
FISH 621 Laboratory #8: Distance Methods and More

Curry Cunningham 2022

Instructions

The purpose of this lab is to:

- Explore distance estimators for line transect sampling of duck nests and simulated data for minke whale.
- Explore of simple Bayesian abundance estimators for simple Petersen mark-recapture experiments.
- Continue developing familiarity with vector autoregressive spatio-temporal (VAST) models.

If you have a question during the lab, please un-mute yourself and ask, or type it into the chat box. There is a high likelihood that someone else has the same question. It is more fun if we all learn together in our distance-learning world.

I have posted the lecture slides to the **Canvas site**, so you can reference this material as you work through the lab.

This and all other labs will be graded based on your attendance and participation.

Lab Contents

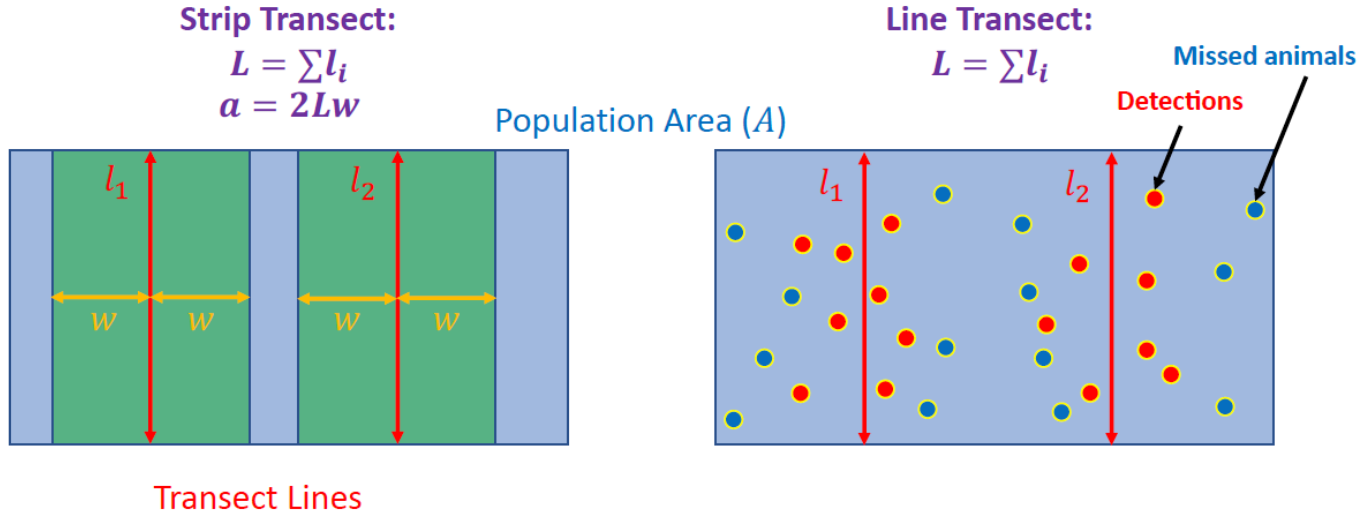
- **621_Lab 7_Distance Methods and More.pdf** (this file)
- **621_Lab 7_Distance Methods and More.R** R script with exercises.
- **race_cpue_by_haul.csv** Observations from the NOAA-AFSC bottom trawl surveys for select species in the Gulf of Alaska and Bering Sea shelf
- **ducks-area-effort.csv** Distance sampling data for duck nests from the Monte Vista National Wildlife Refuge, Colorado from 1967-1968.
- **Petersen.stan** Stan model file for the Petersen mark-recapture estimator.

Exercise 2: Monte Vista Ducks

Distance sampling methods are used to estimate the abundance (\hat{N}) and/or density (\hat{D}) of objects or animals within a finite population area of size A . Much like a strip transect, an observer traverses a straight line counting the number of objects or animals detected within the sampled area a , and extrapolating this abundance up to the population area A . The major differences between sampling of strip transects and line transects under distance sampling, are:

- With strip transect sampling it is assumed that all items within the sampled area $a = 2Lw$ are detected (i.e. detection probability = 1).
- With line transect sampling and distance methods the detection probability for items is assumed to differ, with detection declining at greater distance from the line transect.
 - As a result, additional information regarding the distance of each observation from the center line of the transect is collected.
- With distance methods we use a sighting or detection function to approximate how detection probability declines with distance, based on our data.

As an initial exploration of distance sampling methods and abundance estimation we will use a dataset from a survey of duck nests in the Monte Vista National Wildlife Refuge in Colorado. Line transect surveys of duck nests occurred in 1967 and 1968. For each detected nest along the transect, the distance from the transect was recorded. Observers searched up to 2.4 meters from the transect on each side.



Example 3: Minke Whales

To further explore distance methods, we will utilize a simulated dataset of minke whale sighting from transects sampled in the Southern Ocean surrounding Antarctica. In this example data were collected in two survey strata (e.g. North and South) with separate total strata areas A . Based on transect sampling in

each stratum we can estimate abundance and uncertainty for each stratum separately. The data represent sighting distances for individual whales, observed on transects (identified by Sample.Label) with strata (identified by Region.Label). Distances to animals and effort (i.e. transect lengths) are in nautical miles, and the area (A) of each stratum is in square nautical miles.



Minke Whales

Simulated, similar to: Branch, T.A. and D.S. Butterworth (2001)

• Data

```
> head(minke)
  Region.Label Area Sample.Label Effort distance
1      South 84734             1  86.75    0.10
2      South 84734             1  86.75    0.22
3      South 84734             1  86.75    0.16
4      South 84734             1  86.75    0.78
5      South 84734             1  86.75    0.21
6      South 84734             1  86.75    0.95
```

Nautical Miles

```
> unique(minke$Region.Label)
[1] "South" "North"
> unique(minke$Area)
[1] 84734 630582
> unique(minke$Sample.Label)
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
> unique(minke$Effort)
[1] 86.75 10.57 16.57 113.13 15.07 9.90 25.71 57.55 7.67 13.02 52.65
[12] 0.19 75.63 83.33 100.93 110.55 97.50 126.32 99.37 137.39 106.76 113.28
[23] 144.90 130.33 107.72
```

Transects within strata

Exercise 4: Bayesian Mark-recapture Estimators

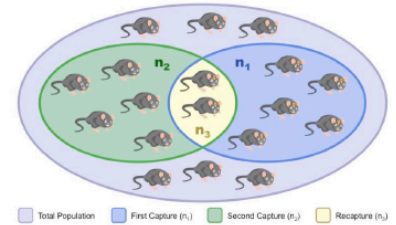
In previous sections of this course we have discussed our simple Petersen mark-recapture experiment in detail. Two-event mark-recapture experiments can be used to estimate abundance. For review of notation the process for the Petersen experiment is:

1. Capture n_1 animals at time t_1 , then mark and release these animals back into the population.
2. Allow the marked n_1 and unmarked $N - n_1$ animals to mix thoroughly.
3. Sample n_2 animals at time t_2 and inspect them for marks.
4. The ratio of the number of marked individuals m_2 observed at time t_2 (i.e. recaptures) to the number sampled n_2 provides information about the true abundance N .

Here we will implement our mark-recapture estimator in a Bayesian context and simulate data from Petersen mark-recapture experiments to evaluate:

- Whether our Bayesian estimator provides reasonable estimate of abundance \hat{N} .
- Whether the uncertainty described the posterior distribution of \hat{N} is appropriate.
- Under what assumptions about the sampling process do we have more or less bias with this estimator.

Bayesian Mark-Recapture Estimators: Petersen



• Petersen.stan

```
1 // Simple Petersen Mark-Recapture Estimator
2
3 data {
4   int<lower=0> n1; // Number marked at t1
5   int<lower=0> n2; // Number captured at t2
6   int<lower=0,upper=min(n1,n2)> m2; // Number of marked animals in second capture
7 }
8 parameters {
9   real<lower=(n2 - m2 + n1)> N;
10 }
11 model {
12   m2 ~ binomial(n2, n1 / N);
13 }
14
```

Exercise 5: VAST Exploration

In this exercise we will explore how adjusting some of our VAST model settings influence predictions across space and time. Specifically, we will explore the impact of changing the level of spatial complexity as controlled by the number of knots, and we will also explore the impact of changing our observation model. With respect to the observation model specifications we will compare models that assume a gamma distribution for the positive catch rate (PCR) component of the delta-model with a lognormal distribution.

make_settings()

```
make_settings(  
  n_x,  
  purpose = "index",  
  Region,  
  fine_scale = TRUE,  
  strata.limits = data.frame(STRATA = "All_areas"),  
  zone = NA,  
  FieldConfig, ← How to define  
  RhoConfig, ← spatial/spatio-temporal  
  OverdispersionConfig, ← components  
  ObsModel, ← See ?make_data  
  bias.correct, ←  
  Options, ← Specify  
  use_anisotropy, ← autocorrelation in  
  vars_to_correct, ← year effects and  
  Version, ← spatio-temporal fields  
  treat_nonencoder_as_zero,  
  n_categories,  
  VamConfig, ← Define observation model  
  max_cells,  
  knot_method  
)
```