# Lab 10: One-way ANOVA and post hoc tests

**Objectives**

1. Perform one-way ANOVA in R, “by hand” and with dedicated functions.
2. Perform Tukey tests for unplanned comparisons.
3. Use the R commands stripchart, tapply, text, TukeyHSD, aov, pf, and ptukey.

**Background**

Male fiddler crabs have one greatly enlarged “major” claw that they use to attract females and defend a burrow. It has also been hypothesized that the major claw acts as a heat sink, keeping males cooler while they are outside their burrows on hot days. To test this, researchers placed crabs into separate plastic cups heated by a 60-watt light bulb. They tested four kinds of crab: intact male crabs, male crabs with the major claw removed, male crabs with the minor claw removed, and intact female crabs. They measured the body temperature of each crab every ten minutes for 1.5 hours. From these measurements they calculated the rate of heat gain (in °C/log minute) for each crab. The data are in file “crabs.csv”. You will use these data to test whether the different crab types have different mean rates of heat gain.

**Exercise**

1. Make a strip chart of the data. Give the plot informative axes labels and titles. Add a symbol to show the average of each group. The average should be shown by a distinct symbol, so that it can be distinguished from the data.
2. Do an ANOVA to test for any differences in mean rate of heat gain among the groups. Do not use the aov command. Instead, do all of the calculations yourself in R. Use your results to fill in an ANOVA table like the one below. In addition, make a clear statement of the null and alternative hypotheses that you tested and the results of your test.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | df | SS | MS | F | P |
| Treatment |  |  |  |  |  |
| Error |  |  |  |  |  |

1. Re-do the ANOVA using R’s aov command.
2. For each group, do the following:
   1. Estimate the group’s mean rate of heat gain.
   2. Calculate the 95% confidence interval of the mean.
3. Now determine perform pairwise comparisons to test which groups differ from which. Perform Tukey-Kramer tests on all possible comparisons. Do these tests by hand and present your results in a table like this:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Comparison | Difference | q | P-value | Conclusion |
|  |  |  |  |  |

1. Repeat these Tukey-Kramer tests using R’s TukeyHSD command.
2. Make an error bar plot, and label it to indicate the Tukey-Kramer test results. For your measure of precision, use the 95% confidence interval of the mean for each group.

**Assignment: Turn in a Microsoft Word document with a heading that includes your name, the lab number, and your section. Include (1) your annotated R script, (2) clear presentations of your output and the results of your analysis, including all figures and tables. In addition to the Word document, you must also submit a separate .R file with all of your R scripts.**

**Question 1 Script :**

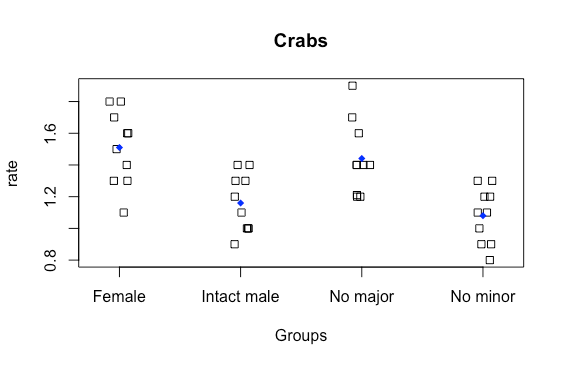
crabs <- read.csv(“/Users/PaigeBranam/Downloads/crabs.csv")

stripchart(rate~group,data=crabs, method="jitter", vertical=TRUE, xlab="Groups", main="Crabs")

avgs <- tapply(crabs$rate, crabs$group, mean)

points(avgs, pch=18, col=“blue")

**Question 1 Output :**



**Question 2 Script :**

#Step 1   
#Hypothesis :   
#H0 : The heat gain among the groups will be the same. (No difference)  
#HA : The heat gain among the groups will not be the same (Different Heat gains)  
#Alpha : 0.5

#Step 2  
#Test Statistic

#F Statistic - MSG/MSE

#Step 3 - Calculate test statistic

#Annotations/key/formulas for calculating ANOVA by hand

#g=individual groups   
#G= grand group   
#k= number of groups : (4)  
#Xbar= average

#n= total number of subjects : (40)  
#N = Length of group : (all groups are the same in length) : (10)

# MSG = sum of squares between groups   
#(Grand Sum of (gXbar(1-4)-GXbar^2)\*N) / k-1

# MSE = sum of squares within groups   
#(Grand sum of (gx(1-4)-gxbar(1-4)^2))/n-k

#Calculations - MSE

n <- 40  
k <- 4  
N <- 10

g\_Xbar <- tapply(crabs$rate, crabs$group, mean)

gXbar1 <- 1.51   
gXbar2 <- 1.16   
gXbar3 <- 1.08   
gXbar4 <- 1.441

G\_Xbar <- mean(g\_Xbar) #1.29775

# Sum of (xg(1-4)-g\_Xbar(1-4))^2

#Female (x1-xbar1)^2  
g1\_x <-((1.3-1.51)^2)+((1.6-1.51)^2)+((1.4-1.51)^2)+((1.1-1.51)^2)+((1.6-1.51)^2)+((1.8-1.51)^2)+((1.3-1.51)^2)+((1.7-1.51)^2)+((1.5-1.51)^2)+((1.8-1.51)^2)

#Intact male (x2-xbar2)^2  
g2\_x <- ((1.3-1.16)^2)+((1.2-1.16)^2)+((1.0-1.16)^2)+((0.9-1.16)^2)+((1.4-1.16)^2)+((1.0-1.16)^2)+((1.3-1.16)^2)+((1.4-1.16)^2)+((1.1-1.16)^2)+((1.0-1.16)^2)

#No Minor (x3-xbar3)^3  
g3\_x <- ((1.2-1.08)^2)+((1.0-1.08)^2)+((.9-1.08)^2)+((.8-1.08)^2)+((1.2-1.08)^2)+((.9-1.08)^2)+((1.1-1.08)^2)+((1.1-1.08)^2)+((1.3-1.08)^2)+((1.3-1.08)^2)

#No Major (x4-xbar4)^2

g4\_x <- ((1.2-1.44)^2)+((1.7-1.44)^2)+((1.4-1.44)^2)+((1.2-1.44)^2)+((1.21-1.44)^2)+((1.6-1.44)^2)+((1.9-1.44)^2)+((1.4-1.444)^2)+((1.4-1.44)^2)+((1.4-1.44)^2)

#Sum of Squares within = gx(1-4)/n-k   
MSE <- (g1\_x + g2\_x + g3\_x + g4\_x) / (n-k)

#MSE <- 0.04301767 (Denominator of F Statistic)

#Calculations MSG   
# Grand Sum of N(xbar(1-4) - G\_Xbar)^2  
n1 <- ((1.51-1.29775)^2)\*10  
n2 <- ((1.16-1.29775)^2)\*10  
n3 <- ((1.08-1.29775)^2)\*10  
n4 <- ((1.441-1.29775)^2)\*10

#Solve for Sums   
N\_Sum <- n1+n2+n3+n4  
#Calculate SSG  
MSG <- N\_Sum/(k-1)

#MSG <- 0.4398692 (Numorator of F Statistic)

FStatistic <- MSG/MSE  
#FStatistic - 10.22531

#Pvalue   
#Pvalue = 1-pf(F,vgroups,verror)

#Vgroups = (n-k) - 3  
#Verror = (k-1) - 36

p <- 1-pf(FStatistic, 3, 36)  
#p = 5.190931e-05

#P value is less than alpha value of .5, reject null hypothesis that the group means are the same.

**Question 2 Output :**

|  | DF | SS | MS | F | P |
| --- | --- | --- | --- | --- | --- |
| Treatment | 3 | 1.319608 | 0.4398692 |  |  |
| Error | 36 | 1.548636 | 0.04301767 | 10.22531 | 5.190931e-05 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**Question 3 Script :**

Crabs\_AOV <- aov(rate~group, data=crabs)

summary(Crabs\_AOV)

**Question 3 Output :**

Df Sum Sq Mean Sq F value Pr(>F)

group 3 1.320 0.4399 10.23 5.18e-05 \*\*\*

Residuals 36 1.548 0.0430

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

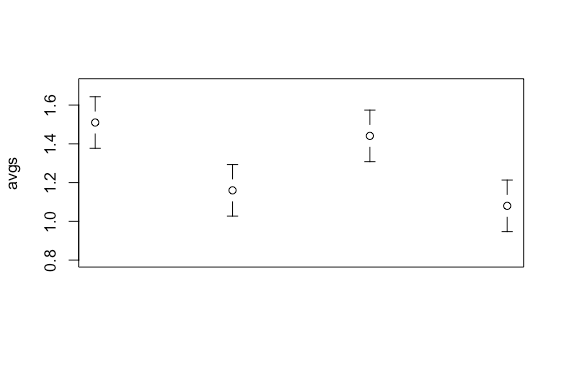
**Question 4 Script :**

# Part A   
#Groups mean rate of heat gain (Calculated above by hand as well as in ANOVA AOV)

#MSG - .4399  
#MSE - .0430

#Part B

#95% Confidence Interval   
tcrit <- qt(0.975, 36)  
se <- sqrt(MSE/N)  
CI <- se\*tcrit   
plotCI(avgs, uiw=CI, xaxt="n", ylim=c(.8,1.7))

**Question 4 Output :**

**Question 5 Script :**

#Tukey-Kramer by Hand

#Group comparison should be k groups = k(k-1)2   
# 4(4-1)/2 = 6   
# 6 Group Comparisons

#Formula :   
# xbar1-xbar2/sqrt((MSE/n)   
#Calculations available from hand calculation of ANOVA above

#all groups have 10 for n and MSE is the same, so this value will not change  
y <- sqrt(MSE/10)

#Group 1 : Female - Intact Male   
(gXbar1-gXbar2)/y = 5.336354  
#pvalue -  
1-ptukey(5.336354, 4, 36) #0.003120495

#Group 2 : Intact Male - No Minor   
(gXbar2-gXbar3)/y = 1.219738  
#pvalue -  
1-ptukey(1.219738, 4, 36) #0.8239841

#Group 3 : No Minor - No Major  
(gXbar4-gXbar3)/y = 5.504068  
#pvalue -  
1-ptukey(5.504068, 4, 36)   
#0.002236915  
#Female - Intact

#Group 4 : Intact - No major   
(gXbar4-gXbar2)/y = 4.28433  
#pvalue -  
1-ptukey(4.28433, 4, 36) #0.02237932  
#Female- Intact / No Major - No minor

#Group 5 : No Minor - Female   
(gXbar1-gXbar3)/y - 6.556092  
#pvalue -  
1-ptukey(6.556092, 4, 36) #0.0002561845   
#Female - Intact / No Major - No minor / Intact - No Major

#Group 6 : Female - No Major   
(gXbar1-gXbar4)/y - 1.052024  
#pvalue -  
1-ptukey(1.052024, 4, 36)   
# 0.8786417  
#Intact Male - No minor

**Question 5 Output :**

Female - Intact Male : 0.003120495 - **A**  
Intact Male - No Minor : 0.8239841 - **B**  
No Minor - No Major : 0.002236915 - **C**  
Intact - No major : 0.02237932 - **C**  
No Minor - Female : 0.0002561845 - **D**  
Female - No Major : 0.8786417 - **B**

**Question 6 Script :**

#Tukey-Kramer Short Cut

TukeyHSD(Crabs\_AOV)

**Question 6 Output :**

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = rate ~ group, data = crabs)

$group

diff lwr upr p adj

Intact male-Female -0.350 -0.59978301 -0.100217 0.0031168

No major-Female -0.069 -0.31878301 0.180783 0.8786067

No minor-Female -0.430 -0.67978301 -0.180217 0.0002558

No major-Intact male 0.281 0.03121699 0.530783 0.0223605

No minor-Intact male -0.080 -0.32978301 0.169783 0.8239361

No minor-No major -0.361 -0.61078301 -0.111217 0.0022342

**Question 7 Script :**

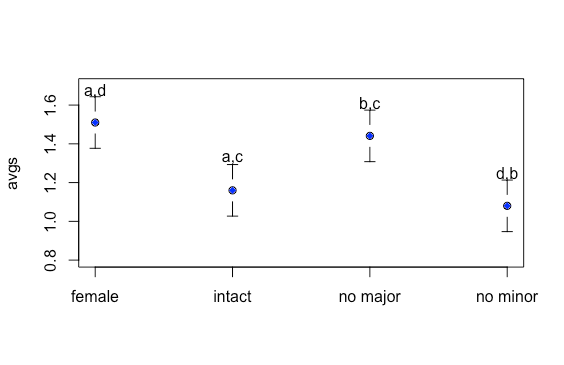
#Make an Error Bar Plot w/ Tukey-Kramer Data

library(gplots)

plotCI(avgs, uiw=CI, xaxt="n", ylim=c(.8,1.7))

axis(1, at=1:4, labels=c("female","intact","no major", "no minor"))

text(x=1:4, y=avgs+CI+.03, labels=c("a,d","a,c","b,c","d,b"))

**Question 7 Output :**