STAT 305 D Homework 11

Due Apr 25, 2013 at 12:40 PM in class

Show all 6 steps in your hypothesis tests.

1. For this problem, use the dataset, `polypolyols.jmp` posted on the materials page of the course website.

   Return to the situation of Exercise 3 of Section 4.1 and the polymer molecular weight study of R. Harris.

   (a) Find s_{LF} for these data. What does this intend to measure in the context of the engineering problem?

   (b) Plot both residuals versus x and the standardized residuals versus x. How much difference is there in the appearance of these two plots?

   (c) Give a 90% two-sided confidence interval for the increase in mean average molecular weight that accompanies a 1°C increase in temperature here.

   (d) Give individual 90% two-sided confidence intervals for the mean average molecular weight at 212°C and also at 250°C.

   (e) Give simultaneous 90% two-sided confidence intervals for the two means indicated in part (d).

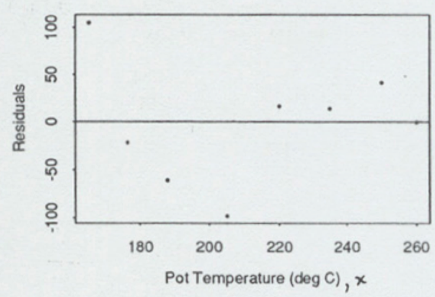
(a) See Exercise 3, Section 1, Chap. 4 for computations. Using equation (9-10),

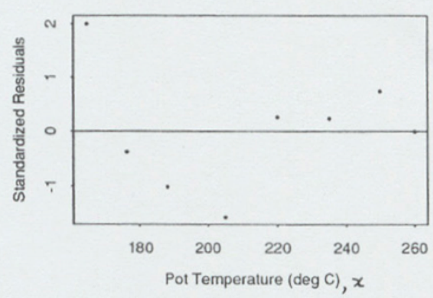
$$s_{\mathrm{LF}}^2 = \frac{1}{8-2}(26940.69) = 4490.116,$$

so $s_{\mathrm{LF}} = \sqrt{4490.116} = 67.01$, with 6 degrees of freedom associated with it. This measures the baseline variation in molecular weight that would be observed for any fixed pot temperature, assuming that model (9-4) is appropriate.

(b) The residuals were computed in Ex. 3, Sec. 1, Ch. 4. Use equation (9-12) to compute the standardized residuals. $\bar{x} = 212.375$, and $\sum(x-\bar{x})^2 = 8469.875$. The rest of the calculations are summarized below.

x	$\sqrt{1-\frac{1}{8}-\frac{(x-212.375)^2}{8469.875}}$	e	e^*
165	.78103	105.35535	2.01306
176	.84781	-21.12558	-.37186
188	.89714	-60.10477	-.99982
205	.93198	-97.57529	-1.56245
220	.93174	16.95072	.27150
235	.90253	14.47673	.23938
250	.84135	42.00275	.74503
260	.77924	.02009	.00038





The plots look almost exactly the same.

(c) This is β_1. Use equation (9-17). For 90% confidence, the appropriate t is $t = Q_6(.95) = 1.943$ from Table B-4. The resulting interval is

$$\begin{aligned} & 23.49827 \pm 1.943\frac{67.01}{\sqrt{8469.875}} \\ = \; & 23.49827 \pm 1.414696 \\ = \; & [22.08, 24.91]. \end{aligned}$$

(d) Use equation (9-24). The appropriate t is the same as the one in part (c). The resulting

interval for the mean at $x = 212$ is

$$\begin{aligned} & 1807.063 \pm 1.943(67.01)\sqrt{\frac{1}{8} + \frac{.140625}{8469.875}} \\ = \; & 1807.063 \pm 46.03471 \\ = \; & [1761.03, 1853.10]. \end{aligned}$$

The resulting interval for the mean at $x = 250$ is

$$\begin{aligned} & 2699.997 \pm 1.943(67.01)\sqrt{\frac{1}{8} + \frac{1415.641}{8469.875}} \\ = \; & 2699.997 \pm 70.37134 \\ = \; & [2629.63, 2770.37]. \end{aligned}$$

(e) Use equation (9-25). The appropriate f is $f = Q_{2,6}(.90) = 3.46$ from Table B-6-B. The resulting interval for the mean at $x = 212$ is

$$\begin{aligned} & 1807.063 \pm \sqrt{2(3.46)}(67.01)\sqrt{\frac{1}{8} + \frac{.140625}{8469.875}} \\ = \; & 1807.063 \pm 62.32548 \\ = \; & [1744.74, 1869.39]. \end{aligned}$$

The resulting interval for the mean at $x = 250$ is

$$\begin{aligned} & 2699.997 \pm \sqrt{2(3.46)}(67.01)\sqrt{\frac{1}{8} + \frac{1415.641}{8469.875}} \\ = \; & 2699.997 \pm \\ = \; & [2604.72, 2795.27]. \end{aligned}$$

2.

Nicholson and Bartle studied the effect of the water/cement ratio on 14-day compressive strength for Portland cement concrete. The water/cement ratios (by volume) and compressive strengths of nine concrete specimens are given next.

Water/Cement Ratio, x	14-Day Compressive Strength, y (psi)
.45	2954, 2913, 2923
.50	2743, 2779, 2739
.55	2652, 2607, 2583

(a) Find estimates of the parameters β_0, β_1, and σ in the simple linear regression model $y = \beta_0 + \beta_1 x + \epsilon$.

(b) Compute residuals and standardized residuals. Plot both against x and $\hat{y}$ and normal-plot them. How much do the appearances of the plots of the standardized residuals differ from those of the raw residuals?

(c) Make a 90% two-sided confidence interval for the increase in mean compressive strength that accompanies a .1 increase in the water/cement ratio. (This is $.1\beta_1$).

(d) Test the hypothesis that the mean compressive strength doesn't depend on the water/cement ratio. What is the p-value?

(e) Make a 95% two-sided confidence interval for the mean strength of specimens with the water/cement ratio .5 (based on the simple linear regression model).

(a) In Ex. 3, Sec. 1, Ch. 4, , $b_1 = -3160$ and $b_0 = 4345.889$. The necessary computations for s_{LF} (the residuals) are also given there. Using equation (9-10),

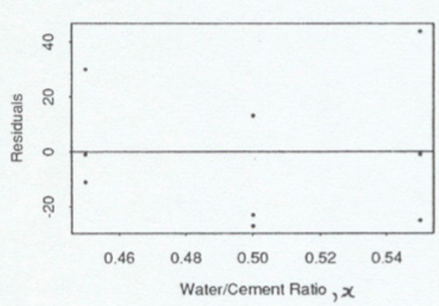
$$s_{\text{LF}}^2 = \frac{1}{9-2}(5010.889) = 715.84,$$

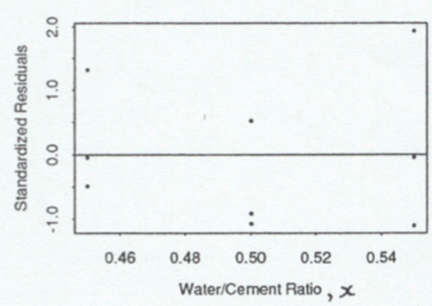
so $s_{\text{LF}} = \sqrt{715.84} = 26.755$ psi, with 7 degrees of freedom associated with it. This measures the baseline variation in 14-day compressive strength that would be observed for any fixed water/cement ratio, assuming that model (9-4) is appropriate. Using equation (7-7), $s_p = 26.890$ psi. These two estimates are very close, giving no indication

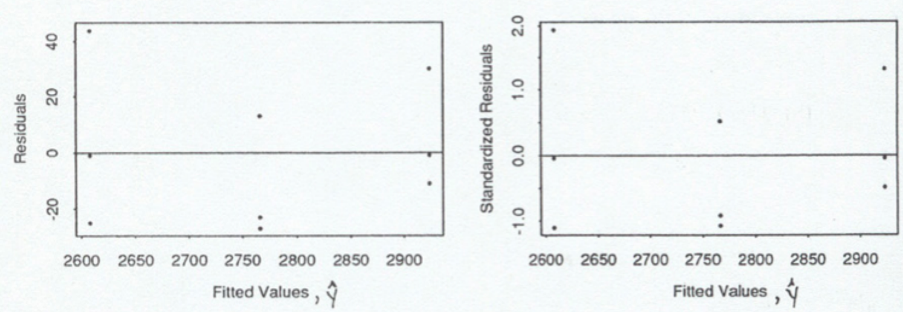
that the model is inappropriate.

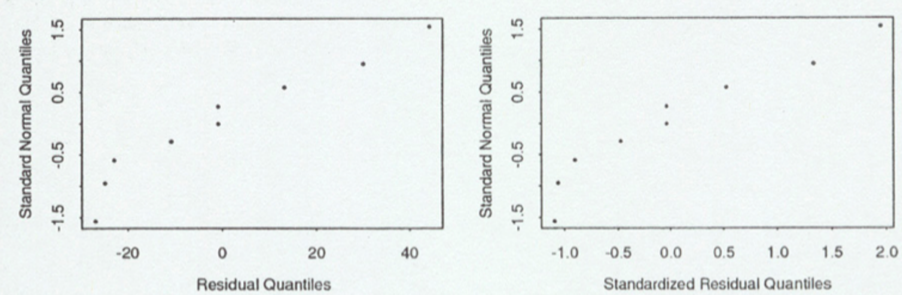
(b) The residuals were computed in Ex. 3, Sec. 1, Ch. 4. Use equation (9-12) to compute the standardized residuals. $\bar{x} = .50$, and $\sum(x - \bar{x})^2 = .015$. The rest of the calculations are summarized below.

x	$\sqrt{1 - \frac{1}{9} - \frac{(x-.50)^2}{.015}}$	e	e^*
.45	.8498	30.1111	1.3243
.45	.8498	-10.8889	-.4789
.45	.8498	-.8889	-.0391
.50	.9428	-22.8889	-.9074
.50	.9428	13.1111	.5198
.50	.9428	-26.8889	-1.0660
.55	.8498	44.1111	1.9400
.55	.8498	-.8889	-.0391
.55	.8498	-24.8889	-1.0946









For each of the three types of plots, the residuals and standardized residuals look almost exactly the same.

(c) First make a confidence interval for β_1, and then multiply the endpoints by .1. Use equation (9-17). For 90% confidence, the appropriate t is $t = Q_7(.95) = 1.895$ from Table B-4. The resulting interval for β_1 is

$$\begin{aligned} & -3160.0 \pm 1.895 \frac{26.755}{\sqrt{.015}} \\ = \; & -3160.0 \pm 413.9729 \\ = \; & [-3573.973, -2746.027] \text{ psi.} \end{aligned}$$

Multiplying each endpoint by .1, the resulting interval for $.1\beta_1$ is $[-357.4, -274.6]$ psi.

(d) This can be done using the test statistic (9-16) or with (9-34). Using (9-16),

1. H_0: $\beta_1 = 0$.
2. H_a: $\beta_1 \neq 0$.
3. The test statistic is given by equation (9-16), with $\# = 0$. The reference distribution is the t_7 distribution. Observed values of t far above or below zero will be considered as evidence against H_0.
4. The samples give

$$t = \frac{-3160.0}{\frac{26.755}{\sqrt{.015}}} = -14.47.$$

5. The observed level of significance is

$$\begin{aligned} & 2P(\text{a } t_7 \text{ random variable} < -14.47) \\ = \ & 2P(\text{a } t_7 \text{ random variable} > 14.47) \\ = \ & 2(\text{less than } .0005) \end{aligned}$$

which is less than .001, according to Table B-4 (14.47 is greater than $Q(.9995) = 5.408$). This is overwhelming evidence that the mean compressive strength is related to the water/cement ratio. (The given model is an improvement over the model $y = \beta_0 + \epsilon$.)

Using (9-34),

1. H_0: $\beta_1 = 0$.
2. H_a: $\beta_1 \neq 0$.
3. The test statistic is given by equation (9-34). The reference distribution is the $F_{1,7}$ distribution. Large observed values of F will be considered as evidence against H_0.
4. The samples give $SSE = (n-2)(s_{LF}) = 5010.889$ and $SSTot = 154794.9$, so $SSR = SSTot - SSE = 149784$.

$$f = \frac{149784}{715.84} = 209.24.$$

5. The observed level of significance is

$$P(\text{an } F_{1,7} \text{ random variable} > 209.24)$$

which is less than .001, according to Tables B-6 (209.24 is greater than $Q(.999) = 29.24$). This is overwhelming evidence that the mean compressive strength is related to the water/cement ratio. (The given model is an improvement over the model $y = \beta_0 + \epsilon$.)

Note that the F statistic is equal to the square of the t statistic.

(e) Use equation (9-24). For 95% confidence, the appropriate t is $t = Q_7(.975) = 2.365$ from Table B-4. The resulting interval for the mean at $x = .5$ is

$$\begin{aligned} & 2765.889 \pm 2.365(26.755)\sqrt{\frac{1}{9} + \frac{0}{.015}} \\ = \ & 2765.889 \pm 21.09202 \\ = \ & [2744.8, 2787.0] \text{ psi.} \end{aligned}$$

3.

Return to the situation of Chapter Exercise 2 of Chapter 4 and the carburetion study of Griffith and Tesdall. Consider an analysis of these data based on the model $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$.

(a) Find s_{SF} for these data. What does this intend to measure in the context of the engineering problem?

(b) Plot both residuals versus x and the standardized residuals versus x. How much difference is there in the appearance of these two plots?

(c) Give 90% individual two-sided confidence intervals for each of β_0, β_1, and β_2.

The dataset for this problem and a description are as follows:

Nicholson and Bartle studied the effect of the water/cement ratio on 14-day compressive strength for Portland cement concrete. The water/cement ratios (by volume) and compressive strengths of nine concrete specimens are given next.

Water/Cement Ratio, x	14-Day Compressive Strength, y (psi)
.45	2954, 2913, 2923
.50	2743, 2779, 2739
.55	2652, 2607, 2583

(a) Using equation (9-40),

$$s_{\text{SF}}^2 = \frac{1}{6-2-1}(.006562857) = .002187619,$$

and so $s_{\text{SF}} = \sqrt{.002187619} = .0468$ sec. This can also be read from the following Minitab Version 9.1 output. s_{SF} measures the variation in elapsed time for any fixed jetting size, assuming that the given model is appropriate.

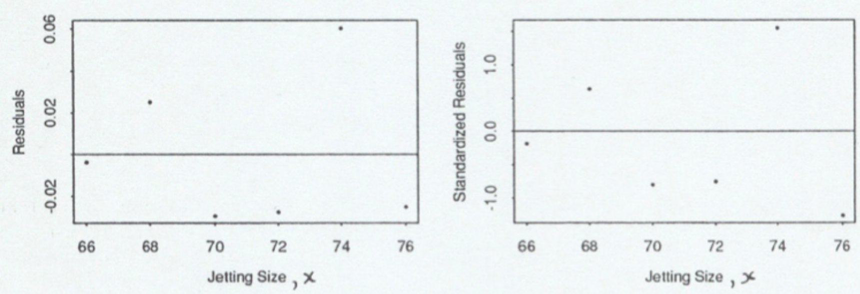
```
MTB > brief = 3
MTB > print c1-c3

 ROW    Time  JetSize    xsq

   1   14.90       66   4356
   2   14.67       68   4624
   3   14.50       70   4900
   4   14.53       72   5184
   5   14.79       74   5476
   6   15.02       76   5776

MTB > regress c1 2 c2 c3;
SUBC> fits c4;
SUBC> residuals c5;
SUBC> sresiduals c6.
* NOTE *  JetSize is highly correlated with other  predictor variables
* NOTE *      xsq is highly correlated with other  predictor variables

The regression equation is
Time = 104 - 2.53 JetSize + 0.0179 xsq

Predictor        Coef        Stdev     t-ratio         p
Constant      103.989        9.633       10.80     0.002
JetSize       -2.5343       0.2718       -9.32     0.003
xsq          0.017946     0.001914        9.38     0.003

s = 0.04677      R-sq = 96.9%     R-sq(adj) = 94.9%
```

s_{SF}

```
Analysis of Variance

SOURCE       DF          SS          MS         F        p
Regression    2     0.20639     0.10319     47.17    0.005
Error         3     0.00656     0.00219
Total         5     0.21295

SOURCE       DF      SEQ SS
JetSize       1     0.01400
xsq           1     0.19239

Obs.  JetSize       Time        Fit  Stdev.Fit   Residual    St.Resid
  1      66.0    14.9000    14.9036     0.0424    -0.0036      -0.18
  2      68.0    14.6700    14.6447     0.0259     0.0253       0.65
  3      70.0    14.5000    14.5294     0.0285    -0.0294      -0.79
  4      72.0    14.5300    14.5577     0.0285    -0.0277      -0.75
  5      74.0    14.7900    14.7296     0.0259     0.0604       1.55
  6      76.0    15.0200    15.0450     0.0424    -0.0250      -1.26
```

```
MTB > name c4 'fits' c5 'resids' c6 'stresids'
MTB > print c1-c6

ROW   Time  JetSize   xsq     fits      resids    stresids

  1  14.90       66   4356  14.9036  -0.0035715  -0.18070
  2  14.67       68   4624  14.6447   0.0252857   0.64948
  3  14.50       70   4900  14.5294  -0.0294285  -0.79361
  4  14.53       72   5184  14.5577  -0.0277147  -0.74739
  5  14.79       74   5476  14.7296   0.0604286   1.55216
  6  15.02       76   5776  15.0450  -0.0249996  -1.26486
```

(b)



There is a slight difference. The large positive residual is less extreme after it has been standardized. One of the negative residuals is more extreme after it has been standardized.

(c) Use equation (9-47) and the Minitab printout. The appropriate t is $t = Q_3(.95) = 2.353$ from Table B-4. The interval for β_0 is

$$\begin{aligned} & 103.989 \pm 2.353(9.633) \\ = \; & 103.989 \pm 22.66645 \\ = \; & [81.3, 126.7]. \end{aligned}$$

The interval for β_1 is

$$\begin{aligned} & -2.5343 \pm 2.353(.2718) \\ = \; & -2.5343 \pm .6395454 \\ = \; & [-3.17, -1.89]. \end{aligned}$$

The interval for β_2 is

$$\begin{aligned} & .017946 \pm 2.353(.001914) \\ = \; & .017946 \pm .004503642 \\ = \; & [.0134, .0225]. \end{aligned}$$

4. For this problem, use the dataset, `pulp.jmp`

posted on the materials page of the course website.

Return to the situation of Exercise 2 of Section 4.2, and the chemithermomechanical pulp study of Miller, Shankar, and Peterson. Consider an analysis of the data there based on the model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$.

(a) Find s_{SF}. What does this intend to measure in the context of the engineering problem?

(b) Plot both residuals and standardized residuals versus x_1, x_2, and $\hat{y}$. How much difference is there in the appearance of these pairs of plots?

(c) Give 90% individual two-sided confidence intervals for all of β_0, β_1, and β_2.

(a) The following output is from Minitab Version 9.1.

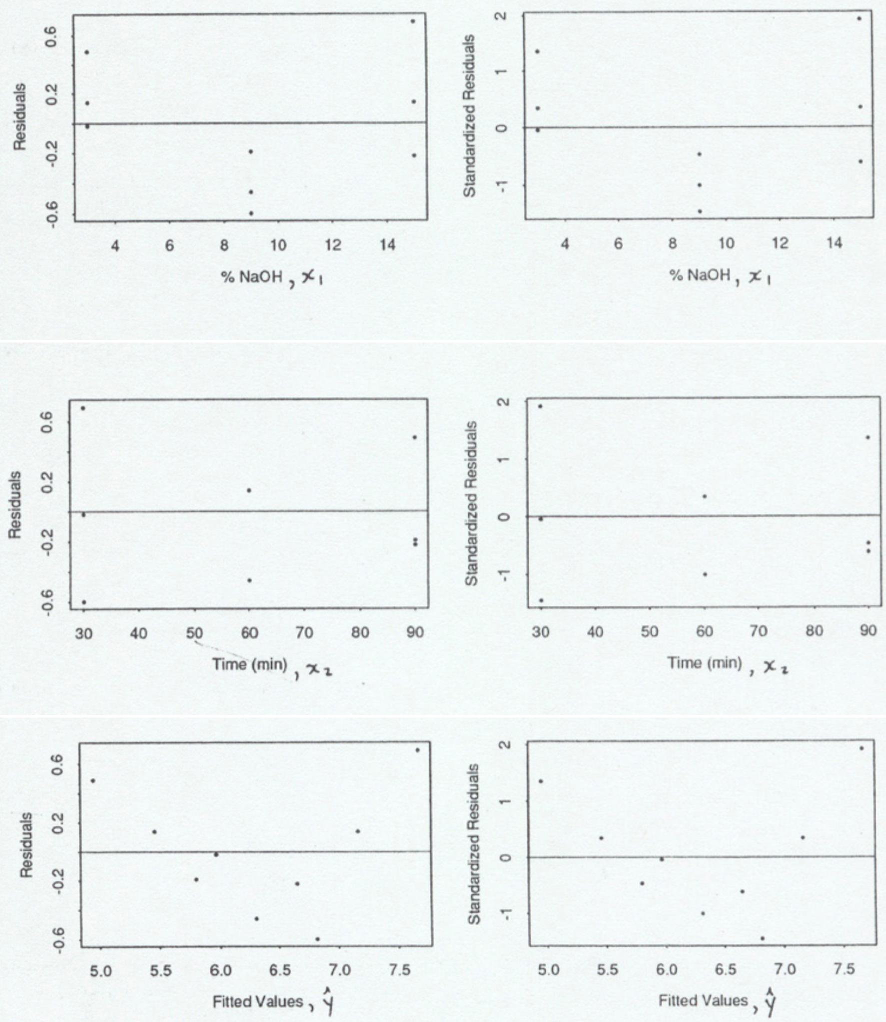
```
MTB > brief = 3
MTB > print c1-c3

 ROW  SurfArea  %NaOH   Time

   1      5.95      3     30
   2      5.60      3     60
   3      5.44      3     90
   4      6.22      9     30
   5      5.85      9     60
   6      5.61      9     90
   7      8.36     15     30
   8      7.30     15     60
   9      6.43     15     90

MTB > regress c1 2 c2 c3;
SUBC> fits c4;
SUBC> residuals c5;
SUBC> sresiduals c6;
SUBC> predict 10 70.

The regression equation is
SurfArea = 6.05 + 0.142 %NaOH - 0.0169 Time

Predictor        Coef       Stdev    t-ratio       p
Constant       6.0483      0.5208      11.61   0.000
%NaOH         0.14167     0.03301       4.29   0.005
Time        -0.016944    0.006601      -2.57   0.043

s = 0.4851        R-sq = 80.7%      R-sq(adj) = 74.2%
```

SSE

```
Analysis of Variance

SOURCE        DF         SS          MS         F        p
Regression     2     5.8854      2.9427     12.51    0.007
Error          6     1.4118      0.2353
Total          8     7.2972

SOURCE        DF     SEQ SS
%NaOH          1     4.3350
Time           1     1.5504

Obs.    %NaOH   SurfArea       Fit  Stdev.Fit   Residual    St.Resid
  1       3.0      5.950     5.965      0.323     -0.015       -0.04
  2       3.0      5.600     5.457      0.256      0.143        0.35
  3       3.0      5.440     4.948      0.323      0.492        1.36
  4       9.0      6.220     6.815      0.256     -0.595       -1.44
  5       9.0      5.850     6.307      0.162     -0.457       -1.00
  6       9.0      5.610     5.798      0.256     -0.188       -0.46
  7      15.0      8.360     7.665      0.323      0.695        1.92
  8      15.0      7.300     7.157      0.256      0.143        0.35
  9      15.0      6.430     6.648      0.323     -0.218       -0.60

     Fit  Stdev.Fit        95% C.I.            95% P.I.
   6.279      0.178   (  5.844,   6.714)  (  5.014,   7.543)

MTB > name c4 'fits' c5 'resids' c6 'stdres'
MTB > print c1-c6

 ROW  SurfArea  %NaOH   Time      fits      resids     stdres

   1      5.95      3     30   5.96500  -0.015000  -0.04149
   2      5.60      3     60   5.45667   0.143333   0.34770
   3      5.44      3     90   4.94833   0.491667   1.35987
   4      6.22      9     30   6.81500  -0.595000  -1.44335
   5      5.85      9     60   6.30667  -0.456666  -0.99854
   6      5.61      9     90   5.79833  -0.188333  -0.45686
   7      8.36     15     30   7.66500   0.695000   1.92226
   8      7.30     15     60   7.15667   0.143333   0.34770
   9      6.43     15     90   6.64833  -0.218334  -0.60388
```

$s_{SF} = .4851\ cm^3/g$. Assuming that the model is appropriate, this measures the variation in Specific Surface Areas for a fixed NaOH/Time condition.

(b)



For each of the three types of plots, the residuals and standardized residuals look almost exactly the same.

(c) Use equation (9-47) and the Minitab printout. The appropriate t is $t = Q_6(.95) = 1.943$

from Table B-4. The interval for β_0 is

$$\begin{aligned} & 6.0483 \pm 1.943(.5208) \\ = \; & 6.0483 \pm 1.011914 \\ = \; & [5.04, 7.06]. \end{aligned}$$

The interval for β_1 is

$$\begin{aligned} & .14167 \pm 1.943(.03301) \\ = \; & .14167 \pm .06413843 \\ = \; & [.078, .206]. \end{aligned}$$

The interval for β_2 is

$$\begin{aligned} & -.016944 \pm 1.943(.006601) \\ = \; & -.016944 \pm .01282574 \\ = \; & [-.0298, -.0041]. \end{aligned}$$

5. Weekly feedback. You get full credit as long as you write something.

   a. Is there any aspect of the subject matter that you currently struggle with? If so, what specifically do you find difficult or confusing? The more detailed you are, the better I can help you.
   You got full credit as long as you wrote something.

   b. Do you have any questions or concerns about the material, class logistics, or anything else? If so, fire away. You got full credit as long as you wrote something.