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:** Chart, histogram

Description automatically generated

**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 adelie penguins

The alternative hypothesis is that there is a difference between the mean flipper length for the chinstrap and adelie 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:**Chart, histogram

Description automatically generated

**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.