

Analysis of Environmental Data - Lab 7

Olivia Dinkelacker

Q1 (1 pt.): What is the sample size, n ? Show the code you used for the calculation and remember to check for missing data.

```
n = sum(!is.na(dat_pen$bill_length_mm))
n = 123
```

Q2 (1 pt.): What is the sample standard deviation? Show the code you used for the calculation.

```
> sd(dat_pen$bill_length_mm, na.rm = TRUE)
[1] 3.081857
```

Q3 (2 pts.): What are the critical t-values? Show the R code you used for the calculation.

```
> t_crit = qt(c(0.025, 0.975), df = length(dat_pen$bill_length_mm)-1)
> t_crit
[1] -1.979439 1.979439
```

Q4 (1 pt.): What is the sample standard error? Show the R code you used for the calculation.

```
> sse = sd(dat_pen$bill_length_mm, na.rm = TRUE) / sqrt(n)
> sse
0.2778817
```

Q5 (2 pts.): Finally, construct the CI and show the R code you used for the calculation.

```
billlength_ci = c(
  lower = mean(dat_pen$bill_length_mm, na.rm = TRUE) - ci_radius,
  upper = mean(dat_pen$bill_length_mm, na.rm = TRUE) + ci_radius)

print(round(billlength_ci, 4))
```

Q6 (1 pt.): What is the CI?

```
2.5% 97.5%
46.96178 48.04759
```

```
m = 10000
result = numeric(m)
head(result)
```

```
for(i in 1:m)
{
  result[i] = mean(sample(dat_pen$bill_length_mm, replace=TRUE), na.rm = TRUE)
}
mean(result)
quantile(result,c(0.025,0.975))
```

Q7 (1 pt.): Show the r code you used to call the `boot()` function.

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

```
myboot =
  boot(
    data = dat_pen$bill_length_mm,
    statistic = boot_mean,
    R = 10000)
print(myboot)
```

Q8 (2 pts.): Show the r code you used to calculate the upper and lower 2.5% quantiles.

```
sd(myboot$t)
quantile(
  myboot$t,
  c(0.025, 0.975))
```

Q9 (5 pts.): Show your completed `rarefaction_sampler()` function.

```
##Debugging tempplate/see if worked
# This clears the current R session's environment
rm(list = ls())

# Re-read my data:
moths = read.csv(here("data", "moths.csv"))

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

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

  # The outer loop: runs once for each bootstrap iteration. index variable is i
  for(i in 1:n_iterations)
  {
    # The inner loop: simulates increasing sampling intensity
    # Sampling intensity ranges from 1 site to the complete count of
    # sites in the input data (n)
    for(j in 1:n_input_rows)
    {
      # sample the input data row indices, with replacement
```

```

rows_j = sample(n_input_rows, size = j, replace=TRUE)

# Creates a new data matrix
t1 = input_dat[rows_j, ]

# Calculates the column sums
t2 = apply(t1, 2, sum)

# Counts the number of columns in which any moths were observed
results_out[i, j] = sum(t2 > 0)
}
}
return(results_out)
}

```

Q10 (1 pt.): What did you find most difficult about building the function?

It was the most difficult to understand what all the arguments mean, especially when the vector names are replaced.

Q11 (4 pts.): Show the code you used to perform the simulations and construct the curve.

```

rarefact = rarefaction_sampler(moths[, -1], 100)
head(rarefact)

```

Q12 (4 pts.): Include your rarefaction curve plot in your report. Show the R-code you used to create your plot.

```

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

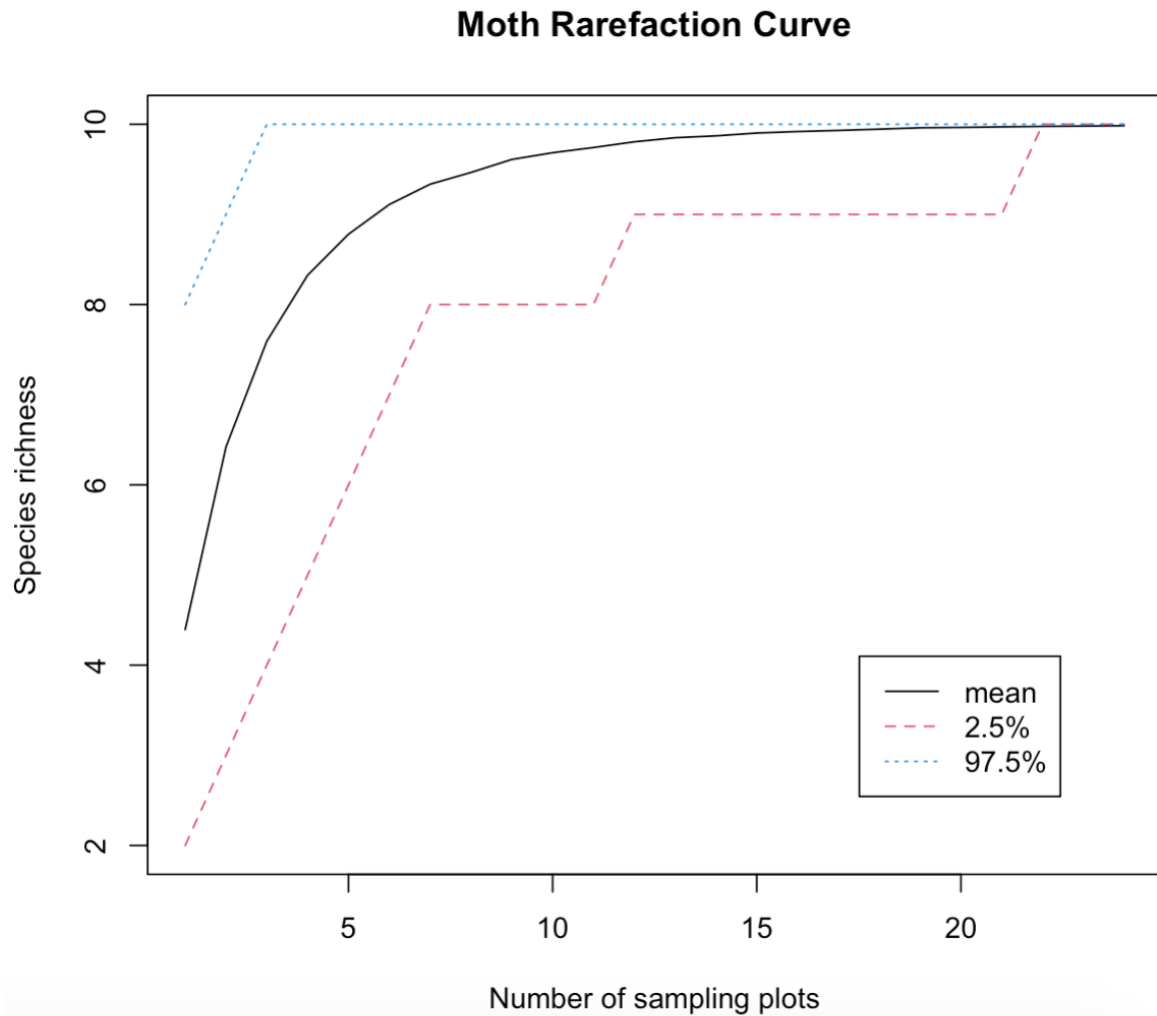
rare_mean = apply(rarefact, 2, mean)
rare_quant = apply(rarefact, 2, quantile, probs=c(0.025, 0.975))
rare = t(rbind(rare_mean, rare_quant))

##plotting the curve
matplot(
  rare,
  type='l',
  xlab='Number of sampling plots',
  ylab='Species richness',
  col = c(1,2,4),
  main="Moth Rarefaction Curve")

legend(

```

```
'bottomright',  
legend=c('mean','2.5%','97.5%'),  
lty=c(1,2,3),col=c(1,2,4), inset=c(.1,.1))
```



Q13 (2 pts.): About how many sites should you visit if you want to see all of the moth species? Explain your reasoning using your rarefaction curve figure.

I should visit as many sampling sites as possible, in this case, based on the plot, it should be at least 20 plots. The more sampling plots we visit, the higher will be the observed species richness.