Ian Eggleston

## Frequency of Salamander Reproduction Catastrophe Rates

1.
2. P- value: 0.04097
   shapiro.test(catrate$cat.rate)
3. The null hypothesis for the Shapiro test asks whether the data is sampled from a normally distributed population.
4. There is strong enough evidence to reject the null hypothesis, suggesting that the sample did not come from a normally distributed population.
5. t.test(catrate$cat.rate, mu = 0.2857143)
6. The null hypothesis asks whether the mean of the catastrophic rate population is different than the mean of the late pond filling rate.
7. This is a 2-tailed test.
8. p-value = 0.01193
   This p-value means that it is unlikely the chance of these results being false positives, due to late pond filling, and instead are due to the catastrophic reproduction rate. The reproduction failures are not just random error due to the late pond filling.
9. 95 percent confidence interval: 0.3526250   0.7261295
   It did not include zero because it was impossible to have a negative rate here.
10. We can reject the null hypothesis based on the p-value, being less than 0.05, and the late pond filling rate is outside of the confidence intervals, compared to the catastrophic mean.
11. wilcox.test(catrate$cat.rate, mu = 2/7)
12. The Wilcox test p-value is 0.006275, which is in agreement with the t-test value of 0.01193. They both test the same null hypothesis, but the Wilcox is more suitable for smaller, non-normally distributed samples.

13. Based on the Wilcox Rank Sum test p-value being less than 0.05, we can reject the null hypothesis.
14. The results of these two tests allow us to reject the claim that the catastrophic rate and late pond-filling rate have overlapping data points. These two datasets are separate, and it strengthens the argument that the catastrophic rate is not a false positive, due to late pond filling rates.
15. The Wilcox test was more appropriate because it specializes in small sample sizes of non-normally distributed data. From the histogram we made, it is clear the data is not normally distributed and the sample size is limited to 13.
16.

```
dat_adelie = subset(penguin_dat, species == "Adelie")
dat_chinstrap = subset(penguin_dat, species == "Chinstrap")
shapiro.test(dat_adelie$flipper_length_mm)
shapiro.test(dat_chinstrap$flipper_length_mm)
```

17. The Adelie and Chinstrap Shapiro tests gave p-values of 0.72 and 0.81 respectively. Neither of these allow us to reject the null hypothesis, suggesting that they come from normally distributed data sets.
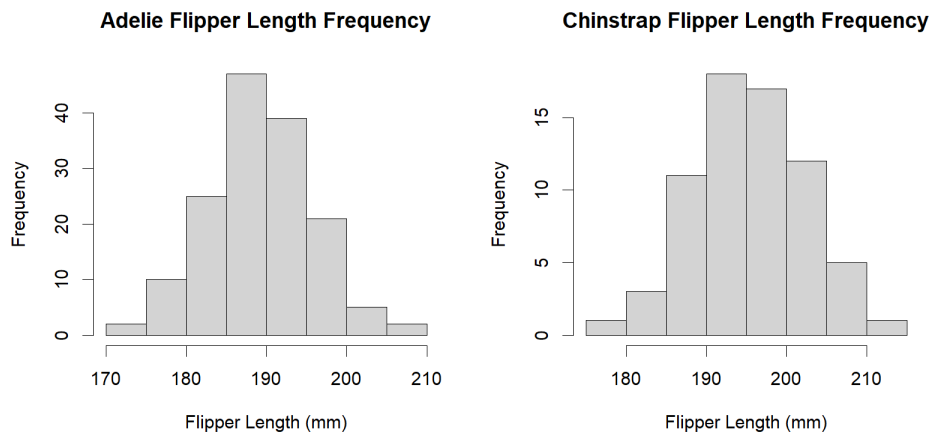
18.
19. The alternative hypothesis is testing whether the mean of Adelie flipper length is significantly different than the mean of Chinstrap flipper length. This is a two-tailed t.test.
20.

```
adelie_flip = na.omit(dat_adelie$flipper_length_mm)
chin_flip = dat_chinstrap$flipper_length_mm
t.test(adelie_flip, chin_flip)
```