**ChE 320\_Spr\_17\_HW 11 Solution**

**6-20**

a) The results from Minitab follow.

**Regression Analysis: Density versus Cont, Loss**

The regression equation is

Density = - 0.110 + 0.407 Cont + 2.11 Loss

Predictor Coef SE Coef T P VIF

Constant -0.1105 0.2501 -0.44 0.670

Cont 0.4072 0.1682 2.42 0.042 390.1

Loss 2.108 5.834 0.36 0.727 390.1

S = 0.00883422 R-Sq = 99.7% R-Sq(adj) = 99.7%

Analysis of Variance

Source DF SS MS F P

Regression 2 0.23563 0.11782 1509.64 0.000

Residual Error 8 0.00062 0.00008

Total 10 0.23626

Source DF Seq SS

Cont 1 0.23562

Loss 1 0.00001

The regression equation iswhere 

b)

-0.0089354

-0.0090847

-0.0030180

0.0025153

0.0074740

0.0087559

0.0079298

0.0098885

0.0008422

-0.0051274

-0.0112404

c) SSE = 0.00062 = 0.00008

d) R-Sq = 99.7%, R-Sq(adj) = 99.7%. R-Sq(adj) is equal to R-Sq.

e) Analysis of Variance

Source DF SS MS F P

Regression 2 0.23563 0.11782 1509.64 0.000

Residual Error 8 0.00062 0.00008

Total 10 0.23626

Based on the P-value from the ANOVA table, the regression model is significant at the 0.05 level of significance.

f) ,

,

and 

g) Predictor Coef SE Coef T P VIF

Constant -0.1105 0.2501 -0.44 0.670

Cont 0.4072 0.1682 2.42 0.042 390.1

Loss 2.108 5.834 0.36 0.727 390.1

Based on the P-values for each coefficient, only Cont is significantly different from zero at the 0.05 level of significance.

h) β0: -0.1105 ± 2.306(0.2501); -0.6872, 0.4662

β1: 0.4072 ± 2.306 (0.1682); 0.01933, 0.7951

β2 : 2.108 ± 2.306 (5.834); -11.345, 15.561

The Cis agree with the t-test results presented in g).

i)

SRES COOK

-1.23832 0.255007

-1.44997 0.692448

-0.37215 0.008618

0.32815 0.011784

0.92827 0.058551

1.08437 0.077203

1.02796 0.109710

1.35664 0.287682

0.11001 0.001337

-0.67570 0.054084

-1.59530 0.485253

The model is accurate.











j) The VIFs are 390.1.

There is an indication of a problem with multicollinearity.

**6-30**

The regression equation is

y = - 171 + 7.03 x1 + 12.7 x2

Predictor Coef SE Coef T P

Constant -171.26 28.40 -6.03 0.001

x1 7.029 1.539 4.57 0.004

x2 12.696 1.539 8.25 0.000

S = 3.07827 R-Sq = 93.7% R-Sq(adj) = 91.6%

Analysis of Variance

Source DF SS MS F P

Regression 2 842.37 421.18 44.45 0.000

Residual Error 6 56.85 9.48

Total 8 899.22

a) 

b) 

,

,



c) Based on the P-values from the t-test for each coefficient, both regressors appear to be significant at the 0.05 level of significance.

d) 





e) 





f) The prediction interval is wider than the confidence interval because it predicts a range for a **future observation** whereas the confidence interval predicts a range for the **mean response.**

**6-32**

a) MINITAB Output

Predictor Coef SE Coef T P

Constant 6.188 2.704 2.29 0.027

x1 9.6864 0.4989 19.42 0.000

x2 -0.3796 0.2339 -1.62 0.112

x3 2.9447 0.2354 12.51 0.000

S = 0.912336 R-Sq = 90.8% R-Sq(adj) = 90.2%

Analysis of Variance

Source DF SS MS F P

Regression 3 363.01 121.00 145.37 0.000

Residual Error 44 36.62 0.83

Total 47 399.63







P-valuex1 = 2P(t > |19.42|): for degrees of freedom of 44 we obtain P-value < 2(0.0005) = P-value < 0.001

P-valuex2 = 2P(t > |-1.62|): for degrees of freedom of 44 we obtain 2(0.05) < P-value < 2(0.1) = 0.1 < P-value < 0.2

P-valuex2 = 2P(t > |12.51|): for degrees of freedom of 44 we obtain P-value < 2(0.0005) = P-value < 0.001











b) 

c) Based on the ANOVA from MINITAB output in part (a), (interpolated) and P-value = 0.000 < α = 0.05. We reject null hypothesis and conclude that the regression is significant.

d) β1:(interpolated) and P-value = 0.000 < α = 0.05. We reject the null hypothesis and conclude that β1 is different from zero.

β2: (interpolated) and P-value = 0.112 > α = 0.05, we fail to reject the null hypothesis that β2 equals zero.

β3: (interpolated) and P-value = 0.000 < α = 0.05, we reject the null hypothesis that β3 equals zero.

e) 95%CI for β1



Because zero is not included in the CI, we reject the null hypothesis that β1 equals zero.

f) 95%CI for β2



Because zero is included in the 95% CI, we fail to reject the null hypothesis that β2 equals zero.

g) 95%CI for β3



Because zero is not included in the CI, we reject the null hypothesis that β3 equals zero.

h) The tests for β1 and β3 are significant while β2 is not significant. Also, the regression is significant (part (c)). A model with x*2* omitted should be considered. Residual analysis should be applied to make sure that model is adequate.

**\*6-33**

a) The regression equation is

y = 643 + 11.4 x1 - 0.933 x2 - 0.0106 x1x2 - 0.0272 x1^2 +0.000471 x2^2

Predictor Coef StDev T P VIF

Constant 642.685 0.000 \* \*

x1 11.3862 0.0000 \* \* 2675.3

x2 -0.933346 0.000000 \* \* 1283.4

x1x2 -0.0106334 0.0000000 \* \* 8342.1

x1^2 -0.0271620 0.0000000 \* \* 502.4

x2^2 0.00047076 0.00000000 \* \* 3301.5

S = \*

Analysis of Variance

Source DF SS MS F P

Regression 5 14112.00 2822.40 \* \*

Residual Error 0 \* \*

Total 5 14112.00

Source DF Seq SS

x1 1 10240.37

x2 1 1921.21

x1x2 1 827.86

x1^2 1 1056.39

x2^2 1 66.17

b) Because VIF’s are much greater than 10, multicollinearity is present in the second-order model.

c) Because SSE(Full Model) is not available the test statistic can not be computed.

**6-35**

a) All possible regressions.

Response is Sat

S

A S u A

Mallows g e r n

Vars R-Sq R-Sq(adj) C-p S e v g x

1 82.1 80.8 4.4 9.3577 X

1 57.0 54.0 27.3 14.487 X

2 87.9 86.1 1.1 7.9723 X X

2 83.0 80.4 5.5 9.4476 X X

3 88.0 85.0 3.0 8.2768 X X X

3 87.9 84.9 3.0 8.2942 X X X

4 88.0 83.6 5.0 8.6446 X X X X

b) Forward selection. Alpha-to-Enter: 0.25

Response is Sat on 4 predictors, with N = 16

Step 1 2

Constant 136.2 146.7

Age -1.43 -1.12

T-Value -8.01 -5.76

P-Value 0.000 0.000

Sev -0.56

T-Value -2.51

P-Value 0.026

S 9.36 7.97

R-Sq 82.07 87.92

R-Sq(adj) 80.79 86.06

Mallows C-p 4.4 1.1

c) Backward elimination. Alpha-to-Remove: 0.1

Response is Sat on 4 predictors, with N = 16

Step 1 2 3

Constant 146.2 146.2 146.7

Age -1.12 -1.12 -1.12

T-Value -5.25 -5.51 -5.76

P-Value 0.000 0.000 0.000

Sev -0.59 -0.59 -0.56

T-Value -2.11 -2.22 -2.51

P-Value 0.058 0.046 0.026

Surg 0.1

T-Value 0.03

P-Value 0.979

Anx 0.5 0.6

T-Value 0.22 0.25

P-Value 0.832 0.809

S 8.64 8.28 7.97

R-Sq 87.98 87.98 87.92

R-Sq(adj) 83.61 84.97 86.06

Mallows C-p 5.0 3.0 1.1

d) Model with only *age* and *severity* seems to be the “best” among all. It has a large R-Sq(adj) and small Cp and values.

**\*6-46**

a) The regression equation is

y = 3829 - 0.215 x3 + 21.2 x4 + 1.66 x5

Predictor Coef SE Coef T P

Constant 3829 2262 1.69 0.099

x3 -0.2149 0.1088 -1.97 0.056

x4 21.2134 0.9050 23.44 0.000

x5 1.6566 0.5502 3.01 0.005

S = 43.66 R-Sq = 99.3% R-Sq(adj) = 99.3%



b) Analysis of Variance

Source DF SS MS F P

Regression 3 9863398 3287799 1724.42 0.000

Residual Error 36 68638 1907

Total 39 9932036

The P-value from the ANOVA table is approximately zero. Therefore, reject H0 and conclude that the regression model is significant at α = 0.01. The test can also be conducted in more detail as follows:



for at least one j

The test statistic is



Reject H0 if f0 > fα,3,36 where f0.01,3,36 = 4.38

Using the results from the ANOVA table



Because 1724.42 > 4.38 reject H0 and conclude that the regression model is significant at α = 0.01. The P-value < 0.00001

c) All at α = 0.01 t0.005,36 = 2.72

  

  

t0 = -1.97 t0 =23.44 t0 = 3.01

  

Fail to reject  Reject  Reject 

Potentially the x3 term can be removed from the model.

d) R2 = 0.993 

The slight decrease in may be reflective of the insignificant x3 term.

e)



The normality assumption appears reasonable. The residuals fall along a line.

f)



The plot is satisfactory. There does not appear to be a nonrandom pattern in the residual vs. predicted plot.

g)



There is a slight indication that variance increases as **x3** increases. There is a “fanning out” appearance of the residuals.

h) Using the equation found in part a):

