

Chloe Lang
Professor Nelson
Lab 8 Report
November 7, 2021

1. Standard deviation of the differences in mean flipper length from my bootstrap simulation:
1.002702.

Code:

```
require(palmerpenguins)
```

```
adelie = subset(penguins, species == "Adelie")
```

```
chinstrap = subset(penguins, species == "Chinstrap")
```

```
boot_mean = function(x, i)
```

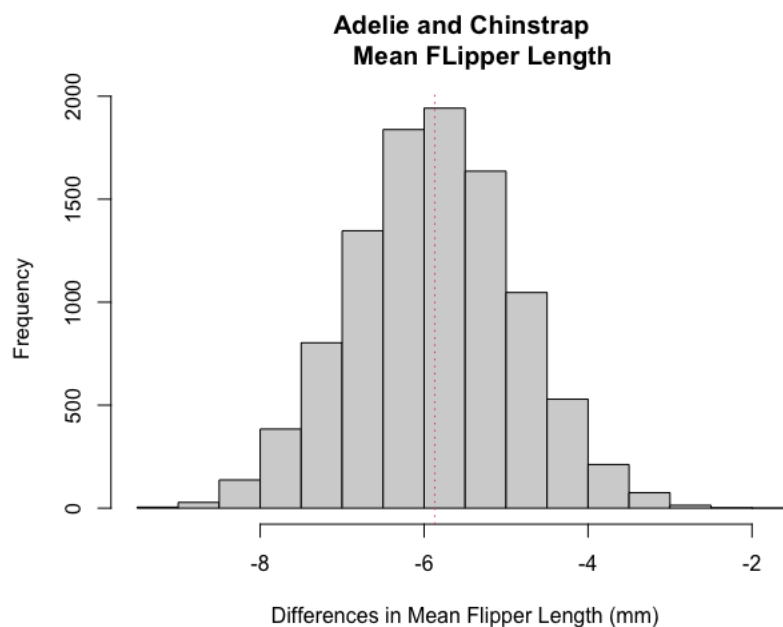
```
{  
  return(mean(x[i], na.rm = TRUE))  
}
```

```
pen_boot =
```

```
  two.boot(  
    adelie$flipper_length_mm,  
    chinstrap$flipper_length_mm,  
    FUN = boot_mean,  
    R = 10000,  
  )
```

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

2. Histogram:



3. Bootstrap CI code:

```
quantile(pen_boot$t, c(0.025, 0.975))  
= -7.84686 and -3.916363
```

4. It does not appear that the resampled differences in means follow a skewed distribution. From initial observation of the histogram, it appears to be an almost perfect bell-curve, i.e. the shape of normally distributed data. Additionally, the means and medians are almost the same, indicating not much of a difference between the two meaning the data are normally distributed.

5. Code for pen_ecdf():

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

6. Probability of 4.5 or greater: **0.0833 or 8.83%.**

Code:

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

7. Probability of -8 or smaller: **0.017 or 1.17%.**

Code:

```
pen_ecdf(-8)
```

8.

Null: There will be no difference in mean flipper length between the two penguin species.

Alternative: There will be a difference in mean flipper length between the two penguin species.

9. p-value: **0.1005**

Code:

```
dat_veg = droplevels(subset(vegdata, treatment %in% c("control", "clipped")))
wilcox.test(pine ~ treatment, data=dat_veg)
```

10. bootstrap CI: **3.875 and 29.75.**

Code:

```
tree_boot =  
  two.boot(  
    subset(dat_tree, treatment == "clipped")$pine,  
    subset(dat_tree, treatment == "control")$pine,  
    FUN = mean,  
    R = 10000,  
    na.rm = TRUE  
  )  
quantile(tree_boot$t, c(0.025, 0.975))
```

11. The observed difference in mean tree counts was **16**. This number does fall within the 95% bootstrap CI.

12. Simpson Diversity Index: This is an index used to measure diversity within nature; typically the diversity in species. It is used to quantify the biodiversity of a habitat. This is obtained by

gathering data on the number of species present and the number of individuals observed for those species.

13. Code to z-standardize the s.sidi column:

```
head(b_sidi_mean)
head(b_sidi_sd)
b_sidi_standarization = (dat_all$b.sidi - b_sidi_mean) / b_sidi_sd

s_sidi_mean = mean(dat_all$s.sidi, na.rm = TRUE)
s_sidi_SD = sd(dat_all$b.sidi, na.rm = TRUE)
s_sidi_standardization = (dat_all$s.sidi - s_sidi_mean) / s_sidi_SD
```

14. Code for completed loop:

```
fit1 = lm(b.sidi ~ s.sidi, data = dat_all)
coef(fit1)

slope = coef(fit1)[2]

dat1 =
  subset(
    dat_all,
    select = c(b.sidi, s.sidi))

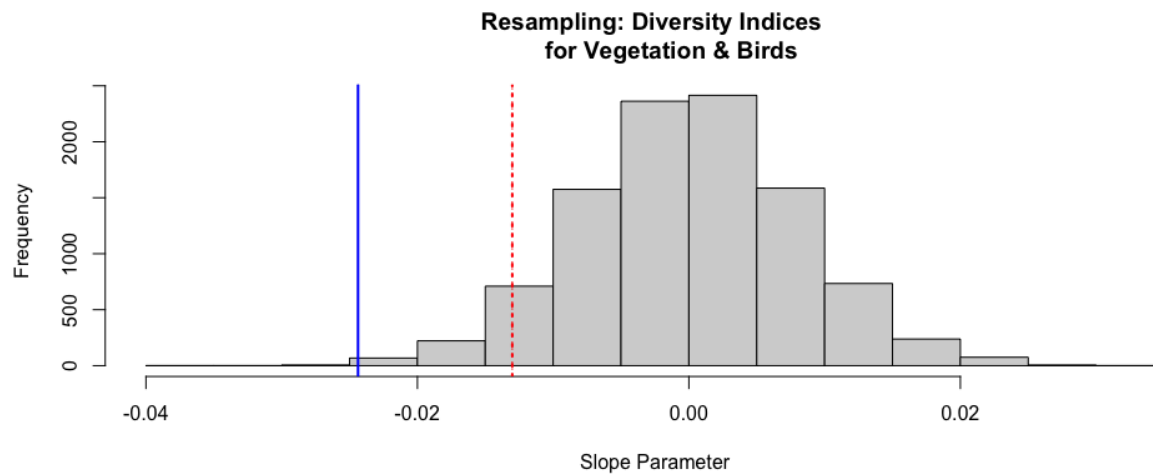
m = 10000
result = numeric(m)

for(i in 1:m)
{
  index_1 = sample(nrow(dat1), replace = TRUE)
  index_2 = sample(nrow(dat1), replace = TRUE)

  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]
}
```

15. Histogram:



16. Critical Value: **-0.0130132**. The observed slope was -0.02437131, so the observed slope was less than the critical value.

17. Conclusion: Based on my calculations and analysis, I would conclude that there appears to be a negative correlation between vegetation density and bird density. This indicates that as vegetation density increases, bird density decreases. Both the critical value calculated and the observed slope being negative, these analyses support this conclusion.