Julian Burgoff

11/8/2022

Analysis of Environmental Data

Using Models 1

**Histogram of catastrophe rates**
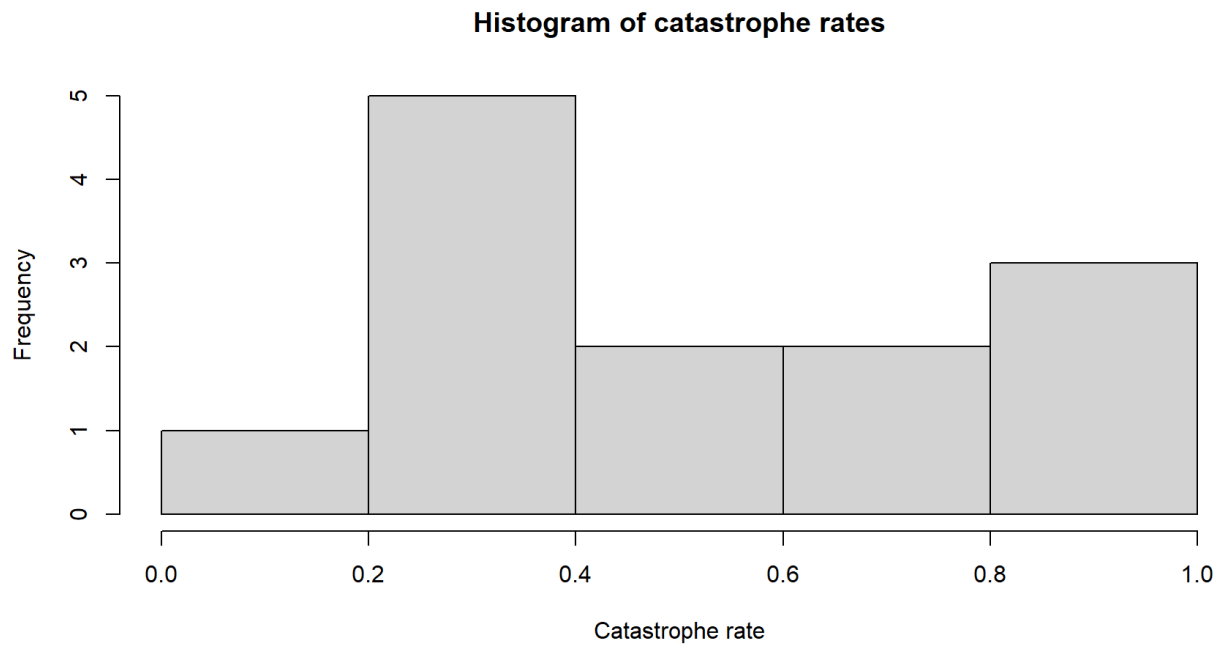


1.

2. shapiro.test(catrate$cat.rate)
   p-value = 0.04097

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

4. The p value is low which suggests good evidence against the null hypothesis, which means that there is strong evidence that the sample came from a non-normally distributed population.

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

6. The null hypothesis is that the mean of the population from which the data were collected is not different from x.

7. This is a two tailed test

8.  p-value = 0.01054. This low p value suggests that the catastrophic rate is significantly different from the pond late-filling rate.

9.
    95 percent confidence interval:
     0.3526250 0.7261295

10. We can safely reject the null hypothesis that there is no difference between the catastrophic rate and the pond late-filling rate because the p value is much lower than 0.05.

11. wilcox.test(catrate$cat.rate, mu = 2 / 7)

12. p-value = 0.006275 smaller p value than the t test

13. There is strong evidence to reject the null hypothesis given the very small p value (0.006) produced from the rank sum test.

14. The two tests led to the same conclusion to reject the null hypothesis, but considering the shape
    15. of the histogram I would say the rank sum test would be better because the data do not look normally distributed.

    16.  # Extract the Adelie penguin data
    dat_adelie = subset(penguin_dat, species == "Adelie")
    shapiro.test(dat_adelie$flipper_length_mm)

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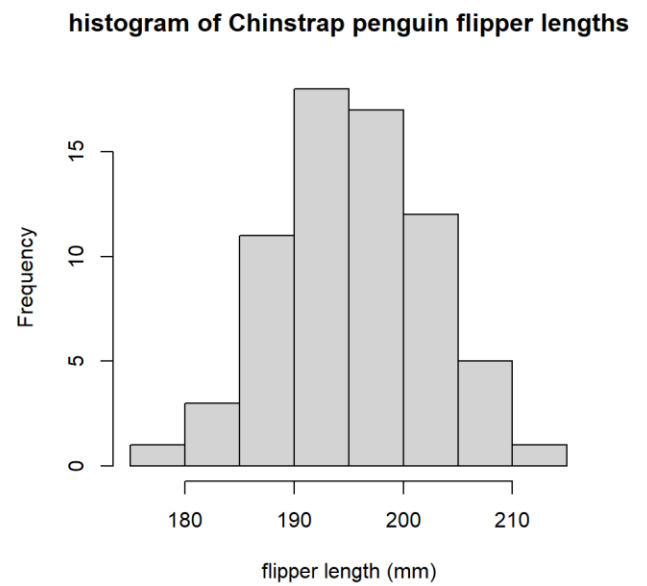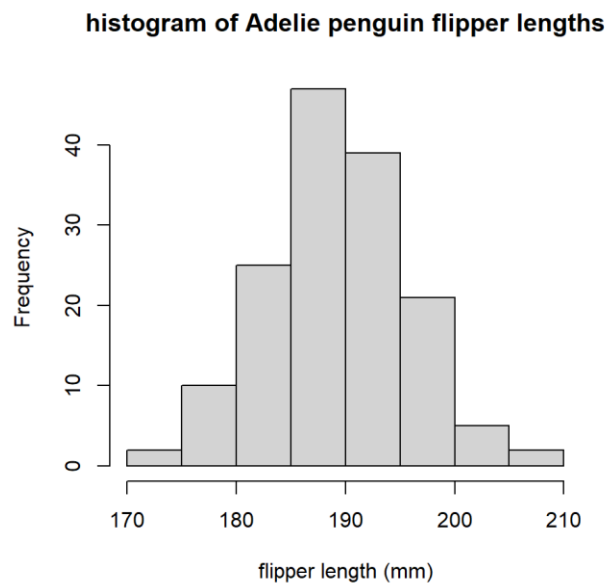    shapiro.test(dat_chinstrap$flipper_length_mm)

    17. Shapiro-Wilk normality test

        data:  dat_adelie$flipper_length_mm
        W = 0.99339, p-value = 0.72

        Shapiro-Wilk normality test

        data:  dat_chinstrap$flipper_length_mm
        W = 0.98891, p-value = 0.8106

        Both tests produced p values that were far greater than 0.05, which suggests that the flipper length data for both species of penguins is normally distributed.

**histogram of Adelie penguin flipper lengths**

**histogram of Chinstrap penguin flipper lengths**



Adelie: Frequency vs flipper length (mm)

Chinstrap: Frequency vs flipper length (mm)

18.


19. The alternative hypothesis is that chinstrap and adelie penguins have significantly different flipper lengths.

20. t.test(flipper_length_mm ~ species, data = penguin_dat)