Design Expert Output

**Response:** **Virus growth**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 691.46 3 230.49 45.12 < 0.0001 significant

*A* *9.38* *1* *9.38* *1.84* *0.1906*

*B* *590.04* *1* *590.04* *115.51* *< 0.0001*

*AB* *92.04* *1* *92.04* *18.02* *0.0004*

Residual 102.17 20 5.11

*Lack of Fit* *0.000* *0*

*Pure Error* *102.17* *20* *5.11*

Cor Total 793.63 23

The Model F-value of 45.12 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case B, AB are significant model terms.



Growth rate is affected by factor *B* (Time) and the *AB* interaction (Culture medium and Time). There is some very slight indication of inequality of variance shown by the small decreasing funnel shape in the plot of residuals versus predicted.



**6.13.** An industrial engineer employed by a beverage bottler is interested in the effects of two different types of 32-ounce bottles on the time to deliver 12-bottle cases of the product. The two bottle types are glass and plastic. Two workers are used to perform a task consisting of moving 40 cases of the product 50 feet on a standard type of hand truck and stacking the cases in a display. Four replicates of a 22 factorial design are performed, and the times observed are listed in the following table. Analyze the data and draw the appropriate conclusions. Analyze the residuals and comment on the model’s adequacy.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Worker | | | |
| Bottle Type | 1 | 1 | 2 | 2 |
| Glass | 5.12 | 4.89 | 6.65 | 6.24 |
|  | 4.98 | 5.00 | 5.49 | 5.55 |
|  |  |  |  |  |
| Plastic | 4.95 | 4.43 | 5.28 | 4.91 |
|  | 4.27 | 4.25 | 4.75 | 4.71 |

Design Expert Output

**Response:** **Times**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 4.86 3 1.62 13.04 0.0004 significant

*A* *2.02* *1* *2.02* *16.28* *0.0017*

*B* *2.54* *1* *2.54* *20.41* *0.0007*

*AB* *0.30* *1* *0.30* *2.41* *0.1463*

Residual 1.49 12 0.12

*Lack of Fit* *0.000* *0*

*Pure Error* *1.49* *12* *0.12*

Cor Total 6.35 15

The Model F-value of 13.04 implies the model is significant. There is only

a 0.04% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B are significant model terms.

There is some indication of non-constant variance in this experiment.





**6.14.** An article in the *AT&T Technical Journal* (March/April 1986, Vol. 65, pp. 39-50) describes the application of two-level factorial designs to integrated circuit manufacturing. A basic processing step is to grow an epitaxial layer on polished silicon wafers. The wafers mounted on a susceptor are positioned inside a bell jar, and chemical vapors are introduced. The susceptor is rotated and heat is applied until the epitaxial layer is thick enough. An experiment was run using two factors: arsenic flow rate (*A*) and deposition time (*B*). Four replicates were run, and the epitaxial layer thickness was measured (in mm). The data are shown below:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Replicate |  |  |  | Factor | Levels |
| *A* | *B* | I | II | III | IV |  | Low (–) | High (+) |
| – | – | 14.037 | 16.165 | 13.972 | 13.907 | *A* | 55% | 59% |
| + | – | 13.880 | 13.860 | 14.032 | 13.914 |  |  |  |
| – | + | 14.821 | 14.757 | 14.843 | 14.878 | *B* | Short | Long |
| + | + | 14.888 | 14.921 | 14.415 | 14.932 |  | (10 min) | (15 min) |

(a) Estimate the factor effects.

Design Expert Output

Term Effect SumSqr % Contribtn

Model Intercept

Error A -0.31725 0.40259 6.79865

Error B 0.586 1.37358 23.1961

Error AB 0.2815 0.316969 5.35274

Error Lack Of Fit 0 0

Error Pure Error 3.82848 64.6525

(b) Conduct an analysis of variance. Which factors are important?

From the analysis of variance shown below, no factors appear to be important. Factor *B* is only marginally interesting with an *F*-value of 4.31.

Design Expert Output

**Response:** **Thickness**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 2.09 3 0.70 2.19 0.1425 not significant

*A* *0.40* *1* *0.40* *1.26* *0.2833*

*B* *1.37* *1* *1.37* *4.31* *0.0602*

*AB* *0.32* *1* *0.32* *0.99* *0.3386*

Residual 3.83 12 0.32

*Lack of Fit* *0.000* *0*

*Pure Error* *3.83* *12* *0.32*

Cor Total 5.92 15

The "Model F-value" of 2.19 implies the model is not significant relative to the noise. There is a

14.25 % chance that a "Model F-value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case there are no significant model terms.

(c) Write down a regression equation that could be used to predict epitaxial layer thickness over the region of arsenic flow rate and deposition time used in this experiment.

Design Expert Output

**Final Equation in Terms of Coded Factors:**

Thickness =

+14.51

-0.16 \* A

+0.29 \* B

+0.14 \* A \* B

**Final Equation in Terms of Actual Factors:**

Thickness =

+37.62656

-0.43119 \* Flow Rate

-1.48735 \* Dep Time

+0.028150 \* Flow Rate \* Dep Time

(d) Analyze the residuals. Are there any residuals that should cause concern? Observation #2 falls outside the groupings in the normal probability plot and the plot of residual versus predicted.



(e) Discuss how you might deal with the potential outlier found in part (d).

One approach would be to replace the observation with the average of the observations from that experimental cell. Another approach would be to identify if there was a recording issue in the original data. The first analysis below replaces the data point with the average of the other three. The second analysis assumes that the reading was incorrectly recorded and should have been 14.165.

The analysis with the run associated with standard order 2 replaced with the average of the remaining three runs in the cell, 13.972, is shown below.

Design Expert Output

**Response:** **Thickness**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 2.97 3 0.99 53.57 < 0.0001 significant

*A 7.439E-003 1 7.439E-003 0.40 0.5375*

*B 2.96 1 2.96 160.29 < 0.0001*

*AB 2.176E-004 1 2.176E-004 0.012 0.9153*

### Pure Error 0.22 12 0.018

Cor Total 3.19 15

The Model F-value of 53.57 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case B are significant model terms.

**Final Equation in Terms of Coded Factors:**

Thickness =

+14.38

-0.022 \* A

+0.43 \* B

+3.688E-003 \* A \* B

**Final Equation in Terms of Actual Factors:**

Thickness =

+13.36650

-0.020000 \* Flow Rate

+0.12999 \* Dep Time

+7.37500E-004 \* Flow Rate \* Dep Time



A new outlier is present and should be investigated.

Analysis with the run associated with standard order 2 replaced with the value 14.165:

Design Expert Output

**Response:** **Thickness**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 2.82 3 0.94 45.18 < 0.0001 significant

*A 0.018 1 0.018 0.87 0.3693*

*B 2.80 1 2.80 134.47 < 0.0001*

*AB 3.969E-003 1 3.969E-003 0.19 0.6699*

*Pure Error 0.25 12 0.021*

Cor Total 3.07 15

The Model F-value of 45.18 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case B are significant model terms.

**Final Equation in Terms of Coded Factors:**

Thickness =

+14.39

-0.034 \* A

+0.42 \* B

+0.016 \* A \* B

**Final Equation in Terms of Actual Factors:**

Thickness =

+15.50156

-0.056188 \* Flow Rate

-0.012350 \* Dep Time

+3.15000E-003 \* Flow Rate \* Dep Time



Another outlier is present and should be investigated.

**6.15. Continuation of Problem 6.14.** Use the regression model in part (c) of Problem 6.14 to generate a response surface contour plot for epitaxial layer thickness. Suppose it is critically important to obtain layer thickness of 14.5 mm. What settings of arsenic flow rate and deposition time would you recommend?

Arsenic flow rate may be set at any of the experimental levels, while the deposition time should be set at 12.4 minutes.



**6.16. Continuation of Problem 6.15.** How would your answer to Problem 6.15 change if arsenic flow rate was more difficult to control in the process than the deposition time?

Running the process at a high level of Deposition Time there is no change in thickness as flow rate changes.

**6.17.** An experimenter has run a single replicate of a 24 design. The following effect estimates have been calculated:

|  |  |  |
| --- | --- | --- |
| *A* = 76.95 | *AB* = -51.32 | *ABC* = -2.82 |
| *B* = -67.52 | *AC* = 11.69 | *ABD* = -6.50 |
| *C* = -7.84 | *AD* = 9.78 | *ACD* = 10.20 |
| *D* = -18.73 | *BC* = 20.78 | *BCD* = -7.98 |
|  | *BD* = 14.74 | *ABCD* = -6.25 |
|  | *CD* = 1.27 |  |

(a) Construct a normal probability plot of these effects.

The plot from Minitab follows.



(b) Identify a tentative model, based on the plot of the effects in part (a).



**6.18.** The effect estimates from a 24 factorial design are as follows: *ABCD* = -1.5138, *ABC* = -1.2661, *ABD* = -0.9852, *ACD* = -0.7566, *BCD* = -0.4842, *CD* = -0.0795, *BD* = -0.0793, *AD* = 0.5988, *BC* = 0.9216, *AC* = 1.1616, *AB* = 1.3266, *D* = 4.6744, *C* = 5.1458, *B* = 8.2469, and *A* = 12.7151. Are you comfortable with the conclusions that all main effects are active?



The upper right 4 dots are the four main effects. Since they do not follow the rest of the data, the normal probability plot shows that they are active.

**6.19.** The effect estimates from a 24 factorial experiment are listed here. Are any of the effects significant?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *ABCD* = | -2.5251 |  | *AD* = | -1.6564 |
| *BCD* = | 4.4054 |  | *AC* = | 1.1109 |
| *ACD* = | -0.4932 |  | *AB* = | -10.5229 |
| *ABD* = | -5.0842 |  | *D* = | -6.0275 |
| *ABC* = | -5.7696 |  | *C* = | -8.2045 |
| *CD* = | 4.6707 |  | *B* = | -6.5304 |
| *BD* = | -4.6620 |  | *A* = | -0.7914 |
| *BC* = | -0.7982 |  |  |  |



No effects appear to be significant.

**6.20S.** Consider a variation of the bottle filling experiment from Example 5.3. Suppose that only two levels of carbonation are used so that the experiment is a 23 factorial design with two replicates. The data are shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Coded Factors | | | Fill Height Deviation | |
| Run | *A* | *B* | *C* | Replicate 1 | Replicate 2 |
| 1 | – | – | – | -3 | -1 |
| 2 | + | – | – | 0 | 1 |
| 3 | – | + | – | -1 | 0 |
| 4 | + | + | – | 2 | 3 |
| 5 | – | – | + | -1 | 0 |
| 6 | + | – | + | 2 | 1 |
| 7 | – | + | + | 1 | 1 |
| 8 | + | + | + | 6 | 5 |

|  |  |  |
| --- | --- | --- |
|  | Factor Levels | |
|  | Low (–1) | High (+1) |
| *A* (%) | 10 | 12 |
| *B* (psi) | 25 | 30 |
| *C* (b/m) | 200 | 250 |

(a) Analyze the data from this experiment. Which factors significantly affect fill height deviation?

The half normal probability plot of effects shown below identifies the factors *A*, *B*, and *C* as being significant and the *AB* interaction as being marginally significant. The analysis of variance in the Design Expert output below confirms that factors *A*, *B*, and *C* are significant and the *AB* interaction is marginally significant.



Design Expert Output

**Response:** **Fill Deviation**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 70.75 4 17.69 26.84 < 0.0001 significant

*A 36.00 1 36.00 54.62 < 0.0001*

*B 20.25 1 20.25 30.72 0.0002*

*C 12.25 1 12.25 18.59 0.0012*

*AB 2.25 1 2.25 3.41 0.0917*

Residual 7.25 11 0.66

### Lack of Fit 2.25 3 0.75 1.20 0.3700 not significant

### Pure Error 5.00 8 0.63

Cor Total 78.00 15

The Model F-value of 26.84 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, C are significant model terms.

Std. Dev. 0.81 R-Squared 0.9071

Mean 1.00 Adj R-Squared 0.8733

C.V. 81.18 Pred R-Squared 0.8033

PRESS 15.34 Adeq Precision 15.424

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

The residual plots below do not identify any violations to the assumptions.







(c) Obtain a model for predicting fill height deviation in terms of the important process variables. Use this model to construct contour plots to assist in interpreting the results of the experiment.

The model in both coded and actual factors are shown below.

Design Expert Output

Final Equation in Terms of Coded Factors

Fill Deviation =

+1.00

+1.50 \* A

+1.13 \* B

+0.88 \* C

+0.38 \* A \* B

Final Equation in Terms of Actual Factors

Fill Deviation =

+9.62500

-2.62500 \* Carbonation

-1.20000 \* Pressure

+0.035000 \* Speed

+0.15000 \* Carbonation \* Pressure

The following contour plots identify the fill deviation with respect to carbonation and pressure. The plot on the left sets the speed at 200 b/m while the plot on the right sets the speed at 250 b/m. Assuming a faster bottle speed is better, settings in pressure and carbonation that produce a fill deviation near zero can be found in the lower left hand corner of the contour plot on the right.



Speed set at 200 b/m Speed set at 250 b/m

(d) In part (a), you probably noticed that there was an interaction term that was borderline significant. If you did not include the interaction term in your model, include it now and repeat the analysis. What difference did this make? If you elected to include the interaction term in part (a), remove it and repeat the analysis. What difference does this make?

The following analysis of variance, residual plots, and contour plots represent the model without the interaction. As in the original analysis, the residual plots do not identify any concerns with the assumptions. The contour plots did not change significantly either. The interaction effect is small relative to the main effects.

Design Expert Output

**Response:** **Fill Deviation**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 68.50 3 22.83 28.84 < 0.0001 significant

*A 36.00 1 36.00 45.47 < 0.0001*

*B 20.25 1 20.25 25.58 0.0003*

*C 12.25 1 12.25 15.47 0.0020*

Residual 9.50 12 0.79

*Lack of Fit 4.50 4 1.13 1.80 0.2221 not significant*

*Pure Error 5.00 8 0.63*

Cor Total 78.00 15

The Model F-value of 28.84 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, C are significant model terms.

Std. Dev. 0.89 R-Squared 0.8782

Mean 1.00 Adj R-Squared 0.8478

C.V. 88.98 Pred R-Squared 0.7835

PRESS 16.89 Adeq Precision 15.735

Final Equation in Terms of Coded Factors

Fill Deviation =

+1.00

+1.50 \* A

+1.13 \* B

+0.88 \* C

Final Equation in Terms of Actual Factors

Fill Deviation =

-35.75000

+1.50000 \* Carbonation

+0.45000 \* Pressure

+0.035000 \* Speed









Speed set at 200 b/m Speed set at 250 b/m

**6.21S.** I am always interested in improving my golf scores. Since a typical golfer uses the putter for about 35-45% of his or her strokes, it seems logical that in improving one’s putting score is a logical and perhaps simple way to improve a golf score (“The man who can putt is a match for any man.” – Willie Parks, 1864-1925, two-time winner of the British Open). An experiment was conducted to study the effects of four factors on putting accuracy. The design factors are length of putt, type of putter, breaking putt vs. straight putt, and level versus downhill putt. The response variable is distance from the ball to the center of the cup after the ball comes to rest. One golfer performs the experiment, a 24 factorial design with seven replicates was used, and all putts were made in random order. The results are as follows.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Design Factors | | | | | Distance from cup (replicates) | | | | | | |
| Length of putt (ft) | Type of putter | Break  of putt | Slope  of putt | 1 | | 2 | 3 | 4 | 5 | 6 | 7 |
| 10 | Mallet | Straight | Level | 10.0 | | 18.0 | 14.0 | 12.5 | 19.0 | 16.0 | 18.5 |
| 30 | Mallet | Straight | Level | 0.0 | | 16.5 | 4.5 | 17.5 | 20.5 | 17.5 | 33.0 |
| 10 | Cavity-back | Straight | Level | 4.0 | | 6.0 | 1.0 | 14.5 | 12.0 | 14.0 | 5.0 |
| 30 | Cavity-back | Straight | Level | 0.0 | | 10.0 | 34.0 | 11.0 | 25.5 | 21.5 | 0.0 |
| 10 | Mallet | Breaking | Level | 0.0 | | 0.0 | 18.5 | 19.5 | 16.0 | 15.0 | 11.0 |
| 30 | Mallet | Breaking | Level | 5.0 | | 20.5 | 18.0 | 20.0 | 29.5 | 19.0 | 10.0 |
| 10 | Cavity-back | Breaking | Level | 6.5 | | 18.5 | 7.5 | 6.0 | 0.0 | 10.0 | 0.0 |
| 30 | Cavity-back | Breaking | Level | 16.5 | | 4.5 | 0.0 | 23.5 | 8.0 | 8.0 | 8.0 |
| 10 | Mallet | Straight | Downhill | 4.5 | | 18.0 | 14.5 | 10.0 | 0.0 | 17.5 | 6.0 |
| 30 | Mallet | Straight | Downhill | 19.5 | | 18.0 | 16.0 | 5.5 | 10.0 | 7.0 | 36.0 |
| 10 | Cavity-back | Straight | Downhill | 15.0 | | 16.0 | 8.5 | 0.0 | 0.5 | 9.0 | 3.0 |
| 30 | Cavity-back | Straight | Downhill | 41.5 | | 39.0 | 6.5 | 3.5 | 7.0 | 8.5 | 36.0 |
| 10 | Mallet | Breaking | Downhill | 8.0 | | 4.5 | 6.5 | 10.0 | 13.0 | 41.0 | 14.0 |
| 30 | Mallet | Breaking | Downhill | 21.5 | | 10.5 | 6.5 | 0.0 | 15.5 | 24.0 | 16.0 |
| 10 | Cavity-back | Breaking | Downhill | 0.0 | | 0.0 | 0.0 | 4.5 | 1.0 | 4.0 | 6.5 |
| 30 | Cavity-back | Breaking | Downhill | 18.0 | | 5.0 | 7.0 | 10.0 | 32.5 | 18.5 | 8.0 |

(a) Analyze the data from this experiment. Which factors significantly affect putting performance?

The half normal probability plot of effects identifies only factors *A* and *B*, length of putt and type of putter, as having a potentially significant effect on putting performance. The analysis of variance with only these significant factors is presented as well and confirms significance.



Design Expert Output with Only Factors *A* and *B*

**Response:** **Distance from cup**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Terms added sequentially (first to last)]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 1305.29 2 652.65 7.69 0.0008 significant

*A 917.15 1 917.15 10.81 0.0014*

*B 388.15 1 388.15 4.57 0.0347*

Residual 9248.94 109 84.85

*Lack of Fit 933.15 13 71.78 0.83 0.6290 not significant*

*Pure Error 8315.79 96 86.62*

Cor Total 10554.23 111

The Model F-value of 7.69 implies the model is significant. There is only

a 0.08% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, are significant model terms.

Std. Dev. 9.21 R-Squared 0.1237

Mean 12.30 Adj R-Squared 0.1076

C.V. 74.90 Pred R-Squared 0.0748

PRESS 9765.06 Adeq Precision 6.266

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

The residual plots for the model containing only the significant factors *A* and *B* are shown below. The normality assumption appears to be violated. Also, as a golfer might expect, there is a slight inequality of variance with regards to the length of putt. A square root transformation is applied which corrects the violations. The analysis of variance and corrected residual plots are also presented. Finally, an effects plot identifies a 10 foot putt and the cavity-back putter reduce the mean distance from the cup.







Design Expert Output with Only Factors *A* and *B* and a Square Rot Transformation

**Response: Distance from cup Transform: Square root Constant: 0**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Terms added sequentially (first to last)]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 37.26 2 18.63 7.85 0.0007 significant

*A 21.61 1 21.61 9.11 0.0032*

*B 15.64 1 15.64 6.59 0.0116*

Residual 258.63 109 2.37

*Lack of Fit 30.19 13 2.32 0.98 0.4807 not significant*

*Pure Error 228.45 96 2.38*

Cor Total 295.89 111

The Model F-value of 7.85 implies the model is significant. There is only

a 0.07% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, are significant model terms.

Std. Dev. 1.54 R-Squared 0.1259

Mean 3.11 Adj R-Squared 0.1099

C.V. 49.57 Pred R-Squared 0.0771

PRESS 273.06 Adeq Precision 6.450









**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* = offer price ($19.95, $24.95). The mailings are made to two groups of 8,000 randomly selected customers, with 1,000 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 is shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Coded Factors | | | Number of Orders | |
| Run | *A* | *B* | *C* | Replicate 1 | Replicate 2 |
| 1 | – | – | – | 50 | 54 |
| 2 | + | – | – | 44 | 42 |
| 3 | – | + | – | 46 | 48 |
| 4 | + | + | – | 42 | 43 |
| 5 | – | – | + | 49 | 46 |
| 6 | + | – | + | 48 | 45 |
| 7 | – | + | + | 47 | 48 |
| 8 | + | + | + | 56 | 54 |

|  |  |  |
| --- | --- | --- |
|  | Factor Levels | |
|  | Low (-1) | High (+1) |
| *A* (class) | 3rd | 1st |
| *B* (type) | BW | Color |
| *C* ($) | $19.95 | $24.95 |

(a) Analyze the data from this experiment. Which factors significantly affect the customer response rate?

The half normal probability plot of effects identifies the two factor interactions, *AB*, *AC*, *BC*, and factors *A* and *C* as significant. Factor *B* is not significant; however, remains in the model to satisfy the hierarchal principle. The analysis of variance confirms the significance of two factor interactions and factor *C*. Factor *A* is marginally significant.



Design Expert Output

**Response:** **Number of orders**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 241.75 6 40.29 12.95 0.0006 significant

A 12.25 1 12.25 3.94 0.0785

B 2.25 1 2.25 0.72 0.4171

C 36.00 1 36.00 11.57 0.0079

AB 42.25 1 42.25 13.58 0.0050

AC 100.00 1 100.00 32.14 0.0003

BC 49.00 1 49.00 15.75 0.0033

Residual 28.00 9 3.11

Lack of Fit 4.00 1 4.00 1.33 0.2815 not significant

Pure Error 24.00 8 3.00

Cor Total 269.75 15

The Model F-value of 12.95 implies the model is significant. There is only

a 0.06% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case C, AB, AC, BC are significant model terms.

Std. Dev. 1.76 R-Squared 0.8962

Mean 47.63 Adj R-Squared 0.8270

C.V. 3.70 Pred R-Squared 0.6719

PRESS 88.49 Adeq Precision 10.286

**Coefficient Standard 95% CI 95% CI**

**Factor Estimate DF Error Low High VIF**

Intercept 47.63 1 0.44 46.63 48.62

A-Class -0.88 1 0.44 -1.87 0.12 1.00

B-Type 0.37 1 0.44 -0.62 1.37 1.00

C-Price 1.50 1 0.44 0.50 2.50 1.00

AB 1.63 1 0.44 0.63 2.62 1.00

AC 2.50 1 0.44 1.50 3.50 1.00

BC 1.75 1 0.44 0.75 2.75 1.00

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

The residual plots below do not identify model inadequacy.







(c) What would you recommend to the company?

Based on the interaction plots below, we recommend 3rd class mail, black-and-white brochures, and an offered price of $19.95 would achieve the greatest number of orders. If the offered price must be $24.95, then the 1st class mail with color brochures is recommended.





**6.23.** Consider the single replicate of the 24 design in Example 6.2. Suppose we had arbitrarily decided to analyze the data assuming that all three- and four-factor interactions were negligible. Conduct this analysis and compare your results with those obtained in the example. Do you think that it is a good idea to arbitrarily assume interactions to be negligible even if they are relatively high-order ones?

Design Expert Output

**Response:** **Filtration Rate** **in A/min**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 5603.13 10 560.31 21.92 0.0016 significant

*A 1870.56 1 1870.56 73.18 0.0004*

*B 39.06 1 39.06 1.53 0.2713*

*C 390.06 1 390.06 15.26 0.0113*

*D 855.56 1 855.56 33.47 0.0022*

*AB 0.063 1 0.063 2.445E-003 0.9625*

*AC 1314.06 1 1314.06 51.41 0.0008*

*AD 1105.56 1 1105.56 43.25 0.0012*

*BC 22.56 1 22.56 0.88 0.3906*

*BD 0.56 1 0.56 0.022 0.8879*

*CD 5.06 1 5.06 0.20 0.6749*

Residual 127.81 5 25.56

Cor Total 5730.94 15

The Model F-value of 21.92 implies the model is significant. There is only

a 0.16% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, C, D, AC, AD are significant model terms.

This analysis of variance identifies the same effects as the normal probability plot of effects approach used in Example 6.2. In general, it is not a good idea to arbitrarily pool interactions. Use the normal probability plot of effect estimates as a guide in the choice of which effects to tentatively include in the model.

**6.24S.** An experiment was run in a semiconductor fabrication plant in an effort to increase yield. Five factors, each at two levels, were studied. The factors (and levels) were *A* = aperture setting (small, large), *B* = exposure time (20% below nominal, 20% above nominal), *C* = development time (30 s, 45 s), *D* = mask dimension (small, large), and *E* = etch time (14.5 min, 15.5 min). The unreplicated 25 design shown below was run.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| (1) = | 7 | *d =* | 8 | *e =* | 8 | *de =* | 6 |
| *a =* | 9 | *ad =* | 10 | *ae =* | 12 | *ade =* | 10 |
| *b =* | 34 | *bd =* | 32 | *be =* | 35 | *bde =* | 30 |
| *ab =* | 55 | *abd =* | 50 | *abe =* | 52 | *abde =* | 53 |
| *c =* | 16 | *cd =* | 18 | *ce =* | 15 | *cde =* | 15 |
| *ac =* | 20 | *acd =* | 21 | *ace =* | 22 | *acde =* | 20 |
| *bc =* | 40 | *bcd =* | 44 | *bce =* | 45 | *bcde =* | 41 |
| *abc =* | 60 | *abcd =* | 61 | *abce =* | 65 | *abcde =* | 63 |

(a) Construct a normal probability plot of the effect estimates. Which effects appear to be large?

From the normal probability plot of effects shown below, effects *A*, *B*, *C*, and the *AB* interaction appear to be large.



(b) Conduct an analysis of variance to confirm your findings for part (a).

Design Expert Output

**Response:** **Yield**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 11585.13 4 2896.28 991.83 < 0.0001 significant

*A* *1116.28* *1* *1116.28* *382.27* *< 0.0001*

*B* *9214.03* *1* *9214.03* *3155.34* *< 0.0001*

*C* *750.78* *1* *750.78* *257.10* *< 0.0001*

*AB* *504.03* *1* *504.03* *172.61* *< 0.0001*

Residual 78.84 27 2.92

Cor Total 11663.97 31

The Model F-value of 991.83 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, C, AB are significant model terms.

(c) Write down the regression model relating yield to the significant process variables.

Design Expert Output

**Final Equation in Terms of Actual Factors:**

Aperture small

Yield =

+0.40625

+0.65000 \* Exposure Time

+0.64583 \* Develop Time

Aperture large

Yield =

+12.21875

+1.04688 \* Exposure Time

+0.64583 \* Develop Time

(d) Plot the residuals on normal probability paper. Is the plot satisfactory?



There is nothing unusual about this plot.

(e) Plot the residuals versus the predicted yields and versus each of the five factors. Comment on the plots.







The plot of residual versus exposure time shows some very slight inequality of variance. There is no strong evidence of a potential problem.

(f) Interpret any significant interactions.



Factor *A* does not have as large an effect when *B* is at its low level as it does when *B* is at its high level.

(g) What are your recommendations regarding process operating conditions?

To achieve the highest yield, run *B* at the high level, *A* at the high level, and *C* at the high level.

(h) Project the 25 design in this problem into a 2*k* design in the important factors. Sketch the design and show the average and range of yields at each run. Does this sketch aid in interpreting the results of this experiment?



This cube plot aids in interpretation. The strong *AB* interaction and the large positive effect of *C* are clearly evident.

**6.25. Continuation of Problem 6.24S.** Suppose that the experimenter had run four runs at the center points in addition to the 32 trials in the original experiment. The yields obtained at the center point runs were 68, 74, 76, and 70.

(a) Reanalyze the experiment, including a test for pure quadratic curvature.

Because aperture and mask dimension are not continuous variables, the four center points were split amongst these two factors as follows.

|  |  |  |
| --- | --- | --- |
| Aperture | Mask Dimension | Yield |
| Small | Small | 68 |
| Large | Small | 74 |
| Small | Large | 76 |
| Large | Large | 70 |

The sum of squares for the curvature can be estimated with the following equation and is confirmed with the analysis of variance shown in the Design Expert output.



Design Expert Output

**Response:** **Yield**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 11461.09 4 2865.27 353.92 < 0.0001 significant

*A* *992.25* *1* *992.25* *122.56* *< 0.0001*

*B* *9214.03* *1* *9214.03* *1138.12* *< 0.0001*

*C* *750.78* *1* *750.78* *92.74* *< 0.0001*

*AB* *504.03* *1* *504.03* *62.26* *< 0.0001*

Curvature 6114.34 1 6114.34 755.24 < 0.0001 significant

Residual 242.88 30 8.10

Cor Total 17818.31 35

The Model F-value of 353.92 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, C, AB are significant model terms.

(b) Discuss what your next step would be.

Add axial points for factors *B* and *C* along with four more center points to fit a second-order model and satisfy blocking concerns.

**6.26S.** In a process development study on yield, four factors were studied, each at two levels: time (*A*), concentration (*B*), pressure (*C*), and temperature (*D*). A single replicate of a 24 design was run, and the resulting data are shown in the following table:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Actual |  |  |  |  |  |  |  |  |
| Run | Run |  |  |  |  | Yield |  | Factor | Levels |
| Number | Order | *A* | *B* | *C* | *D* | (lbs) |  | Low (-) | High (+) |
| 1 | 5 | – | – | – | – | 12 | *A* (h) | 2.5 | 3.0 |
| 2 | 9 | + | – | – | – | 18 | *B* (%) | 14 | 18 |
| 3 | 8 | – | + | – | – | 13 | *C* (psi) | 60 | 80 |
| 4 | 13 | + | + | – | – | 16 | *D* (°C) | 225 | 250 |
| 5 | 3 | – | – | + | – | 17 |  |  |  |
| 6 | 7 | + | – | + | – | 15 |  |  |  |
| 7 | 14 | – | + | + | – | 20 |  |  |  |
| 8 | 1 | + | + | + | – | 15 |  |  |  |
| 9 | 6 | – | – | – | + | 10 |  |  |  |
| 10 | 11 | + | – | – | + | 25 |  |  |  |
| 11 | 2 | – | + | – | + | 13 |  |  |  |
| 12 | 15 | + | + | – | + | 24 |  |  |  |
| 13 | 4 | – | – | + | + | 19 |  |  |  |
| 14 | 16 | + | – | + | + | 21 |  |  |  |
| 15 | 10 | – | + | + | + | 17 |  |  |  |
| 16 | 12 | + | + | + | + | 23 |  |  |  |

(a) Construct a normal probability plot of the effect estimates. Which factors appear to have large effects?



*A, C, D* and the *AC* and *AD* interactions appear to have large effects.

(b) Conduct an analysis of variance using the normal probability plot in part (a) for guidance in forming an error term. What are your conclusions?

Design Expert Output

**Response:** **Yield**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 275.50 5 55.10 33.91 < 0.0001 significant

*A* *81.00* *1* *81.00* *49.85* *< 0.0001*

*C* *16.00* *1* *16.00* *9.85* *0.0105*

*D* *42.25* *1* *42.25* *26.00* *0.0005*

*AC* *72.25* *1* *72.25* *44.46* *< 0.0001*

*AD* *64.00* *1* *64.00* *39.38* *< 0.0001*

Residual 16.25 10 1.62

Cor Total 291.75 15

The Model F-value of 33.91 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, C, D, AC, AD are significant model terms.

(c) Write down a regression model relating yield to the important process variables.

Design Expert Output

**Final Equation in Terms of Coded Factors:**

Yield =

+17.38

+2.25 \*A

+1.00 \*C

+1.63 \*D

-2.13 \*A\*C

+2.00 \*A\*D

**Final Equation in Terms of Actual Factors:**

Yield =

+209.12500

-83.50000 \* Time

+2.43750 \* Pressure

-1.63000 \* Temperature

-0.85000 \* Time \* Pressure

+0.64000 \* Time \* Temperature

(d) Analyze the residuals from this experiment. Does your analysis indicate any potential problems?





There is nothing unusual about the residual plots.

(e) Can this design be collapsed into a 23 design with two replicates? If so, sketch the design with the average and range of yield shown at each point in the cube. Interpret the results.



**6.27S. Continuation of Problem 6.26S.** Use the regression model in part (c) of Problem 6.26S to generate a response surface contour plot of yield. Discuss the practical purpose of this response surface plot.

The response surface contour plot shows the adjustments in the process variables that lead to an increasing or decreasing response. It also displays the curvature of the response in the design region, possibly indicating where robust operating conditions can be found. Two response surface contour plots for this process are shown below. These were formed from the model written in terms of the original design variables.



**6.28S.** **The scrumptious brownie experiment.** The author is an engineer by training and a firm believer in learning by doing. I have taught experimental design for many years to a wide variety of audiences and have always assigned the planning, conduct, and analysis of an actual experiment to the class participants. The participants seem to enjoy this practical experience and always learn a great deal from it. This problem uses the results of an experiment performed by Gretchen Krueger at Arizona State University.

There are many different ways to bake brownies. The purpose of this experiment was to determine how the pan material, the brand of brownie mix, and the stirring method affect the scrumptiousness of brownies. The factor levels were

|  |  |  |
| --- | --- | --- |
| Factor | Low (–) | High (+) |
| *A* = pan material | Glass | Aluminum |
| *B* = stirring method | Spoon | Mixer |
| *C* = brand of mix | Expensive | Cheap |

The response variable was scrumptiousness, a subjective measure derived from a questionnaire given to the subjects who sampled each batch of brownies. (The questionnaire dealt with such issues as taste, appearance, consistency, aroma, and so forth.) An eight-person test panel sampled each batch and filled out the questionnaire. The design matrix and the response data are shown below:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Brownie |  |  |  |  | Test | Panel | Results |  |  |  |  |
| Batch | *A* | *B* | *C* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | – | – | – | 11 | 9 | 10 | 10 | 11 | 10 | 8 | 9 |
| 2 | + | – | – | 15 | 10 | 16 | 14 | 12 | 9 | 6 | 15 |
| 3 | – | + | – | 9 | 12 | 11 | 11 | 11 | 11 | 11 | 12 |
| 4 | + | + | – | 16 | 17 | 15 | 12 | 13 | 13 | 11 | 11 |
| 5 | – | – | + | 10 | 11 | 15 | 8 | 6 | 8 | 9 | 14 |
| 6 | + | – | + | 12 | 13 | 14 | 13 | 9 | 13 | 14 | 9 |
| 7 | – | + | + | 10 | 12 | 13 | 10 | 7 | 7 | 17 | 13 |
| 8 | + | + | + | 15 | 12 | 15 | 13 | 12 | 12 | 9 | 14 |

(a) Analyze the data from this experiment as if there were eight replicates of a 23 design. Comment on the results.

Design Expert Output

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 93.25 7 13.32 2.20 0.0475 significant

*A* *72.25* *1* *72.25* *11.95* *0.0010*

*B* *18.06* *1* *18.06* *2.99* *0.0894*

*C* *0.063* *1* *0.063* *0.010* *0.9194*

*AB* *0.062* *1* *0.062* *0.010* *0.9194*

*AC* *1.56* *1* *1.56* *0.26* *0.6132*

*BC* *1.00* *1* *1.00* *0.17* *0.6858*

*ABC* *0.25* *1* *0.25* *0.041* *0.8396*

Residual 338.50 56 6.04

*Lack of Fit* *0.000* *0*

*Pure Error* *338.50* *56* *6.04*

Cor Total 431.75 63

The Model F-value of 2.20 implies the model is significant. There is only

a 4.75% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A are significant model terms.

In this analysis, *A*, the pan material and *B*, the stirring method, appear to be significant. There are 56 degrees of freedom for the error, yet only eight batches of brownies were cooked, one for each recipe.

(b) Is the analysis in part (a) the correct approach? There are only eight batches; do we really have eight replicates of a 23 factorial design?

The different rankings by the taste-test panel are not replicates, but repeat observations by different testers on the same batch of brownies. It is not a good idea to use the analysis in part (a) because the estimate of error may not reflect the batch-to-batch variation.

(c) Analyze the average and standard deviation of the scrumptiousness ratings. Comment on the results. Is this analysis more appropriate than the one in part (a)? Why or why not?



Design Expert Output

**Response:** **Average**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 11.28 2 5.64 76.13 0.0002 significant

*A* *9.03* *1* *9.03* *121.93* *0.0001*

*B* *2.25* *1* *2.25* *30.34* *0.0027*

Residual 0.37 5 0.074

Cor Total 11.65 7

The Model F-value of 76.13 implies the model is significant. There is only

a 0.02% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B are significant model terms.

Design Expert Output

**Response:** **Stdev**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 6.05 3 2.02 9.77 0.0259 significant

*A* *0.24* *1* *0.24* *1.15* *0.3432*

*C* *0.91* *1* *0.91* *4.42* *0.1034*

*AC* *4.90* *1* *4.90* *23.75* *0.0082*

Residual 0.82 4 0.21

Cor Total 6.87 7

The Model F-value of 9.77 implies the model is significant. There is only

a 2.59% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case AC are significant model terms.

Variables *A* and *B* affect the mean rank of the brownies. Note that the *AC* interaction affects the standard deviation of the ranks. This is an indication that both factors *A* and *C* have some effect on the variability in the ranks. It may also indicate that there is some inconsistency in the taste test panel members. For the analysis of both the average of the ranks and the standard deviation of the ranks, the mean square error is now determined by pooling apparently unimportant effects. This is a more accurate estimate of error than obtained assuming that all observations were replicates.

**6.29S.** An experiment was conducted on a chemical process that produces a polymer. The four factors studied were temperature (*A*), catalyst concentration (*B*), time (*C*), and pressure (*D*). Two responses, molecular weight and viscosity, were observed. The design matrix and response data are shown below:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Actual |  |  |  |  |  |  |  |  |  |
| Run | Run |  |  |  |  | Molecular |  |  | Factor | Levels |
| Number | Order | *A* | *B* | *C* | *D* | Weight | Viscosity |  | Low (–) | High (+) |
| 1 | 18 | – | – | – | – | 2400 | 1400 | *A* (°C) | 100 | 120 |
| 2 | 9 | + | – | – | – | 2410 | 1500 | *B* (%) | 4 | 8 |
| 3 | 13 | – | + | – | – | 2315 | 1520 | *C* (min) | 20 | 30 |
| 4 | 8 | + | + | – | – | 2510 | 1630 | *D* (psi) | 60 | 75 |
| 5 | 3 | – | – | + | – | 2615 | 1380 |  |  |  |
| 6 | 11 | + | – | + | – | 2625 | 1525 |  |  |  |
| 7 | 14 | – | + | + | – | 2400 | 1500 |  |  |  |
| 8 | 17 | + | + | + | – | 2750 | 1620 |  |  |  |
| 9 | 6 | – | – | – | + | 2400 | 1400 |  |  |  |
| 10 | 7 | + | – | – | + | 2390 | 1525 |  |  |  |
| 11 | 2 | – | + | – | + | 2300 | 1500 |  |  |  |
| 12 | 10 | + | + | – | + | 2520 | 1500 |  |  |  |
| 13 | 4 | – | – | + | + | 2625 | 1420 |  |  |  |
| 14 | 19 | + | – | + | + | 2630 | 1490 |  |  |  |
| 15 | 15 | – | + | + | + | 2500 | 1500 |  |  |  |
| 16 | 20 | + | + | + | + | 2710 | 1600 |  |  |  |
| 17 | 1 | 0 | 0 | 0 | 0 | 2515 | 1500 |  |  |  |
| 18 | 5 | 0 | 0 | 0 | 0 | 2500 | 1460 |  |  |  |
| 19 | 16 | 0 | 0 | 0 | 0 | 2400 | 1525 |  |  |  |
| 20 | 12 | 0 | 0 | 0 | 0 | 2475 | 1500 |  |  |  |

(a) Consider only the molecular weight response. Plot the effect estimates on a normal probability scale. What effects appear important?



*A, C* and the *AB* interaction appear to be important.

(b) Use an analysis of variance to confirm the results from part (a). Is there an indication of curvature? *A,C* and the *AB* interaction are significant. While the main effect of *B* is not significant, it could be included to preserve hierarchy in the model. There is no indication of quadratic curvature.

Design Expert Output

**Response:** **Molecular Wt**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 2.809E+005 3 93620.83 73.00 < 0.0001 significant

*A* *61256.25* *1* *61256.25* *47.76* *< 0.0001*

*C* *1.620E+005* *1* *1.620E+005* *126.32* *< 0.0001*

*AB* *57600.00* *1* *57600.00* *44.91* *< 0.0001*

Curvature 3645.00 1 3645.00 2.84 0.1125 not significant

Residual 19237.50 15 1282.50

*Lack of Fit* *11412.50* *12* *951.04* *0.36* *0.9106* *not significant*

*Pure Error* *7825.00* *3* *2608.33*

Cor Total 3.037E+005 19

The Model F-value of 73.00 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, C, AB are significant model terms.

(c) Write down a regression model to predict molecular weight as a function of the important variables.

Design Expert Output

**Final Equation in Terms of Coded Factors:**

Molecular Wt =

+2506.25

+61.87 \* A

+100.63 \* C

+60.00 \* A \* B

(d) Analyze the residuals and comment on model adequacy.



There are two residuals that appear to be large and should be investigated.

(e) Repeat parts (a) – (d) using the viscosity response.



Design Expert Output

**Response:** **Viscosity**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 70362.50 2 35181.25 35.97 < 0.0001 significant

*A* *37056.25* *1* *37056.25* *37.88* *< 0.0001*

*B* *33306.25* *1* *33306.25* *34.05* *< 0.0001*

Curvature 61.25 1 61.25 0.063 0.8056 not significant

Residual 15650.00 16 978.13

*Lack of Fit* *13481.25* *13* *1037.02* *1.43* *0.4298* *not significant*

*Pure Error* *2168.75* *3* *722.92*

Cor Total 86073.75 19

The Model F-value of 35.97 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B are significant model terms.

**Final Equation in Terms of Coded Factors:**

Viscosity =

+1500.62

+48.13 \* A

+45.63 \* B



There is one large residual that should be investigated.

**6.30S. Continuation of Problem 6.29S.** Use the regression models for molecular weight and viscosity to answer the following questions.

(a) Construct a response surface contour plot for molecular weight. In what direction would you adjust the process variables to increase molecular weight? Increase temperature, catalyst and time.



(b) Construct a response surface contour plot for viscosity. In what direction would you adjust the process variables to decrease viscosity?



Decrease temperature and catalyst.

(c) What operating conditions would you recommend if it was necessary to produce a product with a molecular weight between 2400 and 2500, and the lowest possible viscosity?



Set the temperature between 100 and 105, the catalyst between 4 and 5%, and the time at 24.5 minutes. The pressure was not significant and can be set at conditions that may improve other results of the process such as cost.

**6.31.** Consider the single replicate of the 24 design in Example 6.2. Suppose that we ran five points at the center (0, 0, 0, 0) and observed the following responses: 93, 95, 91, 89, and 96. Test for curvature in this experiment. Interpret the results.

Design Expert Output

**Response:** **Filtration Rate**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares – Type III]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 5535.81 5 1107.16 68.01 < 0.0001 significant

*A* *1870.56* *1* *1870.56* *114.90* *< 0.0001*

*C* *390.06* *1* *390.06* *23.96* *0.0002*

*D* *855.56* *1* *855.56* *52.55* *< 0.0001*

*AC* *1314.06* *1* *1314.06* *80.71* *< 0.0001*

*AD* *1105.56* *1* *1105.56* *67.91* *< 0.0001*

Curvature 1969.50 1 1969.50 120.97 < 0.0001 significant

Residual 227.92 14 16.28

*Lack of Fit* *195.12* *10* *19.51* *2.38* *0.2093* *not significant*

*Pure Error* *32.80* *4* *8.20*

Cor Total 7733.24 20

The Model F-value of 68.01 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, C, D, AC, AD are significant model terms.

With an *F*-value of 120.97 and a corresponding *P*-value of < 0.0001, curvature is significant.

**6.32.** An engineer has performed an experiment to study the effect of four factors on the surface roughness of a machined part. The factors (and their levels) are *A* = tool angle (12 degrees, 15 degrees), *B* = cutting fluid viscosity (300, 400), *C* = feed rate (10 in/min, 15 in/min), and *D* = cutting fluid cooler used (no, yes). The data from this experiment (with the factors coded to the usual -1, +1 levels) are shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run | *A* | *B* | *C* | *D* | Surface Roughness |
| 1 | – | – | – | – | 0.00340 |
| 2 | + | – | – | – | 0.00362 |
| 3 | – | + | – | – | 0.00301 |
| 4 | + | + | – | – | 0.00182 |
| 5 | – | – | + | – | 0.00280 |
| 6 | + | – | + | – | 0.00290 |
| 7 | – | + | + | – | 0.00252 |
| 8 | + | + | + | – | 0.00160 |
| 9 | – | – | – | + | 0.00336 |
| 10 | + | – | – | + | 0.00344 |
| 11 | – | + | – | + | 0.00308 |
| 12 | + | + | – | + | 0.00184 |
| 13 | – | – | + | + | 0.00269 |
| 14 | + | – | + | + | 0.00284 |
| 15 | – | + | + | + | 0.00253 |
| 16 | + | + | + | + | 0.00163 |

(a) Estimate the factor effects. Plot the effect estimates on a normal probability plot and select a tentative model.



(b) Fit the model identified in part (a) and analyze the residuals. Is there any indication of model inadequacy?

Design Expert Output

**Response:** **Surface Roughness**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 6.406E-006 4 1.601E-006 114.97 < 0.0001 significant

*A* *8.556E-007* *1* *8.556E-007* *61.43* *< 0.0001*

*B* *3.080E-006* *1* *3.080E-006* *221.11* *< 0.0001*

*C* *1.030E-006* *1* *1.030E-006* *73.96* *< 0.0001*

*AB* *1.440E-006* *1* *1.440E-006* *103.38* *< 0.0001*

Residual 1.532E-007 11 1.393E-008

Cor Total 6.559E-006 15

The Model F-value of 114.97 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, C, AB are significant model terms.





The plot of residuals versus predicted shows a slight “u-shaped” appearance in the residuals, and the plot of residuals versus tool angle shows an outward-opening funnel.

(c) Repeat the analysis from parts (a) and (b) using 1/*y* as the response variable. Is there and indication that the transformation has been useful?

The plots of the residuals are more representative of a model that does not violate the constant variance assumption.



Design Expert Output

**Response:** **Surface Roughness** **Transform:** **Inverse**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 2.059E+005 4 51472.28 1455.72 < 0.0001 significant

*A* *42610.92* *1* *42610.92* *1205.11* *< 0.0001*

*B* *89386.27* *1* *89386.27* *2527.99* *< 0.0001*

*C* *18762.29* *1* *18762.29* *530.63* *< 0.0001*

*AB* *55129.62* *1* *55129.62* *1559.16* *< 0.0001*

Residual 388.94 11 35.36

Cor Total 2.063E+005 15

The Model F-value of 1455.72 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, C, AB are significant model terms.



(d) Fit a model in terms of the coded variables that can be used to predict the surface roughness. Convert this prediction equation into a model in the natural variables.

Design Expert Output

**Final Equation in Terms of Coded Factors:**

1.0/(Surface Roughness) =

+397.81

+51.61 \* A

+74.74 \* B

+34.24 \* C

+58.70 \* A \* B

**6.33.** Resistivity on a silicon wafer is influenced by several factors. The results of a 24 factorial experiment performed during a critical process step is shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run | *A* | *B* | *C* | *D* | Resistivity |
| 1 | – | – | – | – | 1.92 |
| 2 | + | – | – | – | 11.28 |
| 3 | – | + | – | – | 1.09 |
| 4 | + | + | – | – | 5.75 |
| 5 | – | – | + | – | 2.13 |
| 6 | + | – | + | – | 9.53 |
| 7 | – | + | + | – | 1.03 |
| 8 | + | + | + | – | 5.35 |
| 9 | – | – | – | + | 1.60 |
| 10 | + | – | – | + | 11.73 |
| 11 | – | + | – | + | 1.16 |
| 12 | + | + | – | + | 4.68 |
| 13 | – | – | + | + | 2.16 |
| 14 | + | – | + | + | 9.11 |
| 15 | – | + | + | + | 1.07 |
| 16 | + | + | + | + | 5.30 |

(a) Estimate the factor effects. Plot the effect estimates on a normal probability plot and select a tentative model.



(b) Fit the model identified in part (a) and analyze the residuals. Is there any indication of model inadequacy?

The normal probability plot of residuals is not satisfactory. The plots of residual versus predicted, residual versus factor *A*, and the residual versus factor *B* are funnel shaped indicating non-constant variance.

Design Expert Output

**Response:** **Resistivity**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 214.22 3 71.41 148.81 < 0.0001 significant

*A* *159.83* *1* *159.83* *333.09* *< 0.0001*

*B* *36.09* *1* *36.09* *75.21* *< 0.0001*

*AB* *18.30* *1* *18.30* *38.13* *< 0.0001*

Residual 5.76 12 0.48

Cor Total 219.98 15

The Model F-value of 148.81 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, AB are significant model terms.





(c) Repeat the analysis from parts (a) and (b) using ln(*y*) as the response variable. Is there any indication that the transformation has been useful?



Design Expert Output

**Response:** **Resistivity** **Transform:** **Natural log** **Constant:** **0.000**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 12.15 2 6.08 553.44 < 0.0001 significant

*A* *10.57* *1* *10.57* *962.95* *< 0.0001*

*B* *1.58* *1* *1.58* *143.94* *< 0.0001*

Residual 0.14 13 0.011

Cor Total 12.30 15

The Model F-value of 553.44 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B are significant model terms.

The transformed data no longer indicates that the *AB* interaction is significant. A simpler model has resulted from the log transformation.





The residual plots are much improved.

(d) Fit a model in terms of the coded variables that can be used to predict the resistivity.

Design Expert Output

**Final Equation in Terms of Coded Factors:**

Ln(Resistivity) =

+1.19

+0.81 \* A

-0.31 \* B

**6.34. Continuation of Problem 6.33.** Suppose that the experiment had also run four center points along with the 16 runs in Problem 6.33. The resistivity measurements at the center points are: 8.15, 7.63, 8.95, 6.48. Analyze the experiment again incorporating the center points. What conclusions can you draw now?



Design Expert Output

**Response:** **Resistivity**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 214.22 3 71.41 119.35 < 0.0001 significant

*A* *159.83* *1* *159.83* *267.14* *< 0.0001*

*B* *36.09* *1* *36.09* *60.32* *< 0.0001*

*AB* *18.30* *1* *18.30* *30.58* *< 0.0001*

Curvature 31.19 1 31.19 52.13 < 0.0001 significant

Residual 8.97 15 0.60

*Lack of Fit* *5.76* *12* *0.48* *0.45* *0.8632* *not significant*

*Pure Error* *3.22* *3* *1.07*

Cor Total 254.38 19

The Model F-value of 119.35 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B, AB are significant model terms.



Because of the funnel shaped residual versus predicted plot, the analysis was repeated with the natural log transformation.



Design Expert Output

**Response:** **Resistivity** **Transform:** **Natural log** **Constant:** **0.000**

**ANOVA for Selected Factorial Model**

**Analysis of variance table [Partial sum of squares]**

**Sum of** **Mean** **F**

**Source** **Squares** **DF** **Square** **Value** **Prob > F**

Model 12.15 2 6.08 490.37 < 0.0001 significant

*A* *10.57* *1* *10.57* *853.20* *< 0.0001*

*B* *1.58* *1* *1.58* *127.54* *< 0.0001*

Curvature 2.38 1 2.38 191.98 < 0.0001 significant

Residual 0.20 16 0.012

*Lack of Fit* *0.14* *13* *0.011* *0.59* *0.7811* *not significant*

*Pure Error* *0.056* *3* *0.019*

Cor Total 14.73 19

The Model F-value of 490.37 implies the model is significant. There is only

a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, B are significant model terms.

The "Curvature F-value" of 191.98 implies there is significant curvature (as measured by

difference between the average of the center points and the average of the factorial points) in

the design space. There is only a 0.01% chance that a "Curvature F-value" this large

could occur due to noise.

The curvature test indicates that the model has significant pure quadratic curvature.

**6.35.** An article in *Quality and Reliability Engineering International* (2010, Vol. 26, pp. 223-233) presents a 25 factorial design. The experiment is shown in the following table:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| A | B | C | D | E | y |
| -1 | -1 | -1 | -1 | -1 | 8.11 |
| 1 | -1 | -1 | -1 | -1 | 5.56 |
| -1 | 1 | -1 | -1 | -1 | 5.77 |
| 1 | 1 | -1 | -1 | -1 | 5.82 |
| -1 | -1 | 1 | -1 | -1 | 9.17 |
| 1 | -1 | 1 | -1 | -1 | 7.8 |
| -1 | 1 | 1 | -1 | -1 | 3.23 |
| 1 | 1 | 1 | -1 | -1 | 5.69 |
| -1 | -1 | -1 | 1 | -1 | 8.82 |
| 1 | -1 | -1 | 1 | -1 | 14.23 |
| -1 | 1 | -1 | 1 | -1 | 9.2 |
| 1 | 1 | -1 | 1 | -1 | 8.94 |
| -1 | -1 | 1 | 1 | -1 | 8.68 |
| 1 | -1 | 1 | 1 | -1 | 11.49 |
| -1 | 1 | 1 | 1 | -1 | 6.25 |
| 1 | 1 | 1 | 1 | -1 | 9.12 |
| -1 | -1 | -1 | -1 | 1 | 7.93 |
| 1 | -1 | -1 | -1 | 1 | 5 |
| -1 | 1 | -1 | -1 | 1 | 7.47 |
| 1 | 1 | -1 | -1 | 1 | 12 |
| -1 | -1 | 1 | -1 | 1 | 9.86 |
| 1 | -1 | 1 | -1 | 1 | 3.65 |
| -1 | 1 | 1 | -1 | 1 | 6.4 |
| 1 | 1 | 1 | -1 | 1 | 11.61 |
| -1 | -1 | -1 | 1 | 1 | 12.43 |
| 1 | -1 | -1 | 1 | 1 | 17.55 |
| -1 | 1 | -1 | 1 | 1 | 8.87 |
| 1 | 1 | -1 | 1 | 1 | 25.38 |
| -1 | -1 | 1 | 1 | 1 | 13.06 |
| 1 | -1 | 1 | 1 | 1 | 18.85 |
| -1 | 1 | 1 | 1 | 1 | 11.78 |
| 1 | 1 | 1 | 1 | 1 | 26.05 |

(a) Analyze the data from this experiment. Identify the significant factors and interactions.

The half normal plot of effects below identifies the significant factors and interactions.



Design Expert Output

**Response 1 y  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 879.62 11 79.97 49.15 < 0.0001 significant

*A-A* *83.56* *1* *83.56* *51.36* *< 0.0001*  
  *B-B* *0.060* *1* *0.060* *0.037* *0.8492*  
  *D-D* *285.78* *1* *285.78* *175.66* *< 0.0001*  
  *E-E* *153.17* *1* *153.17* *94.15* *< 0.0001*  
  *AB* *48.93* *1* *48.93* *30.08* *< 0.0001*  
  *AD* *88.88* *1* *88.88* *54.63* *< 0.0001*  
  *AE* *33.76* *1* *33.76* *20.75* *0.0002*  
  *BE* *52.71* *1* *52.71* *32.40* *< 0.0001*  
  *DE* *61.80* *1* *61.80* *37.99* *< 0.0001*  
  *ABE* *44.96* *1* *44.96* *27.64* *< 0.0001*  
  *ADE* *26.01* *1* *26.01* *15.99* *0.0007*  
 Residual 32.54 20 1.63  
 Cor Total 912.16 31  
  
 The Model F-value of 49.15 implies the model is significant. There is only  
 a 0.01% chance that a "Model F-Value" this large could occur due to noise.  
  
 Values of "Prob > F" less than 0.0500 indicate model terms are significant.   
 In this case A, D, E, AB, AD, AE, BE, DE, ABE, ADE are significant model terms.

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

The residual plots below do not identify any concerns with model adequacy or the violations of the assumptions.















(c) One of the factors from this experiment does not seem to be important. If you drop this factor, what type of design remains? Analyze the data using the full factorial model for only the four active factors. Compare your results with those obtained in part (a).

The resulting experimental design is a replicated 24 full factorial design. The ANOVA is shown below. The factor names in the output below were modified to match the factor names in the original problem. The same factors are significant below as were significant in the original analysis.

Design Expert Output

**Response 1 y  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 888.80 15 59.25 40.58 < 0.0001 significant

*A-A* *83.56* *1* *83.56* *57.23* *< 0.0001*  
  *B-B* *0.060* *1* *0.060* *0.041* *0.8414*  
  *D-D* *285.78* *1* *285.78* *195.74* *< 0.0001*  
  *E-E* *153.17* *1* *153.17* *104.91* *< 0.0001*  
  *AB* *48.93* *1* *48.93* *33.51* *< 0.0001*  
  *AD* *88.88* *1* *88.88* *60.88* *< 0.0001*  
  *AE* *33.76* *1* *33.76* *23.13* *0.0002*  
  *BD* *5.778E-003* *1* *5.778E-003* *3.958E-003* *0.9506*  
  *BE* *52.71* *1* *52.71* *36.10* *< 0.0001*  
  *DE* *61.80* *1* *61.80* *42.33* *< 0.0001*  
  *ABD* *3.82* *1* *3.82* *2.61* *0.1255*  
  *ABE* *44.96* *1* *44.96* *30.79* *< 0.0001*  
  *ADE* *26.01* *1* *26.01* *17.82* *0.0006*  
  *BDE* *0.050* *1* *0.050* *0.035* *0.8549*  
  *ABDE* *5.31* *1* *5.31* *3.63* *0.0747*  
 Pure Error 23.36 16 1.46  
 Cor Total 912.16 31  
  
 The Model F-value of 40.58 implies the model is significant. There is only  
 a 0.01% chance that a "Model F-Value" this large could occur due to noise.  
  
 Values of "Prob > F" less than 0.0500 indicate model terms are significant.   
 In this case A, C, D, AB, AC, AD, BD, CD, ABD, ACD are significant model terms.

(d) Find the settings of the active factors that maximize the predicted response.

The cube plot below, with factors *E* set at +1 and *C* set at 0, identifies the maximum predicted response with the remaining factors, *A*, *B*, and *D* all set at +1.



**6.36.** A paper in the *Journal of Chemical Technology and Biotechnology* (“Response Surface Optimization of the Critical Media Components for the Production of Surfactin,” 1997, Vol. 68, pp. 263-270) describes the use of a designed experiment to maximize surfactin production. A portion of the data from this experiment is shown in the table below. Surfactin was assayed by an indirect method, which involves measurement of surface tension of the diluted broth samples. Relative surfactin concentrations were determined by serially diluting the broth until the critical micelle concentration (CMC) was reached. The dilution at which the surface tension starts rising abruptly was denoted by CMC-1 and was considered proportional to the amount of surfactant present in the original sample.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run | Glucose (g dm-3) | NH4NO3 (g dm-3) | FeSO4 (g dm-3 x 10-4) | MnSO4 (g dm-3 x 10-2) | y (CMC)-1 |
| 1 | 20.00 | 2.00 | 6.00 | 4.00 | 23 |
| 2 | 60.00 | 2.00 | 6.00 | 4.00 | 15 |
| 3 | 20.00 | 6.00 | 6.00 | 4.00 | 16 |
| 4 | 60.00 | 6.00 | 6.00 | 4.00 | 18 |
| 5 | 20.00 | 2.00 | 30.00 | 4.00 | 25 |
| 6 | 60.00 | 2.00 | 30.00 | 4.00 | 16 |
| 7 | 20.00 | 6.00 | 30.00 | 4.00 | 17 |
| 8 | 60.00 | 6.00 | 30.00 | 4.00 | 26 |
| 9 | 20.00 | 2.00 | 6.00 | 20.00 | 28 |
| 10 | 60.00 | 2.00 | 6.00 | 20.00 | 16 |
| 11 | 20.00 | 6.00 | 6.00 | 20.00 | 18 |
| 12 | 60.00 | 6.00 | 6.00 | 20.00 | 21 |
| 13 | 20.00 | 2.00 | 30.00 | 20.00 | 36 |
| 14 | 60.00 | 2.00 | 30.00 | 20.00 | 24 |
| 15 | 20.00 | 6.00 | 30.00 | 20.00 | 33 |
| 16 | 60.00 | 6.00 | 30.00 | 20.00 | 34 |

(a) Analyze the data from this experiment. Identify the significant factors and interactions.

The half normal probability plot of effects, followed by the ANOVA, identify the significant factors and interactions. Although factor *B* is not significant, the *AB* interaction is.



Design Expert Output

**Response 1 y  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 680.50 6 113.42 20.73 < 0.0001 significant

*A-Glucose* *42.25* *1* *42.25* *7.72* *0.0214*  
  *B-NH4NO3* *0.000* *1* *0.000* *0.000* *1.0000*  
  *C-FeSO4* *196.00* *1* *196.00* *35.82* *0.0002*  
  *D-MnSO4* *182.25* *1* *182.25* *33.30* *0.0003*  
  *AB* *196.00* *1* *196.00* *35.82* *0.0002*  
  *CD* *64.00* *1* *64.00* *11.70* *0.0076*  
 Residual 49.25 9 5.47  
 Cor Total 729.75 15  
  
 The Model F-value of 20.73 implies the model is significant. There is only  
 a 0.01% chance that a "Model F-Value" this large could occur due to noise.  
  
 Values of "Prob > F" less than 0.0500 indicate model terms are significant.   
 In this case A, C, D, AB, CD are significant model terms.

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

The residual plots below do not identify any concerns with model adequacy or the violations of the assumptions.













(c) What conditions would optimize the surfactin production?

The response, y, is maximized when factor *A* is 20, *B* is 2, *C* is 30, and *D* is 20.



**6.37. Continuation of Problem 6.36.** The experiment in Problem 6.36 actually included six center points. The responses at these conditions were 35, 35, 35, 36, 36, and 34. Is there any indication of curvature in the response function? Are additional experiments necessary? What would you recommend doing now?

Curvature appears to be very significant with a *p* value less than 0.0001. Axial runs, along with additional center point runs to identify blocking effects, should be run.

Design Expert Output

**Response 1 y  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 713.00 8 89.12 54.61 < 0.0001 significant  
  *A-Glucose* *42.25* *1* *42.25* *25.89* *0.0003*  
  *B-NH4NO3* *0.000* *1* *0.000* *0.000* *1.0000*  
  *C-FeSO4* *196.00* *1* *196.00* *120.10* *< 0.0001*  
  *D-MnSO4* *182.25* *1* *182.25* *111.68* *< 0.0001*  
  *AB* *196.00* *1* *196.00* *120.10* *< 0.0001*  
  *AD* *12.25* *1* *12.25* *7.51* *0.0179*  
  *BC* *20.25* *1* *20.25* *12.41* *0.0042*  
  *CD* *64.00* *1* *64.00* *39.22* *< 0.0001*  
 Curvature 659.28 1 659.28 403.98 < 0.0001 significant  
 Residual 19.58 12 1.63  
 *Lack of Fit* *16.75* *7* *2.39* *4.22* *0.0660* *not significant*  
 *Pure Error* *2.83* *5* *0.57*  
 Cor Total 1391.86 21

**6.38.** An article in the *Journal of Hazardous Materials* (“Feasibility of Using Natural Fishbone Apatite as a Substitute for Hydroxyapatite in Remediating Aqueous Heavy Metals,” Vol. 69, Issue 2, 1999, pp. 187-197) describes an experiment to study the suitability of fishbone, a natural, apatite rich substance, as a substitute for hydroxyapatite in the sequestering of aqueous divalent heavy metal ions. Direct comparison of hydroxyapatite and fishbone apatite was performed using a three-factor two-level full factorial design. Apatite (30 or 60 mg) was added to 100mL deionized water and gently agitated overnight in a shaker. The pH was then adjusted to 5 or 7 using nitric acid. Sufficient concentration of lead nitrate solution was added to each flask to result in a final volume of 200 mL and a lead concentration of 0.483 or 2.41 mM, respectively. The experiment was a 23 replicated twice and it was performed for both fishbone and synthetic apatite. Results are shown below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Fishbone | | Hydroxyapatite | |
| Apatite | pH | Pb | Pb,mM | pH | Pb,mM | pH |
| + | + | + | 1.82 | 5.22 | 0.11 | 3.49 |
| + | + | + | 1.81 | 5.12 | 0.12 | 3.46 |
| + | + | – | 0.01 | 6.84 | 0.00 | 5.84 |
| + | + | – | 0.00 | 6.61 | 0.00 | 5.90 |
| + | – | + | 1.11 | 3.35 | 0.80 | 2.70 |
| + | – | + | 1.04 | 3.34 | 0.76 | 2.74 |
| + | – | – | 0.00 | 5.77 | 0.03 | 3.36 |
| + | – | – | 0.01 | 6.25 | 0.05 | 3.24 |
| – | + | + | 2.11 | 5.29 | 1.03 | 3.22 |
| – | + | + | 2.18 | 5.06 | 1.05 | 3.22 |
| – | + | – | 0.03 | 5.93 | 0.00 | 5.53 |
| – | + | – | 0.05 | 6.02 | 0.00 | 5.43 |
| – | – | + | 1.70 | 3.39 | 1.34 | 2.82 |
| – | – | + | 1.69 | 3.34 | 1.26 | 2.79 |
| – | – | – | 0.05 | 4.50 | 0.06 | 3.28 |
| – | – | – | 0.05 | 4.74 | 0.07 | 3.28 |

(a) Analyze the lead response for fishbone apatite. What factors are important?

As shown below, all main effects and interactions are significant.

Design Expert Output

**Response 1 Fishbone Pb  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 12.19 7 1.74 2629.41 < 0.0001 significant

*A-Apatite* *0.27* *1* *0.27* *400.34* *< 0.0001*  
  *B-pH* *0.35* *1* *0.35* *525.43* *< 0.0001*  
  *C-Pb* *10.99* *1* *10.99* *16587.51* *< 0.0001*  
  *AB* *0.023* *1* *0.023* *33.96* *0.0004*  
  *AC* *0.19* *1* *0.19* *285.62* *< 0.0001*  
  *BC* *0.36* *1* *0.36* *543.40* *< 0.0001*  
  *ABC* *0.020* *1* *0.020* *29.58* *0.0006*  
 Pure Error 5.300E-003 8 6.625E-004  
 Cor Total 12.20 15

**Coefficient** **Standard** **95% CI** **95% CI**  
 **Factor** **Estimate** **df** **Error** **Low** **High** **VIF**

Intercept 0.85 1 6.435E-003 0.84 0.87

A-Apatite -0.13 1 6.435E-003 -0.14 -0.11 1.00  
 B-pH 0.15 1 6.435E-003 0.13 0.16 1.00  
 C-Pb 0.83 1 6.435E-003 0.81 0.84 1.00  
 AB 0.038 1 6.435E-003 0.023 0.052 1.00  
 AC -0.11 1 6.435E-003 -0.12 -0.094 1.00  
 BC 0.15 1 6.435E-003 0.14 0.16 1.00  
 ABC 0.035 1 6.435E-003 0.020 0.050 1.00

(b) Analyze the residuals from this response and comment on model adequacy.

The normal plot identifies slightly thicker tails in the distribution of the residuals. The plots of residuals vs. predicted and residuals vs. the Pb effect identify nonconstant variance.











(c) Analyze the pH response for fishbone apatite. What factors are important?

The *AB* and *ABC* interactions are only moderately significant; all other main effects and interactions are significant.

Design Expert Output

**Response 1 Fishbone pH  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 21.09 7 3.01 116.29 < 0.0001 significant

*A-Apatite* *1.12* *1* *1.12* *43.17* *0.0002*  
  *B-pH* *8.14* *1* *8.14* *314.08* *< 0.0001*  
  *C-Pb* *9.84* *1* *9.84* *379.98* *< 0.0001*  
  *AB* *0.098* *1* *0.098* *3.77* *0.0881*  
  *AC* *1.17* *1* *1.17* *45.23* *0.0001*  
  *BC* *0.61* *1* *0.61* *23.64* *0.0013*  
  *ABC* *0.11* *1* *0.11* *4.14* *0.0763*  
 Pure Error 0.21 8 0.026  
 Cor Total 21.30 15

**Coefficient** **Standard** **95% CI** **95% CI**  
 **Factor** **Estimate** **df** **Error** **Low** **High** **VIF**

Intercept 5.05 1 0.040 4.96 5.14

A-Apatite 0.26 1 0.040 0.17 0.36 1.00  
 B-pH 0.71 1 0.040 0.62 0.81 1.00  
 C-Pb -0.78 1 0.040 -0.88 -0.69 1.00  
 AB -0.078 1 0.040 -0.17 0.015 1.00  
 AC -0.27 1 0.040 -0.36 -0.18 1.00  
 BC 0.20 1 0.040 0.10 0.29 1.00  
 ABC 0.082 1 0.040 -0.011 0.17 1.00

(d) Analyze the residuals from this response and comment on model adequacy.

Although the normal probability plot is acceptable, the plots of residuals vs. predicted identifies non-constant variance.











(e) Analyze the lead response for hydroxyapatite apatite. What factors are important?

As shown below, all main effects and interactions are significant.

Design Expert Output

**Response 1 Hydroxyapatite Pb  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 4.01 7 0.57 1018.21 < 0.0001 significant

*A-Apatite* *0.54* *1* *0.54* *960.40* *< 0.0001*  
  *B-pH* *0.27* *1* *0.27* *471.51* *< 0.0001*  
  *C-Pb* *2.45* *1* *2.45* *4354.18* *< 0.0001*  
  *AB* *0.036* *1* *0.036* *64.18* *< 0.0001*  
  *AC* *0.50* *1* *0.50* *896.18* *< 0.0001*  
  *BC* *0.17* *1* *0.17* *298.84* *< 0.0001*  
  *ABC* *0.046* *1* *0.046* *82.18* *< 0.0001*  
 Pure Error 4.500E-003 8 5.625E-004  
 Cor Total 4.01 15

**Coefficient** **Standard** **95% CI** **95% CI**  
 **Factor** **Estimate** **df** **Error** **Low** **High** **VIF**

Intercept 0.42 1 5.929E-003 0.40 0.43

A-Apatite -0.18 1 5.929E-003 -0.20 -0.17 1.00  
 B-pH -0.13 1 5.929E-003 -0.14 -0.12 1.00  
 C-Pb 0.39 1 5.929E-003 0.38 0.40 1.00  
 AB -0.048 1 5.929E-003 -0.061 -0.034 1.00  
 AC -0.18 1 5.929E-003 -0.19 -0.16 1.00  
 BC -0.10 1 5.929E-003 -0.12 -0.089 1.00  
 ABC -0.054 1 5.929E-003 -0.067 -0.040 1.00

(f) Analyze the residuals from this response and comment on model adequacy.

The normal plot identifies slightly thicker tails in the distribution of the residuals. The plots of residuals vs. predicted and residuals vs. the effects identifies non-constant variance.











(g) Analyze the pH response for hydroxyapatite apatite. What factors are important?

The *ABC* interaction is not significant; all of the main effects and two factor interactions are significant.

Design Expert Output

**Response 1 Hydroxyapatite pH  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 20.44 7 2.92 1487.66 < 0.0001 significant

*A-Apatite* *0.084* *1* *0.084* *42.85* *0.0002*  
  *B-pH* *8.82* *1* *8.82* *4494.73* *< 0.0001*  
  *C-Pb* *8.15* *1* *8.15* *4153.39* *< 0.0001*  
  *AB* *0.13* *1* *0.13* *64.22* *< 0.0001*  
  *AC* *0.014* *1* *0.014* *7.34* *0.0267*  
  *BC* *3.24* *1* *3.24* *1650.96* *< 0.0001*  
  *ABC* *2.250E-004* *1* *2.250E-004* *0.11* *0.7436*  
 Pure Error 0.016 8 1.963E-003  
 Cor Total 20.45 15

**Coefficient** **Standard** **95% CI** **95% CI**  
 **Factor** **Estimate** **df** **Error** **Low** **High** **VIF**

Intercept 3.77 1 0.011 3.74 3.79

A-Apatite 0.072 1 0.011 0.047 0.098 1.00  
 B-pH 0.74 1 0.011 0.72 0.77 1.00  
 C-Pb -0.71 1 0.011 -0.74 -0.69 1.00  
 AB 0.089 1 0.011 0.063 0.11 1.00  
 AC -0.030 1 0.011 -0.056 -4.461E-003 1.00  
 BC -0.45 1 0.011 -0.48 -0.42 1.00  
 ABC -3.750E-003 1 0.011 -0.029 0.022 1.00

(h) Analyze the residuals from this response and comment on model adequacy.

The only potential concern with the residual plots is the non-constant variance shown in the plot of residuals vs. pH.











(i) What differences do you see between fishbone and hydroxyapatite apatite? The authors of this paper concluded that fishbone apatite was comparable to hydroxyapatite apatite. Because the fishbone apatite is cheaper, it was recommended for adoption. Do you agree with these conclusions?

The authors of the journal article did not show their analysis for this experiment. When comparing the Fishbone and Hydroxyapatite models main effects and interactions for the Pb and pH responses, we might disagree with the authors.

A more effective approach to understand the differences between Fishbone and Hydroxyapatite would be to include this as a factor in the experimental design. The modified table is shown below followed by the analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Apatite | pH | Pb | Type | Pb,mM | pH |
| + | + | + | Fishbone | 1.82 | 5.22 |
| + | + | + | Fishbone | 1.81 | 5.12 |
| + | + | – | Fishbone | 0.01 | 6.84 |
| + | + | – | Fishbone | 0.00 | 6.61 |
| + | – | + | Fishbone | 1.11 | 3.35 |
| + | – | + | Fishbone | 1.04 | 3.34 |
| + | – | – | Fishbone | 0.00 | 5.77 |
| + | – | – | Fishbone | 0.01 | 6.25 |
| – | + | + | Fishbone | 2.11 | 5.29 |
| – | + | + | Fishbone | 2.18 | 5.06 |
| – | + | – | Fishbone | 0.03 | 5.93 |
| – | + | – | Fishbone | 0.05 | 6.02 |
| – | – | + | Fishbone | 1.70 | 3.39 |
| – | – | + | Fishbone | 1.69 | 3.34 |
| – | – | – | Fishbone | 0.05 | 4.50 |
| – | – | – | Fishbone | 0.05 | 4.74 |
| + | + | + | Hydroxyapatite | 0.11 | 3.49 |
| + | + | + | Hydroxyapatite | 0.12 | 3.46 |
| + | + | – | Hydroxyapatite | 0.00 | 5.84 |
| + | + | – | Hydroxyapatite | 0.00 | 5.90 |
| + | – | + | Hydroxyapatite | 0.80 | 2.70 |
| + | – | + | Hydroxyapatite | 0.76 | 2.74 |
| + | – | – | Hydroxyapatite | 0.03 | 3.36 |
| + | – | – | Hydroxyapatite | 0.05 | 3.24 |
| – | + | + | Hydroxyapatite | 1.03 | 3.22 |
| – | + | + | Hydroxyapatite | 1.05 | 3.22 |
| – | + | – | Hydroxyapatite | 0.00 | 5.53 |
| – | + | – | Hydroxyapatite | 0.00 | 5.43 |
| – | – | + | Hydroxyapatite | 1.34 | 2.82 |
| – | – | + | Hydroxyapatite | 1.26 | 2.79 |
| – | – | – | Hydroxyapatite | 0.06 | 3.28 |
| – | – | – | Hydroxyapatite | 0.07 | 3.28 |

The ANOVA below identifies factor *D*, the type of apatite, as being very significant.

Design Expert Output

**Response 1 Pb Response**

**ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 17.73 15 1.18 1929.32 < 0.0001 significant

*A-Apatite* *0.78* *1* *0.78* *1275.51* *< 0.0001*  
  *B-pH* *2.813E-003* *1* *2.813E-003* *4.59* *0.0478*  
  *C-Pb* *11.91* *1* *11.91* *19440.33* *< 0.0001*  
  *D-Type* *1.52* *1* *1.52* *2485.73* *< 0.0001*  
  *AB* *8.000E-004* *1* *8.000E-004* *1.31* *0.2699*  
  *AC* *0.66* *1* *0.66* *1070.22* *< 0.0001*  
  *AD* *0.024* *1* *0.024* *39.51* *< 0.0001*  
  *BC* *0.018* *1* *0.018* *29.47* *< 0.0001*  
  *BD* *0.61* *1* *0.61* *996.76* *< 0.0001*  
  *CD* *1.53* *1* *1.53* *2500.00* *< 0.0001*  
  *ABC* *2.813E-003* *1* *2.813E-003* *4.59* *0.0478*  
  *ABD* *0.058* *1* *0.058* *94.37* *< 0.0001*  
  *ACD* *0.038* *1* *0.038* *61.73* *< 0.0001*  
  *BCD* *0.51* *1* *0.51* *832.73* *< 0.0001*  
  *ABCD* *0.063* *1* *0.063* *102.88* *< 0.0001*  
 Pure Error 9.800E-003 16 6.125E-004  
 Cor Total 17.74 31

The residual plots below identify concerns, so a power transformation with lambda of 0.7 was applied to the Pb response.













Design Expert Output

**Response 1 Pb Response**

**Transform: Power Lambda: 0.7 Constant: 0  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 12.10 15 0.81 1844.15 < 0.0001 significant

*A-Apatite* *0.49* *1* *0.49* *1123.40* *< 0.0001*  
  *B-pH* *0.014* *1* *0.014* *32.86* *< 0.0001*  
  *C-Pb* *9.02* *1* *9.02* *20612.51* *< 0.0001*  
  *D-Type* *0.75* *1* *0.75* *1712.43* *< 0.0001*  
  *AB* *3.973E-003* *1* *3.973E-003* *9.08* *0.0082*  
  *AC* *0.29* *1* *0.29* *661.73* *< 0.0001*  
  *AD* *0.024* *1* *0.024* *55.70* *< 0.0001*  
  *BC* *5.077E-003* *1* *5.077E-003* *11.61* *0.0036*  
  *BD* *0.38* *1* *0.38* *877.65* *< 0.0001*  
  *CD* *0.73* *1* *0.73* *1670.33* *< 0.0001*  
  *ABC* *0.011* *1* *0.011* *25.97* *0.0001*  
  *ABD* *0.049* *1* *0.049* *112.97* *< 0.0001*  
  *ACD* *0.067* *1* *0.067* *152.27* *< 0.0001*  
  *BCD* *0.21* *1* *0.21* *472.40* *< 0.0001*  
  *ABCD* *0.057* *1* *0.057* *131.40* *< 0.0001*  
 Pure Error 6.999E-003 16 4.374E-004  
 Cor Total 12.11 31

There are no concerns with the residual plots shown below.











****

The analysis for the pH response is shown below.

Design Expert Output

**Response 2 pH Response**

**ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 54.62 15 3.64 261.32 < 0.0001 significant

*A-Apatite* *0.91* *1* *0.91* *65.15* *< 0.0001*  
  *B-pH* *16.95* *1* *16.95* *1216.47* *< 0.0001*  
  *C-Pb* *17.96* *1* *17.96* *1288.54* *< 0.0001*  
  *D-Type* *13.09* *1* *13.09* *939.72* *< 0.0001*  
  *AB* *9.031E-004* *1* *9.031E-004* *0.065* *0.8023*  
  *AC* *0.72* *1* *0.72* *51.89* *< 0.0001*  
  *AD* *0.29* *1* *0.29* *21.14* *0.0003*  
  *BC* *0.52* *1* *0.52* *37.15* *< 0.0001*  
  *BD* *6.903E-003* *1* *6.903E-003* *0.50* *0.4916*  
  *CD* *0.040* *1* *0.040* *2.86* *0.1100*  
  *ABC* *0.049* *1* *0.049* *3.50* *0.0796*  
  *ABD* *0.22* *1* *0.22* *15.99* *0.0010*  
  *ACD* *0.46* *1* *0.46* *33.24* *< 0.0001*  
  *BCD* *3.33* *1* *3.33* *239.31* *< 0.0001*  
  *ABCD* *0.059* *1* *0.059* *4.21* *0.0569*  
 Pure Error 0.22 16 0.014  
 Cor Total 54.84 31

Non-constant variance is identified in the residual plots below, so an inverse square root transformation was applied.

****

****

****

****

****

****

Design Expert Output

**Response 2 pH Response**

**Transform: Inverse Sqrt Constant: 0  
 ANOVA for selected factorial model  
 Analysis of variance table [Partial sum of squares - Type III]**

**Sum of** **Mean** **F** **p-value**  
 **Source** **Squares** **df** **Square** **Value** **Prob > F**

Model 0.17 15 0.011 483.43 < 0.0001 significant

*A-Apatite* *1.418E-003* *1* *1.418E-003* *60.44* *< 0.0001*  
  *B-pH* *0.055* *1* *0.055* *2346.06* *< 0.0001*  
  *C-Pb* *0.054* *1* *0.054* *2303.65* *< 0.0001*  
  *D-Type* *0.045* *1* *0.045* *1918.39* *< 0.0001*  
  *AB* *1.417E-005* *1* *1.417E-005* *0.60* *0.4484*  
  *AC* *9.441E-004* *1* *9.441E-004* *40.23* *< 0.0001*  
  *AD* *3.287E-004* *1* *3.287E-004* *14.00* *0.0018*  
  *BC* *4.687E-005* *1* *4.687E-005* *2.00* *0.1768*  
  *BD* *8.263E-004* *1* *8.263E-004* *35.21* *< 0.0001*  
  *CD* *3.302E-004* *1* *3.302E-004* *14.07* *0.0017*  
  *ABC* *3.443E-004* *1* *3.443E-004* *14.67* *0.0015*  
  *ABD* *7.072E-004* *1* *7.072E-004* *30.14* *< 0.0001*  
  *ACD* *7.593E-004* *1* *7.593E-004* *32.36* *< 0.0001*  
  *BCD* *0.010* *1* *0.010* *437.95* *< 0.0001*  
  *ABCD* *4.034E-005* *1* *4.034E-005* *1.72* *0.2083*  
 Pure Error 3.755E-004 16 2.347E-005  
 Cor Total 0.17 31

There are no concerns with the residual plots shown below.













In summary, there is a difference between the Fishbone Apatite and the synthetic Hydroxyapatite.

**6.39S.** Often the fitted regression model from a 2*k* factorial design is used to make predictions at points of interest in the design space.

(a) Find the variance of the predicted response at the point *x*1, *x*2, . . . , *xk* in the design space. Hint: Remember that the *x*’s are coded variables, and assume a 2*k* design with an equal number of replicates *n* at each design point so that the variance of a regression coefficient  is  and that the covariance between any pair of regression coefficients is zero.

Let’s assume that the model can be written as follows:



where **x**’ = [ *x*1, *x*2, . . . , *xk* ] are the values of the original variables in the design at the point of interest where a prediction is required, and the variables in the model *x*1, *x*2, . . . , *xp* potentially include interaction terms among the original *k* variables. Now the variance of the predicted response is



This result follows because the design is orthogonal and all model parameter estimates have the same variance. Remember that some of the *x*’s involved in this equation are potentially interaction terms.

(b) Use the result of part (a) to find an equation for a 100(1–**)% confidence interval on the true mean response at the point ,,… , in the design space.

The confidence interval is



where *dfE* is the number of degrees of freedom used to estimate  and the estimate of  has been used in computing the variance of the predicted value of the response at the point of interest.

**6.40.** Suppose that you want to run a 23 factorial design. The variance of an individual observation is expected to be about 4. Suppose that you want the length of a 95% confidence interval on any effect to be less than or equal to 1.5. How many replicates of the design do you need to run?

With the equations for the *se*(Effect) and 100(1–*α*)% confidence interval on the effects shown below, we can iteratively estimate the number of replicates. From the table of iterations, 14 replicates are required.





|  |  |  |  |
| --- | --- | --- | --- |
| *n* | se(Effect) | *t*(0.025,*N*–*p*) | 95% CI Length |
| 12 | 0.408 | 1.987 | 1.623 |
| 13 | 0.392 | 1.985 | 1.557 |
| 14 | 0.378 | 1.983 | 1.499 |
| 15 | 0.365 | 1.981 | 1.447 |

**6.41.** Suppose that a full 24 factorial uses the following levels:

|  |  |  |
| --- | --- | --- |
| Factor | Low (–) | High(+) |
| A: Acid strength (%) | 85 | 95 |
| B: Reaction time (min) | 15 | 35 |
| C: Amount of acid (ml) | 35 | 45 |
| D: Reaction temperature (⁰C) | 60 | 80 |

The fitted model from this experiment is



Predict the response at the following points:

(a) *A* = 89, *B* = 20, *C* = 38, *D* = 66

The coded values are: *x*1 = –0.2, *x*2 = –0.5, *x*3 = –0.4, *x*4 = –0.4



(b) A = 90, B = 16, C = 40, D = 70

The coded values are: *x*1 = 0, *x*2 = –0.9, *x*3 = 0, *x*4 = 0



(c) A = 87, B = 28, C = 42, D = 61

The coded values are: *x*1 = –0.6, *x*2 = 0.3, *x*3 = 0.4, *x*4 = –0.9



(d) A = 90, B = 27, C = 37, D = 69

The coded values are: *x*1 = 0, *x*2 = 0.2, *x*3 = –0.6, *x*4 = –0.1



**6.42.** An article in *Quality and Reliability Engineering International* (2010, Vol. 26, pp. 223-233) presents a factorial design. The experiment is shown in Table P6.15.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *A* | *B* | *C* | *D* | *E* | *y* |
| -1 | -1 | -1 | -1 | -1 | 8.11 |
| 1 | -1 | -1 | -1 | -1 | 5.56 |
| -1 | 1 | -1 | -1 | -1 | 5.77 |
| 1 | 1 | -1 | -1 | -1 | 5.82 |
| -1 | -1 | 1 | -1 | -1 | 9.17 |
| 1 | -1 | 1 | -1 | -1 | 7.8 |
| -1 | 1 | 1 | -1 | -1 | 3.23 |
| 1 | 1 | 1 | -1 | -1 | 5.69 |
| -1 | -1 | -1 | 1 | -1 | 8.82 |
| 1 | -1 | -1 | 1 | -1 | 14.23 |
| -1 | 1 | -1 | 1 | -1 | 9.2 |
| 1 | 1 | -1 | 1 | -1 | 8.94 |
| -1 | -1 | 1 | 1 | -1 | 8.68 |
| 1 | -1 | 1 | 1 | -1 | 11.49 |
| -1 | 1 | 1 | 1 | -1 | 6.25 |
| 1 | 1 | 1 | 1 | -1 | 9.12 |
| -1 | -1 | -1 | -1 | 1 | 7.93 |
| 1 | -1 | -1 | -1 | 1 | 5 |
| -1 | 1 | -1 | -1 | 1 | 7.47 |
| 1 | 1 | -1 | -1 | 1 | 12 |
| -1 | -1 | 1 | -1 | 1 | 9.86 |
| 1 | -1 | 1 | -1 | 1 | 3.65 |
| -1 | 1 | 1 | -1 | 1 | 6.4 |
| 1 | 1 | 1 | -1 | 1 | 11.61 |
| -1 | -1 | -1 | 1 | 1 | 12.43 |
| 1 | -1 | -1 | 1 | 1 | 17.55 |
| -1 | 1 | -1 | 1 | 1 | 8.87 |
| 1 | 1 | -1 | 1 | 1 | 25.38 |
| -1 | -1 | 1 | 1 | 1 | 13.06 |
| 1 | -1 | 1 | 1 | 1 | 18.85 |
| -1 | 1 | 1 | 1 | 1 | 11.78 |
| 1 | 1 | 1 | 1 | 1 | 26.05 |

(a) Analyze the data from this experiment. Identify the significant factors and interactions.

Minitab Output



The half-normal plot shows that no 4-factor or 5-factor interactions are significant, so they can be dropped from the set of terms and be used to estimate error. Factors *A, D, E*, and the *AD, DE, BE, AB, AE, ABE*, *ADE* interactions appear to be important. The model can now be fit. But notice that there are several of the 2 and 3-factor interactions that are not significant. They can be removed for a better model fit; see part (c).

Minitab Output

**Factorial Regression: y versus A, B, C, D, E** Analysis of Variance

Source DF Adj SS Adj MS F-Value P-Value

Model 25 900.849 36.034 19.11 0.001

Linear 5 522.578 104.516 55.43 0.000

A 1 83.560 83.560 44.31 0.001

B 1 0.060 0.060 0.03 0.864

C 1 0.005 0.005 0.00 0.962

D 1 285.784 285.784 151.56 0.000

E 1 153.169 153.169 81.23 0.000

2-Way Interactions 10 290.219 29.022 15.39 0.002

A\*B 1 48.931 48.931 25.95 0.002

A\*C 1 0.000 0.000 0.00 0.995

A\*D 1 88.878 88.878 47.13 0.000

A\*E 1 33.764 33.764 17.91 0.005

B\*C 1 1.221 1.221 0.65 0.452

B\*D 1 0.006 0.006 0.00 0.958

B\*E 1 52.711 52.711 27.95 0.002

C\*D 1 0.000 0.000 0.00 0.989

C\*E 1 2.910 2.910 1.54 0.260

D\*E 1 61.799 61.799 32.77 0.001

3-Way Interactions 10 88.051 8.805 4.67 0.036

A\*B\*C 1 2.005 2.005 1.06 0.342

A\*B\*D 1 3.816 3.816 2.02 0.205

A\*B\*E 1 44.959 44.959 23.84 0.003

A\*C\*D 1 0.129 0.129 0.07 0.803

A\*C\*E 1 2.148 2.148 1.14 0.327

A\*D\*E 1 26.010 26.010 13.79 0.010

B\*C\*D 1 2.983 2.983 1.58 0.255

B\*C\*E 1 0.935 0.935 0.50 0.508

B\*D\*E 1 0.050 0.050 0.03 0.875

C\*D\*E 1 5.017 5.017 2.66 0.154

Error 6 11.314 1.886

Total 31 912.162

Model Summary

S R-sq R-sq(adj) R-sq(pred)

1.37319 98.76% 93.59% 64.72%

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

The residual plots look good; no indication of model inadequacy.

Minitab Output



(c) One of the factors from this experiment does not seem to be important. If you drop this factor, what type of design remains? Analyze the data using the full factorial model for only the four active factors. Compare your results with those obtained in part (a).

Factor *C* and all of its interactions are not important. The design becomes a replicated 24. The degrees of freedom increase to 17 when Factor *C* has been removed and *R*2 adjusted and *R*2 predicted have increased. There are a couple of 2 and 3-factor interaction that can be removed at this point; *BD, ABD* and *BDE*. *B* is not significant by as a main effect, but is part of several interactions, so leave it in the model.

Minitab Output

**Factorial Regression: y versus A, B, D, E**

Analysis of Variance

Source DF Adj SS Adj MS F-Value P-Value

Model 11 879.625 79.966 49.15 0.000

Linear 4 522.574 130.643 80.30 0.000

A 1 83.560 83.560 51.36 0.000

B 1 0.060 0.060 0.04 0.849

D 1 285.784 285.784 175.66 0.000

E 1 153.169 153.169 94.15 0.000

2-Way Interactions 5 286.082 57.216 35.17 0.000

A\*B 1 48.931 48.931 30.08 0.000

A\*D 1 88.878 88.878 54.63 0.000

A\*E 1 33.764 33.764 20.75 0.000

B\*E 1 52.711 52.711 32.40 0.000

D\*E 1 61.799 61.799 37.99 0.000

3-Way Interactions 2 70.969 35.484 21.81 0.000

A\*B\*E 1 44.959 44.959 27.64 0.000

A\*D\*E 1 26.010 26.010 15.99 0.001

Error 20 32.538 1.627

Total 31 912.162

Model Summary

S R-sq R-sq(adj) R-sq(pred)

1.27549 96.43% 94.47% 90.87%

(d) Find settings of the active factors that maximize the predicted response.

The optimum is found at Factors *A*, *B*, *D* and *E* all set to the high level. The predicted response at these settings is 25.7063

Minitab Output



**6.43.** Consider the 23 experiment shown in the figure below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Process Variables | | | Coded Variables | | |  |
| Run | Temp (⁰C) | Pressure (psing) | Conc (g/l) | *x*1 | *x*2 | *x*3 | Yield, *y* |
| 1 | 120 | 40 | 15 | -1 | -1 | -1 | 32 |
| 2 | 160 | 40 | 15 | 1 | -1 | -1 | 46 |
| 3 | 120 | 80 | 15 | -1 | 1 | -1 | 57 |
| 4 | 160 | 80 | 15 | 1 | 1 | -1 | 65 |
| 5 | 120 | 40 | 30 | -1 | -1 | 1 | 36 |
| 6 | 160 | 40 | 30 | 1 | -1 | 1 | 48 |
| 7 | 120 | 80 | 30 | -1 | 1 | 1 | 57 |
| 8 | 160 | 80 | 30 | 1 | 1 | 1 | 68 |
| 9 | 140 | 60 | 22.5 | 0 | 0 | 0 | 50 |
| 10 | 140 | 60 | 22.5 | 0 | 0 | 0 | 44 |
| 11 | 140 | 60 | 22.5 | 0 | 0 | 0 | 53 |
| 12 | 140 | 60 | 22.5 | 0 | 0 | 0 | 56 |

When running a designed experiment, it is sometimes difficult to reach and hold the precise factors required by the design. Small discrepancies are not important, but large ones are potentially of more concern. To illustrate, the experiment presented in Table P6.16 below shows a variation of the 23 design above, where many of the test combinations are not exactly the ones specified in the design. Most of the difficulty seems to have occurred with the temperature variable.

Table P6.16

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Process Variables | | | Coded Variables | | |  |
| Run | Temp  (⁰C) | Pressure (psing) | Conc (g/l) | *x*1 | *x*2 | *x*3 | Yield, *y* |
| 1 | 120 | 41 | 14 | -0.75 | -0.95 | -1.133 | 32 |
| 2 | 158 | 40 | 15 | 0.9 | -1 | -1 | 46 |
| 3 | 121 | 82 | 15 | -0.95 | 1.1 | -1 | 57 |
| 4 | 160 | 80 | 15 | 1 | 1 | -1 | 65 |
| 5 | 118 | 39 | 33 | -1.1 | -1.05 | 1.14 | 36 |
| 6 | 163 | 40 | 30 | 1.15 | -1 | 1 | 48 |
| 7 | 122 | 80 | 30 | -0.9 | 1 | 1 | 57 |
| 8 | 165 | 83 | 30 | 1.25 | 1.15 | 1 | 68 |
| 9 | 140 | 60 | 22.5 | 0 | 0 | 0 | 50 |
| 10 | 140 | 60 | 22.5 | 0 | 0 | 0 | 44 |
| 11 | 140 | 60 | 22.5 | 0 | 0 | 0 | 53 |
| 12 | 140 | 60 | 22.5 | 0 | 0 | 0 | 56 |

Fit a first-order model to both the original data and the data in Table P6.16. Compare the inference from the two models. What conclusions can you draw from this simple example?

Minitab Output

**Regression Analysis: yield versus X1, X2, X3**

Analysis of Variance

Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value

Regression 3 1166.38 92.86% 1166.38 388.792 34.70 0.000

X1 1 253.13 20.15% 253.13 253.125 22.59 0.001

X2 1 903.12 71.90% 903.12 903.125 80.61 0.000

X3 1 10.13 0.81% 10.13 10.125 0.90 0.370

Error 8 89.63 7.14% 89.63 11.203

Lack-of-Fit 5 10.88 0.87% 10.88 2.175 0.08 0.990

Pure Error 3 78.75 6.27% 78.75 26.250

Total 11 1256.00 100.00%

Model Summary

S R-sq R-sq(adj) PRESS R-sq(pred)

3.34711 92.86% 90.19% 130.230 89.63%

Coefficients

Term Coef SE Coef 95% CI T-Value P-Value VIF

Constant 51.000 0.966 (48.772, 53.228) 52.78 0.000

X1 5.63 1.18 ( 2.90, 8.35) 4.75 0.001 1.00

X2 10.62 1.18 ( 7.90, 13.35) 8.98 0.000 1.00

X3 1.13 1.18 ( -1.60, 3.85) 0.95 0.370 1.00

Regression Equation

yield = 51.000 + 5.63 X1 + 10.62 X2 + 1.13 X3

Minitab Output for Table XXX (2nd set of data)

**Regression Analysis: yield versus X1, X2, X3**

Analysis of Variance

Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value

Regression 3 1153.38 91.83% 1153.38 384.460 29.97 0.000

X1 1 270.47 21.53% 234.78 234.780 18.30 0.003

X2 1 872.70 69.48% 874.31 874.315 68.16 0.000

X3 1 10.21 0.81% 10.21 10.210 0.80 0.398

Error 8 102.62 8.17% 102.62 12.828

Lack-of-Fit 5 23.87 1.90% 23.87 4.774 0.18 0.951

Pure Error 3 78.75 6.27% 78.75 26.250

Total 11 1256.00 100.00%

Model Summary

S R-sq R-sq(adj) PRESS R-sq(pred)

3.58156 91.83% 88.77% 170.624 86.42%

Coefficients

Term Coef SE Coef 95% CI T-Value P-Value VIF

Constant 50.52 1.04 (48.13, 52.91) 48.76 0.000

X1 5.37 1.26 ( 2.48, 8.27) 4.28 0.003 1.00

X2 10.13 1.23 ( 7.30, 12.96) 8.26 0.000 1.00

X3 1.09 1.22 (-1.73, 3.91) 0.89 0.398 1.00

Regression Equation

yield = 50.52 + 5.37 X1 + 10.13 X2 + 1.09 X3

The Minitab regression output for both experiments will model the first order model. There is not much difference between the regression equations. Note that *x*2 is the most important variable and there is very little deviation from the original design in *x*2. Because of the deviations in *x*1, there are now 9 distinct levels of *x*1 and not enough degrees of freedom to model the interactions.

**6.44.** In two-level designs, the expected value of a non-significant factor effect is zero.

**True** False

**6.45.** A half-normal plot of factor effects plots the expected normal percentile versus the effect estimate.

True **False**

**6.46.** In an unreplicated design, the degrees of freedom associated with the ‘pure error’ component of error are zero.

**True**  False

**6.47.** In a replicated design (16 runs) the estimated of the model intercept is equal to one-half of the total of all 16 runs.

True **False**

**6.48.** Adding center runs to a design affects the estimate of the intercept term but not the estimates of any other factor effects.

**True** False

**6.49.** The mean square for pure error in a replicated factorial design can get smaller if non-significant terms are added to a model.

True **False**

**6.50.** A factorial design is a D-optimal design for fitting a first-order model.

**True** False

**6.51.** If a D-optimal design algorithm is used to create a 12-run design for fitting a first-order model in three variables with all three two-factor interactions, the algorithm will construct a factorial with 4 center runs.

True **False**

**6.52.** Suppose that you want to replicate 2 of the 8 runs in a 23 factorial design. How many ways are there to choose the 2 runs to replicate?

8 choose 2 = 8!/(8-2)!2!=28. However, the better way to choose is to pick a fraction.

Suppose that you decide to replicate the run with all three factors at the high level and the run with all three factors at the low level.

(a) Is the resulting design orthogonal?

No. In order to be orthogonal, all of the column of +1 and –1 sum to zero. The column of the *AB*, *AC* and *BC* interaction no longer sum to zero.

(b) What are the relative variance of the model coefficients if the main effects plus two-factor interaction model is to fit the data from this design?

Per the table below from Design Expert, the variance inflation of the main effects is now 1.07, versus 1.00 for the 8 run, orthogonal design, and 1.05 versus 1.00 for the two-factor interactions.

Design Expert Output

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Term** | **StdErr1** | **VIF** | **Ri-Squared** | **2 Std. Dev.** |
| A | 0.33 | 1.07 | 0.0667 | 54.6 % |
| B | 0.33 | 1.07 | 0.0667 | 54.6 % |
| C | 0.33 | 1.07 | 0.0667 | 54.6 % |
| AB | 0.33 | 1.05 | 0.0476 | 53.9 % |
| AC | 0.33 | 1.05 | 0.0476 | 53.9 % |
| BC | 0.33 | 1.05 | 0.0476 | 53.9 % |

(c) What is the power for detecting effects of two standard deviations in magnitude?

From the Design Expert output, the power is 54.6% for the main effects and 53.9 for the two-factor interaction.

**6.53.** The display below summarizes the results of analyzing a factorial design.

|  |  |  |  |
| --- | --- | --- | --- |
| Term | Effect Estimate | Sum of Squares | % Contribution |
| *A* | ? | 6.25 | 3.25945 |
| *B* | 5.25 | 110.25 | 57.4967 |
| *C* | 3.5 | 49 | 25.5541 |
| *D* | 0.75 | ? | 1.1734 |
| *AB* | 0.75 | 2.25 | 1.1734 |
| *AC* | -0.5 | 1 | 0.521512 |
| *AD* | 0.75 | 2.25 | 1.1734 |
| *BC* | 1.5 | 9 | ? |
| *BD* | 0.25 | 0.25 | 0.130378 |
| *CD* | 0.5 | 1 | 0.521512 |
| *ABC* | -1 | 4 | 2.08605 |
| *ABD* | ? | 2.25 | 1.1734 |
| *ACD* | -0.5 | ? | 0.521512 |
| *BCD* | 0 | 0 | 0 |
| *ABCD* | -0.5 | 1 | ? |

(a) Fill in the missing information in this table.

|  |  |  |  |
| --- | --- | --- | --- |
| Term | Effect Estimate | Sum of Squares | % Contribution |
| *A* | 1.25 | 6.25 | 3.25945 |
| *B* | 5.25 | 110.25 | 57.4967 |
| *C* | 3.5 | 49 | 25.5541 |
| *D* | 0.75 | **2.25** | 1.1734 |
| *AB* | 0.75 | 2.25 | 1.1734 |
| *AC* | -0.5 | 1 | 0.521512 |
| *AD* | 0.75 | 2.25 | 1.1734 |
| *BC* | 1.5 | 9 | **4.6936** |
| *BD* | 0.25 | 0.25 | 0.130378 |
| *CD* | 0.5 | 1 | 0.521512 |
| *ABC* | -1 | 4 | 2.08605 |
| *ABD* | **0.75** | 2.25 | 1.1734 |
| *ACD* | -0.5 | **1** | 0.521512 |
| *BCD* | 0 | 0 | 0 |
| *ABCD* | -0.5 | 1 | **0.521512** |

(b) Construct a normal probability plot of the effects. Which factors seem to be active? Factors *B* and *C* appear to be important.

Minitab Output

