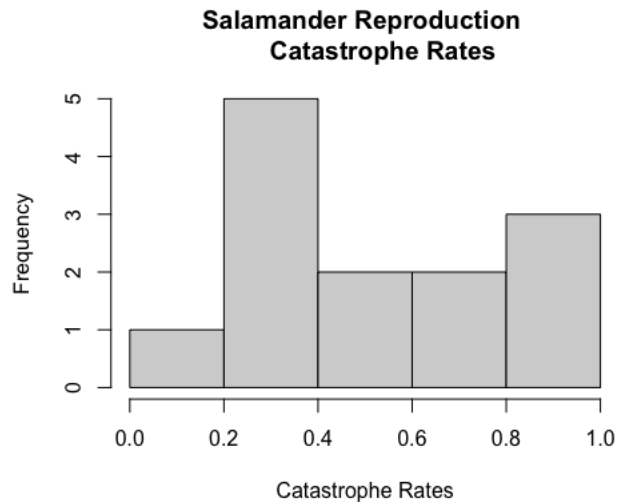Chloe Lang
Professor Nelson
Using Models 1 Assignment
November 7, 2021

1.


```
hist(catrate$cat.rate,
    main = "Salamander Reproduction
    Catastrophe Rates",
    xlab="Catastrophe Rates")
```

2. The p-value is **0.04097.** Code:
shapiro.test(catrate$cat.rate)

3. Null Hypothesis: For this test, the data are normally distributed.

4. Based on the Shapiro test results; the results appear to not be normally distributed as the p-value is less than 0.05.

5. t-test code:
t.test(catrate$cat.rate, mu = 2/7)
=0.01193

6. Null hypothesis: The salamander catastrophe rate and the pond-late filling rate are the same (28%).

7. This is a two-tailed test because you are only trying to observe if there is a difference between the two rates.

8. The p-value from my t-test was **0.01193**. Interpretation: This is meaning that there appears to be no correlation between the two rates. The p-value statistically indicates that the rates would be the same approximately 1.193% of the time.

9. Confidence interval: **0.3526250 and 0.7261295.** It did not include zero.

10. Conclusion: Based on the results of the t-test, I concluded that there was strong evidence to reject the null hypothesis. This is because the p-value calculated was less than the 0.05 point of significance in rejecting the null hypothesis. The value of 0.02293 is much lower than 0.05.

11. Wilcoxon rank sum test code:
wilcox.test(x= catrate$cat.rat, mu = 2/7)
p-value=**0.006275**

12. The p-value of the Wilcoxon test was 0.006, lower than p-value of the t-test. (0.01193), indicating there is a different between the two rates.

13. Based on my results from my rank sum test, and previous tests conducted I can conclude that there was a strong evidence based on the p-values to reject the null hypothesis.

14. Based on the results of the tests conducted, all of my results were consistent in showing their p-values were significantly lower than the 0.05 value used to reject or fail to reject a hypothesis. By both results being much lower than 0.05, I have strong evidence to conclude the two rates were of different values.

15. Based on the numerical and graphical data exploration, the Wilcoxon test would be more appropriate for these data because we know that the salamander data are not normally distributed.

16. Code used:
require(palmerpenguins)

penguin_dat = droplevels(subset(penguins, species != "Gentoo"))
summary(penguin_dat)
dat_adelie = subset(penguin_dat, species =="Adelie")
dat_chinstrap = subset(penguin_dat, species == "Chinstrap")

shapiro.test(dat_adelie$flipper_length_mm)
=0.72
shapiro.test(dat_chinstrap$flipper_length_mm)
=0.816

17. Interpretation: Based on the p-values calculated using the shapiro-wilk test, both data sets are likely normally distributed. This is because the values observed for both species were well above 0.05, indicating that we fail to reject the null hypothesis. The null hypothesis of the shapiro- wilk nest is that the data are normally distributed; thus based on the test, our data is likely normally distributed.

18. See file attachment. Code:

```
png(
  filename = here("assignments", "Using_Models_1", "flipper_histogram_species.png"),
  width =1000,
  height =750)

par(mfrow=c(1,2))
hist(
  dat_adelie$flipper_length_mm,
  main = "Adelie
  Flipper Length",
  xlab = "Flipper Length (mm)",
  col = "lightpink"
)

hist(
  dat_chinstrap$flipper_length_mm,
  main = "Chinstrap
  Flipper Length",
  xlab = "Flipper Length (mm)",
  col = "lightblue1"
)
```

19. Alternative hypothesis: the mean flipper lengths between the two species, Adelie and Chinstrap would be of different lengths.

20. Code used:
```
t.test(flipper_length_mm ~ species, data = penguin_dat)
```
= 6.049e-08