

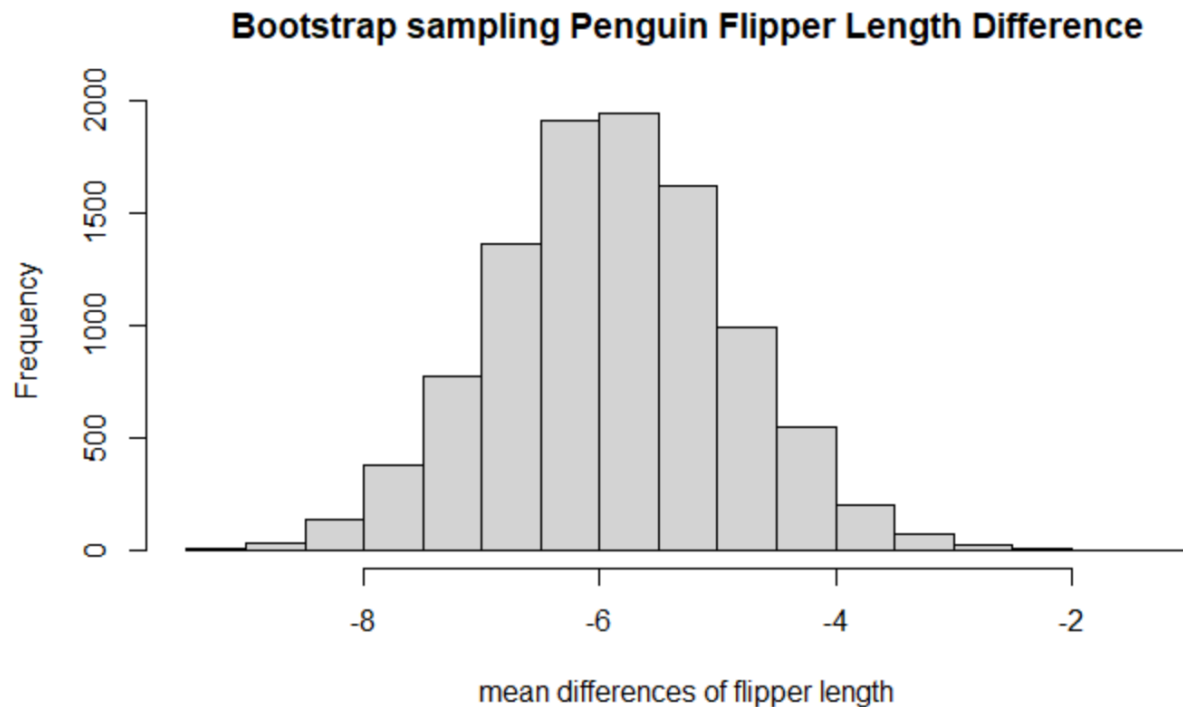
Heather Siart
ECO – 634 Analysis of Environmental Data – LAB
11/8/2021
Partners: Jahiya, Matt, Many, Sonja

Lab 8: Modeling 1

Q1: Standard deviation = 1.002915

```
Require(palmerpenguins)
penguin_dat = droplevels(subset(penguins, species != "Gentoo"))
t.test(flipper_length_mm ~ species, data = penguin_dat, alternative = "less")
pen_boot =
  two.boot(
    subset(penguin_dat, species == "Adelie")$flipper_length_mm,
    subset(penguin_dat, species == "Chinstrap")$flipper_length_mm,
    FUN = mean,
    R = 10000,
    na.rm = TRUE
  )
mean(pen_boot$t)
sd(pen_boot$t)
```

Q2:



Q3: To get 95% quantile we used 2.5% and 97.5% and got 2.5% = -7.855263 and 97.5% = -3.919560

```
quantile(pen_boot$t, c(0.025))
```

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

Q4: The resampled difference of mean does not look skewed. It looks to be close to a normal distribution.

Q5: `pen_ecdf = ecdf(pen_boot$t)`

Q6: 0.0846

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

Q7: 0.0178

```
pen_ecdf(-8)
```

Q8: The null hypothesis is that there is no difference between the mean flipper length for the chinstrap and adelinie penguins

The alternative hypothesis is that there is a difference between the mean flipper length for the chinstrap and adelinie penguins.

Q9: The p-value is 0.1005

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

Q10: The endpoints are 30 for the 97.5% and 4.25 for the 2.5%

```
quantile(tree_boot$t, 0.975)
```

```
quantile(tree_boot$t, 0.025)
```

Q11: The observed difference in mean tree counts is 16 and it does fall within the 95% bootstrap CI.

Q12: The Simpson diversity index is a way to measure the diversity within a community. It uses species abundance and richness to determine this.

Q13:

```
dat_1 = subset(dat_all, select = c(b.sidi, s.sidi))
```

```
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:

```
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)

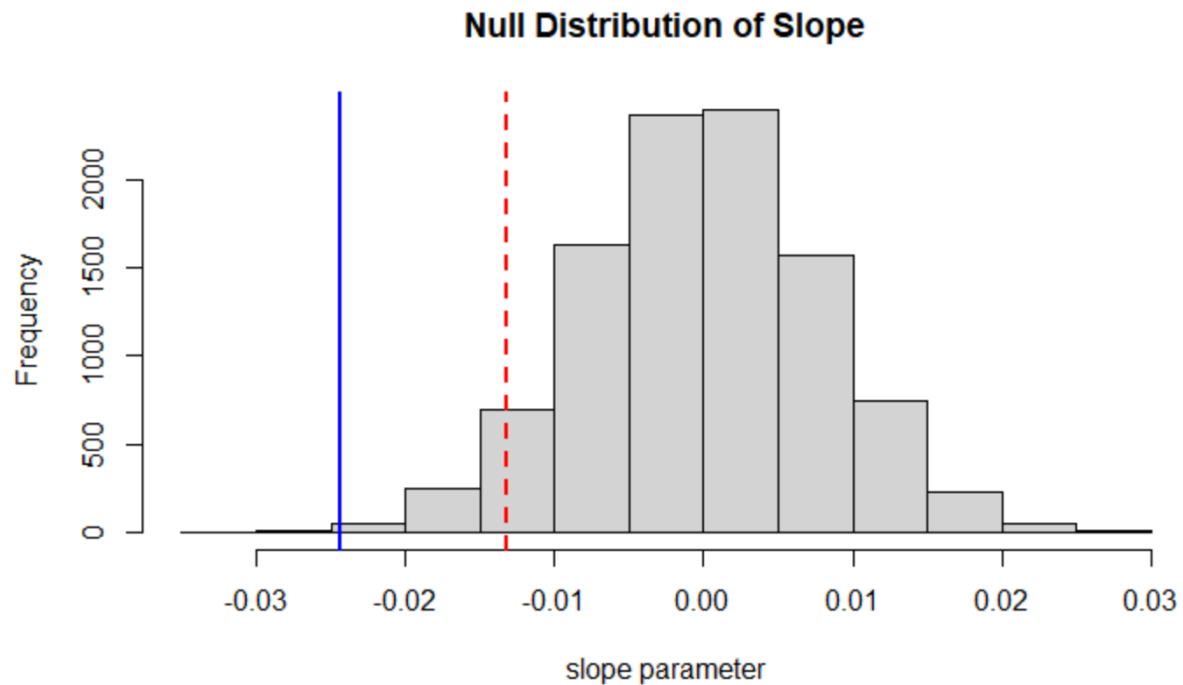
  dat_resampled_i =
  data.frame(b.sidi = dat_1$b.sidi[index_1], s.sidi = dat_1$s.sidi[index_2])

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

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

```

Q15:



Q16: The critical value is -0.01322608 which is greater than the slope observed which is -0.02437131.

Q17: Based off the Simpson Diversity index there appears to be a significant difference between habitat and bird diversity. Because the crit value is greater than the observed slope it shows that there is some sort of correlation between them.