Module 3 Assignment 2

Ellen Bledsoe

2023-04-10

Estimating Abundances

Assignment Details

Purpose

The goal of this assignment is to understand, apply, and interpret detection probability in estimating abundance through capture-mark-recapture.

Task

Write R code to successfully answer each question below or write text to successfully answer the question.

Criteria for Success

- Code is within the provided code chunks
- Code chunks run without errors
- Code produces the correct result

    - Code attempts will get half points
    - Code that produces the correct answer will receive full points

- Text answers correctly address the question asked

Due Date

April 27 at 2pm MST

Assignment Questions

Set-Up

In this assignment, we will be exploring methods for estimating abundances when detection is less than perfect. We will be using some data we create as well as some simulated data.

Capture-Mark-Recapture

Data   We are running a capture-mark-recapture study on pocket mice. To do this, we have set out a 10 x 10 grid of Sherman live traps.

We catch 37 desert pocket mice, *Chaetodipus penicillatus*, during our first night of trapping. To mark them, we use a green Sharpie to put a dot on their belly (and giggle because it now looks like they have a Leprechaun's belly button).

The next night, we set out the same number of traps. In the morning, we find that we have caught 42 desert pocket mice, and 30 of them have green dots on their bellies.

We want to estimate the total abundance of desert pocket mice in our population using the Lincoln-Peterson estimator.

1. Create objects with the correct values from the scenario above to represent the following:

   - `n1`: # individuals captured (and marked) in the first sample
   - `n2`: # individuals captured in the second sample (total)
   - `m2`: # individuals in the second sample which were marked

```
n1 <- 37
n2 <- 42
m2 <- 30
```

2. Use the Lincoln-Peterson estimator to estimate the population total (`N_hat`)

```
N_hat <- n1 * n2 / m2
N_hat
```

```
## [1] 51.8
```

3. Remember, we can consider the Lincoln-Peterson estimator equation as a version of the "canonical abundance estimator," meaning we can calculate an estimate of detection probability. Calculate our detection probability estimate for desert pocket mice.

```
p_hat <- m2 / n1
p_hat
```

```
## [1] 0.8108108
```

During this same trapping experiment, we also caught 12 Bailey's pocket mice, *Chaetodipus baileyi*, on the first night and marked them. The second night, we caught another 12 Bailey's, 4 of which were previously marked.

4. Assign the correct values for `n1`, `n2`, and `m2` for the Bailey's pocket mice.

```
n1_Baileys <- 12
n2_Baileys <- 12
m2_Baileys <- 4
```

5. Using those values, use the Lincoln-Peterson estimator to estimate the population total of Bailey's.

```
N_hat_Baileys <- (n1_Baileys * n2_Baileys) / m2_Baileys
N_hat_Baileys
```

```
## [1] 36
```

6. Calculate the estimated detection probability of Bailey's pocket mouse.

```
p_hat_Baileys <- m2_Baileys / n1_Baileys
p_hat_Baileys
```

```
## [1] 0.3333333
```

7. In your own words, briefly explain what the differences in estimated detection probabilities and abundances mean in the context of this example. (2 points)

   Answer: Baileyi have a lower detection probability (0.33) than penicillatus (0.81). Baileyi have a smaller estimated abundance than penicillatus.

How Precision Changes with Detection Probability

We can use data simulation to explore what happens to our estimates when detection probability changes through time.

First, we'll walk through an example of the code.

To be very clear, I don't expect you to know how this code is working! What I want you to be able to do, though, is interpret our results from the code. Lower abundances could potentially impact detection rates, but that is not necessarily the case–other factors about species life history and behavior could influence detection rates, as well.

Building the Simulation   Let's run a simulation of capture-mark-recapture data. We will stick with pocket mice as our example species.

In this case, we will tell R the *true* abundance, N. Remember, though, this is the parameter we never know in real life; it is the parameter we are trying to estimate by calculating `N_hat`.

We will set our true abundance of pocket mice to be 250 individuals.

```
N <- 250
```

We want to create a simulation where we give R the detection and recapture probabilities and determine the *estimate* of abundance (`N_hat`) that we will get based on those probabilities. We want to do this over and over and over again to help us get an idea of how good our estimates are overall. In fact, let's do it 1000 times.

```
nRuns <- 1000
```

Below, I've written a custom function that will create `N_hat` estimates for us. We need to provide the function with values in order for it to run:

- `p1`: the *detection* probability on the first survey occasion

- `p2`: the *detection* probability on the second survey occasion
- `c2`: the *recapture* probability on the second survey occasion

Run the code chunk below to add this function into your work space. You will see it pop up in the environment tab.

Now we can use this function as we would use any other function in R! Again, I don't expect you to know how it works—just know that we will provide the function with our 3 values as arguments and the result will be our simulated data.

```r
LP_estimator <- function(p1, p2, c2) {

  # make sure we get the same numbers every time
  set.seed(4)

  # create an empty data frame to hold simulated values
  Nhat <- matrix(NA, nRuns)  # create a matrix to hold estimates

  for(i in 1:nRuns) {
    # simulate data
    n1 <- rbinom(1, N, p1)      # no. individuals captured on survey occasion 1
    u2 <- rbinom(1, N-n1, p2)   # no. marked individuals recaptured on survey occasion 2
    m2 <- rbinom(1, n1, c2)     # no. marked individuals recaptured on survey occasion 2
    n2 <- u2 + m2               # total no. individuals captured on survey occasion 2

    # calculate estimated abundance via (modified) Lincoln-Peterson
    Nhat[i] <- (n1 )*(n2) / (m2)
  }

  return(Nhat) # the function will return our estimate of the population
}
```

Let's Practice! Let's say that our detection probabilities (`p1` and `p2`) and recapture probability (`c2`) are consistent, always at 0.4.

We can input those values into our `LP_estimator` function. Doing this will run our scenario 1000 times to see how close our estimate of mice, `N_hat`, is to the true abundance, `N`, which we have set to be 250 mice.

```r
N_hat_demo <- LP_estimator(p1 = 0.4, p2 = 0.4, c2 = 0.4)
```

Now, let's plot a histogram of our estimates.

```
hist(N_hat_demo)
```

Histogram of N_hat_demo
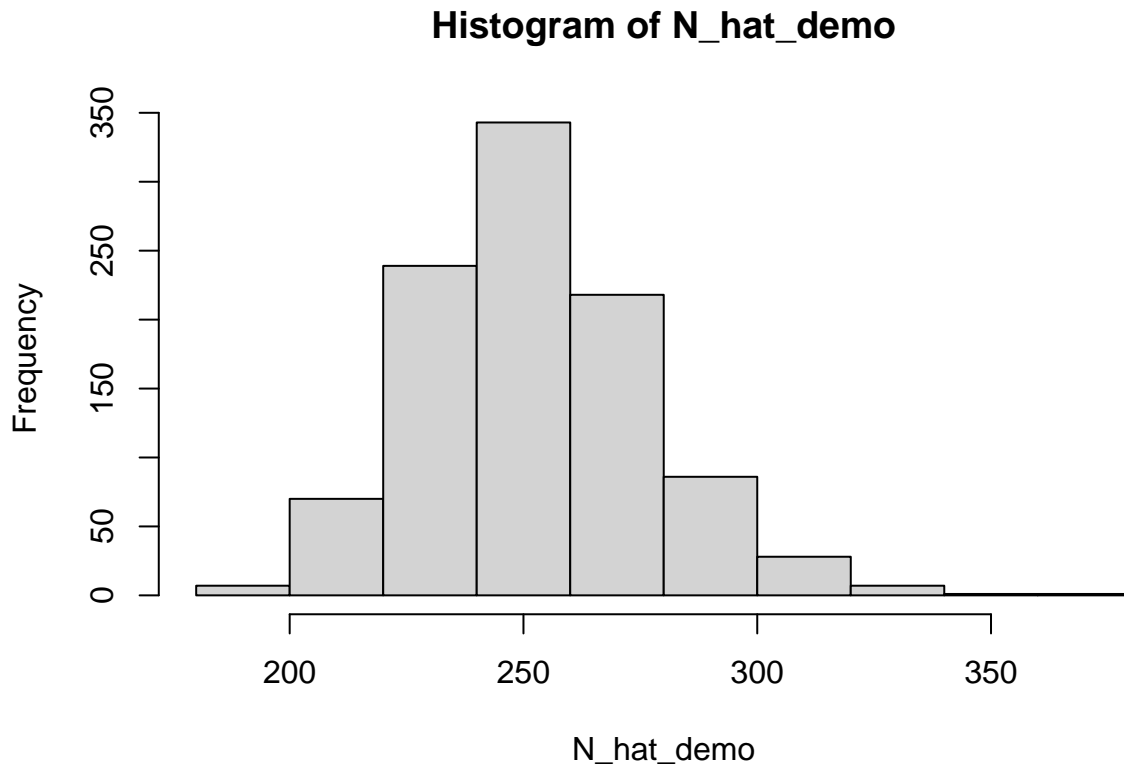
We can also calculate the bias and the precision of our estimates.

```
# Remember, our true abundance is 250

# What is the mean of all of our N_hat estimates?
mean(N_hat_demo)              # mean of all N-hat estimates
```

```
## [1] 252.0373
```

```
# On average, how biased are our estimates?
mean(N_hat_demo) - N          # bias
```

```
## [1] 2.037264
```

```
# How precise are our estimates?
sd(N_hat_demo)                # precision
```

```
## [1] 24.14464
```

Now it is your turn! Use the code above (only from the "Let's Practice" section; nothing else needs to change) to compare our estimates, bias, and precision in difference scenarios.

Scenario 1: Detection Probability is Higher or Lower

What happens to bias and precision at higher or lower rates of detection across the board?
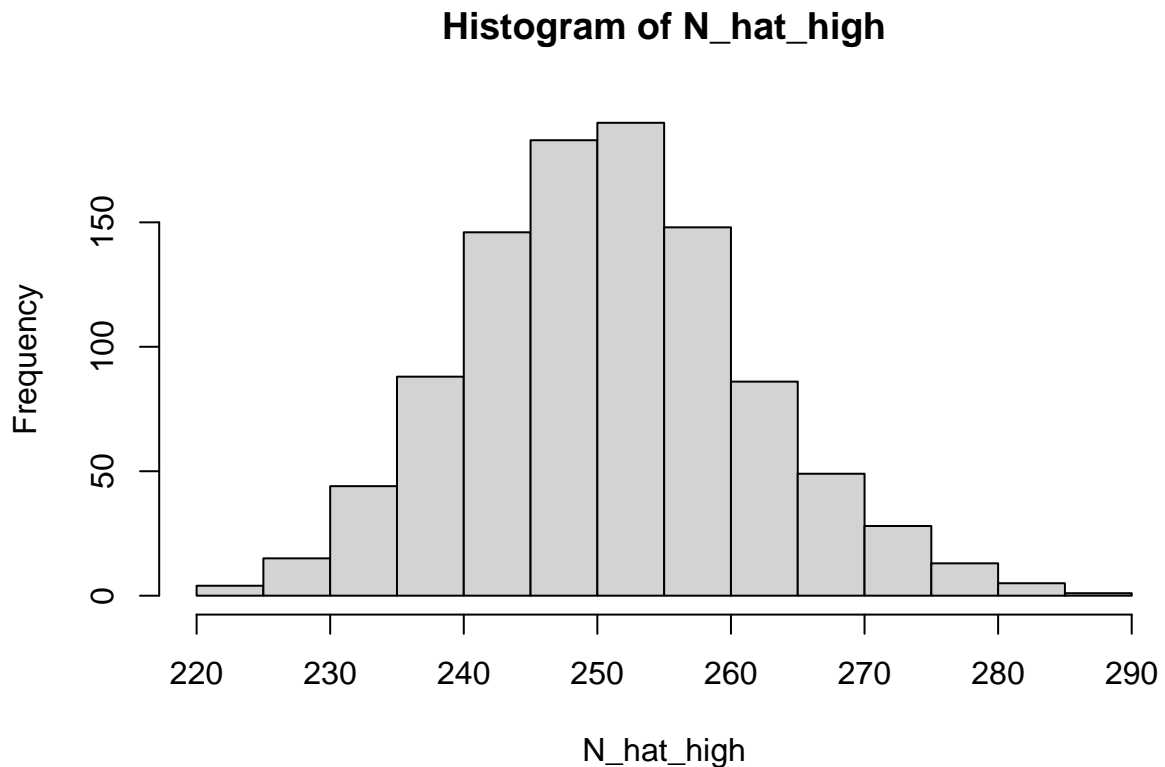
Higher

8. Our detection probabilities (`p1` and `p2`) and recapture probability (`c2`) are consistent and *higher* than the demo: `p1 = 0.6`, `p2 = 0.6`, `c2 = 0.6` (2 points)

   a. Run the simulation.

```
N_hat_high <- LP_estimator(p1 = 0.6, p2 = 0.6, c2 = 0.6)
```

   b. Plot the histogram of this data.

```
hist(N_hat_high)
```


Histogram of N_hat_high

   c. Calculate the mean, bias, and precision of the estimate.

```
mean(N_hat_high)
```

```
## [1] 250.935
```

```
# On average, how biased are our estimates?
mean(N_hat_high) - N        # bias
```

```
## [1] 0.9350213
```

```
# How precise are our estimates?
sd(N_hat_high)              # precision
```
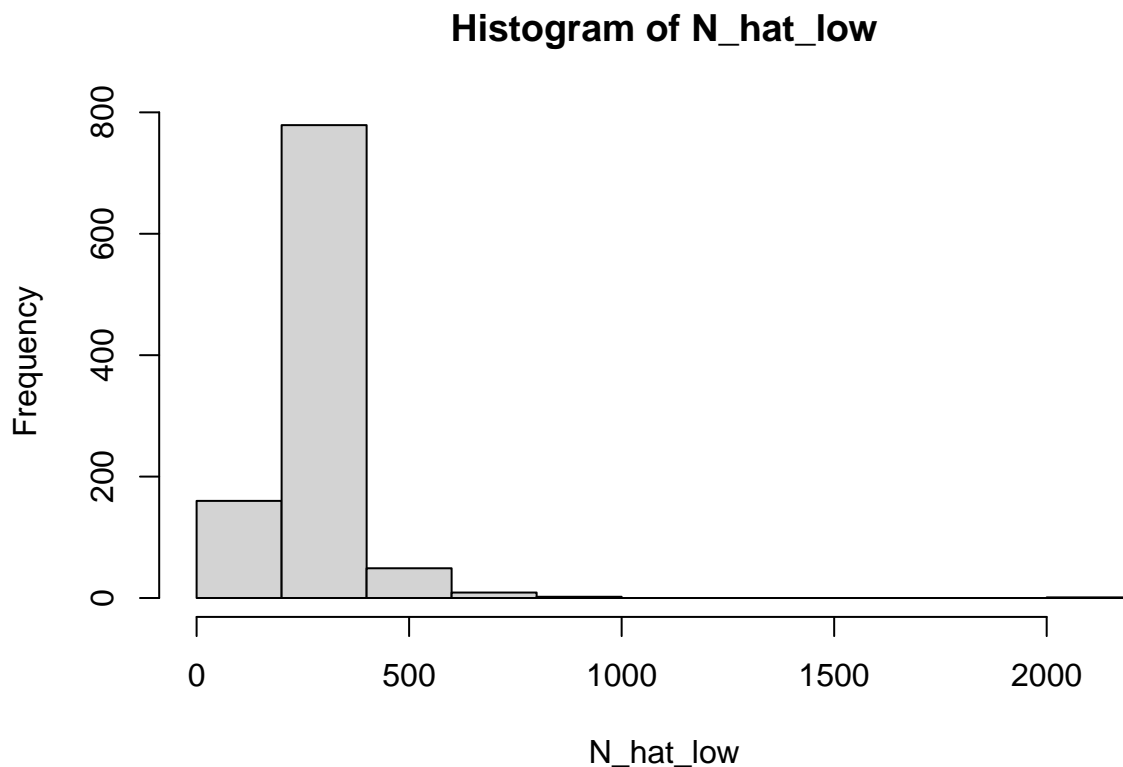
```
## [1] 10.64647
```

Lower

9. Our detection probabilities (`p1` and `p2`) and recapture probability (`c2`) are consistent and *lower* than the demo: `p1` = 0.2, `p2` = 0.2, `c2` = 0.2 (2 points)

   a. Run the simulation.

```
N_hat_low <- LP_estimator(p1 = 0.2, p2 = 0.2, c2 = 0.2)
```

   b. Plot the histogram of this data.

```
hist(N_hat_low)
```

Histogram of N_hat_low

   c. Calculate the mean, bias, and precision of the estimate.

```
mean(N_hat_low)
```

```
## [1] 270.927
```

```
# On average, how biased are our estimates?
mean(N_hat_low) - N       # bias
```

```
## [1] 20.92698
```

```
# How precise are our estimates?
sd(N_hat_low)            # precision
```

```
## [1] 102.697
```

Interpretation

10. Compare the bias and precision of our total abundance estimates given the different detection probabilities and write 2-3 sentences explaining the results. Does higher detection probability give us more or less bias? More or less precision? Does this make sense to you? (2 points)

    Answer: As detection probability increases, the bias decreases and the precision increases. This is because higher detection probability means there is likely less error, so estimates are going to be more accurate. Interestingly, the increase is not proportional—a jump from 0.2 to 0.4 detection prob makes a substantial difference in both bias and precision, while the increase from 0.4 to 0.6 increases accuracy but not as much.

Scenario 2: Detection Probability Changes with Behavior

What happens to the bias and precision of our estimates when animals are trap happy or trap shy?
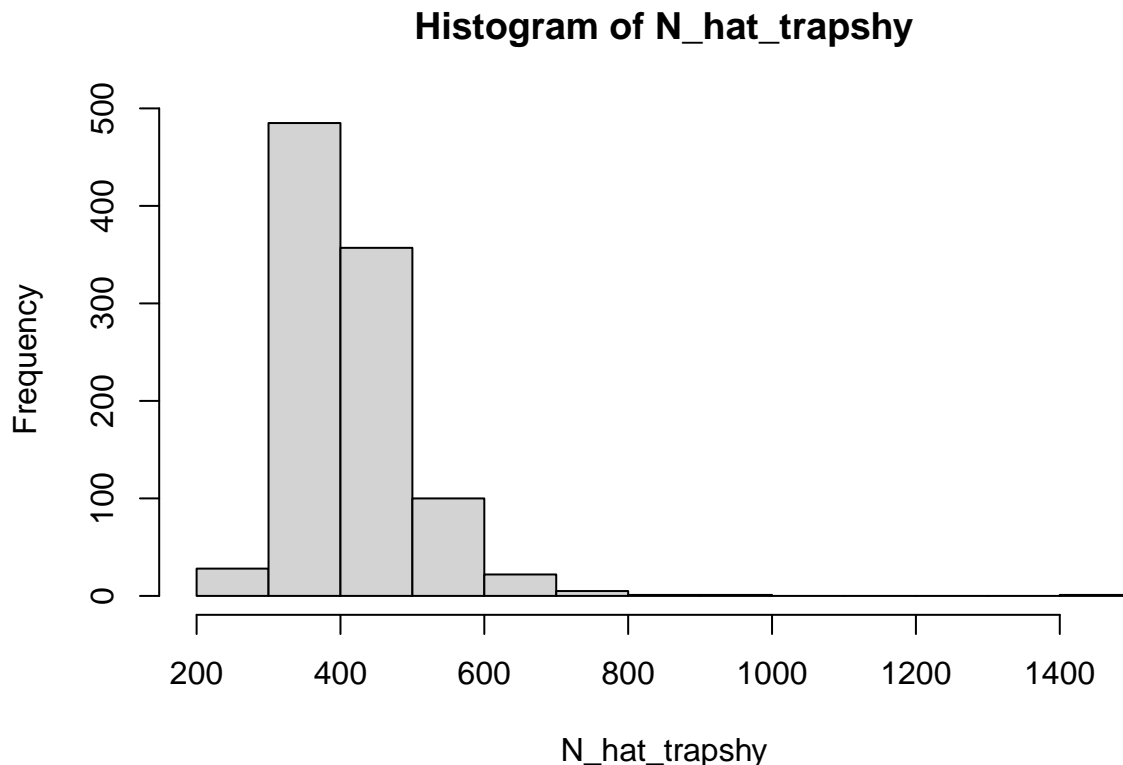
Trap Shy

11. We have a population that is trap shy, meaning that the probability of recapture is lower than the probability of initial capture ($p > c$): $p1 = 0.4$, $p2 = 0.4$, $c2 = 0.2$ (2 points)

    a. Run the simulation.

    ```
    N_hat_trapshy <- LP_estimator(p1 = 0.4, p2 = 0.4, c2 = 0.2)
    ```

    b. Plot the histogram of this data.

    ```
    hist(N_hat_trapshy)
    ```


Histogram of N_hat_trapshy

c. Calculate the mean, bias, and precision of the estimate.

```
mean(N_hat_trapshy)
```

```
## [1] 414.7512
```

```
# On average, how biased are our estimates?
mean(N_hat_trapshy) - N        # bias
```

```
## [1] 164.7512
```

```
# How precise are our estimates?
sd(N_hat_trapshy)              # precision
```
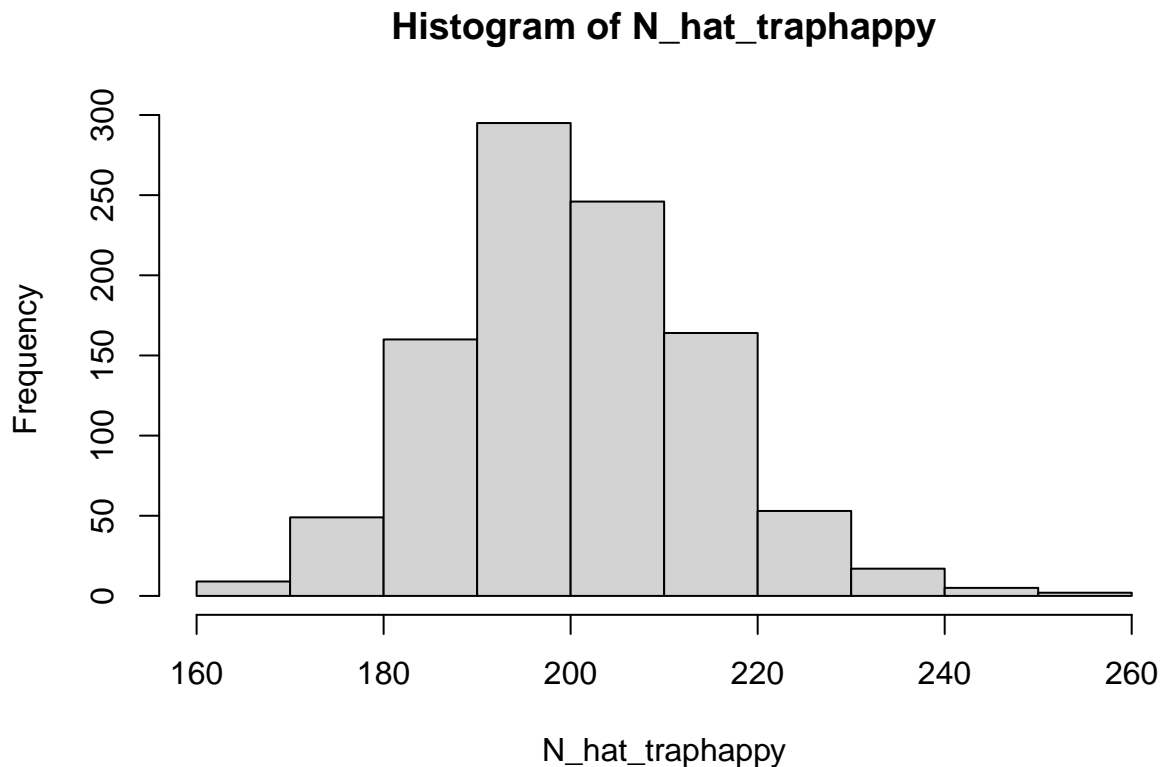
```
## [1] 88.06007
```

Trap Happy

12. We have a population that is trap happy, meaning that the probability of recapture is higher than the probability of initial capture (p < c): p1 = 0.4, p2 = 0.4, c2 = 0.6 (2 points)

   a. Run the simulation.

```
N_hat_traphappy <- LP_estimator(p1 = 0.4, p2 = 0.4, c2 = 0.6)
```

   b. Plot the histogram of this data.

```
hist(N_hat_traphappy)
```


Histogram of N_hat_traphappy

c. Calculate the bias and precision of the estimate.

```
mean(N_hat_traphappy)
```

```
## [1] 200.692
```

```
# On average, how biased are our estimates?
mean(N_hat_traphappy) - N        # bias
```

```
## [1] -49.30801
```

```
# How precise are our estimates?
sd(N_hat_traphappy)              # precision
```

```
## [1] 13.85927
```

Interpretation

13. Compare the bias and precision of our total abundance estimates for the demo, trap happy, and trap shy populations. Write 2-3 sentences explaining the results. How do animals being trap happy or trap shy affect the accuracy of our estimate? Does this make sense to you? (2 points)

    Answer: When animals are trap shy, we end up with an under-estimate (bias towards under counting) and imprecise estimates; when animals are trap happy, we end up over-estimating the population.