Bonnie Turek

Eco 634 – Lab 10

11/17/21

Q1. ANOVA by hand code:

```
#Bonnie Turek
#ECo 634 - Lab 10
require(here)
rm(list = ls())
rope = read.csv(
  here("data", "rope.csv"))


rope$rope.type <- factor(rope$rope.type)
class(rope$rope.type)
levels(rope$rope.type)
length(rope$blade)
nrow(rope)
n_obs = nrow(rope)
length(levels(rope$rope.type))
n_groups = length(levels(rope$rope.type))
#calculate the total sum of squares
#SST = Σ(yi – ybar)^2
ss_tot = sum((rope$p.cut - mean(rope$p.cut))^2)
df_tot = n_obs-1
par(mfrow= c(1,2))
boxplot(rope$p.cut, main = "Percent Rope Cut
    of all rope types", ylab = "Percent Rope Cut", xlab = "All rope types")
boxplot(rope$p.cut ~ rope$rope.type, main = "Percent Rope Cut
of different rope types", ylab = "Percent Rope Cut", xlab = "")
```

```
#define the function to get residuals

resid_fun = function(x)

  {

  (x - mean(x))

}

#create aggregate of residuals within group

agg_resids = aggregate(

 x = rope$p.cut,

 by = list(rope$rope.type),

 FUN = resid_fun)

str(agg_resids)

sum_sq_resid = function(x)

{

  sum((x - mean(x))^2)

}

agg_sq_resids = aggregate(

 x = rope$p.cut,

 by = list(rope$rope.type),

 FUN = sum_sq_resid)

str(agg_sq_resids)

ss_within = sum(agg_sq_resids$x)

df_within = n_obs - n_groups

ss_among = ss_tot - ss_within

df_among = n_groups - 1


ms_among  =  ss_among / (n_groups - 1)

ms_within = ss_within / (n_obs - n_groups)
```

```
# F-ratio, defined as the among-group variance divided by

#the within-group variance.

f_ratio = ms_among / ms_within

f_pval = pf(f_ratio, (n_groups - 1), (n_obs - n_groups), lower.tail = FALSE)


fit_1 = lm(p.cut ~ rope.type, data=rope)

anova(fit_1)

anova_fit_1 = anova(fit_1)

str(anova_fit_1)

#check our sum of squared resids (ss_among and ss_within)

anova_fit_1$"Sum Sq"
```

Q2. Examine the conditional boxplot in the Partitioning Variance: Within-Group section of the walkthrough. Based on the figure, do you think there are equal variances among the groups?
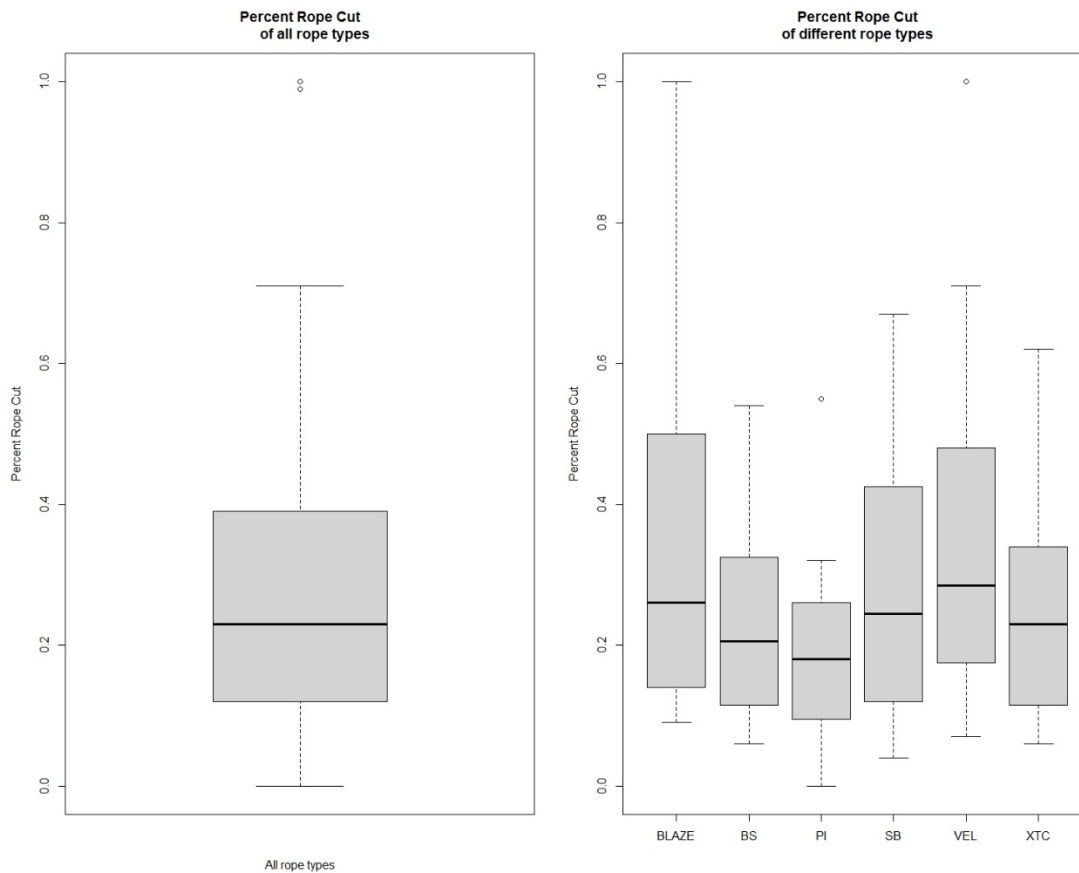
In looking at the boxplot on the right above, I believe there is **NOT** equal variances in percent rope cut among the groups of rope types. We can see this in the widths of the boxes. They are all different widths, and therefore have different variances.

Q3. Conduct a Bartlett test to assess the homogeneity of variances in the rope type groups. Report the p-value.

#Bartlett Test

bartlett.test(p.cut ~ rope.type, data = rope)

Bartlett test of homogeneity of variances

Bartlett's K-squared = 19.687, df = 5, **p-value = 0.00143**

Q4. Given your graphical assessment (question 2) and the Bartlett test, do you think an ANOVA-type analysis is appropriate on the raw data? Explain why or why not.

Given both the graphical analysis from the boxplot and the p-value resulting from the Bartlett test, we can see that there is an issue with heterogeneity in the variance of our percent rope cut data based on rope type.

Given the resulting p-value of 0.00143 in the Bartlett test, this means we have strong evidence to reject the null hypothesis that the variances of the percent cut rope (by rope type) are homogenous. We can also see that the variances between rope types are not homogeneous in the boxplot.

Homogeneity of the variances is one of the assumptions of group 1 general linear models, which the ANOVA-type analysis falls within. In that case, the ANOVA-type analysis may not be appropriate with the raw data given here, since the Group 1 model homogeneity assumption is violated. However, we could consider performing data transformations like a log-transform to make it work.

Q5. Which rope type is the base case?

BLAZE is the base case / intercept.

Q6. What is the mean percent cut of the base case rope? Show your calculation using value(s) from the model coefficient table.

0.36714

This is directly from the model coefficient table for the INTERCEPT, which is also the Base Case, BLAZE.

Q7. What is the mean percent cut rope type XTC? Show your calculation using value(s) from the model coefficient table.

$0.36714 + (0 * -0.13014) + (0 * -0.18014) + (0 * -0.09514) + (0 * -0.01714) + (1 * -0.10164)$

= **0.2655**

We can also check our answers to both Q6 and Q7 using the following code:

#check q6 and q7

mean(rope$p.cut[which(rope$rope.type == "XTC")])

mean(rope$p.cut[which(rope$rope.type == "BLAZE")])

OUTPUT:

0.2655

0.3671429