Module 3 Assignment 1

Ellen Bledsoe

2022-11-01

Occupancy Modeling

Assignment Details

Purpose

The goal of this assignment is to understand, apply, and interpret detection probability and occupancy models.

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 8 before lab

Assignment Questions

For this assignment, we have 2 different data sets of occupancy from the Santa Cruz River.

There are a number of trail cameras that have been set up along the river to look at which species are present along various points of the Santa Cruz. Today, we will be working with some presence-absence surveys from these trail cameras for two different species: one for javelina (collared pecaries) and one for coyotes.

Following the structure of the tortoise exercise we did in lab, create models for each species. Our ultimate goal is to compare the estimated occupancy and detection probabilities of our two species.

All questions are worth 1 point unless otherwise specified.

Set-Up

Let's go ahead and get ourselves set up for the rest of the assignment.

Before you start anything, I highly recommend going over to the "Environment" tab and clicking on the broom. It will clean up the clutter in your environment and also prevent you from making some unintentional errors.

Load in the package we will need to use.

```
library(unmarked)
```

Using the `read.csv()` function, read in the two data sets we will use. One is for javelina and the other is for coyotes. Create objects containing these two data sets called `javelina` and `coyote`, respectively.

```
javelina <- read.csv("../data_raw/javelina.csv")
coyote <- read.csv("../data_raw/coyote.csv")
```

Javelina

We are going to calculate naive occupancy and then run occupancy models for the javelina.

Naive Occupancy First, let's calculate the naive occupancy for javelina.

1. Calculate the naive occupancy for javelina.

```
# count up the number of times each site was marked as occupied
# returns a vector that we can now use
num_detections <- rowSums(javelina)
num_detections
```

```
##  [1] 0 2 0 0 0 1 2 0 0 3 0 1 0 0 0 0 2 0 0 0 0 2 2 0 2 1 0 0 0 2 0 0 2 0 0 0 0 0
## [39] 2 1
```

```
# total number of sites that were surveyed
sites_surveyed <- nrow(javelina)
sites_surveyed
```

```
## [1] 40
```

```
# number of sites with detection of greater than zero
sites_detection <- length(num_detections[num_detections > 0])
sites_detection
```

```
## [1] 14
```

```
# divide the number of sites with detection
naive_occupancy <- sites_detection / sites_surveyed
naive_occupancy
```

```
## [1] 0.35
```

Data Preparation

2. Explore the `javelina` data frame, either in the environment or through code.

   a. How many sites were surveyed? 40
   b. How many times was each individual site surveyed? 5

```
# optional: space for code if you want it
```

3. Like we did in the tortoise example, create a data frame of observation co-variates. It should have the number of survey occasions this data set.

   Then, create a list out of this one data frame, setting the data frame equal to the "time" argument.

   (*Hint: Be sure to create a data frame with the correct number of rows based on your answer to Question 2*)

```
# create the data frame with 40 rows and each column with the survey number
javelina_surveys <- data.frame(time_1 = rep("1", 40),
                               time_2 = rep("2", 40),
                               time_3 = rep("3", 40),
                               time_4 = rep("4", 40),
                               time_5 = rep("5", 40))

# make this into a list of one dataframe
javelina_survey_time <- list(time = javelina_surveys)
```

4. Create an unmarked data frame for the javelina data

```
javelinaUMF <- unmarkedFrameOccu(javelina, obsCovs = javelina_survey_time)
```

```
## Warning: obsCovs contains characters. Converting them to factors.
```

```
head(javelinaUMF) # check out the data
```

```
## Data frame representation of unmarkedFrame object.
##    y.1 y.2 y.3 y.4 y.5 time.1 time.2 time.3 time.4 time.5
## 1    0   0   0   0   0      1      2      3      4      5
## 2    0   0   1   1   0      1      2      3      4      5
## 3    0   0   0   0   0      1      2      3      4      5
## 4    0   0   0   0   0      1      2      3      4      5
## 5    0   0   0   0   0      1      2      3      4      5
## 6    0   0   1   0   0      1      2      3      4      5
## 7    1   0   0   0   1      1      2      3      4      5
## 8    0   0   0   0   0      1      2      3      4      5
## 9    0   0   0   0   0      1      2      3      4      5
## 10   1   0   0   1   1      1      2      3      4      5
```

Occupancy Models

5. Fit an occupancy model where occupancy and detection probability are constant.

```r
javelina_pdot_psidot <- occu(~1 ~1, data = javelinaUMF)
```

6. Fit an occupancy model where occupancy stays constant but detection probability varies with survey.

```r
javelina_ptime_psidot <- occu(~time-1 ~1, data = javelinaUMF)
```

Model Selection

7. Use AIC to determine which model is the best.

```r
javelina_models <- fitList(javelina_pdot_psidot  = javelina_pdot_psidot,
                           javelina_ptime_psidot = javelina_ptime_psidot)
modSel(javelina_models)
```

```
##                        nPars    AIC delta AICwt cumltvWt
## javelina_pdot_psidot       2 142.94  0.00 0.969     0.97
## javelina_ptime_psidot      6 149.80  6.86 0.031     1.00
```

8. Which model should we choose? Why?

   Answer: pdot_psidot, lower AIC

9. Calculate and report the estimate of detection probability and the 95% CI for the estimate from our *best* model. (Report means type them out)

   Detection Probability: 0.29 (0.183 - 0.439)

```r
plogis(coef(javelina_pdot_psidot))  # estimates of psi and p
```

```
##  psi(Int)    p(Int)
## 0.4238748 0.2948984
```

```r
plogis(confint(javelina_pdot_psidot, type = "det", level = 0.95))
```

```
##            0.025     0.975
## p(Int) 0.1829776 0.4385327
```

10. Calculate and report the estimate of occupancy and the 95% CI for the estimate from our *best* model.

    Occupancy: 0.42 (0.247 - 0.622)

    ```r
    plogis(coef(javelina_pdot_psidot))  # estimates of psi and p
    ```

    ```
    ##  psi(Int)    p(Int)
    ## 0.4238748 0.2948984
    ```

    ```r
    plogis(confint(javelina_pdot_psidot, type = "state", level = 0.95))
    ```

    ```
    ##              0.025     0.975
    ## psi(Int) 0.2471549 0.6224758
    ```

Comparison

14. Compare the naive occupancy estimate to the occupancy estimate from our model. Which one is higher? Why? Your answer should be 2-3 sentences. (2 points)

Answer: The naive estimate of occupancy is an underestimate (lower) because we are not taking into account individuals which are present but we did not detect.

Coyotes

Let's do the same as above for the coyotes.

Naive Occupancy First, let's calculate the naive occupancy for coyotes.

15. Calculate the naive occupancy for coyotes.

```
# count up the number of times each site was marked as occupied
# returns a vector that we can now use
num_detections <- rowSums(coyote)
num_detections
```

```
##  [1] 0 0 1 0 0 2 0 2 2 0 0 0 0 1 1 2 0 0 0 1 1 5 1 1 1 0 1 4 0 1 2 2 1 2 1 2 2 3
## [39] 0 1
```

```
# total number of sites that were surveyed
sites_surveyed <- nrow(coyote)
sites_surveyed
```

```
## [1] 40
```

```
# number of sites with detection of greater than zero
sites_detection <- length(num_detections[num_detections > 0])
sites_detection
```

```
## [1] 25
```

```
# divide the number of sites with detection
naive_occupancy <- sites_detection / sites_surveyed
naive_occupancy
```

```
## [1] 0.625
```

Data Preparation

13. Like we did in the tortoise example (and in Question 3 above), create a data frame of observation co-variates. It should have the number of survey occasions this data set.

    Then, create a list out of this one data frame, setting the data frame equal to the "time" argument.

```r
# create the data frame with 20 rows and each column with the survey number
coyote_survey <- data.frame(time_1 = rep("1", 40),
                            time_2 = rep("2", 40),
                            time_3 = rep("3", 40),
                            time_4 = rep("4", 40),
                            time_5 = rep("5", 40))

# make this into a list of one dataframe
coyote_survey_time <- list(time = coyote_survey)
```

14. Create an unmarked data frame for the coyote data

```r
coyoteUMF <- unmarkedFrameOccu(coyote, obsCovs = coyote_survey_time)
```

```
## Warning: obsCovs contains characters. Converting them to factors.
```

```r
head(coyoteUMF) # check out the data
```

```
## Data frame representation of unmarkedFrame object.
##    y.1 y.2 y.3 y.4 y.5 time.1 time.2 time.3 time.4 time.5
## 1    0   0   0   0   0      1      2      3      4      5
## 2    0   0   0   0   0      1      2      3      4      5
## 3    0   0   0   1   0      1      2      3      4      5
## 4    0   0   0   0   0      1      2      3      4      5
## 5    0   0   0   0   0      1      2      3      4      5
## 6    0   1   1   0   0      1      2      3      4      5
## 7    0   0   0   0   0      1      2      3      4      5
## 8    0   1   0   1   0      1      2      3      4      5
## 9    0   0   1   1   0      1      2      3      4      5
## 10   0   0   0   0   0      1      2      3      4      5
```

Occupancy Models

15. Fit our two occupancy models: one where occupancy and detection probability are constant and one where occupancy stays constant but detection probability varies with survey.

```r
coyote_pdot_psidot <- occu(~1 ~1, data = coyoteUMF)
coyote_ptime_psidot <- occu(~time-1 ~1, data = coyoteUMF)
```

Model Selection

16. Use AIC to determine which model is the best.

```r
coyote_models <- fitList(coyote_pdot_psidot  = coyote_pdot_psidot,
                         coyote_ptime_psidot = coyote_ptime_psidot)
modSel(coyote_models)
```

```
##                     nPars    AIC delta AICwt cumltvWt
## coyote_pdot_psidot      2 209.51  0.00 0.964     0.96
## coyote_ptime_psidot     6 216.06  6.55 0.036     1.00
```

17. Calculate and report the estimates and the 95% CI for the estimate for both detection probability and occupancy from our *best* model.

    Detection Probability: 0.275 (0.19 - 0.38)

    Occupancy: 0.781 (0.48 - 0.93)

```r
plogis(coef(coyote_pdot_psidot))  # estimates of psi and p
```

```
##  psi(Int)    p(Int)
## 0.7813023 0.2751816
```

```r
plogis(confint(coyote_pdot_psidot, type = "det", level = 0.95))
```

```
##            0.025     0.975
## p(Int) 0.1895566 0.3812877
```

```r
plogis(confint(coyote_pdot_psidot, type = "state", level = 0.95))
```

```
##              0.025     0.975
## psi(Int) 0.4842553 0.9314732
```

Comparing Species

18. Write 3-4 sentences comparing the javelina and coyote populations that we surveyed. (2 points)

- Which species has higher occupancy?
- Which species has higher detectability?
- In your answer, be sure to explain the difference between detection probability and occupancy.
- Speculate what could be driving the differences between the two species (I don't expect you to know, but you should put forward an educated guess that fits with your answers above)

Answer: While the two species have similar detection probabilities (javelina is slightly higher), coyotes have much higher occupancy (0.781) than javelina (0.42). This means that when each species is present at a site, they have a similar chance of being observed. However, coyotes are present at far more sites than javelina. This could be because coyotes have a higher population and/or that the area we are surveying provides better habitat for them than it does for javelina.