STAT 2020 PROJECT

Labs 7 and 8

Andrew Warrington

On my honor as a student, I have neither given nor received unauthorized aid on this assessment.

Andrew Warrington

Introduction

The data used in this project was compiled in the Cox Evolutionary Biology laboratory at the University of Virginia. The researchers were interested in elucidating the mechanisms by which reproductive investment seems to hinder an organism's capability of survival. Female brown anole lizards were separated into two treatment groups: one group of lizards was ovariectomized (OVX), and the other group of lizards served as a reproductive control (SHAM). The experimenters tracked four physiological metrics thought to influence survival, as well as changes in length (mm) and mass (g) across both groups of lizards. Understanding the changes in these measurements as a result of reproductive status could illuminate the ways reproductive investment can decrease survival rates (Cox et al., 2010).

Our team questioned whether changes in body length (Growth..mm.), as measured by the change in snout-vent length (SVL), was statistically different between the OVX and SHAM treatment groups. We hypothesized that the mean change in body length for the OVX lizards would be greater than the mean change in body length for SHAM lizards (H_A). Our data exploration and statistical analyses sought to determine if there was sufficient evidence to support this claim over the null hypothesis, which states that the mean change in body length for both groups is equal (H_0).

Data Exploration

Summary statistics and graphs were generated for changes in OVX body length first (Figure 1). The histogram and boxplot visualizations indicated a unimodal, slight right-skewed distribution, but can still be described as normal due to the large sample size and central limit theorem. There are at least three high outliers and two low outliers. The center appears to be around 2.0 mm of growth, with a mean of 2.309 mm and a median of 2.0 mm. The mean being of higher value than the median is also indicative of a right-skewed distribution. The values ranged from -1.0 mm to 7.00 mm, and the standard deviation was calculated to be 1.274607 mm.

Next, summary statistics and graphs were generated for changes in SHAM body length (Figure 2). The histogram and boxplot visualizations indicate the growth for the SHAM group follows a unimodal distribution with a possible slight-right skew. There is at least one high outlier and one low outlier. Although the graph has a spacing issue around the center, the data is approximately normal. There is a wide center, ranging roughly from 1.0 mm to 2.0 mm. The growth for the sham group has a mean of 1.426 mm and a median of 1.0 mm. The data ranges from -1.0 mm to 4.0 mm, and the standard deviation was calculated as 1.042156 mm.

Hypothesis Test and Confidence Interval

Our team hypothesized the mean change in body length for OVX lizards would be significantly greater than the mean change in body length for SHAM lizards ($Mu_{ovx} > Mu_{sham}$), as opposed to the null hypothesis ($Mu_{ovx} = Mu_{sham}$). In order to test this hypothesis, our team chose to use a two-sample, right-sided t test for the following reasons: We have two independent random samples and population parameters mu_1, $sigma_1$, mu_2, and $sigma_2$ are unknown. We also met all the assumptions required of a t test: This data came from simple random samples of Anoles. Although the samples are slightly right-skewed, there is a large enough sample size, even after subtracting the missing observations (ending with $n_{ovx} = 139$ & $n_{sham} = 121$), to assume approximate normality—due to the robustness of two-sample t procedures. Lastly, the observations in each anole treatment group are independent.

The t test statistic was found to be 6.1426 using the following calculations in the R programming software:

$$(2.309 - 1.426) = 0.883 \quad \& \quad \sqrt{\frac{1.274607^2}{139} + \frac{1.042156^2}{121}} = 0.14374936$$

$$\frac{0.883}{0.14374936} = 6.142636044$$

Subsequently, the p-value was calculated as 5.445467e-09 using the “pt” function for a one-sided, right-tailed t test: pt(6.142636044, df=120, lower.tail = FALSE). The calculated p-value is less than the set significance level of 5%, thus we reject the null hypothesis and accept the alternative hypothesis.

Next, a 95% confidence interval was calculated for change in body length. A -t* value of -1.97993 was obtained using the following code: qt(.025, df=120). The lower and upper bounds of the 95% confidence interval were determined to be (0.5984, 1.1676).

$$-1.97993 \; * \; \sqrt{\frac{1.274607^2}{139} + \frac{1.042156^2}{121}} \; = \; -0.2846136703, m\ (margin\ of\ error)$$

$$(2.309 - 1.426) + \; (-0.2846136703) = 0.5983863297, lower\ bound\ of\ 95\%\ CI$$

$$(2.309 - 1.426) - \; (-0.2846136703) = 1.1676136703, upper\ bound\ of\ 95\%\ CI$$

Conclusion

We reject the null hypothesis, and accept the alternative which states that the mean change in body length for the OVX treatment groups is *significantly* greater than the mean change in length for the SHAM group. Further, we are 95% confident that the true mean *difference* of change in body length for both groups falls between 0.5984 mm and 1.1676 mm. These findings are relevant because they suggest that there is a reproductive cost to female brown anole growth. When reproductive mechanisms are removed in OVX, they experience a significantly greater growth than SHAM lizards. The lack of growth that occurs in reproducing anoles could contribute to their lower annual survival rate.

Appendix

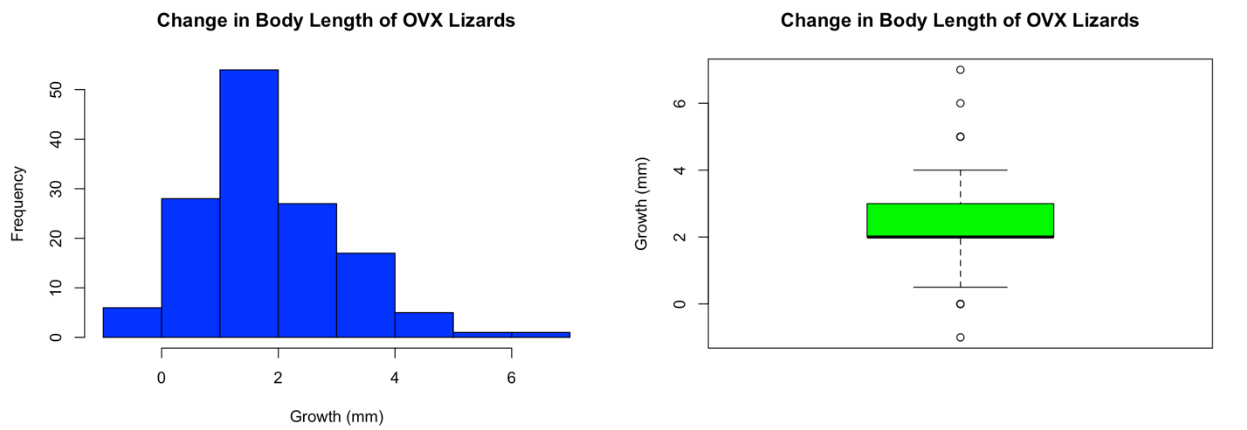

Figure 1. Generated histogram and boxplot for the change in body length of OVX lizards.

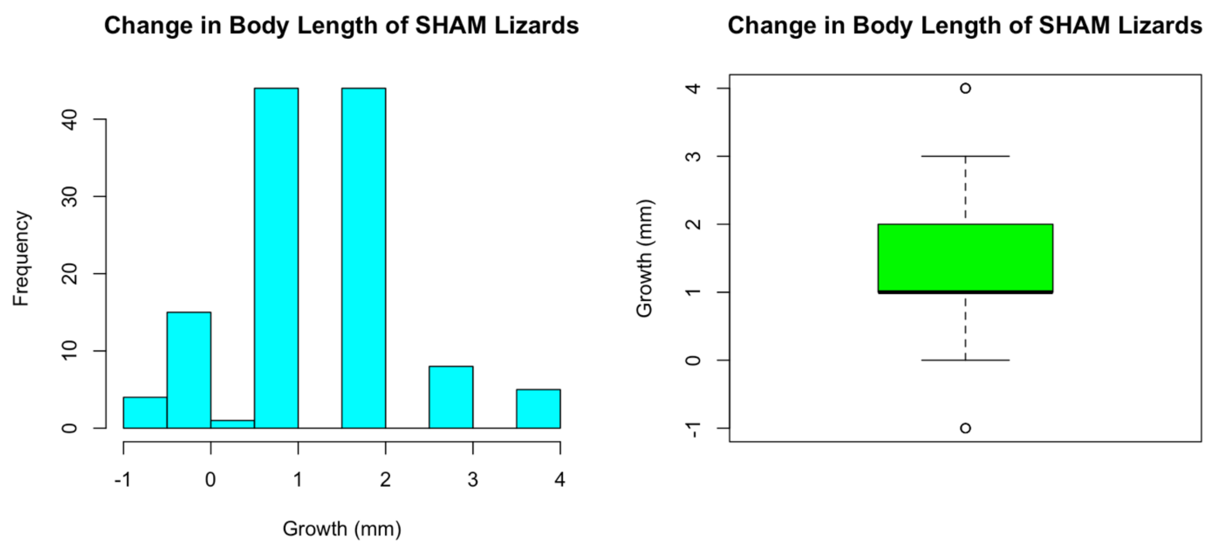

Figure 2. Generated histogram and boxplot for the change in body length of SHAM lizards.

Full R Script:

```
#Lab 7 and 8 Project Script
anole<-read.csv("Cox+Functional+Ecology+data.csv")
ovx <- subset(anole, Treat == "OVX")
sham <- subset(anole, Treat == "SHAM")

#Data Exploration
summary(ovx)
```

```
summary(sham)

#Growth (mm) 5 number Summary for ovx:
#min: -1.0
#1st Quart: 2.0
#Median: 2.0
#3rd Quart: 3.0
#Max: 7.0
#mean is 2.309
#sd is 1.274607

sd(ovx$Growth..mm.,na.rm = TRUE)
par(mfrow=c(1,2))
hist(ovx$Growth..mm., main="Change in Body Length of OVX Lizards",
xlab="Growth (mm)", col="blue")
boxplot(ovx$Growth..mm., main="Change in Body Length of OVX Lizards",
ylab="Growth (mm)", col="green")

#The histogram and boxplot visualizations indicate a unimodal, right-
skewed distribution. There are at least 3 high outliers and at least 2 low
outliers.
#The center appears to be around 2.0 mm of growth, with a mean value of
2.309 mm, and a median value of 2.0 mm of growth.
#The mean being of higher value than the median is also indicative of a
right-skewed distribution.


#Growth (mm) 5 number summar for sham:
#Min -1.0
#1st Quart: 1.0
#Median: 1.0
#3rd Quart: 2.0
#Max: 4.0
#mean is 1.426
#sd is 1.042156

sd(sham$Growth..mm.,na.rm=TRUE)
hist(sham$Growth..mm., main="Change in Body Length of SHAM Lizards",
xlab="Growth (mm)", col="cyan")
boxplot(sham$Growth..mm., main="Change in Body Length of SHAM Lizards",
ylab="Growth (mm)", col="green")

#The histogram and boxplot visualizations indicate the growth for the sham
group follows a bimodal distribution.
#There is at least one high outlier and one low outlier. There are two
centers, one around 1.0 mm and a center closer to 2.0 mm.
#The growth for sham group has a mean growth of 1.426 mm and a median
growth of 1.0 mm.

#Hypothesis Test and Confidence Interval

# H_0: mu_ovx = mu_sham
# H_a: mu_ovx > mu_sham
```

```r
# Is there sufficient evidence to conclude that the mean growth of the ovx
group is significantly higher than the mean sham growth?
#Choosing a two-sample t test, one-sided, right-tailed:
#Because we have two independent random samples and Population parameters
m1 , s1 , m2 , s2 are unknown.
#Our H_a is that the mean ovx growth in mm is significantly greater than
the mean sham growth.

#Assumptions:

#Sample Mean is Approximately Normal.
#Data Comes from SRS (simple random sample) or random experiment
#Observations are independent.

#This data came from a random sample of Anoles.
#Even though the data itself is slightly right skewed (with outliers, both
high and low), we have a large enough sample size to assume that the
sample mean and standard deviation will be approximately Normal.
#The anole observations are independent.

#Calculating T test statistic:

#Calling ovx group 1 and sham group 2
#T_ts
#ovx: n = 139, (140 observations - 1 NA)
#sham: n = 121, (122 observations - 1 NA)
(2.309-1.426) # 0.883
(1.274607^2)/(139) + (1.042156^2)/(121) #0.02066388
sqrt(0.02066388) #0.1688002
0.883/0.1437494 # 6.142634
# T_ts = 6.142634
#Df = 120

#P-value Calculation
pt(6.142634,df=120, lower.tail = FALSE)
#one-sided, right-tailed t test
#P-value = 5.445521e-09
#significance level of 5% or 0.05 (alpha)
#p-value < 0.05
#5.445521e-09 < 0.05, reject the null hypothesis, which states that the
mean growth of ovx and sham group are the same.

#P-value interpretation:
#If the true mean growth for the ovx and sham group were equal (H_0 is
true), then there is a 5.445521e-09 chance that
#we would have obtained the sample means we observed from both groups
based on the data.

#There is sufficient evidence to reject the null hypothesis (H_0), which
states
#that the mean growth for the ovx and sham groups of anoles is the same,
and conclude H_a, which states that the mean growth in mm ofr ovx group is
greater than the sham group
```

```
#Confidence interval calculation: 95% CI

#Calculate t*
qt(.025,df=120)
# t* = -1.97993

-1.97993 * 0.1437494 # equals m (margin of error) -> -0.2846137

(2.309-1.426) + (-0.2846137) #  0.5983863, lower bound of 95% CI

(2.309-1.426) - (-0.2846137) #1.167614, upper bound of 95% CI


#We are 95% confident that the difference in true mean growth for the ovx
and shame group is between
#0.5983863 and 1.167614 mm.

#Conclusion:

#There is sufficient evidence to reject the null hypothesis (H_0), which
states
#that the mean growth for the ovx and sham groups of anoles is the same,
and conclude H_a, which states that the mean growth in mm ofr ovx group is
greater than the sham group
#We are 95% confident that the difference in true mean growth for the ovx
and shame group is between
#0.5983863 and 1.167614 mm.

#These findings are relevant because they suggest that there is a
reproductive cost that is associated with the growth of female anoles.
#When reproductive mechanisms are removed in OVX, they experience a
greater growth relative to the sham group that reproduces normally.
#The lack of growth that occurs in reproducing anoles could contribute to
their lower annual survival rate.
```

Works Cited

Cox, RM. (2010) *Coxetal.FunctionalEcologyData.* Rawdata.

Cox, RM. Biostats Bob Cox 100115. Retrieved from UVA Collab.