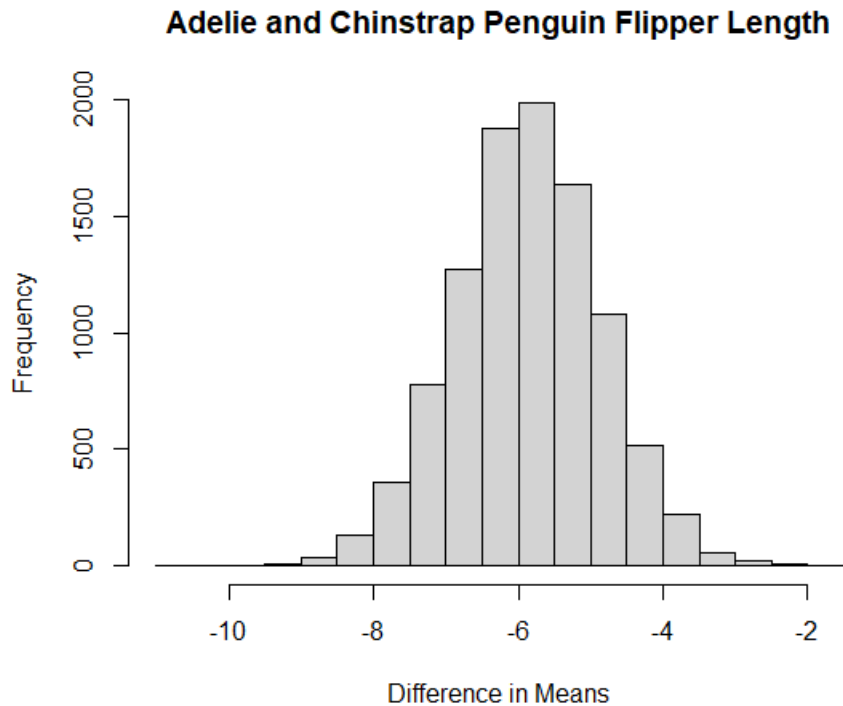


Julia Vineyard

Worked with: Andrew Gordon, Juliana Berube, Jessica Bonin

Lab 8: Modeling 1

1. `sd= 1.002531`
`sd(pen_boot$t)`



3. `2.5%= -7.838642, 97.5%= -3.891121`
`quantile(pen_boot$t, c(0.025, 0.975))`
4. No, the resampled differences in means does not follow a skewed distribution. The histogram resembles a normal distribution, and when observing the mean and median the data is centered around both. (mean= -5.871841, median= -5.867501)
5. `pen_ecdf= ecdf(pen_boot$t)`
6. The probability of observing a mean difference of -4.5 or greater is 0.0822.
`pen_ecdf(-4.5)`
7. The probability of observing a mean difference of -8 or smaller is 0.0177.
`pen_ecdf(-8)`

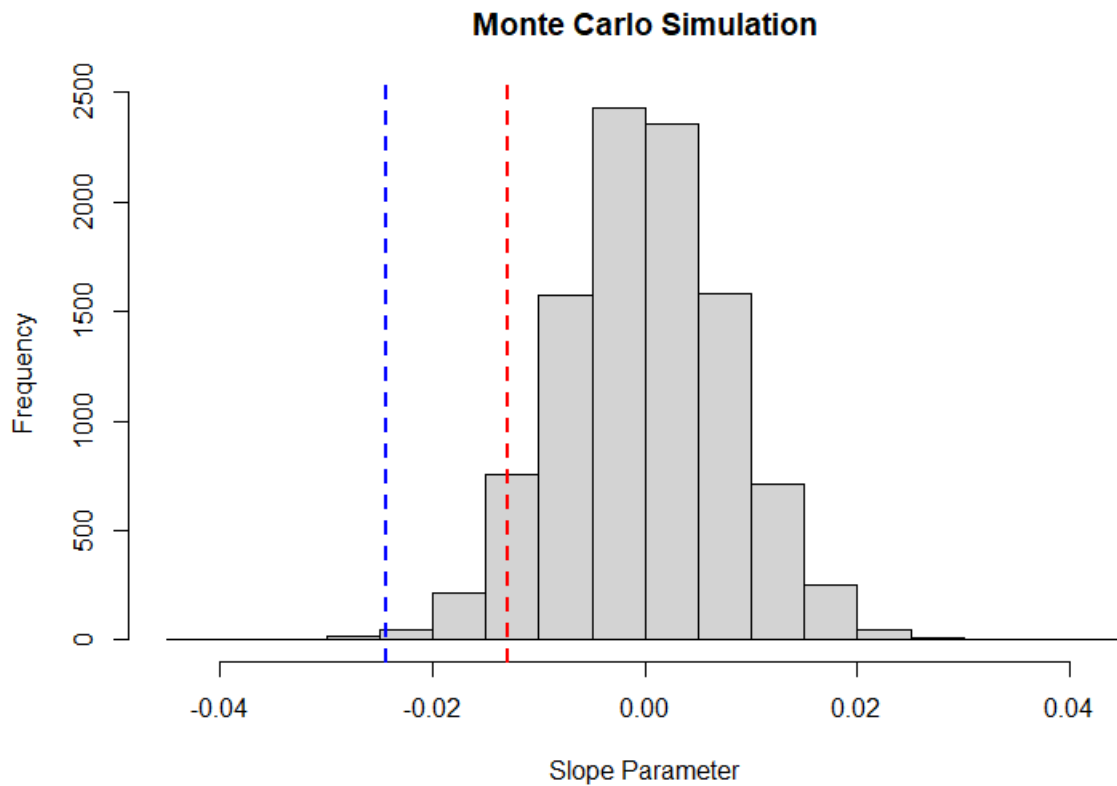
8. The null hypothesis: the difference in mean flipper length between Adeline and Chinstrap penguins is equal to 0. The alternative hypothesis: the difference in mean flipper length between Adelie and Chinstrap penguins is not equal to 0.
9. $p\text{-value}=0.1005$

```
veg= read.csv(here("data", "vegdata.csv"))  
dat_tree= droplevels(subset(veg, treatment %in% c("control", "clipped")))  
wilcox.test(pine ~ treatment, data=dat_tree)
```
10. 3.871875. 30.250000

```
tree_boot= two.boot(subset(dat_tree, treatment == "clipped")$pine,  
                    subset(dat_tree, treatment == "control")$pine,  
                    FUN=mean, R=1000, na.rm=TRUE)  
  
quantile(tree_boot$t, c(0.025, 0.975))
```
11. 16.26138 Yes it falls within the 95% bootstrap CI
12. Simpson's Diversity Index is a measure of biodiversity that takes into account the number of species present and the relative abundance of each species.
13.

```
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
```
14. $m = 10000$

```
result = numeric(m)  
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))  
  
for(i in 1:m)  
{  
  index_1=sample(nrow(dat_1), replace = TRUE)  
  index_2=sample(nrow(dat_1), 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]  
}  
print(result[i])
```



16. Critical value= -0.01338145. The observed slope is lower than the critical value. Slope observed= -0.02437131
17. Since our critical value is negative we can infer that there is a negative relationship between vegetation cover diversity and bird diversity, meaning that as vegetation cover diversity increases the bird diversity decreases.