STAT 571B, Homework 2
Ken Youens-Clark

1. Montgomery 3.3

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	3	36.15	<MST>	?	?
Error	?	<SSE>	<MSE>	?	
Total	19	196.04			

```
DF(Error) = N (20) - a (4) = 16
SS(Error) = SS(Total) - SST = 196.04 - 36.15 = 157.91
SSE = SS(Total) / SST = 196.04 / 36.15 = 5.423
MST = SST / a - 1 = 196.04 / 3 = 10.318
MSE = SSE / N - a = 5.423 / (20 - 4) = 0.3389
F = MST / MSE = 10.318 / 0.3389 = 30.446
P = F (3, 19) = 1.49
```

2. Montgomery 3.22 (skip part d)

a) Do the three circuits have the same response time?

```
> dat = read.delim("~/work/stat571/hw02/3.22.dat")
> dat$circuit = factor(dat$circuit)
> amod = aov(response ~ circuit, data=dat)
> amod.sum = unlist(summary(amod))
> amod.sum['Pr(>F)1']
     Pr(>F)1
0.0004023258
```

P << α = 0.01 ∴ reject H0 (no difference), response times are different

b) Tukey's test

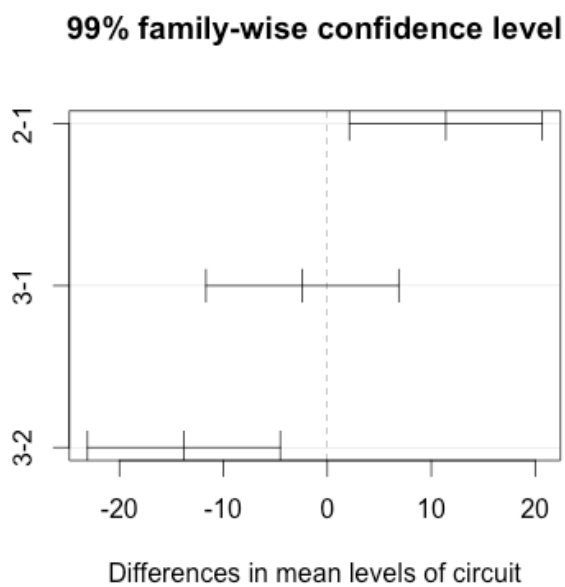
```
> TukeyHSD(amod, conf.level=0.99)
  Tukey multiple comparisons of means
    99% family-wise confidence level

Fit: aov(formula = response ~ circuit, data = dat)

$circuit
     diff        lwr       upr     p adj
2-1  11.4   2.123163 20.676837 0.0023656
3-1  -2.4 -11.676837  6.876837 0.6367043
3-2 -13.8 -23.076837 -4.523163 0.0005042

> library("multcomp")
```

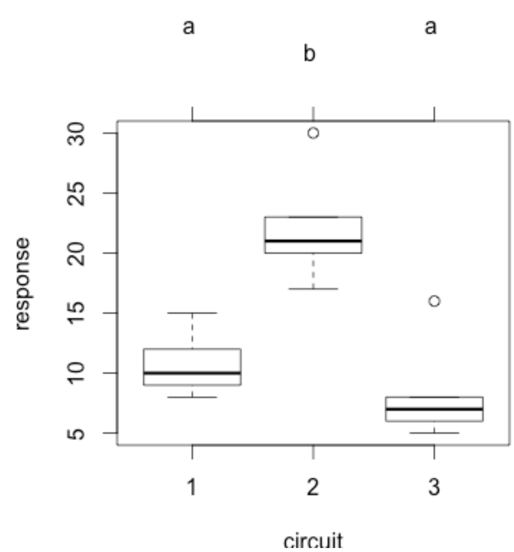
```
> tmod <- glht(amod, linfct = mcp(circuit="Tukey"))
> summary(tmod)

       Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts


Fit: aov(formula = response ~ circuit, data = dat)

Linear Hypotheses:
           Estimate Std. Error t value Pr(>|t|)
2 - 1 == 0     11.4        2.6   4.385   0.0024 **
3 - 1 == 0     -2.4        2.6  -0.923   0.6367
3 - 2 == 0    -13.8        2.6  -5.308   <0.001 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Adjusted p values reported -- single-step method)
```

The combination 3-1 is the most significantly similar. The others are different. (???)

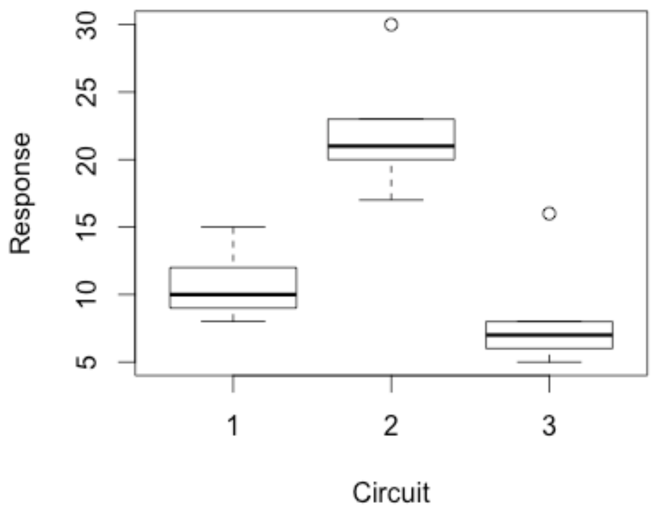
c) graphical comparison

The means of the three groups do not overlap.

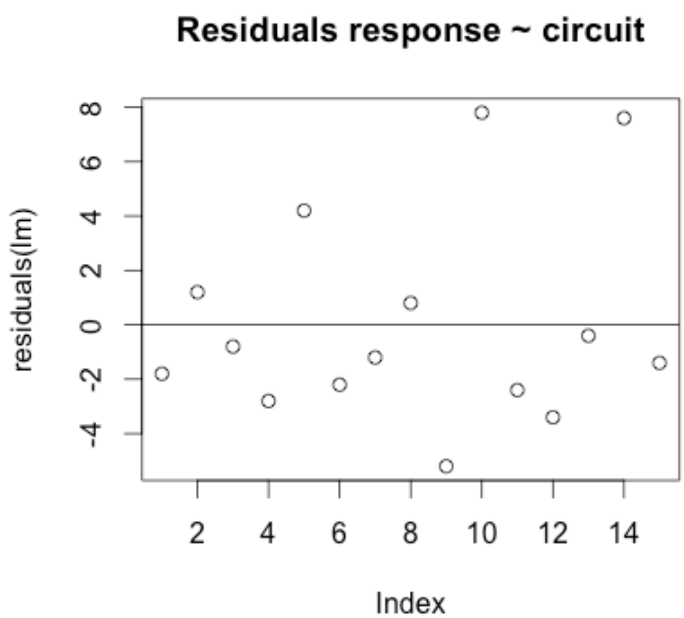




e) Circuit 3 seems to have the lowest response time which would be desirable for a shutoff valve.



f) The plot of residuals shows a random distribution of residuals, so basic analysis of variance assumptions are satisfied.



3. Montgomery 3.23

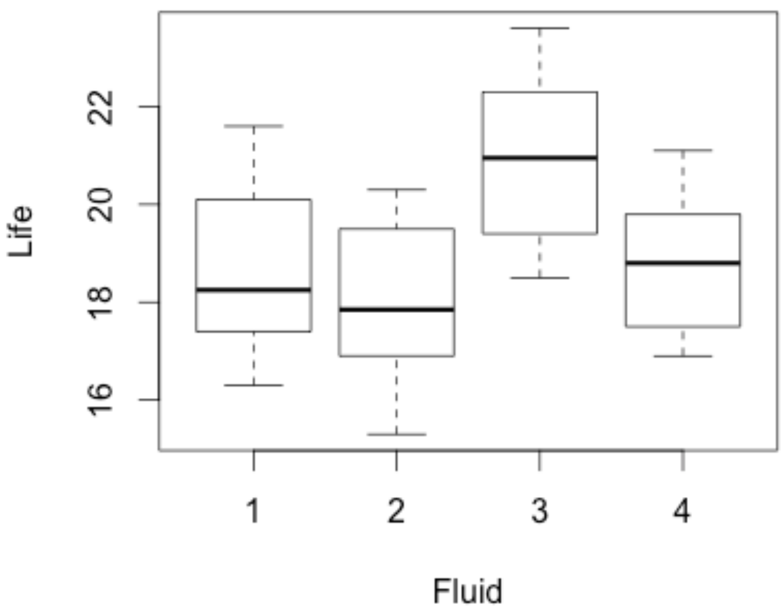
a) Do the fluids differ?

```
> dat = read.delim("~/work/stat571/hw02/3.23.long.dat")
> dat$fluid = factor(dat$fluid)
> amod = aov(life ~ fluid, data=dat)
> amod.sum = unlist(summary(amod))
```

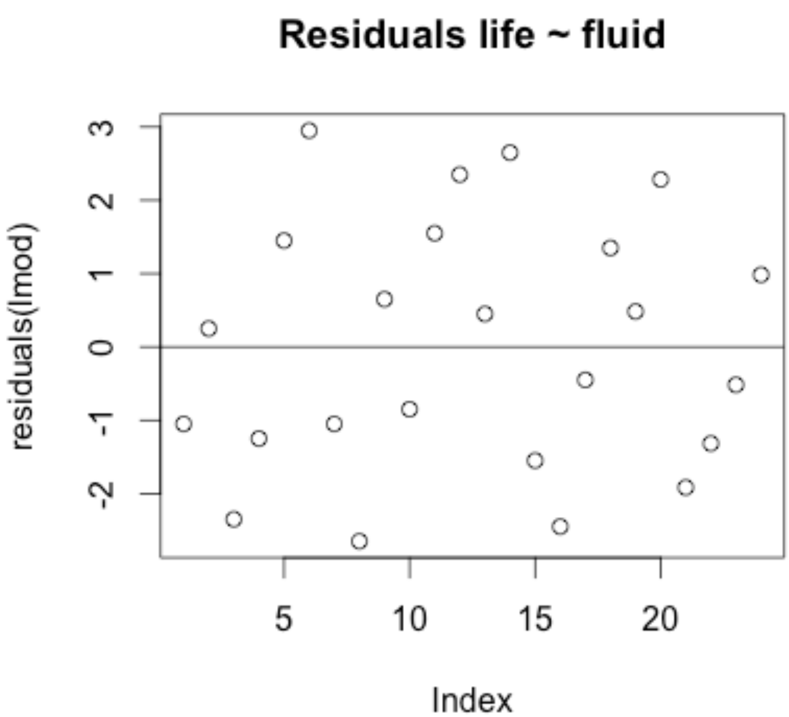
```
> amod.sum['Pr(>F)1']
   Pr(>F)1
0.05246316
```

$P > \alpha = 0.05$ ∴ reject null hypothesis, fluids are different

b) Fluid 3 has the longest life



c) Plot of residuals is random, so basic analysis of variance satisfied.



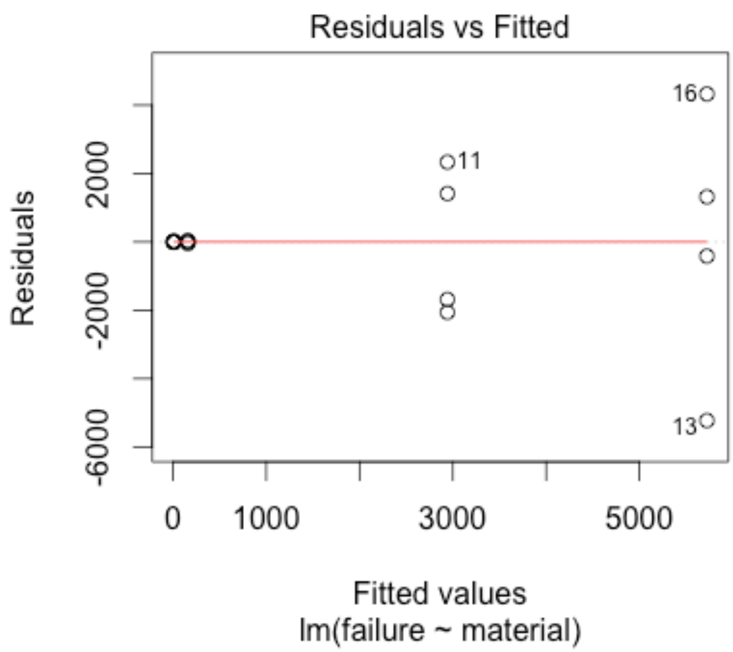
4. Montgomery 3.28

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

```
> dat = read.delim("~/work/stat571/hw02/3.28.dat")
> dat$material = factor(dat$material)
> amod = aov(failure ~ material, data=dat)
> amod.sum = unlist(summary(amod))
> amod.sum['Pr(>F)1']
    Pr(>F)1
0.003785956
```

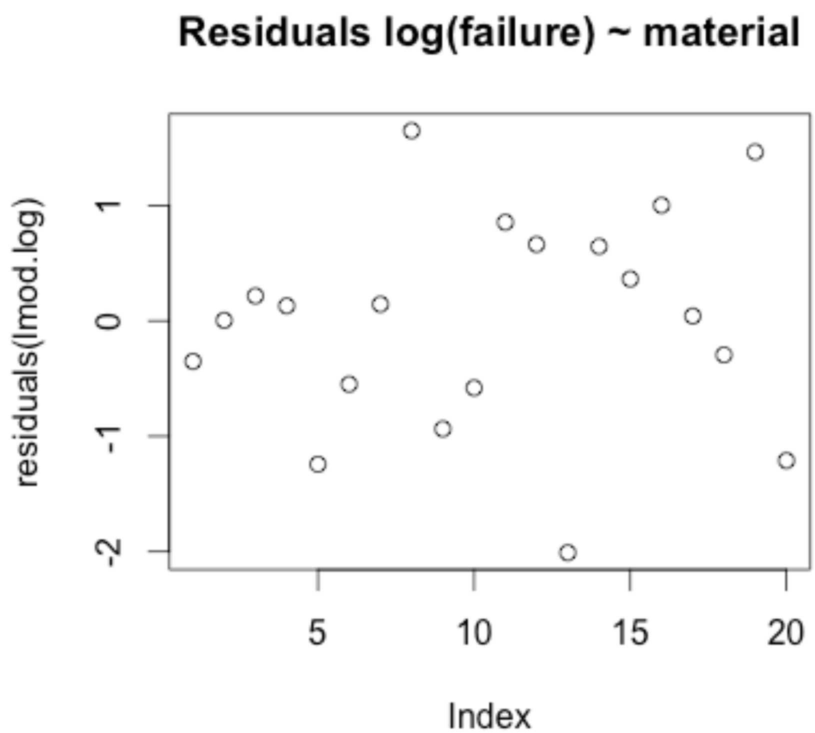
Very small p-value to support null hypothesis ∴ there is a difference.

b) Plot of residuals vs predicted shows poor variance (opening funnel to right).



c) Transform "failure" by log

```
> dat$log.failure = log(dat$failure)
> boxplot(log.failure ~ material, data=dat)
> amod2 = aov(log.failure ~ material, data=dat)
> amod2.sum = unlist(summary(amod2))
> amod2.sum['Pr(>F)1']
     Pr(>F)1
1.176093e-07
```



5. Montgomery 3.51

```
> dat = read.delim("~/work/stat571/hw02/3.23.long.dat")
> dat$fluid = factor(dat$fluid)
> kruskal.test(life ~ fluid, data=dat)

      Kruskal-Wallis rank sum test

data:  life by fluid
Kruskal-Wallis chi-squared = 6.2177, df = 3, p-value = 0.1015
```

Low p-value, so still reject null hypothesis.