Julian Burgoff

10/17/22

Analysis of Environmental Data

Lab 6

1.
```
rm(list = ls())
sse_mean = function(x)
{
 sd(x, na.rm= TRUE)/sqrt(length(x))
}
sse_mean(penguins$body_mass_g)
sse_mean(mtcars$mpg)
```

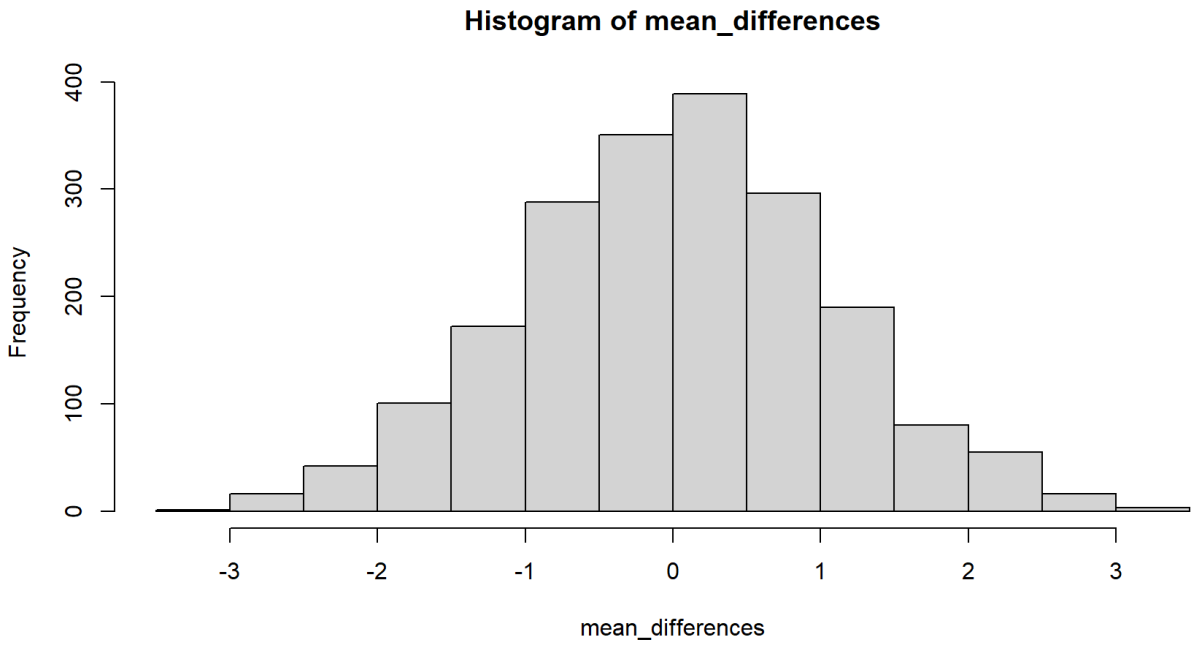
2.
```
two_group_resample_diff = function(x, n_1, n_2)
{
 dat_1 = sample(x, n_1, replace = TRUE)
 dat_2 = sample(x, n_2, replace = TRUE)

 difference_in_means =
  mean(dat_1, na.rm = TRUE) - mean(dat_2, na.rm = TRUE)

 return(difference_in_means)
}
```

3. MC

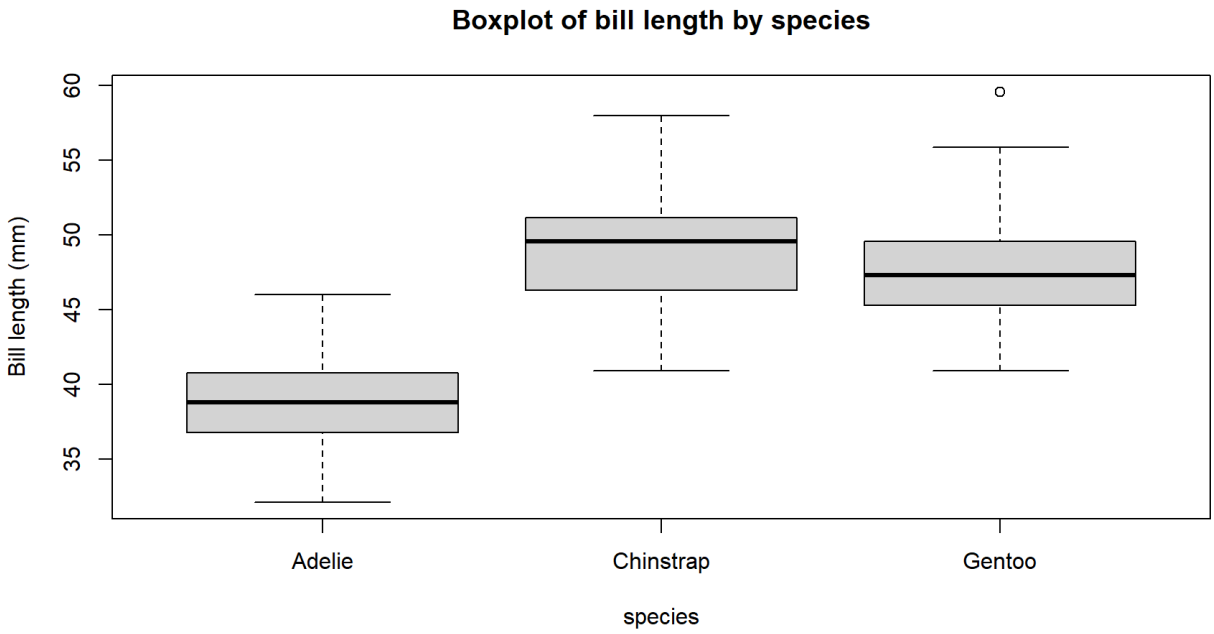
4.



5.
```
n = 2000
mean_differences = c()
for (i in 1:n)
{
 mean_differences = c(
  mean_differences,
  two_group_resample_diff(dat_pen$flipper_length_mm, 68, 152)
 )
}
hist(mean_differences)

sum(abs(mean_differences)>5.8)
```

6. Over 10 million



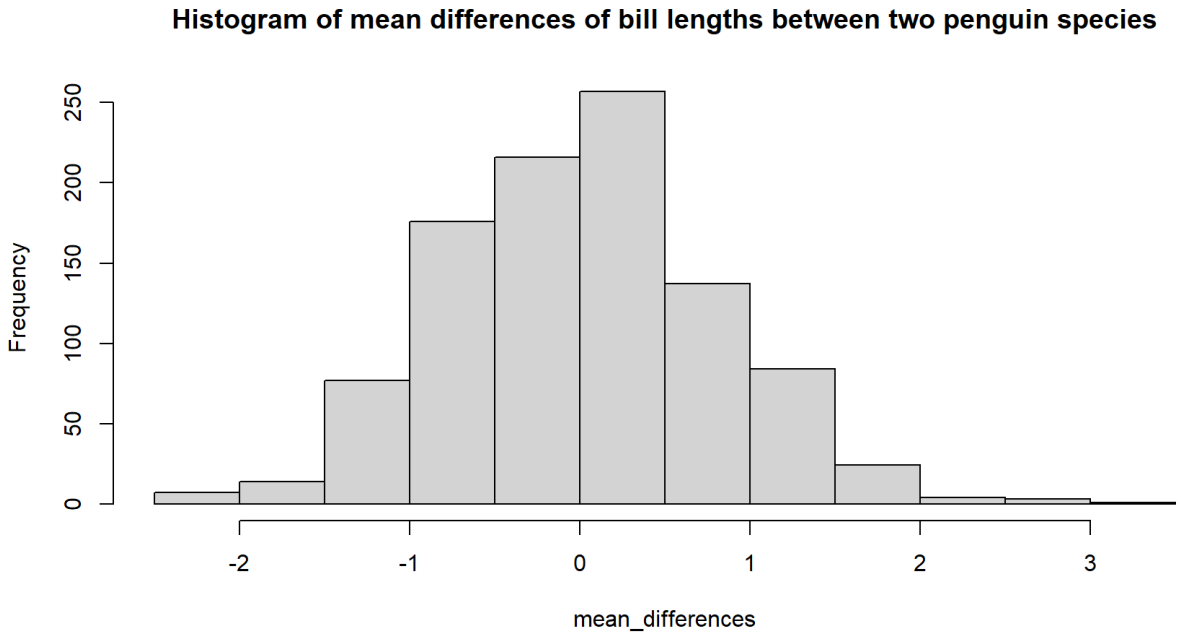
7.

8. agg_means
   species bill_length_mm
   1 Adelie 38.79139
   2 Chinstrap 48.83382
   diff_crit
   [1] 10.04243

9. p-value < 2.2e-16
   The p value is extremely close to zero, meaning the means of the two groups are very different from one another. There is less than one trillionth of one percent of a chance that there is no difference between the means of the two groups.

10. sum(abs(mean_differences) >= diff_crit)
    [1] 0

11.


Histogram of mean differences of bill lengths between two penguin species
Frequency
0
50
100
150
200
250
-2
-1
0
1
2
3
mean_differences