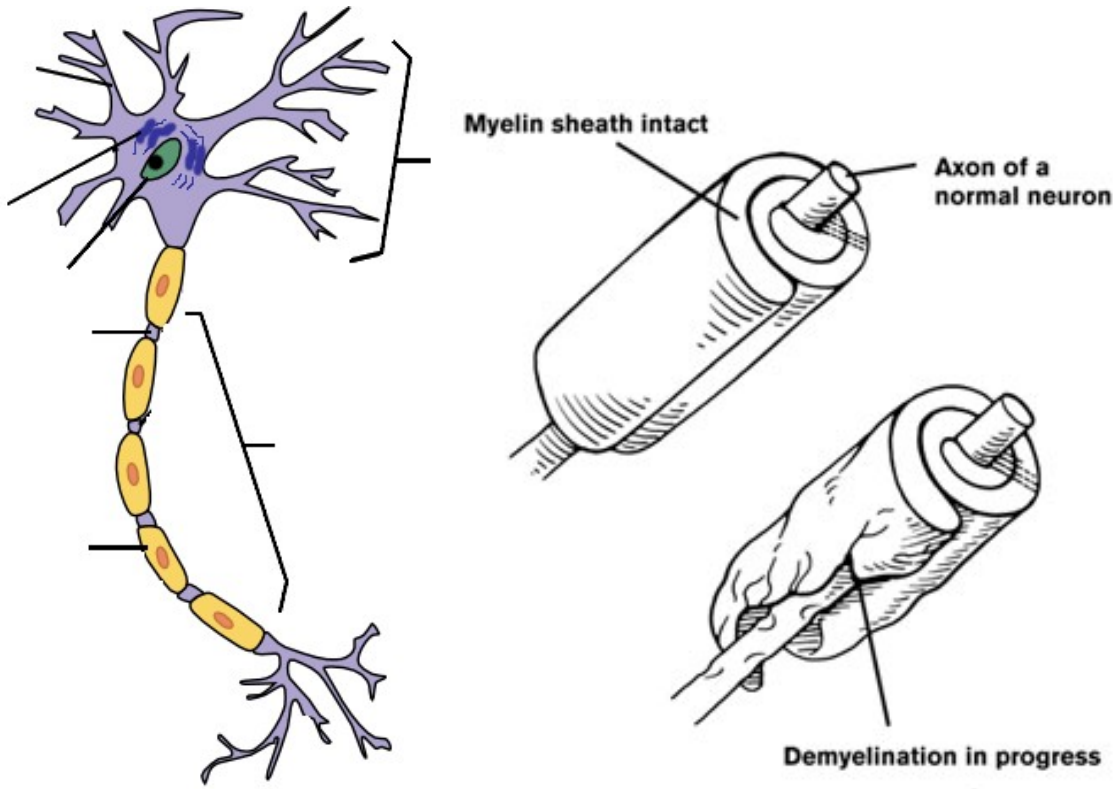
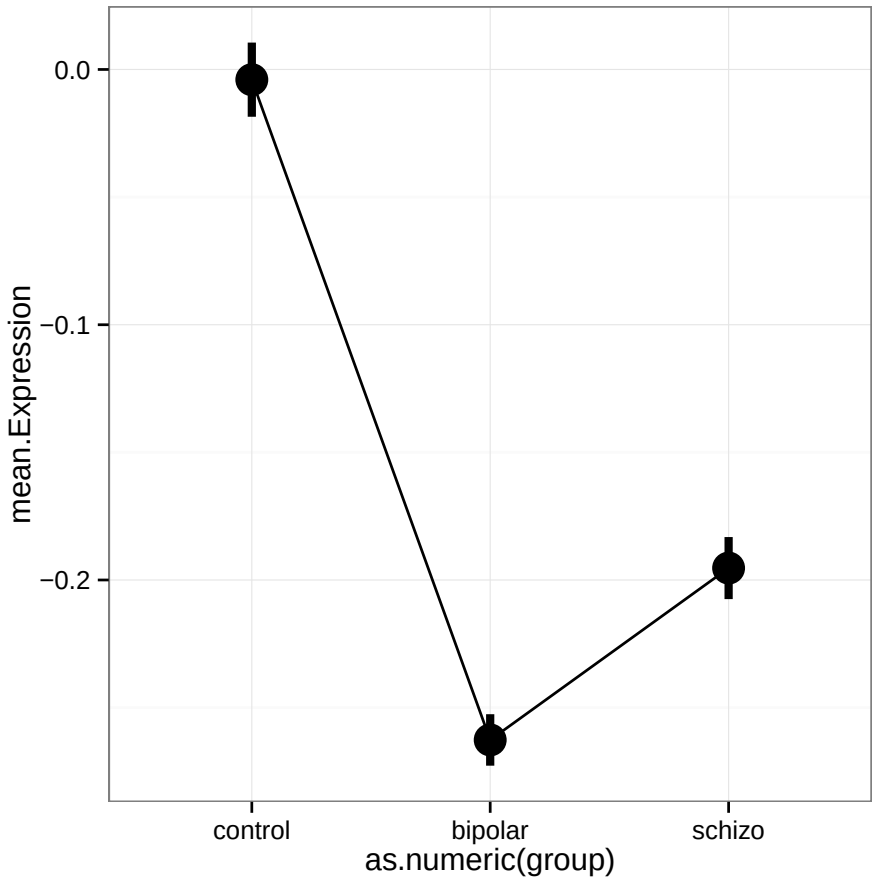
After the ANOVA

Categorical Predictors: Gene Expression and Mental Disorders



The Data


0.0
−0.1
−0.2
mean.Expression
control
bipolar
schizo
as.numeric(group)

Fit the Data with a Linear Model

```
bg.sub.lm <- lm(PLP1.expression ~ group, data=brainGene)
```

F-Test to Compare Variation Within versus Between Groups

$$SS_{Total} = SS_{Between} + SS_{Within}$$

$$SS_{Between} = \sum_i \sum_j (\bar{Y}_i - \bar{Y})^2, \text{ df=k-1}$$

$$SS_{Within} = \sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2, \text{ df=n-k}$$

F-Test to Compare Variation Within versus Between Groups

$$SS_{Total} = SS_{Between} + SS_{Within}$$

$$SS_{Between} = \sum_i \sum_j (\bar{Y}_i - \bar{Y})^2, \text{ df=k-1}$$

$$SS_{Within} = \sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2, \text{ df=n-k}$$

MS = SS/DF, e.g, $MS_W = \frac{SS_W}{n-k}$

$$F = \frac{MS_B}{MS_W} \text{ with DF=k-1,n-k}$$

ANOVA

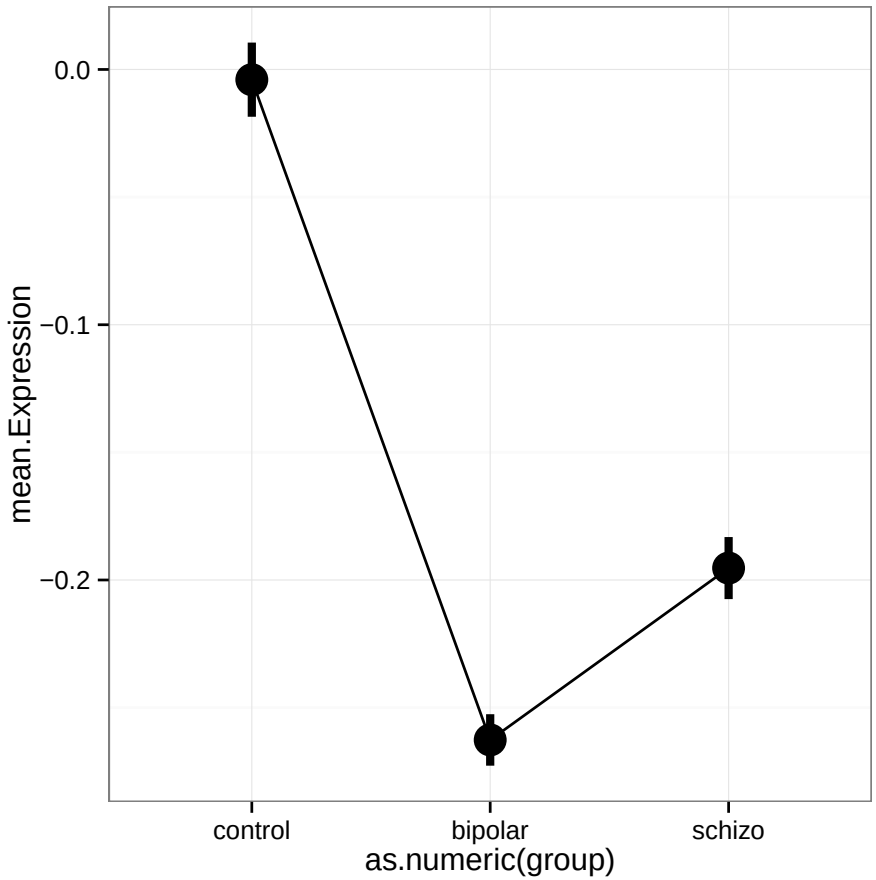
```
anova(bg.sub.lm)

# Analysis of Variance Table
#
# Response: PLP1.expression
#           Df Sum Sq Mean Sq F value Pr(>F)
# group      2   0.54  0.2701    7.82 0.0013
# Residuals 42   1.45  0.0345
```

ANOVA

```
anova(bg.sub.lm)

# Analysis of Variance Table
#
# Response: PLP1.expression
#           Df Sum Sq Mean Sq F value Pr(>F)
# group      2   0.54  0.2701    7.82 0.0013
# Residuals 42   1.45  0.0345
```

Which groups are different from one another?

The Data


0.0
−0.1
−0.2
mean.Expression
control
bipolar
schizo
as.numeric(group)

The Coefficients

```
summary(bg.sub.lm)

#
# Call:
# lm(formula = PLP1.expression ~ group, data = brainGene)
#
# Residuals:
#     Min      1Q  Median      3Q     Max
# -0.2960 -0.1273 -0.0347  0.0753  0.4840
#
# Coefficients:
#             Estimate Std. Error t value Pr(>|t|)
# (Intercept)   -0.0040     0.0480   -0.08  0.93395
# groupbipolar  -0.2587     0.0678   -3.81  0.00044
# groupschizo   -0.1913     0.0678   -2.82  0.00730
#
# Residual standard error: 0.186 on 42 degrees of freedom
# Multiple R-squared: 0.271, Adjusted R-squared: 0.237
```

Default "Treatment" Contrasts

```
contrasts(brainGene$group)

#         bipolar schizo
# control       0      0
# bipolar       1      0
# schizo        0      1
```

The Coefficients

```
summary(lm(PLP1.expression ~ group -1, data=brainGene))

#
# Call:
# lm(formula = PLP1.expression ~ group - 1, data = brainGene)
#
# Residuals:
#     Min      1Q  Median      3Q     Max
# -0.2960 -0.1273 -0.0347  0.0753  0.4840
#
# Coefficients:
#             Estimate Std. Error t value Pr(>|t|)
# groupcontrol   -0.004      0.048   -0.08   0.9340
# groupbipolar   -0.263      0.048   -5.47  2.3e-06
# groupschizo    -0.195      0.048   -4.07   0.0002
#
# Residual standard error: 0.186 on 42 degrees of freedom
# Multiple R-squared: 0.526, Adjusted R-squared: 0.492
```

OK, but WHICH GROUPS ARE DIFFERENT?

ANOVA is an Omnibus Test

Remember your Null:

$$H_0 = \mu_1 = \mu 2 = \mu 3 = ...$$

This had nothing to do with specific comparisons of means.

A priori contrasts

Specific sets of *a* priori null hypotheses:

$$\mu_1 = \mu 2$$

$$\mu_1 = \mu 3 = ...$$

Use t-tests.

A priori contrasts

```
library(contrast)

contrast(bg.sub.lm, list(group="control"),
         list(group="schizo"))

# lm model parameter contrast
#
#   Contrast    S.E.  Lower  Upper    t df Pr(>|t|)
# 1   0.1913 0.06785 0.0544 0.3283 2.82 42   0.0073
```

A priori contrasts

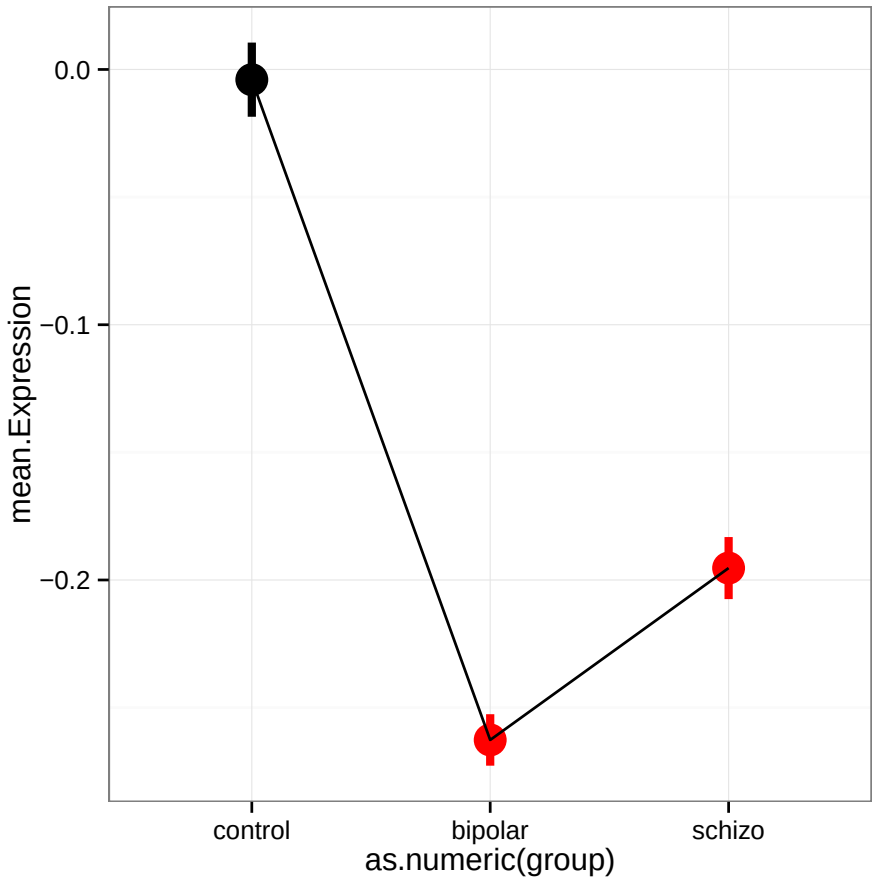
```
contrast(bg.sub.lm, list(group="control"),
         list(group=c("schizo", "bipolar")))

# lm model parameter contrast
#
#  Contrast    S.E.  Lower  Upper    t df Pr(>|t|)
#    0.1913 0.06785 0.0544 0.3283 2.82 42   0.0073
#    0.2587 0.06785 0.1217 0.3956 3.81 42   0.0004
```

A priori contrasts

```
contrast(bg.sub.lm, list(group="control"),
         list(group=c("schizo", "bipolar")))

# lm model parameter contrast
#
#  Contrast    S.E.  Lower  Upper    t df Pr(>|t|)
#    0.1913 0.06785 0.0544 0.3283 2.82 42   0.0073
#    0.2587 0.06785 0.1217 0.3956 3.81 42   0.0004
```

Note: can only do k-1, as each takes 1df

The Data


0.0
−0.1
−0.2
mean.Expression
control
bipolar
schizo
as.numeric(group)

Orthogonal A priori contrasts

Sometimes you want to test very specific hypotheses about the structure of your groups

```
#                       control bipolar schizo
# Control v. Disorders        1    -0.5   -0.5
# Bipolar v. Schizo           0     1.0   -1.0
```

Orthogonal A priori contrasts

Sometimes you want to test very specific hypotheses about the structure of your groups

```
#                       control bipolar schizo
# Control v. Disorders        1    -0.5   -0.5
# Bipolar v. Schizo           0     1.0   -1.0
```

Note: can only do k-1, as each takes 1df

Orthogonal A priori contrasts with multcomp

```
library(multcomp)
#
bg_orthogonal <- glht(bg.sub.lm, linfct=contrast_mat,
                      test=adjusted("none"))
#
summary(bg_orthogonal)
```

Orthogonal A priori contrasts with multcomp

```
library(multcomp)
#
bg_orthogonal <- glht(bg.sub.lm, linfct=contrast_mat,
                      test=adjusted("none"))
#
summary(bg_orthogonal)
```

Note adjusted p-value is set to none...

Orthogonal A priori contrasts

```
#
#   Simultaneous Tests for General Linear Hypotheses
#
# Fit: lm(formula = PLP1.expression ~ group, data = brainGene)
#
# Linear Hypotheses:
#                            Estimate Std. Error t value
# Control v. Disorders == 0   0.2210     0.1018    2.17
# Bipolar v. Schizo == 0     -0.0673     0.0679   -0.99
#                            Pr(>|t|)
# Control v. Disorders == 0     0.07
# Bipolar v. Schizo == 0        0.54
# (Adjusted p values reported -- single-step method)
```

Post hoc contrasts

I want to test all possible comparisons!

Post hoc contrasts

Only to be done if you reject Ho

Post hoc contrasts

Only to be done if you reject Ho

- All possible comparisons via t-test
- But...with many comparisons, does type I error rate increase?

Post hoc contrasts

Only to be done if you reject Ho

- All possible comparisons via t-test
- But...with many comparisons, does type I error rate increase?
- Consider adjusted alpha
- But, adjusting alpha also may increase type II error rate!

Post hoc contrasts

Only to be done if you reject Ho

- All possible comparisons via t-test
- But...with many comparisons, does type I error rate increase?
- Consider adjusted alpha
- But, adjusting alpha also may increase type II error rate!
- Additional multiple comparison methods calulate family-wise critical values of differences.

All Possible T-Tests

```
with( brainGene, pairwise.t.test(PLP1.expression, group,
                                 p.adjust.method ="none") )

#
#  Pairwise comparisons using t tests with pooled SD
#
# data:  PLP1.expression and group
#
#         control bipolar
# bipolar 0.00044 -
# schizo  0.00730 0.32671
#
# P value adjustment method: none
```

P-Value Adjustments

Bonferroni : $\alpha_{adj} = \frac{\alpha}{m}$ where m $= \#$ of tests
- VERY conservative

P-Value Adjustments

Bonferroni : $\alpha_{adj} = \frac{\alpha}{m}$ where m $= \#$ of tests
- VERY conservative

False Discovery Rate: $\alpha_{adj} = \frac{k\alpha}{m}$
- Order your p values from smallest to largest, rank $=$ k,
- Adjusts for small v. large p values
- Less conservative

P-Value Adjustments

Bonferroni : $\alpha_{adj} = \frac{\alpha}{m}$ where m = # of tests
- VERY conservative

False Discovery Rate: $\alpha_{adj} = \frac{k\alpha}{m}$
- Order your p values from smallest to largest, rank = k,
- Adjusts for small v. large p values
- Less conservative

Other Methods: Sidak, Dunn, Holm, etc.

P-Value Adjustments

Bonferroni : $\alpha_{adj} = \frac{\alpha}{m}$ where m = # of tests
- VERY conservative

False Discovery Rate: $\alpha_{adj} = \frac{k\alpha}{m}$
- Order your p values from smallest to largest, rank = k,
- Adjusts for small v. large p values
- Less conservative

Other Methods: Sidak, Dunn, Holm, etc.
We're very focused on p here!

Bonferroni Correction

```
with( brainGene, pairwise.t.test(PLP1.expression, group,
                                 p.adjust.method ="bonferroni") )

#
#  Pairwise comparisons using t tests with pooled SD
#
# data:  PLP1.expression and group
#
#         control bipolar
# bipolar 0.0013  -
# schizo  0.0219  0.9801
#
# P value adjustment method: bonferroni
```

False Discovery Rate

```
with( brainGene, pairwise.t.test(PLP1.expression, group,
                                 p.adjust.method ="fdr") )

#
#  Pairwise comparisons using t tests with pooled SD
#
# data:  PLP1.expression and group
#
#         control bipolar
# bipolar 0.0013  -
# schizo  0.0110  0.3267
#
# P value adjustment method: fdr
```

Other Methods Use Critical Values

- Tukey's Honestly Significant Difference
- Dunnet's Test for Comparison to Controls
- Ryan's Q (sliding range)
- etc...

Tukey Test

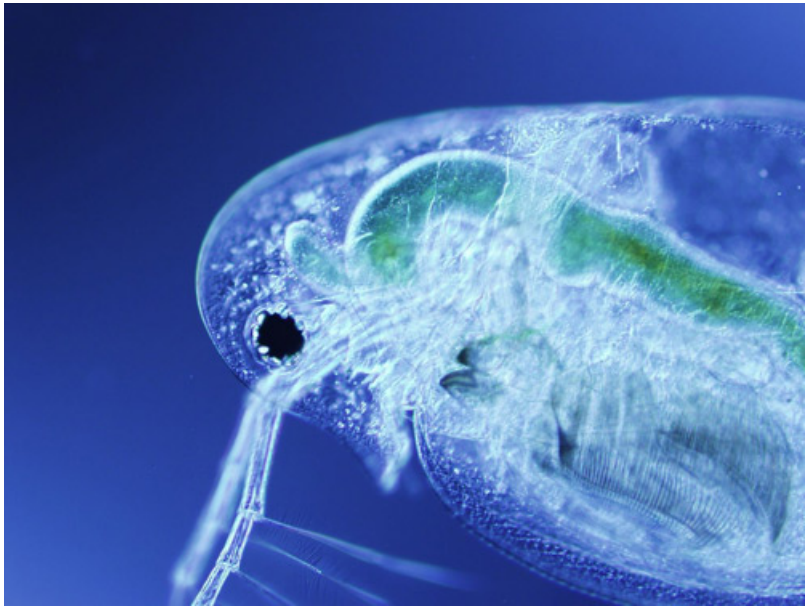
```
bg.sub.aov <- aov(PLP1.expression ~ group, data=brainGene)
TukeyHSD(bg.sub.aov)

#   Tukey multiple comparisons of means
#     95% family-wise confidence level
#
# Fit: aov(formula = PLP1.expression ~ group, data = brainGene)
#
# $group
#                     diff      lwr      upr  p adj
# bipolar-control -0.25867 -0.42351 -0.09382 0.0013
# schizo-control  -0.19133 -0.35618 -0.02649 0.0196
# schizo-bipolar   0.06733 -0.09751  0.23218 0.5857
```

Final Notes of Caution

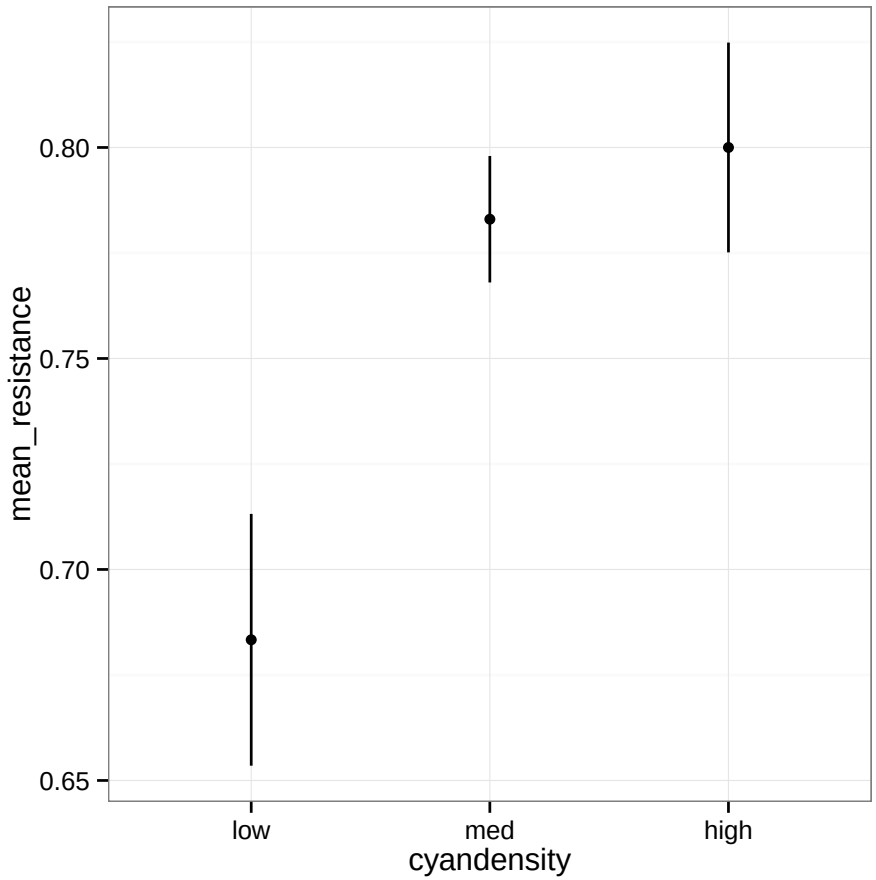
- Often you DO have a priori contrasts in mind
- If you reject Ho with ANOVA, differences between groups exist
- Consider Type I v. Type II error before correcting

Exercise: Daphnia Resistance

- Fit an ANOVA
- Which groups are different?



Daphnia Data



ANOVA shows an Effect

```
daphniaLM <- lm(resistance ~ cyandensity, data=daphnia)
anova(daphniaLM)

# Analysis of Variance Table
#
# Response: resistance
#             Df Sum Sq Mean Sq F value Pr(>F)
# cyandensity  2 0.0892  0.0446    6.69 0.0041
# Residuals   29 0.1933  0.0067
```

High and Med Not Different

```
summary( glht(daphniaLM, linfct=mcp(cyandensity="Tukey")),
         test=adjusted("none"))


#
#   Simultaneous Tests for General Linear Hypotheses
#
# Multiple Comparisons of Means: Tukey Contrasts
#
#
# Fit: lm(formula = resistance ~ cyandensity, data = daphnia)
#
# Linear Hypotheses:
#                  Estimate Std. Error t value Pr(>|t|)
# med - low == 0    0.0997     0.0350    2.85   0.0079
# high - low == 0   0.1167     0.0350    3.34   0.0023
# high - med == 0   0.0170     0.0365    0.47   0.6450
# (Adjusted p values reported -- none method)
```