

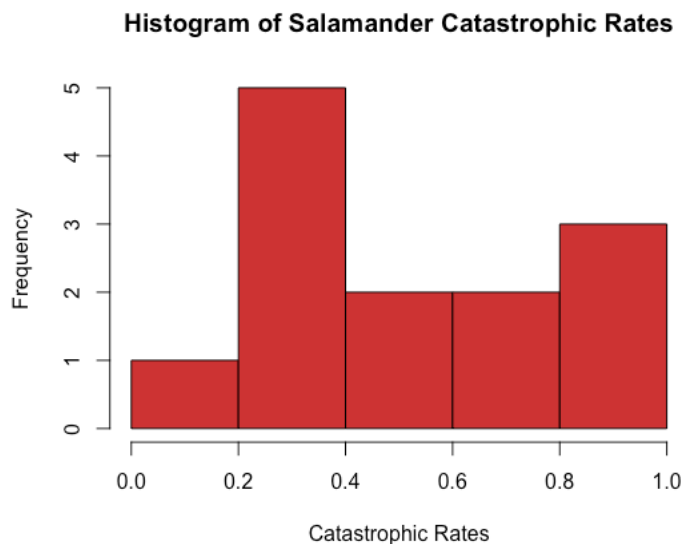
Juliana Berube

11/4/21

Using Models 1

*Worked with Julia Vineyard and Jessica Bonin

1.



2. 0.04097

`shapiro.test(catrate$cat.rate)`

3. The null hypothesis is that the data was sampled from a normally distributed population

4. Yes, the null hypothesis can be rejected because the p-value is less than 0.05, so we can accept the alternative that the data is non-normal

5. `t.test(catrate$cat.rate, mu=0.28)`

6. The null hypothesis is that the true mean is equal to 0.28

7. This is a two tailed test

8. P-value = 0.01054

If this were a false positive, then 1% of the time we would see a value that is different from the expected rate of 0.28.

9. 0.3526250 0.7261295

This confidence interval does not include 0.

10. Yes, we can reject the null hypothesis and accept the alternative. The t-test gave us a p-value of 0.01054, indicating that the result is significant, and the true mean is not 0.28. We can double check this against the true mean of the catastrophic rates, which is 0.5393773. We can clearly see that this is greater than 0.28, so we can confidently reject the null, which is that the true mean is equal to 0.28.
11. `wilcox.test(catrate$cat.rate, mu = 0.28, exact=FALSE, alternative = "g")`
12. The p-value for the Wilcoxon test is 0.003137, and the p-value for the t-test is 0.005271. The value from the Wilcoxon test is smaller than that from the t-test, although both are less than 0.05, suggesting significance.
13. Yes, there is strong evidence to reject the null given that the p-value is 0.003, which is much smaller than the 0.05 significance threshold.
14. Both tests suggest a significant relationship, which allows you to reject the null hypothesis and accept the alternative. This tells us that the true mean is greater than 0.28.
15. The Wilcoxon test is a better fit for the data, because this test is non-parametric and the data is non-normal. If the data were normally distributed, the t-test would be a better fit.
16. `dat_adelie = subset(penguin_dat, species == "Adelie")
dat_chinstrap = subset(penguin_dat, species == "Chinstrap")
shapiro.test(dat_adelie$flipper_length_mm)
shapiro.test(dat_chinstrap$flipper_length_mm)`
17. Both tests result in a p-value of greater than 0.05. With this result, we can accept the null hypothesis that the data is normally distributed.
18. Graph attached
19. The alternative hypothesis is that the average flipper length for Adelie penguins is different than the average flipper length for Chinstrap penguins. This would be considered a one-tailed test.
20. `t.test(flipper_length_mm ~ species,
data = penguin_dat)`