Lab 1: Estimating Population Size

Timeline

This lab will take place over five weeks.

- ➔ **September 14:** Part 1 | Introduction and initial planning
- ➔ **September 28:** Part 2 | Independent work, ***proposal due!***
- ➔ **October 12:** Part 3 | Independent work
- ➔ **October 19:** ***Final report due!***

Objectives

1. Learn how to use **mark-recapture** methods to estimate population size
2. Get experience designing a field study

Introduction

Population size, defined as the number of individuals in a given population, is a fundamental quantity in population ecology. However, measuring population size by identifying and enumerating each individual within a population is easier said than done. Not all individuals are easily observed, and when observed, not all individuals are easily distinguished from one another. Further, uneven species distributions, large ranges, and complicating life history traits can make measuring population size extremely difficult. In this lab, we'll use the mark-recapture method to estimate population sizes.

The mark-recapture method involves capturing and marking some subset of individuals within a population, releasing them, and re-sampling the population. We can then use the proportion of resampled individuals who were marked to estimate population size.

In this lab we won't actually capture and mark animals. Instead, we'll use sight-resight methods similar to those used by biologists to study whales and other animals that have unique markings (e.g. scars, the shape of a dorsal fin/fluke). We will apply these methods to study dog populations in dog parks in Toronto.

The Petersen Method

The simplest mark-recapture estimate of population size is the Petersen method, which was developed to estimate marine fish population sizes. **The Petersen method reasons that (1) if a proportion of the population was marked, (2) returned to the original population, and (3) after complete mixing, (4) a second sample was taken, (5) the proportion of marked individuals in the second sample would be the same as was initially marked.** We can express the Petersen method mathematically as follows:

$$\frac{k}{n} = \frac{K}{N}$$

Where N is the total population size to be estimated, K is the number of individuals collected and marked in the first sample, n is the total number of individuals in the 2nd sample, and k is the number of individuals in the 2nd sample that had marks. Given that N is the quantity we want to estimate, and that we can “measure” K, n, and k, we can estimate total population size as:

$$N = \frac{Kn}{k}$$

For example, suppose you took 200 mice out of a forest having an unknown number of mice, put leg bands on them, returned them to the forest and let them mix thoroughly. If you then take 250 mice from the forest and find 50 of them to have leg bands, then $K = 200$, $n = 250$, $k = 50$, and the estimated total number of mice (N) is:

$$N = \frac{200 \cdot 250}{50}$$
$$N = 1000$$

It’s important to remember that the accuracy and precision of the Petesen mark-recapture estimate is based on a few assumptions:

Assumption 1: During the time between releasing animals from the first capture (and marking) and the second sample, there is no change to the proportion of marked animals.

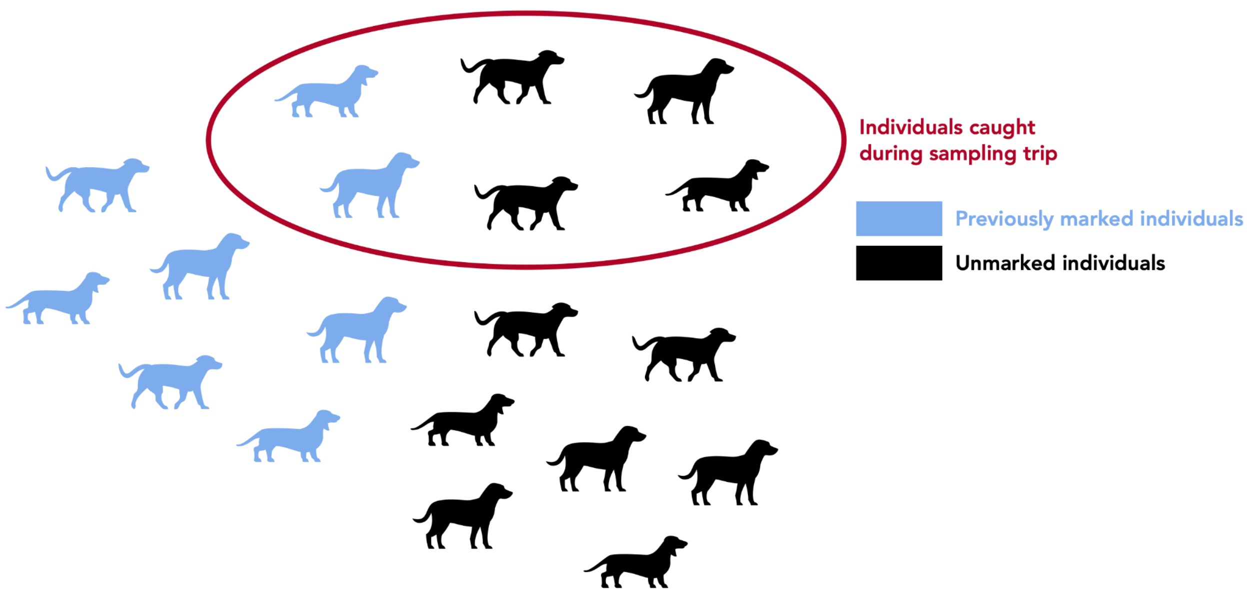
Assumption 2: The probability that each individual is captured is equal and does not change between the initial capture and the subsequent recapture.

Assumption 3: Marked individuals don't lose their marks.

The Schnabel Method

The Schnabel method extends the Petersen method to incorporate a series of recapture events. As described above, individuals caught during the first sampling trip are marked. During each subsequent sampling trip, individuals are examined for marks, and, if found to be unmarked, are given a mark before being released back into the population.

During each recapture event, we need to determine:
C_t = Number of individuals captured
R_t = Number of previously-marked individuals captured during the sampling trip
M_t = Number of marked individuals in the population prior to the sampling trip



For example, in the above image:
$C_t = 6$
$R_t = 2$
$M_t = 8$

If the Petersen method uses single recaptures to calculate a population size estimate, or, a Petersen estimate, then the Schnabel method treats recapture events as a series of single recaptures, and calculates a population size estimate ($\widehat{N}$) as a weighted average of Petersen estimates:

$$\widehat{N} = \frac{\Sigma M_t C_t}{\Sigma R_t}$$

Sampling variance and a 95% confidence interval can be calculated as follows:

$$Var\left(\frac{1}{\widehat{N}}\right) = \frac{\Sigma R_t}{\left(\Sigma M_t C_t\right)^2}$$

$$\frac{1}{\widehat{N}} \pm 1.965 \cdot \sqrt{Var\left(\frac{1}{\widehat{N}}\right)}$$

The accuracy and precision of the Schnabel mark-recapture estimate is based on the same assumptions as the Peterson estimate.

The Cormack-Jolly-Seber Model

Alternatively, we can use a series of recapture events to construct a sight-resight history for each dog, and use a Cormack-Jolly-Seber model to estimate (re)capture probability, apparent survival, and population size.

For example, if you observe four dogs over five sampling trips, your data may look something like this ("0" denotes absence; "1" denotes presence):
Dog 1: 10011
Dog 2: 01011
Dog 3: 11101
Dog 4: 00110

We can use these sight-resight data to construct the CJS model:

$P_1 = p \cdot \varphi(1 - p) \cdot \varphi(1 - p) \cdot \varphi p \cdot \varphi p$

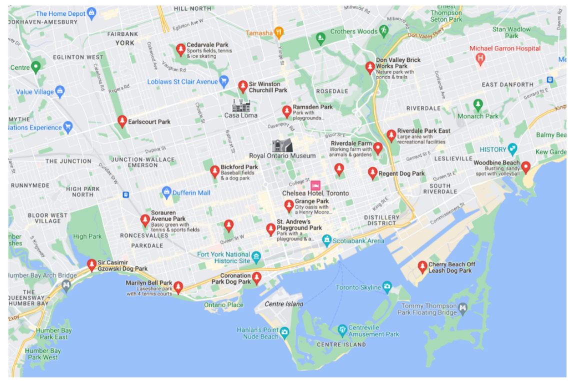
$P_2 = (1 - p) \cdot \varphi p \cdot \varphi(1 - p) \cdot \varphi p \cdot \varphi p$

$P_3 = p \cdot \varphi p \cdot \varphi p \cdot \varphi(1 - p) \cdot \varphi p$

$$P_4 = (1 - p) \cdot \varphi(1 - p) \cdot \varphi p \cdot \varphi p \cdot \varphi(1 - p) + (1 - \varphi)$$
$$= (1 - p) \cdot \varphi(1 - p) \cdot \varphi p \cdot \varphi p \cdot (1 - \varphi p)$$

Where p is the probability a dog is observed, and φ is the probability that a dog survives from one observation period to the next. Using maximum likelihood methods, we can find the values of p and φ that maximize the likelihood of observing the data.

Methods



In this lab, we'll use the Petersen and Schnabel methods to estimate dog population sizes in Toronto. In small groups, you'll visit a dog park from [this list]() ("Parks" sheet), describe the dogs present (initial capture and mark), and return to look for previously marked individuals (recapture). The list includes 22 dog parks. When your group decides which dog park they'd like to use as their study site, add your names and lab section to the sheet. **No two groups should be using the same study site.**

	A	B	C	D	E	F	G
1	FOR A MAP OF THESE PARKS, CLICK HERE: https://goo.gl/maps/R1yd7yg3pSDp8v3z9						
2	**Name**	**Type**	**General Location**	**Coordinates**	**Features**	**Group**	**Lab Section**
3	Allan Gardens	Botanical garden	East Toronto	43.66187481719904, -79.3747596603748	Small, human traffic, no car traffic, indoors, wide variety of plant species, dog park	Your group members' names.	Your group's lab section.
4	Bickford Park	Park	Central Toronto	43.661266596038566, -79.41865828151654	Mid-size, human traffic, surrounded by light car traffic, sports fields, dog park		
5	Cedarvale Park	Park	North-central Toronto	43.6924849954334, -79.43063880084858	Very large, human traffic, sports fields, trails, skating rink, dog park		
6	Cherry Beach	Beach	South-central Toronto	43.636897201744524, -79.34529617382694	Heavy human traffic, access to Lake Ontario, dog		

Proposal

100 Points, 5% of total grade; max. four double-spaced pages.

After deciding where to conduct your study, you should begin writing your proposal. Your proposal should:

1) **10 Points:** Motivate your study
2) **5 Points:** Describe your study site
3) **10 Points:** Describe how you plan to conduct your study
    - ➢ Example 1: How many recaptures should you do? (NOTE: The data sheet includes space for five recaptures. If your group chooses to do more than five recaptures, edit the sheet to accommodate them.)
    - ➢ Example 2: How should you decide when to do your recaptures?
4) **20 Points:** Explain the reasoning behind your design
5) **20 Points:** Describe how you plan to analyze your data
6) **20 Points:** Describe any shortcomings of your study design
7) **15 Points:** Describe what you expect to find and why

While your group can work together to design your study, each group member should submit their own proposal.

Data Collection

The Petersen Method

Your group can ***copy*** [this data sheet] ("DataSheet_Template_Petersen" sheet) to record your data.

	A	B	C	D	E	F	G
1	**Group Members:**						
2	**Lab Section:**						
3	**Park:**						
4		**Date of initial sampling trip:**	**Date of 2nd sampling trip:**	**Date of 3rd sampling trip:**	**Date of 4th sampling trip:**	**Date of 5th sampling trip:**	**Date of 6th sampling trip:**
5		**Number of individuals "captured" during initial sampling trip:**	**Number of individuals observed during 2nd sampling trip:**	**Number of individuals observed during 3rd sampling trip:**	**Number of individuals observed during 4th sampling trip:**	**Number of individuals observed during 5th sampling trip:**	**Number of individuals observed during 6th sampling trip:**
6	**Individual ID Number**	**Desriptions of individuals "captured" during initial sampling trip**	**Individual Recaptured? (Y/N)**	**Individual Recaptured? (Y/N)**	**Individual Recaptured? (Y/N)**	**Individual Recaptured? (Y/N)**	**Individual Recaptured? (Y/N)**
7	1						
8	2						
9	3						
10	4						
11	5						
12	6						
13	7						
14	8						
15	9						
16	10						
17	11						
18	12						
19	13						

After collecting your data, add the values of N, K, n, and k for each recapture trip to the collaborative data sheet ("CollaborativeDataSheet_Petersen" sheet).

	A		B	C	D	E
1			*N*	*K*	*n*	*k*
2			Estimated population size	Number of individuals "captured" during first sampling trip	Number of individuals "recaptured"	Number of individuals "recaptured" during given sampling trip
3	Group Members:	Sampling Trip 2	#DIV/0!	Should be same value		
4			#DIV/0!			
5	Lab Section:		#DIV/0!			
6			#DIV/0!			
7	Park:	Sampling Trip 6	#DIV/0!			
8	Group Members:		#DIV/0!			
9			#DIV/0!			
10	Lab Section:		#DIV/0!			
11			#DIV/0!			
12	Park:		#DIV/0!			
13	Group Members:		#DIV/0!			
14			#DIV/0!			
15	Lab Section:		#DIV/0!			

The Schnabel Method

Your group can ***copy*** this data sheet ("DataSheet_Template_Schnabel" sheet) to record your data.

	A	B	C	D	
1	Group Members:				
2	Lab Section:				
3	Park:				
4		Date of initial sampling trip:	Date of 2nd sampling trip:	Date of 3rd sampling trip:	Date of 4t
5		Number of individuals "captured" during initial sampling trip:	Number of individuals observed during 2nd sampling trip:	Number of individuals observed during 3rd sampling trip:	Number o
6	Individual ID Number	Desriptions of individuals "captured" during initial sampling trip	(1) Enter "Recovered" if previously described and caught. (2) Enter a description of individual if previously undescribed. (3) Enter "Unrecovered" if previously described but not caught.	(1) Enter "Recovered" if previously described and caught. (2) Enter a description of individual if previously undescribed. (3) Enter "Unrecovered" if previously described but not caught.	(1) Enter " (2) Enter a (3) Enter "
7	1				
8	2				
9	3				
10	4				
11	5				
12	6				
13	7				
14	8				
15	9				
16	10				
17	11				
18	12				
19	13				
20	14				

After collecting your data, add the values of $\widehat{N}$, C_t, R_t, and M_t for each recapture trip to the collaborative data sheet ("CollaborativeDataSheet_Schnabel" sheet).

	A	B	C	D	E	F
1		*N_hat*	*C_t*	*R_t*	*M_t*	*Variance*
2		Estimated population size	Number of individuals captured	Number of previously-marked individuals captured	Number of marked individuals in the population prior to the sampling trip	A measure of dispersion
3	Group Members:					
4						
5	Lab Section:					
6						
7	Park:					
8	Group Members:					
9						
10	Lab Section:					
11						
12	Park:					
13	Group Members:					
14						
15	Lab Section:					

Final Report

100 points total, 5% of total grade.

Using your own data and the collaborative data sheet, answer the following questions. Make sure to explain your reasoning! Additionally, please include a link to your data collection sheets in your report.

1) **5 Points:** Explain five assumptions about a population that are needed to be met for Assumption 1 to be true.
   5 Points: Explain three assumptions about the population that must be true for Assumption 2 to be true.

2) **20 Points:** Describe two species, their life histories, and fieldwork context for which the above assumptions are met for a Peterson mark-recapture population estimate.

3) **5 Points:** Discuss the pros and cons of doing fewer/more recaptures.

4) **4 Points:** Using your own data in the "CollaborativeDataSheet_Petersen" sheet, calculate the mean estimated population size and variance. Interpret your results. Repeat these steps with three other groups' data. Make sure to show your work.
   6 Points: Compare and contrast these mean population size estimates and variances with the corresponding weighted averages and variances you/other groups calculated using the Schnabel method.

5) **10 Points:** Discuss how the features of the park you chose may have affected your data. Use the collaborative data sheet**s** to contextualize your answer.

6) **10 Points:** Thinking about your answers to Questions 5 and 6, use your results to compare and contrast different dog parks and our ability to use them to estimate population sizes.

7) **10 Points:** How might increasing the value of K change the accuracy and/or precision of your estimates? Use both conceptual and mathematical reasoning in your answer.

8) **10 Points:** With reference to your study design, discuss the generalizability of your results, and the shortcomings of using dog parks to estimate dog population sizes.

9) **5 Points:** Go to the CJS Model section of the lab, and take a look at the equation for P_4. Why might the final term read $\varphi(1 - p) + (1 - \varphi)$, rather than $\varphi(1 - p)$?

10) **4 Points:** Compare and contrast the Peterson method, the Schnabel method, and the Cormack-Jolly-Seber model.
6 Points: Imagine that you're conducting a study on the population dynamics of a threatened species, and are interested in the long-term viability of the population. Why might you feel inclined to employ the CJS model, over the Peterson and Schnabel methods?