

One-Way ANOVA Handout

1 Initialization

```
> library(NCStats)
```

2 Raspberry Example

You must change the directory to where the following file is located.

```
> rasp <- read.table("Raspberry.txt",head=TRUE)
> 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)
> attach(rasp)
```

2.1 Fitting the Linear Model

```
> lm1 <- lm(weight~fwater)
> anova(lm1)
```

Analysis of Variance Table

```
Response: weight
      Df Sum Sq Mean Sq F value    Pr(>F)
fwater  3 115.042   38.347   10.793 0.001004
Residuals 12  42.635    3.553
Total    15 157.677
```

```
> summary(lm1)
```

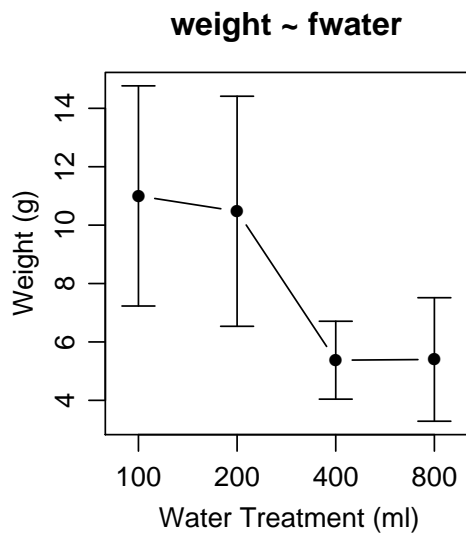
```
Call:
lm(formula = weight ~ fwater)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-3.6750 -0.5500  0.1125  1.0500  2.9000
```

```
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
```

```
> fit.plot(lm1,xlab="Water Treatment (ml)",ylab="Weight (g)")
```

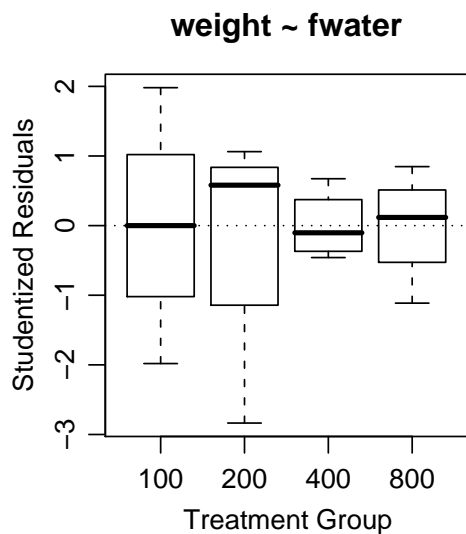


2.2 Checking the Assumptions

```
> levene.test(lm1)
```

```
Levene's Test for Homogeneity of Variance
  Df F value Pr(>F)
group 3  0.3256 0.8069
  12
```

```
> residual.plot(lm1)
```

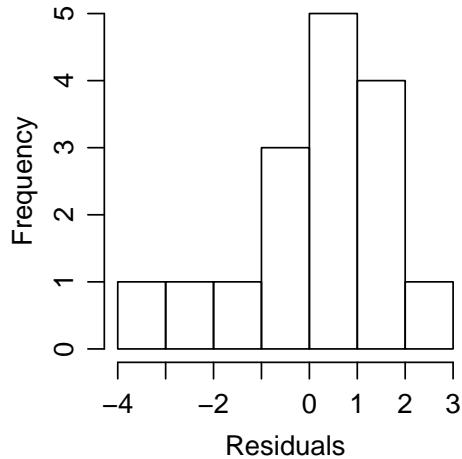


```
> ad.test(lm1$residuals)
```

Anderson-Darling normality test

```
data: lm1$residuals
A = 0.4308, p-value = 0.2688
```

```
> hist(lm1$residuals,xlab="Residuals",main="")
```



```
> outlier.test(lm1)
```

```
max|rstudent| = 2.836044, degrees of freedom = 11,
unadjusted p = 0.0161961, Bonferroni p = 0.2591376
```

```
Observation: 8
```

2.3 Multiple Comparison Tests

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

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: lm(formula = weight ~ fwater)

Linear Hypotheses:

| | Estimate | Std. Error | t value | p value |
|----------------|----------|------------|---------|---------|
| 200 - 100 == 0 | -0.525 | 1.333 | -0.394 | 0.97832 |
| 400 - 100 == 0 | -5.625 | 1.333 | -4.220 | 0.00581 |
| 800 - 100 == 0 | -5.600 | 1.333 | -4.202 | 0.00587 |
| 400 - 200 == 0 | -5.100 | 1.333 | -3.826 | 0.01110 |
| 800 - 200 == 0 | -5.075 | 1.333 | -3.808 | 0.01153 |
| 800 - 400 == 0 | 0.025 | 1.333 | 0.019 | 1.00000 |

(Adjusted p values reported -- single-step method)

```
> confint(rasp.mc)
```

Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts

Fit: $\text{lm}(\text{formula} = \text{weight} \sim \text{fwater})$

Estimated Quantile = 2.9675

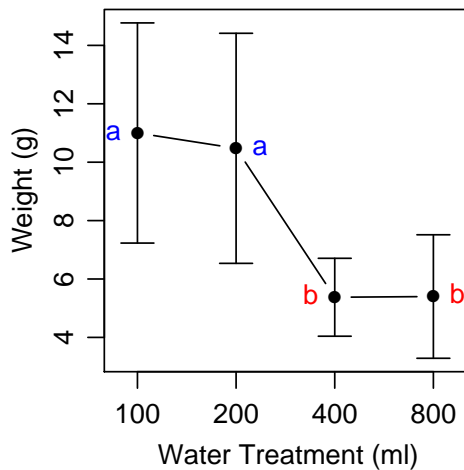
95% family-wise confidence level

Linear Hypotheses:

| | Estimate | lwr | upr |
|----------------|----------|---------|---------|
| 200 - 100 == 0 | -0.5250 | -4.4802 | 3.4302 |
| 400 - 100 == 0 | -5.6250 | -9.5802 | -1.6698 |
| 800 - 100 == 0 | -5.6000 | -9.5552 | -1.6448 |
| 400 - 200 == 0 | -5.1000 | -9.0552 | -1.1448 |
| 800 - 200 == 0 | -5.0750 | -9.0302 | -1.1198 |
| 800 - 400 == 0 | 0.0250 | -3.9302 | 3.9802 |

```
> fit.plot(lm1,xlab="Water Treatment (ml)",ylab="Weight (g)",main="")
```

```
> add.sig.letters(lm1,lets=c("a", "a", "b", "b"),pos=c(2,4,2,4),col=c("blue", "blue", "red", "red"))
```



```
> detach(rasp)
```

3 Benthic Infaunal Example

It is assumed that the initialization steps shown in Section 1 have already been followed and that the working directory has been changed to where the external data file is located.

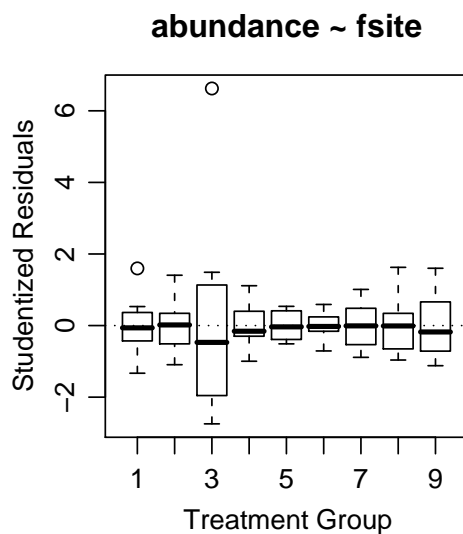
```
> ben <- read.table("BenthicInfaunal.txt", head=TRUE)
> ben$fsite <- factor(ben$site)
> attach(ben)
> lm2 <- lm(abundance~fsite)
```

3.1 Assumption Checking with Possible Transformations

```
> levene.test(lm2)
```

```
Levene's Test for Homogeneity of Variance
      Df F value Pr(>F)
group  8  3.2452 0.003726
      63
```

```
> residual.plot(lm2)
```

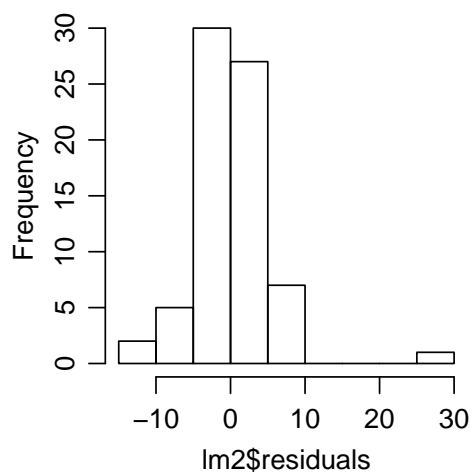


```
> ad.test(lm2$residuals)
```

```
Anderson-Darling normality test
```

```
data: lm2$residuals
A = 1.6389, p-value = 0.0002996
```

```
> hist(lm2$residuals, main="")
```

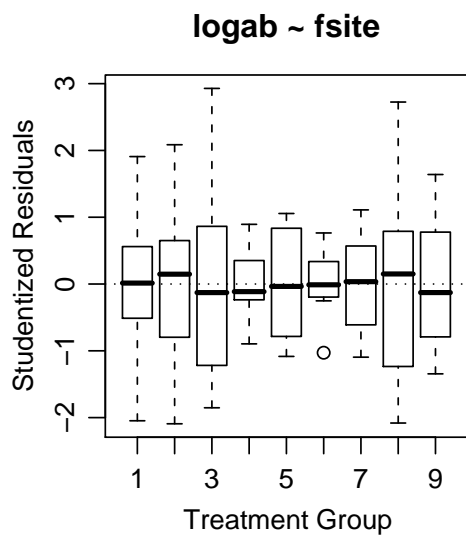


The following function was used to determine that a log transformation would most likely lead to the assumptions being met. This function cannot be illustrated in a handout because it requires interactions from the user.

```
> trans.chooser(lm2)
> detach(ben)
> ben$logab <- log(ben$abundance)
> attach(ben)
> lm3 <- lm(logab~fsite)
> levene.test(lm3)
```

```
Levene's Test for Homogeneity of Variance
  Df F value Pr(>F)
group 8  1.5339 0.1636
 63
```

```
> residual.plot(lm3)
```



```
> ad.test(lm3$residuals)
```

Anderson-Darling normality test

```
data: lm3$residuals
A = 0.3323, p-value = 0.5062
```

```
> outlier.test(lm3)
```

```
max|rstudent| = 2.928889, degrees of freedom = 62,
unadjusted p = 0.004753989, Bonferroni p = 0.3422872
```

```
Observation: 20
```

3.2 Model Summarization

```
> anova(lm3)
```

Analysis of Variance Table

Response: logab

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|-----------|----|---------|---------|---------|-----------|
| fsite | 8 | 8.6683 | 1.0835 | 29.066 | < 2.2e-16 |
| Residuals | 63 | 2.3485 | 0.0373 | | |
| Total | 71 | 11.0168 | | | |

```
> ben.mc <- glht(lm3, mcp(fsites = "Dunnett"))
> summary(ben.mc)
```

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Dunnett Contrasts

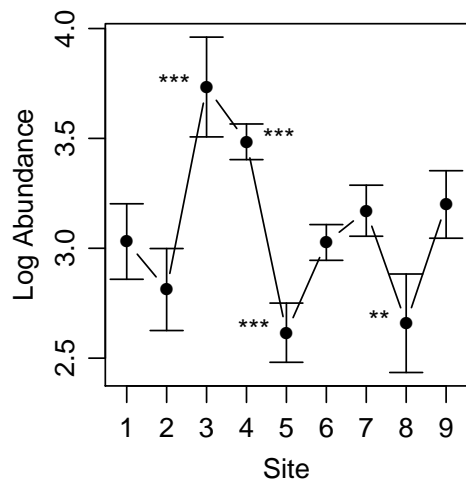
Fit: lm(formula = logab ~ fsites)

Linear Hypotheses:

| | Estimate | Std. Error | t value | p value |
|------------|-----------|------------|---------|---------|
| 2 - 1 == 0 | -0.218435 | 0.096537 | -2.263 | 0.14558 |
| 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 | 0.57962 |
| 8 - 1 == 0 | -0.371867 | 0.096537 | -3.852 | 0.00196 |
| 9 - 1 == 0 | 0.168668 | 0.096537 | 1.747 | 0.37950 |

(Adjusted p values reported -- single-step method)

```
> fit.plot(lm3, ylab="Log Abundance", xlab="Site", main="")
> add.sig.letters(lm3, lets=c("", "", "***", "***", "***", "", "", "**", ""), pos=c(2,4,2,4,2,2,4,2,4))
```



```
> logdiff <- confint(ben.mc)$confint
> exp(logdiff)
```

```
      Estimate      lwr      upr
2 - 1 0.8037761 0.6176613 1.0459714
3 - 1 2.0201841 1.5524093 2.6289097
4 - 1 1.5743404 1.2098009 2.0487236
5 - 1 0.6604332 0.5075095 0.8594361
6 - 1 0.9957713 0.7651999 1.2958190
7 - 1 1.1505965 0.8841752 1.4972963
8 - 1 0.6894457 0.5298041 0.8971908
9 - 1 1.1837272 0.9096344 1.5404100
attr(,"conf.level")
[1] 0.95
attr(,"calpha")
[1] 2.728279
attr(,"error")
[1] 8.446623e-05
```

```
> detach(ben)
```