Homework 2 Solutions

1. Montgomery 3.3

3.3. A computer ANOVA output is shown below. Fill in the blanks. You may give bounds on the *P*-value.

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	3	36.15	?	?	?
Error	?	?	?		
Total	19	196.04			

Completed table is:

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	3	36.15	12.05	1.21	0.3395
Error	16	159.89	9.99		
Total	19	196.04			

2. Montgomery 3.22 (skip part d)

3.22. The response time in milliseconds was determined for three different types of circuits that could be used in an automatic valve shutoff mechanism. The results are shown in the following table.

Circuit Type	Response Time				
1	9	12	10	8	15
2	20	21	23	17	30
3	6	5	8	16	7

(a) Test the hypothesis that the three circuit types have the same response time. Use $\alpha = 0.01$.

Hypothesis test: $H_0 : \mu_1 = \mu_2 = \mu_3$ against the alternative H_a : not all are equal. SAS output is shown as below. Since the P-value (0.0004) is less than alpha = 0.01, we reject the null that the circuits have the same response time. Therefore, there is at least one circuit that is different.

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	2	543.6000000	271.8000000	16.08	0.0004
Error	12	202.8000000	16.9000000		
Corrected Total	14	746.4000000			

(b) Use Tukey's test to compare pairs of treatment means. Use $\alpha = 0.01$.

From the SAS output, it appears that the means for circuits 1 and 3 are not significantly different from each other, but the means of both circuits 1 and 3 differ from the mean of circuit 2.

The GLM Procedure
Tukey's Studentized Range (HSD) Test for response

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.01
Error Degrees of Freedom	12
Error Mean Square	16.9
Critical Value of Studentized Range	5.04582
Minimum Significant Difference	9.2766

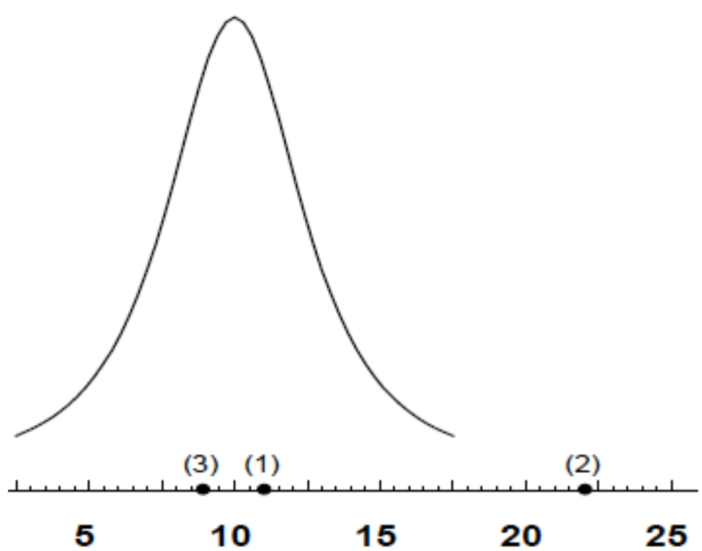
Means with the same letter are not significantly different.			
Tukey Grouping	Mean	N	circuit
A	22.200	5	2
B	10.800	5	1
B			
B	8.400	5	3

(c) Use the graphical procedure in Section 3.5.3 to compare the treatment means. What conclusions can you draw? How do they compare with the conclusions from part (a).

scale factor is :

$$\sqrt{\frac{MS_E}{n}} = \sqrt{\frac{16.9}{5}} = \sqrt{3.38} \approx 1.8385$$

use hand drawing:



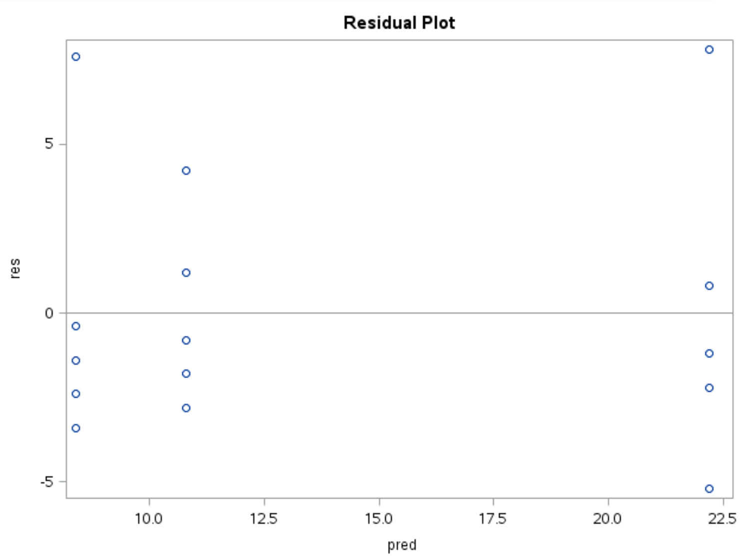
(e) If you were a design engineer and you wished to minimize the response time, which circuit type would you select?

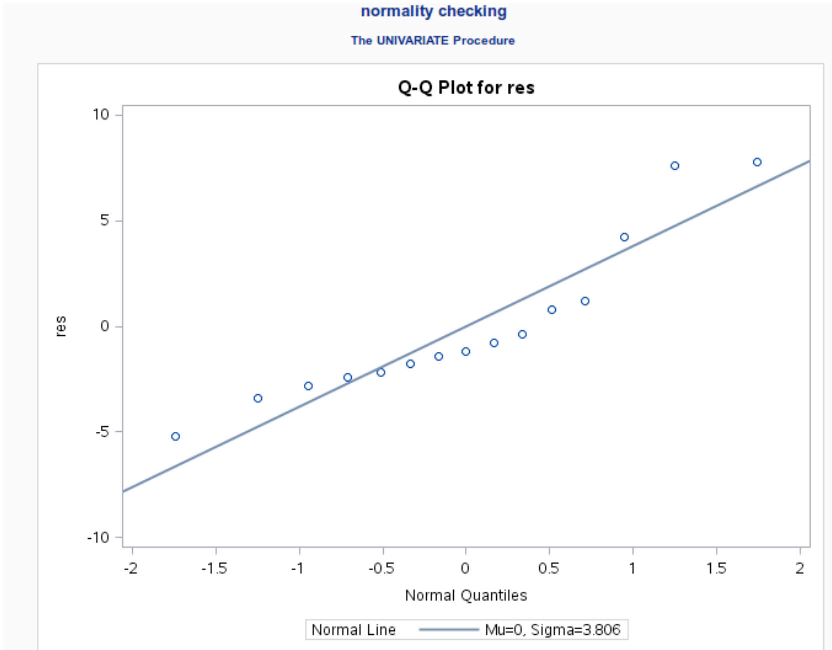
Either type 1 or type 3 as they are not different from each other and have the lowest response time.

(f) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied?

The residual plots do not suggest that the basic analysis of variance assumptions are violated as the spread of the residuals does not have apparent patterns. However, the normal probability plot indicates somehow non-normal pattern, which is confirmed by the results of formal tests of normality (p-values are around 0.05). Outlier checking shows no outliers (the values are between -1.26 and 1.84).

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.873524	Pr < W	0.0380
Kolmogorov-Smirnov	D	0.208517	Pr > D	0.0788
Cramer-von Mises	W-Sq	0.139029	Pr > W-Sq	0.0299
Anderson-Darling	A-Sq	0.815908	Pr > A-Sq	0.0267





SAS code:

Part a)

```
data Q322;
   input circuit $ response @@;
   datalines;
1 9 1 12 1 10 1 8 1 15
2 20 2 21 2 23 2 17 2 30
3 6 3 5 3 8 3 16 3 7
;

proc glm ;
   class circuit;
   model response=circuit;
   output out=diag p=pred r=res;
run;
```

Part b)

```
proc glm data=Q322;
   class circuit;
   model response=circuit;
   means circuit /alpha=0.01 lines tukey;
run;
```

Part e)

```
 proc glm data=Q322;
   class circuit;
   model response=circuit;
   means circuit /alpha=0.01 lines lsd;
   run;
```

Part f)

```
proc sgplot data=diag;
   scatter y=res x=pred;
```

```
run;

proc univariate data=diag normal;
 var res;
 qqplot res/normal(mu=est sigma=est color=red L=1);
 run;
```

3. Montgomery 3.23

3.23. The effective life of insulating fluids at an accelerated load of 35 kV is being studied. Test data have been obtained for four types of fluids. The results from a completely randomized experiment were as follows:

Fluid Type	Life (in h) at 35 kV Load					
1	17.6	18.9	16.3	17.4	20.1	21.6
2	16.9	15.3	18.6	17.1	19.5	20.3
3	21.4	23.6	19.4	18.5	20.5	22.3
4	19.3	21.1	16.9	17.5	18.3	19.8

(a) Is there any indication that the fluids differ? Use $\alpha = 0.05$.

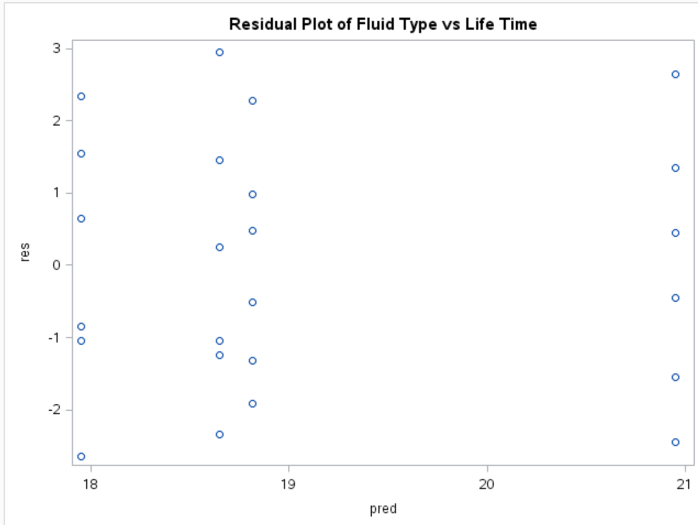
At $\alpha = 0.05$ there is no difference, but since the *P*-value is just slightly above 0.05, there is probably a difference in means.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	30.16500000	10.05500000	3.05	0.0525
Error	20	65.99333333	3.29966667		
Corrected Total	23	96.15833333			

(b) Which fluid would you select, given that the objective is long life?

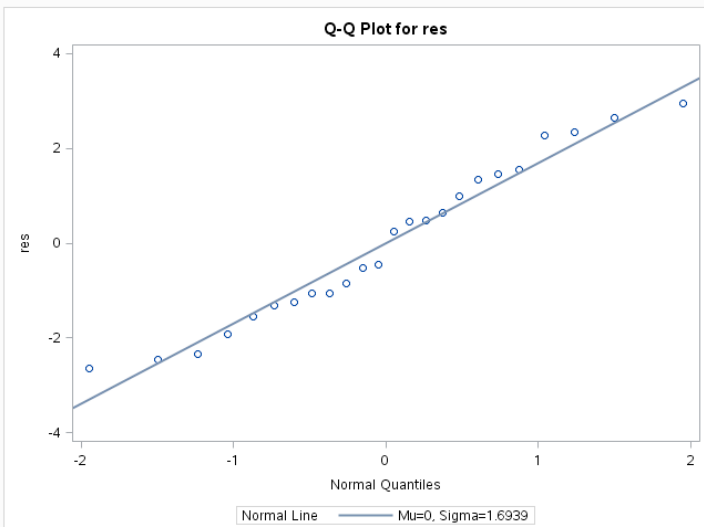
Treatment 3. The Fisher LSD procedure in the computer output indicates that the fluid 3 is different from the others, and it's average life also exceeds the average lives of the other three fluids.

(c) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied'
There is nothing unusual in the residual plots.



Normality Checking

The UNIVARIATE Procedure



SAS code:
Part a)

```
data Q323;
  input fluid $ life @@;
  datalines;
1 17.6 1 18.9 1 16.3 1 17.4 1 20.1 1 21.6
2 16.9 2 15.3 2 18.6 2 17.1 2 19.5 2 20.3
3 21.4 3 23.6 3 19.4 3 18.5 3 20.5 3 22.3
4 19.3 4 21.1 4 16.9 4 17.5 4 18.3 4 19.8
;

proc glm data=Q323;
  class fluid;
  model life=fluid;
  output out=diag p=pred r=res;
  run;
```

part b)

```
proc glm data=Q323;
  class fluid;
  model life=fluid;
   means percent /alpha=.05 lines lsd;
  run;
```

part c)

```
proc sgplot data=diag;
  scatter y=res x=pred;
run;

proc univariate data=diag normal plot;
  var res;
run;
```

4. Montgomery 3.28

3.28. An experiment was performed to investigate the effectiveness of five insulating materials. Four samples of each material were tested at an elevated voltage level to accelerate the time to failure. The failure times (in minutes) is shown below:

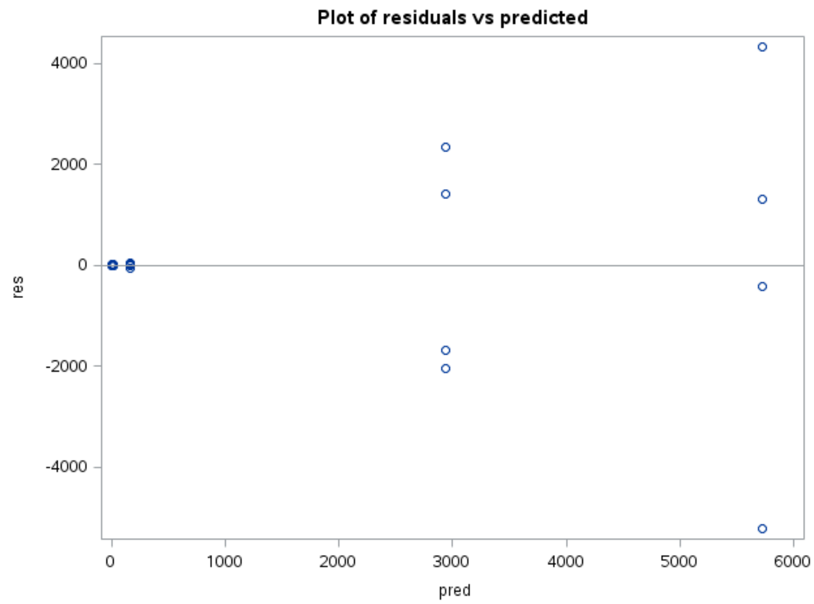
Material	Failure Time (minutes)			
1	110	157	194	178
2	1	2	4	18
3	880	1256	5276	4355
4	495	7040	5307	10050
5	7	5	29	2

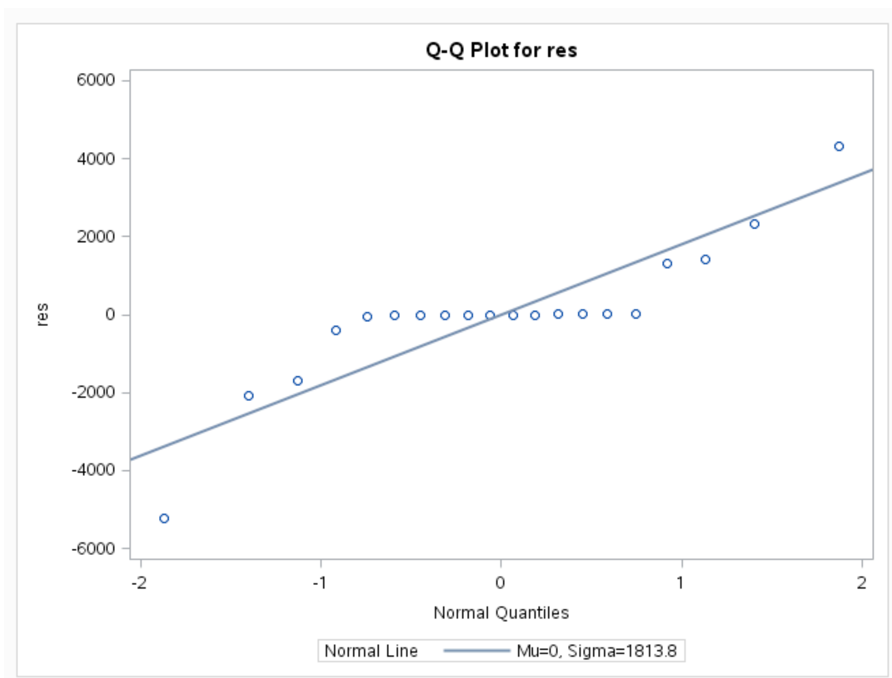
(a) Do all five materials have the same effect on mean failure time?

Since the P-value is small (0.0038, provided in the SAS output that follows), we reject the null hypothesis that the mean failure times are the same. There appears to be a difference in at least one material's effectiveness on mean failure time.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	103191489.2	25797872.3	6.19	0.0038
Error	15	62505657.0	4167043.8		
Corrected Total	19	165697146.2			

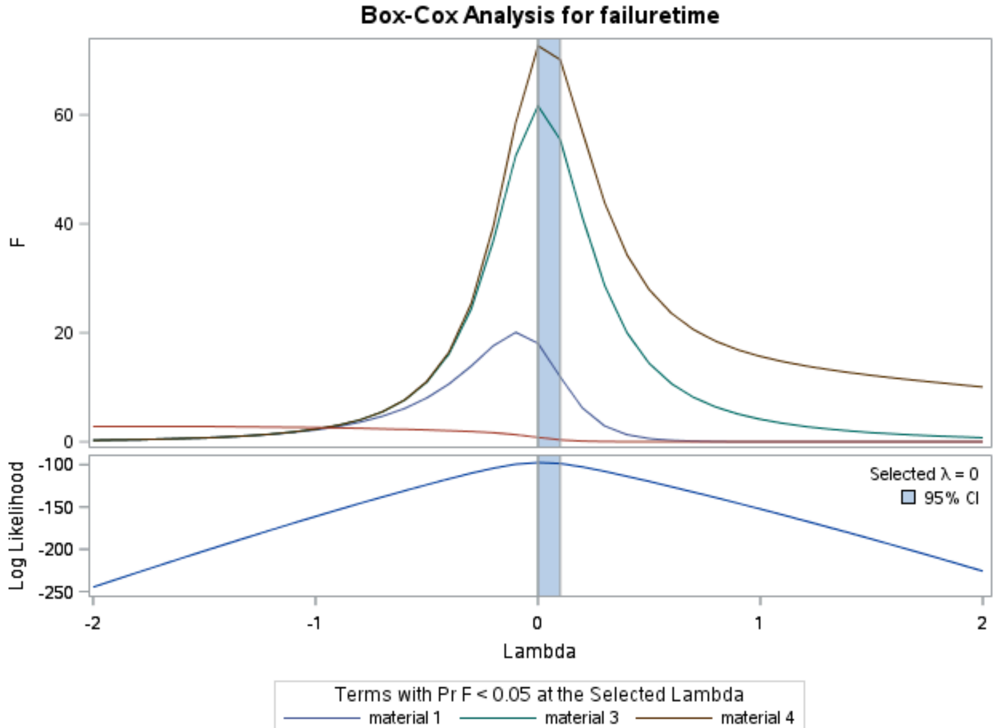
(b) Plot the residuals versus the predicted response. Construct a normal probability plot of the residuals. What information do these plots convey?





The plot of residuals versus predicted has a strong outward-opening funnel shape, which indicates the variance of the original observations is not constant. The normal probability plot also indicates that the normality assumption is not valid. A data transformation is recommended.

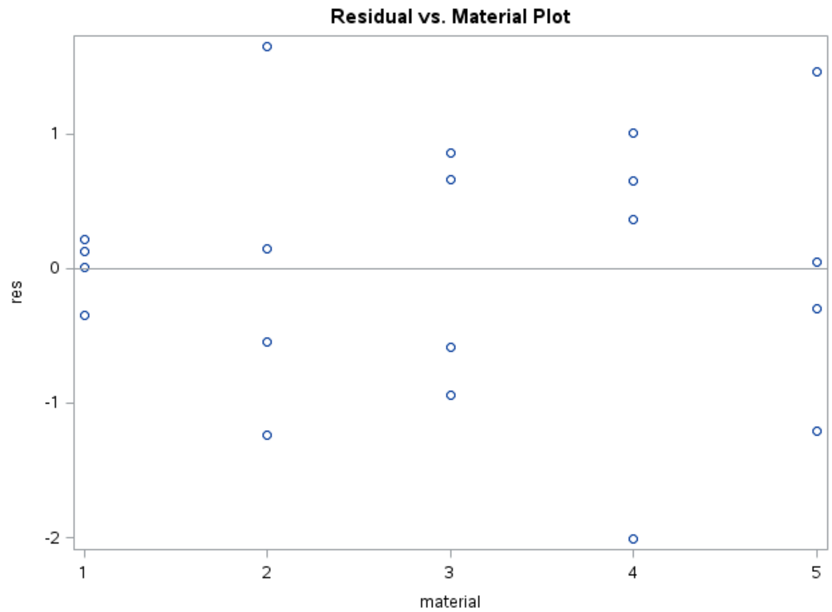
(c) Based on your answer to part (b) conduct another analysis of the failure time data and draw appropriate conclusions.

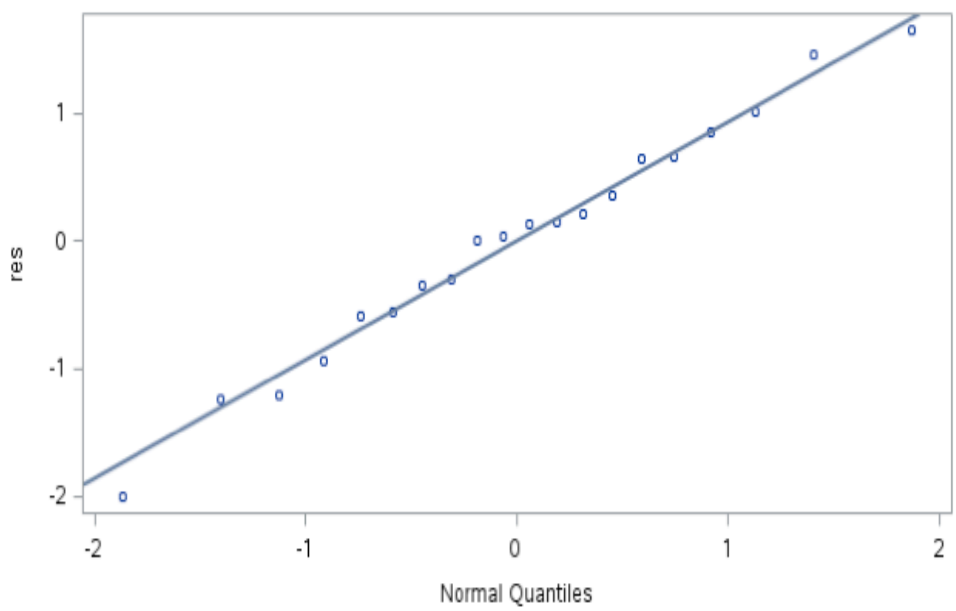


Use boxcox to determine transformation. Since lamda = 0 is recommended in the SAS output plot (see above), a natural log transformation will be used. From the residual and normal QQ plots for the new analysis, it's obvious that the assumptions are more satisfied, and the there exists a significant difference among the material types.

Dependent Variable: NewY

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	165.0564576	41.2641144	37.66	<.0001
Error	15	16.4368913	1.0957928		
Corrected Total	19	181.4933489			





SAS code:

Part a)

```
data Q328;
  input material $ failuretime @@;
  datalines;
1 110 1 157 1 194 1 178
2 1 2 2 2 4 2 18
3 880 3 1256 3 5276 3 4355
4 495 4 7040 4 5307 4 10050
5 7 5 5 5 29 5 2
```

```
 ;

 proc glm data=Q328;
   class material;
   model failuretime=material;
   output out=diag p=pred r=res;
   run;
```

part b)

```
title1 'Plot of residuals vs predicted';

proc sgplot data=diag;
  scatter y=res x=pred;
  refline 0;
run;


proc univariate data=diag normal plot;
  var res;
run;
```

part c)

```
proc transreg data=Q28;
  model boxcox (failuretime/convenient lambda=-2.0 to 2.0 by
          0.1)=class(material);
run;

data trans;
  set Q28;
  NewY=log(failuretime);
run;

proc glm data=trans;
  class material;
  model NewY=material;
  output out=newdiag p=pred r=res;
  run;
  title1 "Residual Plot";

proc sgplot data=newdiag;
  scatter y=res x=pred;
run;

proc univariate data=newdiag normal plot;
  var res;
run;
```

5. Montgomery 3.51

3.51. Use the Kruskal-Wallis test for the experiment in Problem 3.23. Are the results comparable to those found by the usual analysis of variance?

Kruskal-Wallis Test	
Chi-Square	6.2177
DF	3
Pr > Chi-Square	0.1015

Since the P-value of 0.1015 is greater than alpha = 0.05, we fail to reject the null hypothesis and conclude that the fluids do not differ. In the usual analysis of variance we also failed to reject the null, but because the P-value was so close to alpha, our conclusion was not strong.

SAS code:

```
proc npar1way data=Q323;
        class fluid;
        var life;
run;
```