Siddarth Puliyanda

Wasp Reproduction Write-up

Introduction:

Female wasps lay their eggs inside pests eggs and the developing parasitoid larvae kill the pest eggs by feeding inside them. The objective of the experiment is to test if some plants can increase the reproduction of wasps in order to control the pest population. Controlling the pests would be beneficial in many ways to people such as farmers, who can produce more crop if there are less pests to ruin it. Each individual experiment consists of a container with one female wasp, one treatment, and a mealy bug egg mass. The plant treatments being used for testing are buckwheat (BW), vetch (V), and the control of no plant/water only (W). The five variables of interest are number of female offspring, number of male offspring, total offspring, longevity, and sex ratio.

Methodology:

The statistical model being used is $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$, to create a linear model. The hypotheses of interest look to test whether there is a significant difference between buckwheat, vetch, and the control for each variable of interest. All the variables of interest share a common null hypothesis, alternative hypothesis, assumptions, and rejection region.

Null hypothesis: effects of the treatments are the same for the variable of interest

Alternative hypothesis: at least one of the treatments differs for the variable of interest

The assumptions to be checked for each variable are having a normal distribution, common variance, and coming from independent, random samples. If any of the assumptions are not met, a transformation would need to be made in order to properly fit the data. We assume the data was collected from independent, random samples so this assumption is already complete.

Finally the rejection region is {p-value$< \alpha = 0.05$}, which will determine whether or not to reject the null hypothesis at the 95% significance level.

If the null hypothesis is rejected, I will run a post-hoc Tukey's test to find any significant differences between pairs.

Data Analysis and Results:

of Female Offspring

For the variable of number of female offspring, I tested BW=W=V (null) vs. at least one of the treatments differing (alternative). When checking normality assumption the QQ plot displayed variation from the diagonal, showing that the data for female offspring is not normal. Also, the residuals vs. fitted plot did not have the point randomly scattered across the plot, meaning the common variance assumption is violated. In order to fix this, I transformed the data using the Box-Cox log transformation ($\lambda=0$). Rechecking the assumptions, I found the QQ plot was in a near straight line and the points in the residuals vs. fitted plot were more randomly distributed.

Boxplot of Female Offspring
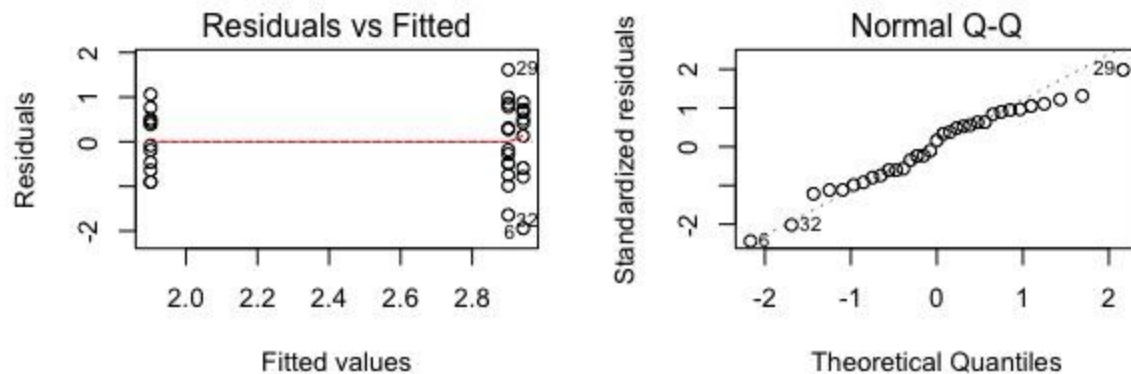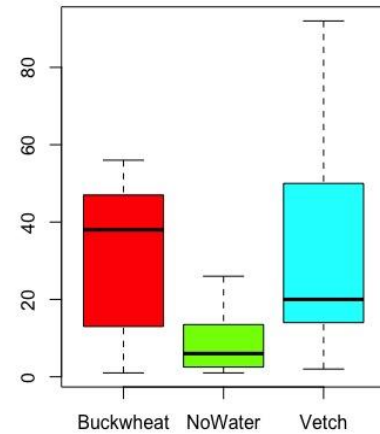
After running the data using the log transformation, the ANOVA table produced a p-value of 0.0107. Since the p-value is less than the α of 0.05, we reject the null hypothesis at the given rejection region.
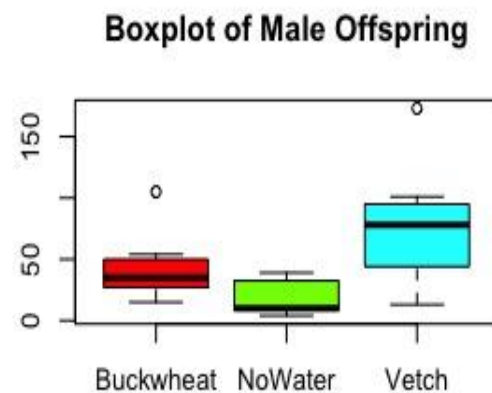
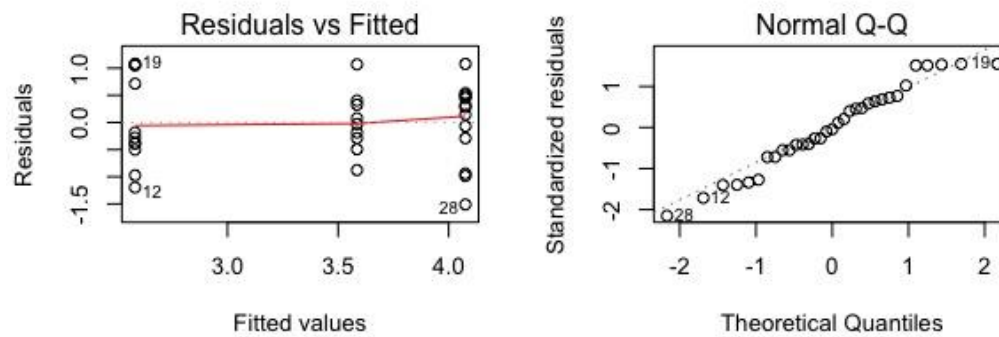|	Df	SS	MS	F-value	P-value
Treatment	2	7.593	3.797	5.305	0.0107
Errors	30	21.470	0.716		

We can conclude that at least one of the plant treatments differs in number of female wasps offspring produced. A post-hoc analysis was then performed with Tukey's Test, which showed there was a significant difference between the pairs of BW-W (p-value of 0.027) and V-W (p-value of 0.019).

	Difference	Lower	Upper	P-value adj
No Water-Buckwheat	-1.04146839	-1.9788578	-0.1040790	0.0269595
Vetch- Buckwheat	-0.04149493	-0.9458543	0.8628644	0.9929725
Vetch-No Water	0.99997345	0.1455754	1.8543715	0.0190948

# of Male Offspring

For the variable of number of male offspring, the same null and alternative hypotheses were tested as before. The normality and common variance assumptions for this set of data were not met, as the QQ plot displayed variation from the diagonal and the residual vs. fitted plot did not show points that were randomly scattered. For this variable, I again used a Box-Cox log transformation ($\lambda=0$). The normality and common variance assumptions were met after retesting. The residuals vs. fitted graph showed randomly scattered points, while the QQ plot showed points close to a straight line.

Boxplot of Male Offspring

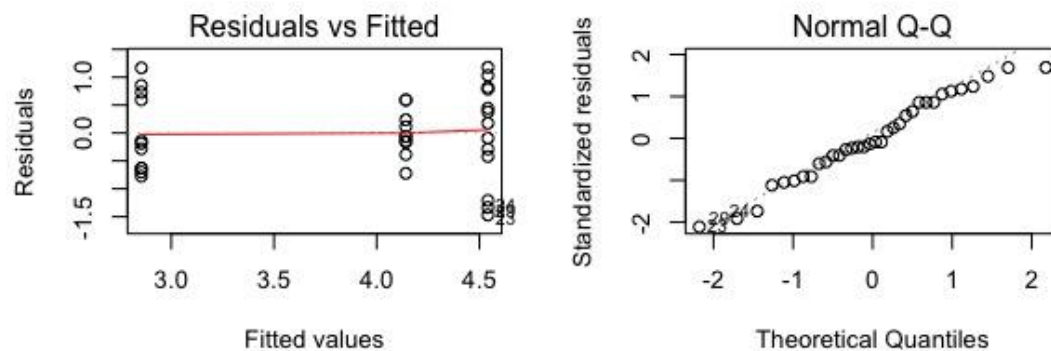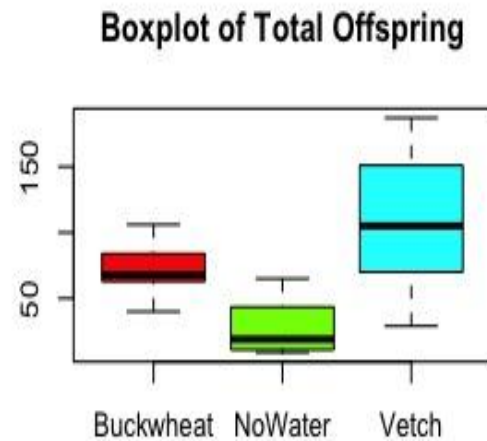The ANOVA table produced a p-value of 0.000104, which is less than α, therefore we reject the null hypothesis.

	Df	SS	MS	F-value	P-value
Treatment	2	13.54	6.768	12.64	0.000104
Errors	30	16.06	0.535		

We can conclude that at least one of the plant treatments differs in number of male wasps produced. Using Tukey's Test for post-hoc analysis, I once again found a significant difference between the pairs of BW-W (p-value of 0.027) and V-W (p-value of 0.019).

	Difference	Lower	Upper	P-value adj
No Water-Buckwheat	-1.0033129	-1.8140487	-0.1925771	0.0127842
Vetch- Buckwheat	0.4900378	-0.2921308	1.2722064	0.2850962
Vetch-No Water	1.4933507	0.7543930	2.2323084	0.0000711

Total # of Offspring

The values of female and male offspring were added together to create the variable of total number of offspring. The null hypothesis being tested is BW=W=V for total offspring, vs. the alternative hypothesis that at least one of the plant treatments differs in total offspring. The normality and common variance assumptions were again not met, as the QQ plot did not display linear points and the the residuals vs. fitted plot did not show a random distribution of points. I used a Box-Cox cubed root

Boxplot of Total Offspring

transformation ($\lambda=0.3$) in order to meet all the assumptions. When rechecking the assumptions, the QQ plot displayed more linearity and the residuals vs. fitted plot showed points with a more randomly scattered distribution, satisfying the normality and common variance assumptions.

The ANOVA table produced a p-value of 0.00000985, meaning we reject the null hypothesis at the $\alpha=0.05$ level of significance.

	Df	SS	MS	F-value	P-value
Treatment	2	17.94	8.970	17.11	0.00000985
Errors	30	16.25	0.524		

We can conclude that at least one of the plant treatments differs in number of total wasp offspring produced. Again, I used Tukey's Test and found a significant difference between the pairs of BW-W (p-value of 0.00086) and V-W (p-value of 0.0000088).

	Difference	Lower	Upper	P-value adj
No Water-Buckwheat	-1.2874965	-2.0661686	-0.5088244	0.0008564
Vetch- Buckwheat	0.3993358	-0.3502711	1.1489427	0.3998139
Vetch-No Water	1.6868323	0.9567380	2.4169265	0.0000088

Longevity

For the variable of longevity, the null of BW=W=V vs. the alternative of at least one plant treatment differing in amount of days lived after birth was tested. Once again the assumptions were not met, as the QQ plot showed clear deviance from a straight line and the residuals vs. fitted plot showed a clear cluster of points. This shows that the

Boxplot of Longevity
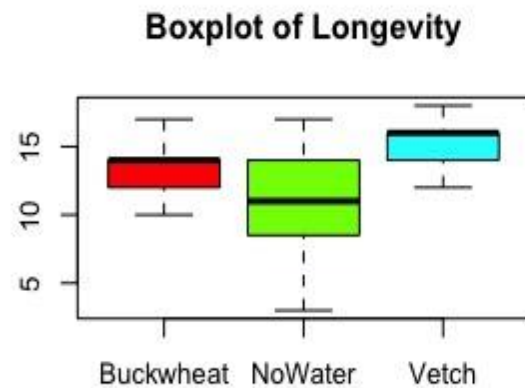
normality and common variance assumptions were violated. For this variable I used a Box-Cox cubed root transformation ($\lambda=0.3$) and rechecked the assumptions. The QQ plot was a little more linear but still showed some deviance; however, it appeared to be the strongest transformed plot to support the normality assumption. Also the residuals vs. fitted plot showed a slight cluster of points near the right side of the graph, but I chose to accept the common variance assumption because this plot showed the most randomly scattered points compared to other Box-Cox transformations.
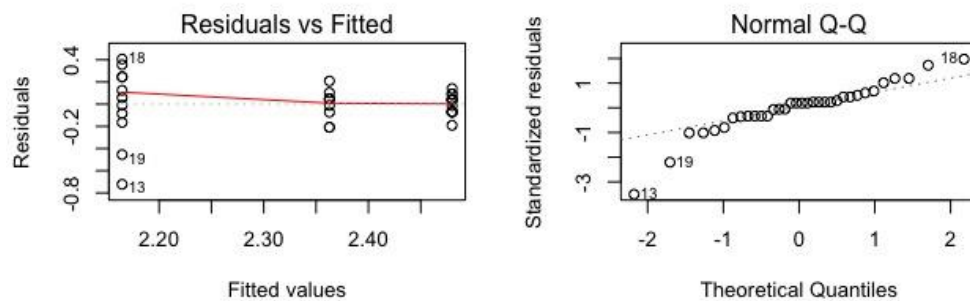
Continuing on, the ANOVA table produced a p-value of 0.00474, which is less than the α of 0.05.

|	Df	SS	MS	F-value	P-value
Treatment	2	0.5973	0.29867	6.392	0.00474
Errors	30	1.4484	0.04672		

We can reject the null hypothesis and conclude that at least one of the plant treatments differs in longevity of wasp offspring. Through Tukey's Test for post-hoc analysis, I found one one significantly different pair with V-W (p-value of 0.0034).

	Difference	Lower	Upper	P-value adj
No Water-Buckwheat	-0.1980822	-0.43052881	0.03436445	0.1068653
Vetch- Buckwheat	0.1171445	-0.10662566	0.34091469	0.4121918
Vetch-No Water	0.3152267	0.09728137	0.53317202	0.0034010

Sex Ratio

The final variable being tested is the sex ratio of female to total wasp offspring. The hypotheses being tested are BW=W=V (null) vs. at least one plant treatment differs in sex ratio of wasp offspring. The QQ plot shows points in roughly a straight line, meaning the normality assumption is met. Also the residuals vs. fitted graph shows the points are randomly scattered throughout the graph, meaning the common variance assumption is met.

Since all the assumptions for this dataset were met no transformation is needed.
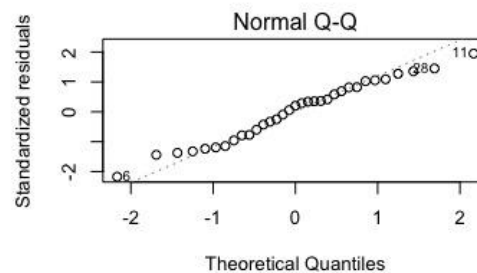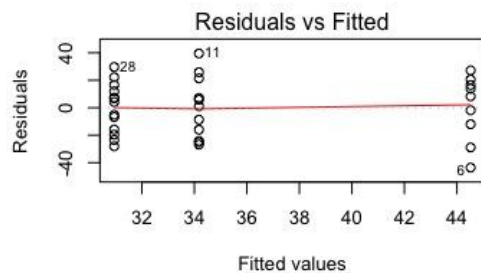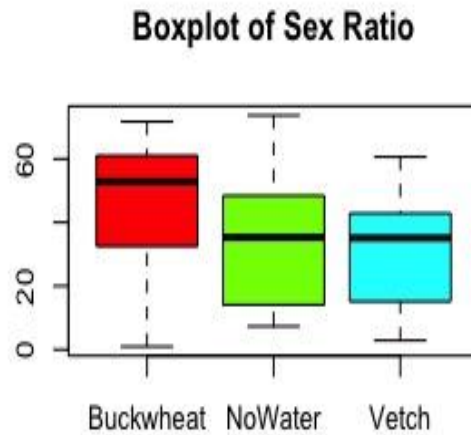
The resulting ANOVA table produced a p-value of 0.337, meaning we fail to reject the null hypothesis because it is higher than the $\alpha=0.05$ significance level. There is sufficient evidence to conclude that buckwheat, vetch, and no plant/water only are equal for sex ratio of wasp offspring. No post-hoc test is required.

	Df	SS	MS	F-value	P-value
Treatment	2	1020	509.9	1.127	0.337
Errors	30	13572	452.4		

Conclusion:

The goal of the experiment was to find the treatment that would increase overall wasp reproduction in order to control the pest population. Based on the results of the study, we can see that every variable of interest involving the number of wasp offspring produced (number of female offspring, number of male offspring, total number of offspring) resulted in a significant difference between the BW-W and V-W pairs. For longevity, however, the only significant pair was V-W, while BW was not significantly different from either of the other two treatments. None of the variables of interest showed any significant difference between the BW-V pair. Vetch would seem to be the best treatment option because it is significantly different from the control in terms of total offspring produced and also the longevity of the offspring, while the buckwheat treatment is not significantly different from the control in terms of offspring longevity.

R Code:

```
treatment <- c("Buckwheat","NoWater", "Vetch") #femaledata

treatment

female <-

c(38,41,47,13,10,1,29,56,49,6,14,5,19,3,13,1,1,12,26,2,10,7,14,32,53,50,59,20,92,14,18,2,33)

female

get.female <-MASS::boxcox(female~treatment1)

get.female

female.model1 <-aov(female^(1/3)~treatment1)

female.model1

summary(female.model1)

par(mfrow=c(2,2))

plot(female.model1)

TukeyHSD(female.model1, conf.level=0.95)

plot(TukeyHSD(female.model1, conf.level = 0.95),las=1, col = "red")

#male data

treatment1 <- c(rep("Buckwheat",9),rep("NoWater",11),rep("Vetch",13))

treatment1

male <-

c(15,22,30,27,54,105,39,50,35,11,5,4,27,38,38,9,10,8,39,9,78,22,23,92,98,55,95,13,81,173,101,6
8,44)

male.data <-data.frame(treatment1,male)

male.data

boxplot(male~treatment1, main="Boxplot of Male Offspring", col= rainbow(4))

male.model <-aov(male~treatment1)

male.model

summary(male.model)

par(mfrow=c(2,2))

plot(male.model) #assumptions not met
```

```
MASS::boxcox(male~treatment1)#boxcox transformation

male.model2 <-aov(log(male)~treatment1)

male.model2

summary(male.model2)

par(mfrow=c(2,2))

plot(male.model2)

TukeyHSD(male.model2, conf.level=0.95)

plot(TukeyHSD(male.model2, conf.level = 0.95),las=1, col = "red")

#total data

treatment1 <- c(rep("Buckwheat",10),rep("NoWater",11),rep("Vetch",13))

treatment1

total <-
c(53,63,77,40,64,106,68,68,106,84,17,19,9,46,41,51,10,11,20,65,11,88,29,37,124,151,105,154,3
3,173,187,119,70,77)

total.data <-data.frame(treatment1,total)

total.data

boxplot(total~treatment1, main="Boxplot of Total Offspring", col= rainbow(4))

total.model <-aov(total~treatment1)

summary(total.model)

par(mfrow=c(2,2))

plot(total.model)

get.total<- MASS::boxcox(total~treatment1) #boxcox transformation

get.total

total.model1 <-aov(total^(1/3)~treatment1)

total.model1

summary(total.model1)

par(mfrow=c(2,2))

plot(total.model1)

TukeyHSD(total.model1, conf.level=0.95)
```

```r
plot(TukeyHSD(total.model1, conf.level = 0.95),las=1, col = "red")
#Longevity data
treatment1 <- c(rep("Buckwheat",10),rep("NoWater",11),rep("Vetch",13))
treatment1
longevity <-
c(10,10,12,14,14,13,17,14,15,14,8,14,3,14,9,12,11,17,5,10,16,14,16,14,16,17,14,16,17,12,18,16,
14,15)
longevity.data <-data.frame(treatment1,longevity)
longevity.data
boxplot(longevity~treatment1, main="Boxplot of Longevity", col= rainbow(4))
longevity.model <-aov(longevity~treatment1)
summary(longevity.model)
par(mfrow=c(2,2))
plot(longevity.model)
get.longevity <- MASS::boxcox(total~treatment1) #boxcox transformation
get.longevity
longevity.model1 <-aov(longevity^(1/3)~treatment1)
longevity.model1
summary(longevity.model1)
par(mfrow=c(2,2))
plot(longevity.model1)
TukeyHSD(longevity.model1, conf.level=0.95)
plot(TukeyHSD(longevity.model1, conf.level = 0.95),las=1, col = "red")
#Sex ratio data
treatment1 <- c(rep("Buckwheat",9),rep("NoWater",11),rep("Vetch",13))
treatment1
sexratio <-
c(71.6981132075472,65.0793650793651,61.038961038961,32.5,15.625,0.943396226415094,42.
6470588235294,52.8301886792453,58.3333333333333,35.2941176470588,73.6842105263158,
```

55.5555555555556,41.304347826087,7.31707317073171,25.4901960784314,10,9.09090909090
909,60,40,18.1818181818182,11.3636363636364,24.1379310344828,37.8378378378378,25.806
4516129032,35.0993377483444,47.6190476190476,38.3116883116883,60.6060606060606,53.1
791907514451,7.48663101604278,15.1260504201681,2.85714285714286,42.8571428571429)

sexratio.data <-data.frame(treatment,sexratio)

sexratio.data

boxplot(sexratio~treatment1, main="Boxplot of Sex Ratio", col= rainbow(4))

sexratio.model <-aov(sexratio~treatment1)

summary(sexratio.model)

par(mfrow=c(2,2))

plot(sexratio.model)