R Handout - One-Way ANOVA	Dr. Derek Ogle
Winter 2015, MTH207 Biometry	Northland College

1 Initialization

```
> library(NCStats)
> library(multcomp)  # glht()
```

2 Raspberry Example

```
> rasp <- read.csv("https://raw.githubusercontent.com/droglenc/NCData/master/Raspberry.csv")
> str(rasp)

'data.frame': 16 obs. of  2 variables:
 $ water : int  100 100 100 100 200 200 200 200 400 400 ...
 $ weight: num  8.1 10.9 11.1 13.9 12.2 11.5 11.4 6.8 6.5 5.5 ...

> rasp$fwater <- factor(rasp$water)
> str(rasp)

'data.frame': 16 obs. of  3 variables:
 $ water : int  100 100 100 100 200 200 200 200 400 400 ...
 $ weight: num  8.1 10.9 11.1 13.9 12.2 11.5 11.4 6.8 6.5 5.5 ...
 $ fwater: Factor w/ 4 levels "100","200","400",..: 1 1 1 1 2 2 2 2 3 3 ...
```

2.1 Fitting the Linear Model

```
> lm1 <- lm(weight~fwater,data=rasp)
> anova(lm1)

          Df  Sum Sq Mean Sq F value   Pr(>F)
fwater     3 115.043  38.348  10.793 0.001004
Residuals 12  42.635   3.553
Total     15 157.678

> summary(lm1)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  11.0000     0.9425  11.672 6.58e-08
fwater200    -0.5250     1.3328  -0.394  0.70057
fwater400    -5.6250     1.3328  -4.220  0.00119
fwater800    -5.6000     1.3328  -4.202  0.00123


Residual standard error: 1.885 on 12 degrees of freedom
Multiple R-squared: 0.7296,Adjusted R-squared: 0.662
F-statistic: 10.79 on 3 and 12 DF,  p-value: 0.001004

> fitPlot(lm1,xlab="Water Treatment (ml)",ylab="Weight (g)",main="")

Loading required namespace:  sciplot
```
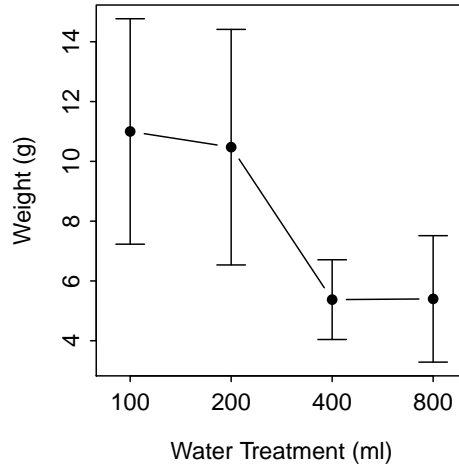
2.2 Multiple Comparison Tests

```
> rasp.mc <- glht(lm1, mcp(fwater = "Tukey"))
> summary(rasp.mc)

               Estimate Std. Error t value Pr(>|t|)
200 - 100 == 0   -0.525      1.333  -0.394  0.97833
400 - 100 == 0   -5.625      1.333  -4.220  0.00564
800 - 100 == 0   -5.600      1.333  -4.202  0.00597
400 - 200 == 0   -5.100      1.333  -3.826  0.01089
800 - 200 == 0   -5.075      1.333  -3.808  0.01167
800 - 400 == 0    0.025      1.333   0.019  1.00000

> confint(rasp.mc)

               Estimate lwr      upr
200 - 100 == 0 -0.5250  -4.4818   3.4318
400 - 100 == 0 -5.6250  -9.5818  -1.6682
800 - 100 == 0 -5.6000  -9.5568  -1.6432
400 - 200 == 0 -5.1000  -9.0568  -1.1432
800 - 200 == 0 -5.0750  -9.0318  -1.1182
800 - 400 == 0  0.0250  -3.9318   3.9818

> fitPlot(lm1,xlab="Water Treatment (ml)",ylab="Weight (g)",main="")
> addSigLetters(lm1,lets=c("a","a","b","b"),pos=c(2,4,2,4))
```
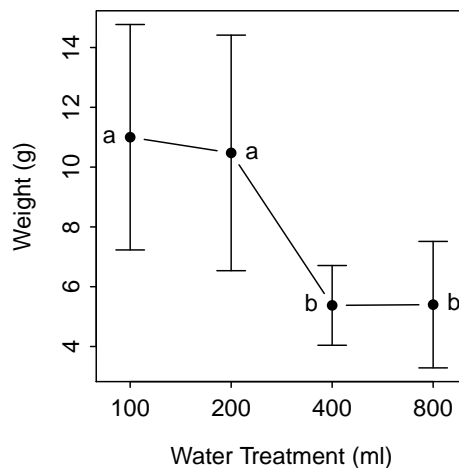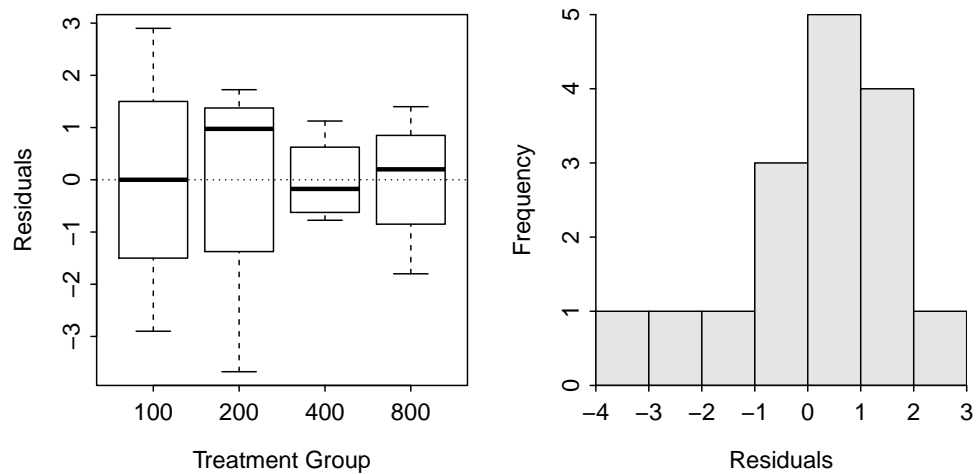
2.3 Checking the Assumptions

```
> leveneTest(lm1)
      Df F value Pr(>F)
group  3  0.3256 0.8069
      12
> residPlot(lm1)
```



```
> adTest(lm1$residuals)

Anderson-Darling normality test with lm1$residuals
A = 0.4308, p-value = 0.2688
> outlierTest(lm1)


No Studentized residuals with Bonferonni p < 0.05
Largest |rstudent|:
   rstudent unadjusted p-value Bonferonni p
8 -2.836044          0.016196      0.25914
```

# 3   Benthic Infaunal Example

```
> ben <- read.csv("https://raw.githubusercontent.com/droglenc/NCData/master/BenthicInfaunal.csv")
> ben$fsite <- factor(ben$site)
> str(ben)

'data.frame': 72 obs. of  3 variables:
 $ site     : int  1 1 1 1 1 1 1 1 2 2 ...
 $ abundance: num  14.4 20.4 21.2 17.6 29 ...
 $ fsite    : Factor w/ 9 levels "1","2","3","4",..: 1 1 1 1 1 1 1 1 2 2 ...
> lm2 <- lm(abundance~fsite,data=ben)
```

3.1   Assumption Checking with Possible Transformations

```
> leveneTest(lm2)
      Df F value   Pr(>F)
group  8  3.2452 0.003726
      63
> residPlot(lm2)
```
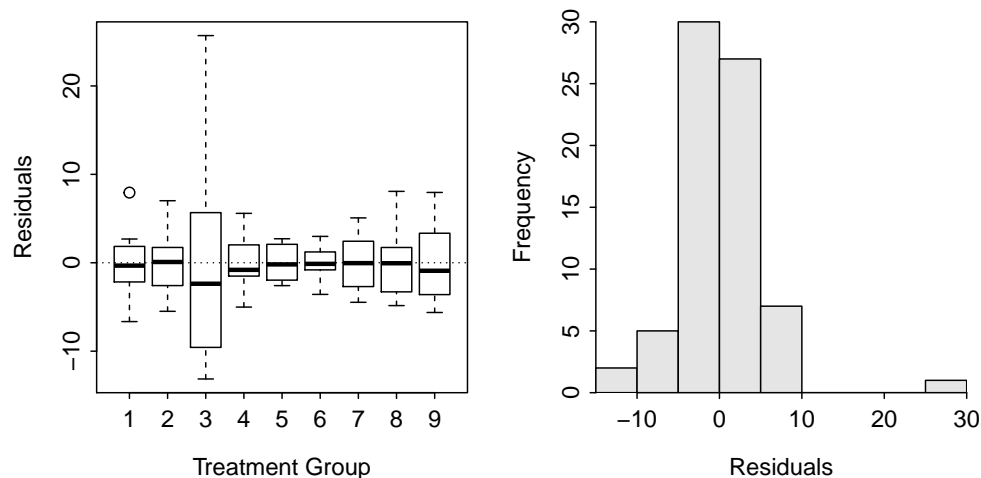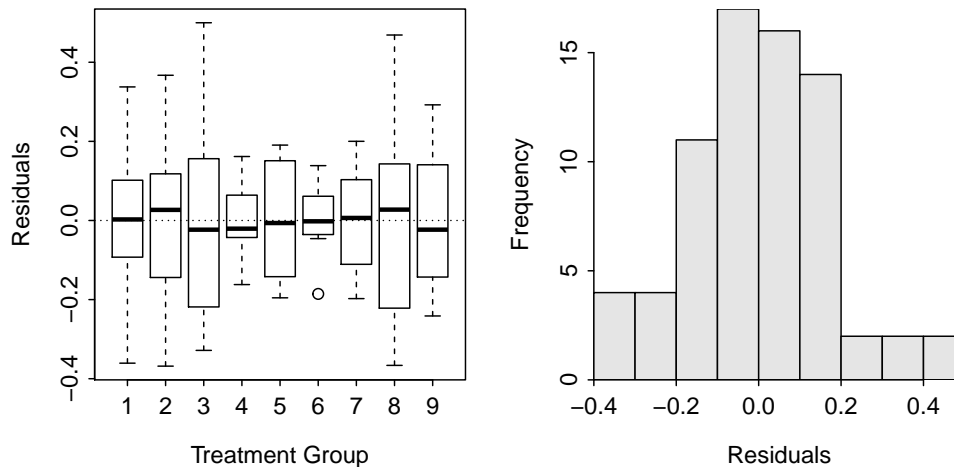


```
> adTest(lm2$residuals)

Anderson-Darling normality test with lm2$residuals
A = 1.6389, p-value = 0.0002996
> outlierTest(lm2)

   rstudent unadjusted p-value Bonferonni p
20 6.624666         9.5554e-09   6.8799e-07
```

```
> ## transChooser(lm2)  # interactive, result not shown
> ben$logab <- log(ben$abundance)
> lm3 <- lm(logab~fsite,data=ben)
> leveneTest(lm3)
      Df F value Pr(>F)
group  8  1.5339 0.1636
      63
> residPlot(lm3)
```

```
> adTest(lm3$residuals)

Anderson-Darling normality test with lm3$residuals
A = 0.3323, p-value = 0.5062
> outlierTest(lm3)


No Studentized residuals with Bonferonni p < 0.05
Largest |rstudent|:
   rstudent unadjusted p-value Bonferonni p
20 2.928889           0.004754      0.34229
```

3.2   Model Summarization

```
> anova(lm3)
          Df  Sum Sq Mean Sq F value    Pr(>F)
fsite      8  8.6683 1.08353  29.066 < 2.2e-16
Residuals 63  2.3485 0.03728
Total     71 11.0168
> ben.mc <- glht(lm3, mcp(fsite = "Dunnett"))
> summary(ben.mc)

            Estimate Std. Error t value Pr(>|t|)
2 - 1 == 0 -0.218435   0.096537  -2.263  0.14568
3 - 1 == 0  0.703189   0.096537   7.284  < 0.001
4 - 1 == 0  0.453836   0.096537   4.701  < 0.001
5 - 1 == 0 -0.414859   0.096537  -4.297  < 0.001
6 - 1 == 0 -0.004238   0.096537  -0.044  1.00000
7 - 1 == 0  0.140280   0.096537   1.453  0.57953
8 - 1 == 0 -0.371867   0.096537  -3.852  0.00199
9 - 1 == 0  0.168668   0.096537   1.747  0.37938

> fitPlot(lm3,ylab="Log Abundance",xlab="Site",main="")
> addSigLetters(lm3,lets=c("","","***","***","***","","","***",""),pos=c(2,4,2,4,2,2,4,2,4))
```
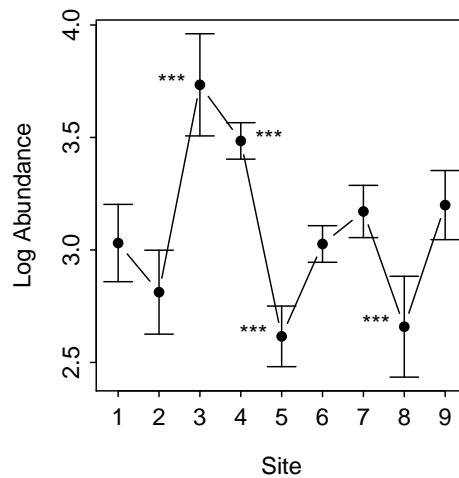
```
> confint(ben.mc)

          Estimate  lwr        upr
2 - 1 == 0 -0.218435 -0.481870  0.045001
3 - 1 == 0  0.703189  0.439753  0.966624
4 - 1 == 0  0.453836  0.190401  0.717272
5 - 1 == 0 -0.414859 -0.678295 -0.151424
6 - 1 == 0 -0.004238 -0.267673  0.259198
7 - 1 == 0  0.140280 -0.123155  0.403716
8 - 1 == 0 -0.371867 -0.635303 -0.108432
9 - 1 == 0  0.168668 -0.094767  0.432103

> exp(confint(ben.mc)$confint)

       Estimate       lwr       upr
2 - 1 0.8037761 0.6176700 1.0459565
3 - 1 2.0201841 1.5524313 2.6288723
4 - 1 1.5743404 1.2098181 2.0486945
5 - 1 0.6604332 0.5075167 0.8594239
6 - 1 0.9957713 0.7652107 1.2958006
7 - 1 1.1505965 0.8841877 1.4972750
8 - 1 0.6894457 0.5298117 0.8971780
9 - 1 1.1837272 0.9096474 1.5403882
attr(,"conf.level")
[1] 0.95
attr(,"calpha")
[1] 2.728132
```