

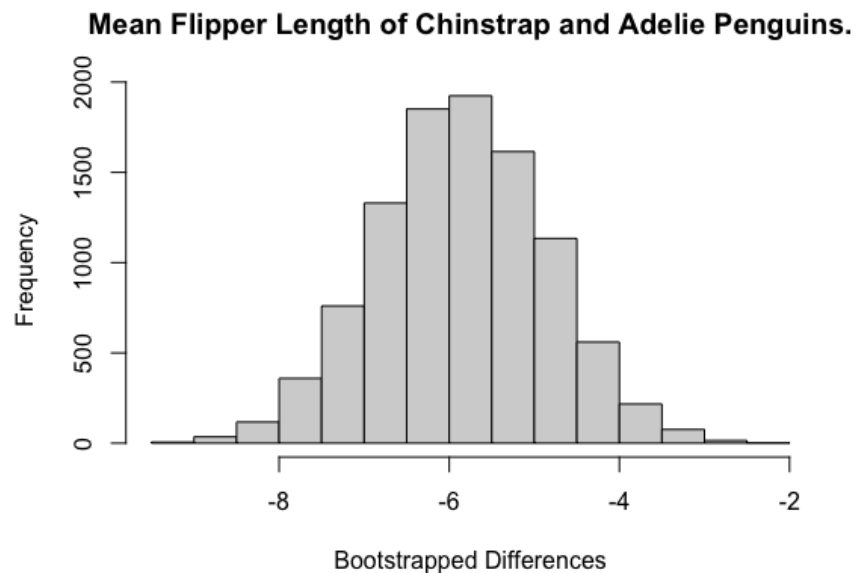
Jessica Bonin
Analysis of Environmental Data
Lab 8 Report
November 7, 2021
Worked with Juliana Berube

Q1: Calculate the standard deviation of the differences in mean flipper length from your bootstrap simulation. Show the R-code you used to find do the calculation.

```
sd(pen_boot$t)
```

1.004902

Q2: Include your histogram in your lab report (you don't need to show the R-code but make sure your plot includes appropriate title, axes, etc.).



Q3: What was the 95% bootstrap CI you calculated using quantile()? Show the R-code you used to answer the question.

```
quantile(pen_boot$t, c(.025, 0.975))
```

2.5% 97.5%
-7.831375 -3.921436

Q4: Do you think the resampled differences in means follow a skewed distribution? Your answer should make reference to the mean, median, and histogram of the differences in means.

I do not think the resampled differences in means follow a skewed distribution. I say this because the mean (-5.874678) and median (-5.882933) of the difference of means are almost exactly equal. This means the middle of the peak of the distribution and the mean are in the same place. This is also shown in the histogram because it follows a relatively normal distribution curve.

Q5: Show the R-code you used to create pen_ecdf()

```
pen_ecdf = ecdf(pen_boot$t)
```

Q6: What is the probability, according to the empirical distribution function, of observing a mean difference of -4.5 or greater? Show the R code you used to perform the calculation.

```
1 - pen_ecdf(-4.5)  
0.0923
```

Q7: What is the probability, according to the empirical distribution function, of observing a mean difference of -8 or smaller? Show the R code you used to perform the calculation.

```
pen_ecdf(-8)  
0.0164
```

Q8: State the null and alternative hypotheses of a two-sample, two-tailed test for the difference in mean flipper lengths between the two penguin species.

Null: The difference between the difference in mean flipper length between two penguin species is equal to 0.

Alternative: The difference between the difference in mean flipper length between two penguin species is not equal to 0.

Q9: What was the p-value? Show the R-code you used to find out.

```
wilcox.test(pine ~ treatment, data = dat_tree, alternative = "two.sided", exact = FALSE)  
0.1005
```

Q10: What were the endpoints of your bootstrap CI? Show the R-code you used to find out.

```
quantile(tree_boot$t, c(0.025, 0.975))  
2.5%      97.5%  
4.246875 29.375000
```

Q11: What is the observed difference in mean tree counts and does it fall within the 95% bootstrap CI?

15.92385, yes

Q12: Briefly describe the Simpson diversity index and explain what it quantifies.

It is a measure of diversity. It quantifies the diversity of species in association with other variables.

Q13: Show the code you used to z-standardize the s.sidi column.

```
s_sidi_mean = mean(dat_all$s.sidi, na.rm = TRUE)
s_sidi_sd    = sd(dat_all$s.sidi, na.rm = TRUE)
dat_all$s.sidi.standardized = (dat_all$s.sidi -
s_sidi_mean)/s_sidi_sd
```

Q14: Show the code for your completed loop.

```
dat_bird = read.csv(here("data", "bird.sub.csv"))
dat_habitat = read.csv(here("data", "hab.sub.csv"))
dat_all = merge(dat_bird, dat_habitat, by = c("basin", "sub"))
dat_1 = subset(dat_all, select = c(b.sidi, s.sidi))

m = 10000
result = numeric(m)

for(i in 1:m)
{
  index_1 = sample(nrow(dat_1), replace = TRUE)
  index_2 = sample(nrow(dat_1), replace = TRUE)

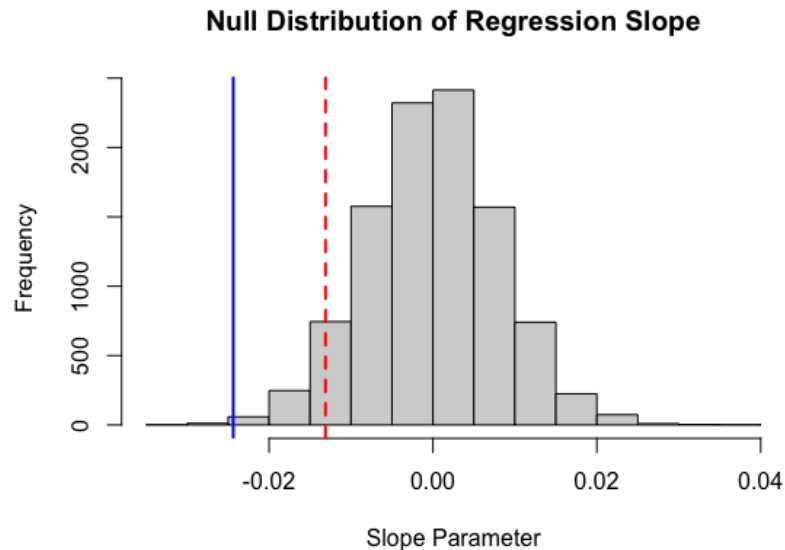
  fit_1 = lm(b.sidi ~ s.sidi, data = dat_all)
  coef(fit_1)
  slope_observed = coef(fit_1)[2]

  dat_resampled_i =
    data.frame(
      b.sidi = dat_1$b.sidi[index_1],
      s.sidi = dat_1$s.sidi[index_2]
    )

  fit_resampled_i = lm(b.sidi ~ s.sidi, data = dat_resampled_i)

  result[i] = coef(fit_resampled_i)[2]
}
```

Q15: In your report, include a plot of your histogram with vertical lines showing the observed slope and the critical value from the resampled MC slopes.



Q16: What was your critical value? Was the observed slope less than the critical value?

-0.01311585. The observed slope (blue line) was less than the critical value (red dashed line).

Q17: What is your conclusion regarding the evidence of a negative relationship between vegetation cover diversity and bird diversity? Make sure to justify your conclusions using the results of your analysis.

We have convincing evidence that bird species declines with an increase in vegetation index. This is because our observed slope is further away from 0 than our critical value. If there were no relationship, the observed slope would be nearly the same as the critical value. If there were a positive relationship, the observed slope would be greater than the critical value.