Method of Moments: Mark-Recapture Simulation

2024-03-05

Set-up

Recall the following scenario: we would like estimate the number of individuals N of a species in a particular location. The mark-recapture sampling scheme proceeds as follows:

1. Sample r individuals from the population and mark/tag them all. Release the individuals and wait some time.
2. After waiting, take a second sample of m individuals from the population. Count how many of the m are marked.

Defining X as the number of individuals in the second sample who were tagged, we obtained the following method of moments estimator: $\hat{N}_{MM} = \frac{rm}{X}$.

Simulation

Let's simulate data to see how well this estimator performs. Note that in simulating data, I get to choose/know the *true* value of N.

```
# set true value
N_true <- 2000
```

Now, I will determine how many to take in the first sample. Then I create a vector of 1's and 0's called `marked`, where I have r 1's representing all the tagged individuals, and the remaining $N - r$ 0's representing all the untagged.

```
# number in first sample
r <- 200

# make vector of "marked" and "unmarked" individuals after initial sample
marked <- c(rep(1, r), rep(0, N_true -r))
```

We write a function that performs one iteration of the second sample: sampling m individuals without replacement from the population again, and counting how many were `marked`.

```
# write function to simulate second sample
mark_recapture <- function(marked, m){
  samp2 <- sample(marked, size = m, replace = F)
  x <- sum(samp2)
  return(x)
}
```

Now we actually simulate! I can choose how many simulations to run. Then I use the `replicate()` function to repeatedly perform the simulations. The outcome to each replicate gets stored in `R`, which I use to obtain the method of moments estimates over repeated samples.

```
# choose number of simulations
n_sims <- 10000
# number in second sample
m <- 300

# simulate!
R <- replicate(n_sims, mark_recapture(marked, m))

# calculate MoM estimate
N_hat_vec <- r * m / R

# plot
library(tidyverse)
data.frame(N_hat = N_hat_vec) %>%
  ggplot(., aes(x = N_hat)) +
  geom_histogram() +
  geom_vline(xintercept = N_true, col = "red", linetype = "dashed")
```