

Kneitel_2010_algae_fdr2.R

SIU850486795

2023-11-14

```
# Kneitel_2010_algae_fdr2.sas
# One-way ANOVA and FDR method
# Load necessary libraries
library(ggplot2)
library(multcomp)

## Warning: package 'multcomp' was built under R version 4.0.4
## Loading required package: mvtnorm
## Loading required package: survival
## Loading required package: TH.data
## Warning: package 'TH.data' was built under R version 4.0.4
## Loading required package: MASS
##
## Attaching package: 'TH.data'
## The following object is masked from 'package:MASS':
##
##      geyser

# Read in data set
kndata <- read.table(header=T,colClasses=c("factor","numeric","numeric","numeric"),
text="
treat richness total algae
Control 8 78 1
Control 5 84 7
Control 10 115 45
Control 7 200 100
Control 6 72 20
Low 8 73 15
Low 7 124 70
Low 8 116 50
Low 8 92 5
Low 7 138 60
Medium 7 124 85
Medium 8 116 80
Medium 8 145 60
Medium 6 154 100
Medium 7 129 90
High 6 134 95
High 7 138 95
High 8 103 70
```

```

High      8  119  75
High      6  132  80
VeryHigh  6  148  95
VeryHigh  5  134  95
VeryHigh  5  119 100
VeryHigh  5  117  90
VeryHigh  5  129  80
")

```

Apply transformations here

```
kndata <- transform(kndata,y=asin(sqrt(algae/100)))
```

Print data

```
kndata
```

```

##      treat richness total algae      y
## 1 Control        8    78     1 0.1001674
## 2 Control        5    84     7 0.2677633
## 3 Control       10   115    45 0.7353145
## 4 Control        7   200   100 1.5707963
## 5 Control        6    72    20 0.4636476
## 6      Low        8    73    15 0.3976994
## 7      Low        7   124    70 0.9911566
## 8      Low        8   116    50 0.7853982
## 9      Low        8    92     5 0.2255134
## 10     Low        7   138    60 0.8860771
## 11   Medium        7   124    85 1.1730969
## 12   Medium        8   116    80 1.1071487
## 13   Medium        8   145    60 0.8860771
## 14   Medium        6   154   100 1.5707963
## 15   Medium        7   129    90 1.2490458
## 16     High        6   134    95 1.3452829
## 17     High        7   138    95 1.3452829
## 18     High        8   103    70 0.9911566
## 19     High        8   119    75 1.0471976
## 20     High        6   132    80 1.1071487
## 21 VeryHigh        6   148    95 1.3452829
## 22 VeryHigh        5   134    95 1.3452829
## 23 VeryHigh        5   119   100 1.5707963
## 24 VeryHigh        5   117    90 1.2490458
## 25 VeryHigh        5   129    80 1.1071487

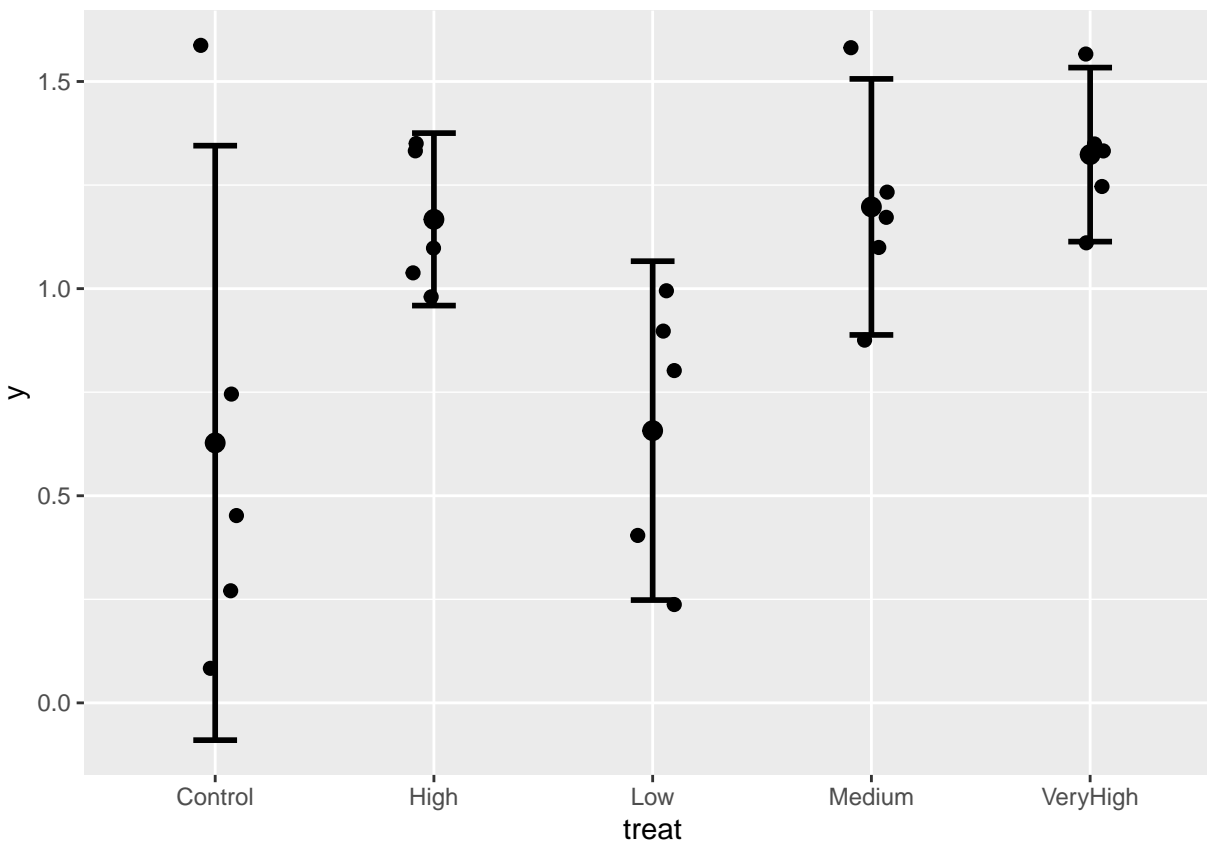
```

Graphics using ggplot2

```

ggplot(kndata,aes(treat,y))+
geom_jitter(size=2,position=position_jitter(width=0.1))+
stat_summary(fun="mean",geom="point",size=3)+
stat_summary(fun.data="mean_cl_normal",geom="errorbar",width=0.2,linewidth=1)

```



```
# One-way ANOVA
aovout <- aov(y~treat,data=kndata)
anova(aovout)

## Analysis of Variance Table
##
## Response: y
##          Df Sum Sq Mean Sq F value    Pr(>F)
## treat      4  2.1382   0.53454    4.7632 0.007295 **
## Residuals 20  2.2444   0.11222
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Obtain multiple comparison object
compout <- glht(aovout,linfct=mcp(treat="Tukey"))

# LSD method - only controls the per comparison error rate
lsdout <- summary(compout,test=univariate())
lsdout

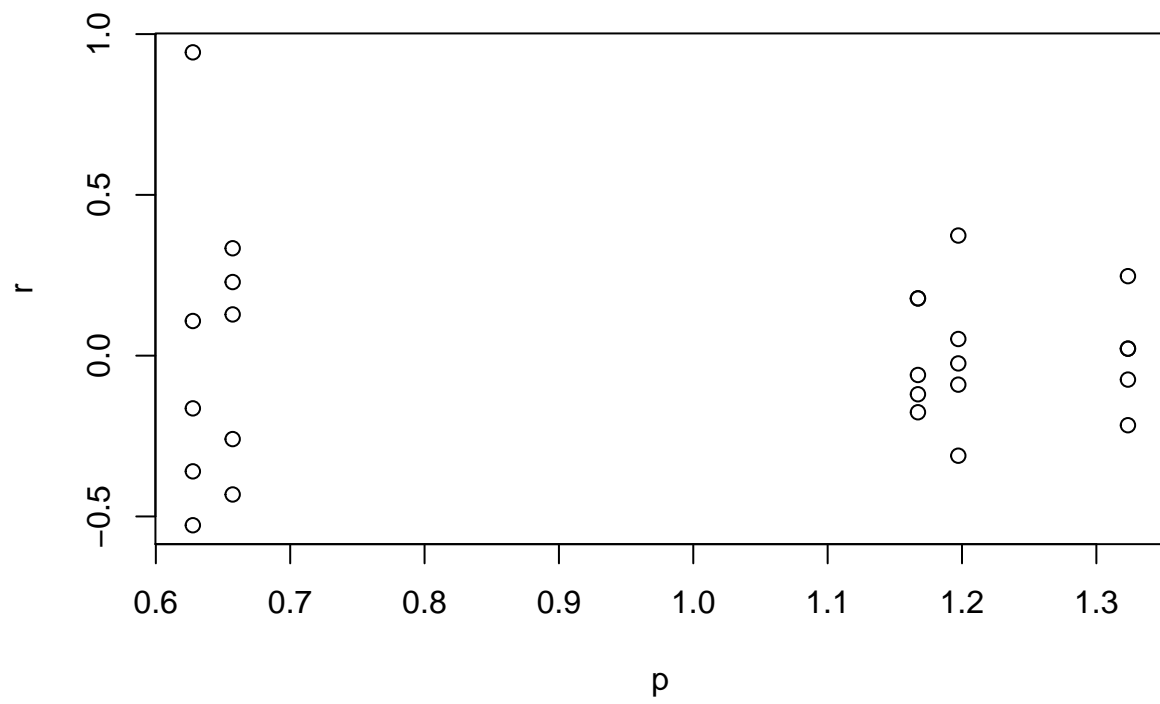
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = y ~ treat, data = kndata)
##
```

```
## Linear Hypotheses:
##               Estimate Std. Error t value Pr(>|t|)
## High - Control == 0    0.53968    0.21187   2.547   0.0192 *
## Low - Control == 0     0.02963    0.21187   0.140   0.8902
## Medium - Control == 0   0.56970    0.21187   2.689   0.0141 *
## VeryHigh - Control == 0 0.69597    0.21187   3.285   0.0037 **
## Low - High == 0        -0.51004    0.21187  -2.407   0.0258 *
## Medium - High == 0      0.03002    0.21187   0.142   0.8887
## VeryHigh - High == 0    0.15630    0.21187   0.738   0.4693
## Medium - Low == 0       0.54006    0.21187   2.549   0.0191 *
## VeryHigh - Low == 0     0.66634    0.21187   3.145   0.0051 **
## VeryHigh - Medium == 0  0.12628    0.21187   0.596   0.5578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Univariate p values reported)
```

```
# FDR method - controls false discovery rate
p.adjust(lsdout$test$pvalues,method="fdr")
```

```
##      High - Control      Low - Control      Medium - Control      VeryHigh - Control
##      0.03838307         0.89017317         0.03838307         0.02547915
##      Low - High        Medium - High        VeryHigh - High        Medium - Low
##      0.04307604         0.89017317         0.67037272         0.03838307
##      VeryHigh - Low    VeryHigh - Medium
##      0.02547915         0.69731081
```

```
# Diagnostic plots to check ANOVA assumptions
p <- predict(aovout)
r <- resid(aovout)
plot(p,r)
```



`qqnorm(r)`

