Jahiya Clark
Prof. Nelson
Lab 7
10/28/2021

<div align="center">Lab 7</div>

1. N= 123

```
dat_pen = subset(penguins, species == "Gentoo")
sum(as.numeric(is.na(dat_pen$bill_length_mm)))
n_pen <- length(dat_pen$bill_length_mm) - 1
```

2. Sample standard deviation is 3.081857

```
sd_pen <- sd(dat_pen$bill_length_mm, na.rm = TRUE)
```

3. T critical is -1.9796,  1.9796

```
pen_crit <- qt(c(0.025, 0.975), df = n_pen-1)
pen_crit
```

4. SSE is 0.2778817

```
sse_pen <- (sd_pen)/sqrt(n_pen)
sse_pen
```

5. The parametric CI is $47.50488 \pm 0.5500946$

```
radius_pen <- pen_crit * sse_pen
gen_pen <- mean(dat_pen$bill_length_mm, na.rm = TRUE)
ci_pen <-gen_pen + radius_pen
ci_pen
```

6. Lower: 47.06523 Upper: 48.00394

7.

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

require(boot)
billboot = boot(
  data = dat_pen$bill_length_mm,
  statistic = boot_mean,
  R = 10000)

print(billboot)
```

8.
```
quantile(
  billboot$t,
  c(0.025, 0.975))
  2.5%   97.5%
46.96195 48.05366
```


9.

```
moths = read.csv(here("data", "moths.csv"))
moth_dat = moths[,-1]

rarefaction_sampler = function(input_dat, n_iterations)
{
  n_input_rows = nrow(input_dat)

  results_out = matrix(
    nrow = n_iterations,
    ncol = n_input_rows)

  for(i in 1:n_iterations)
  {
    for(j in 1:n)
    {
      rows_j = sample(n, size = j, replace=TRUE)

      t1 = input_dat[rows_j, ]

      t2 = apply(t1, 2, sum)

      results_out[i, j] = sum(t2 > 0)
    }
  }
  return(results_out)
}
```
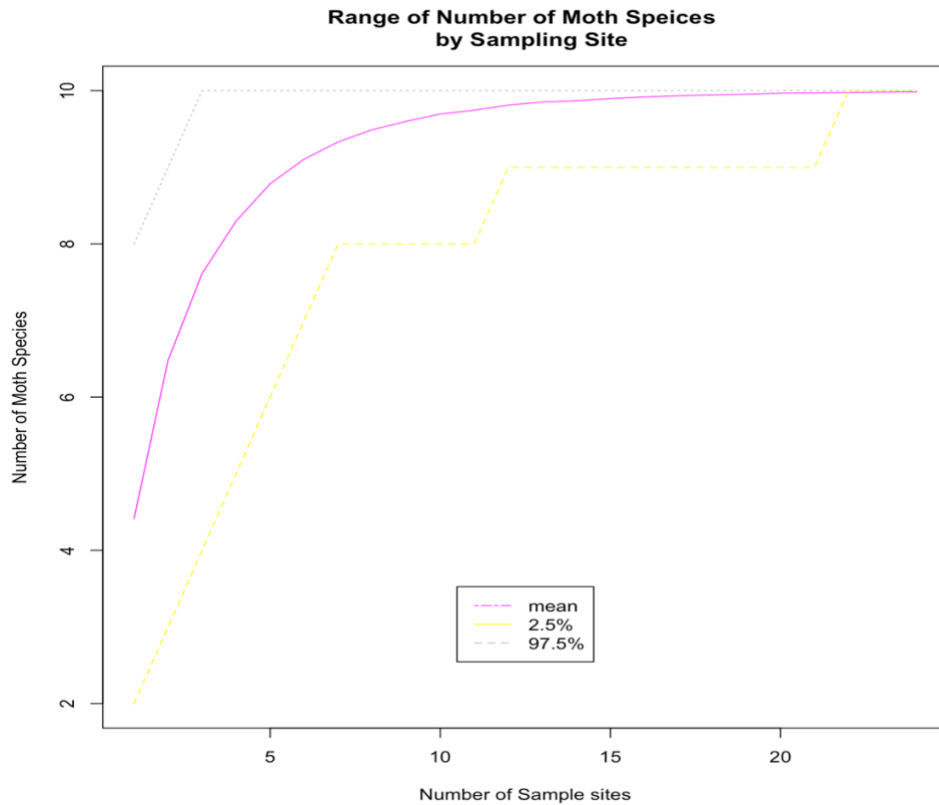
10. I wasn't able to add shading to the confidence envelope.

11.
```
rarefact = rarefaction_sampler(moths[,-1], 10000)

rarefact

rare_mean = apply(rarefact, 2, mean)
rare_quant = apply(rarefact, 2, quantile, probs=c(0.025, 0.975))
rare = t(rbind(rare_mean, rare_quant))
```

**Range of Number of Moth Speices
by Sampling Site**



12.
```
matplot(
 rare,
 type='l',
 xlab='Number of Sample sites',
 ylab='Number of Moth Species',
 main='Range of Number of Moth Speices
 by Sampling Site',
 col = 6:8)


legend(
 'bottom',
 legend=c('mean','2.5%','97.5%'),
 lty=c(6,7,8),col=c(6,7,8), inset=c(.1,.1))
```

13. I would visit 20 sites to guarantee that I will see all the moth species. At 20 sites, the lower and upper quantile meet with the mean.