Estimating Population Size

Part 1: Background

The number of individuals in a population, or ***population size***, is perhaps the most important thing to know about a population. This is most clear in cases where ecologists are working to help endangered species, when an accurate count or estimate of population size is critical to assessing their success. Ideally, population ecologists would have an exact count of all the individuals in a population at all times. Obviously, this would rarely be the case and, in most cases, accurately counting all the individuals in a population is impossible. Imagine trying to count the fire ants that are invading the southeast US. This population would be growing far faster than they could be counted! For this reason ecologist rely on various techniques to estimate population size. There are several established methods for this, and each has advantages and disadvantages that make them appropriate for different situations.

Part 2: Common Population Estimates

Complete counts- There are cases where it is practical to actually count all of the individuals in a population. These are cases where the population is relatively small (maybe thousands of individuals), and contained within a limited area (e.g. national parks). Complete counts are more practical when the species in question is conspicuous, slow moving or sedentary, and has long generation times (so they can't get ahead of you).

A case where exact counts are used is with African elephants on the savannah. These animals live in stable herds, are very hard to miss, and reproduce slowly. Aerial photographs can yield accurate counts of the elephant population on the open savannah. However, the forest-dwelling elephants of equatorial Africa are not easily seen from the air; consequently their populations are not as well known. Another case where direct counts are used is with the giant sequoia population of California. In this case, the location of all the mature trees is known.

Direct sampling- This is the simplest method of population estimation. Here, all of the individuals are counted in a manageable area within the population's ***range***, and this is assumed to represent the population ***density*** across the entire range. This is extrapolated to estimate the overall population size by the equation:

$$N_{est} = N_c \, (A_{tot}/A_c)$$

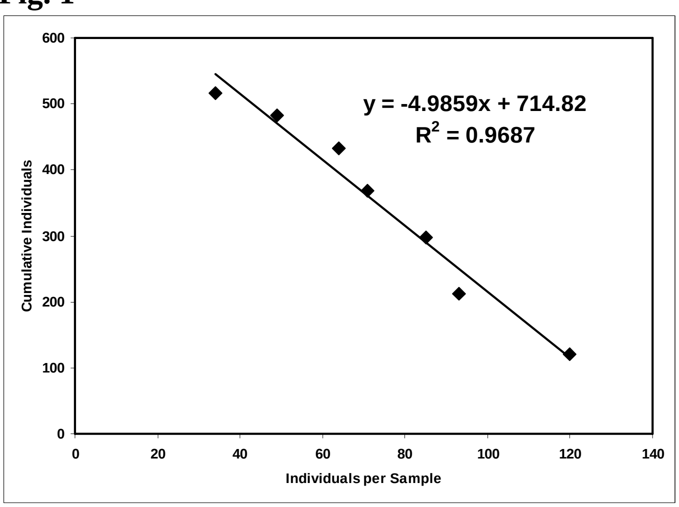
where, N_c is the number counted in the sample, A_{tot} is the total area covered by the population, and A_c is the area covered by the sample. The accuracy and precision of this estimation method depends on several factors: how accurately the individuals can be counted in the area covered, how accurately the sample represents the overall population, and how accurately the range of the overall population can be known. How accurate the count in the sampled area will be depends on how easily individuals are detected, and how ***motile*** they are (if they move around too quickly individuals may get counted multiple times). The ***distribution*** of individuals within their range can drastically affect the overall estimate. ***Uniformly distributed*** individuals will lead to the most accurate population estimates, as the sampled count will more likely represent the overall density of the population. If the population is ***randomly distributed***, accurate estimates can be obtained as long as the sampled area is large enough to 'catch' the overall density. ***Clumped distributions*** are the most problematic for the direct sampling method as a given sample will likely be skewed, having no, or a lot of, individuals. In this case the method would be modified to estimate the mean number of individuals in a clump, and multiply that by the estimated number of clumps.

An example where the direct sampling method could be used is in estimating the population size of a given plant species on the prairie. Accounting for how the plant is distributed, appropriately sized areas are sampled. The overall range of the habitat can be accurately estimated from aerial or satellite images. In practice, ecologists will not rely on a single sample to produce an estimate. Rather they will use several, adjusting the number and size of the samples to best estimate the population.

Sampling with removal- This method, as the name implies, involves repeated sampling of the population and removing the individuals from the population as they are counted. This method can be effective in cases of highly motile animals, and is less affected by the distribution of individuals. It is also effective when organisms can be captured, but not easily seen in their natural habitat. A drawback of this method is that individuals must be held out of the natural population while the sampling continues, potentially disturbing the population. This method assumes that the number of individuals caught in a given sample is directly related the population size. As individuals are removed from the population, the population size will decrease, and fewer individuals will be caught in subsequent samples. The population size is estimated from differences in subsequent samples. Thus, a given sample must be large enough relative to the overall population to have a discernable effect.

There are a few different estimation methods based on this principle. The one we will use is called the Hayne method which plots the cumulative number of individuals caught from the current and all previous samples against the number caught in the current sample (Fig. 1). The Y intercept of the linear regression equation is then the population estimate.

Fig. 1



Sampling with removal is used to estimate the size of fish populations from data collected by the fishing industry. In this case the individuals are being removed from the population anyway so there is no further disturbance. By using the weekly combined hauls from across the fishing fleet as the sample, and then comparing them over the course of a season, problems associated with distribution and ***sampling error*** are reduced.

Mark/recapture- This is another method that can effectively estimate population sizes of hidden and motile populations. The advantage of this method over sampling with removal is that the population is left relatively undisturbed. In this method, individuals are captured from the natural population. Instead of removing them however, the individuals are marked in some way and then released back into the population, recording the number marked. After some time there is another round of capturing, and the numbers of marked and unmarked individuals are recorded. The classic ***Lincoln-Peterson Index*** estimates the population size by the equation:

$$N_{est} = (N_m * N_{2nd}) / N_r$$

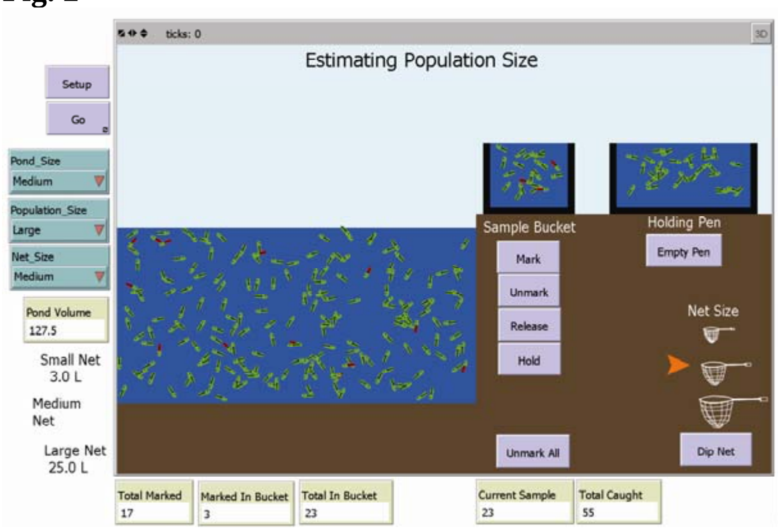
where N_m is the number of individuals marked in the first round of capture, N_{2nd} is the total number of individuals captured in the 2^{nd} round, and N_r is the number of marked individuals caught in the 2^{nd} round. This method usually overestimates the population size. This bias is reduced by ***Bailey's modification*** of the equation:

$$N_{est} = (N_m * (N_{2nd} + 1)) / (N_r + 1).$$

Mark/recapture is commonly used to estimate population sizes of small rodents. These animals are hard to see in their native habitat, but are easily caught in live-traps. Here again, care must be taken to understand the biology of the animal to design appropriate trapping schemes. Animals which are territorial are more likely to be caught twice by a trap in the same location. Also, individuals may learn to avoid traps a second time. Or, if the traps are baited, they may learn that traps are a good source of food and more likely return to them. All of these behaviors could affect population estimates.

Part 3: Using the Model

Fig. 2



This model simulates a pond full of tadpoles which swim around randomly, independent of one another. In this virtual experiment you will sample the population by dipping a net into the pond and emptying it into a bucket. The program reports the number of tadpoles in the bucket. At that point, you can release them back into the pond, mark them, or sequester them in a holding pen. The volume of the pond and the nest are known, so all the data needed for the three estimation methods are provided. Note, the sampling is done by moving all the tadpoles from within a radius around a point in the center of the pond. You should allow enough time between samples for the tadpoles to move around sufficiently to ensure independence of the samples. When working with this model, it will be useful to have a spreadsheet open in which to record data directly.

Table 1: Model controls and parameters

Control	Action
Setup	Sets the model ready to go with the assigned parameters
Go	Puts the tadpoles in motion
Pond Size	Sets the volume of the pond in liters (Small = 76.5, Medium = 127.5, = Large = 178.5)
Population Size	Sets the population size (Small = 25, Medium = 100, Large = 250)
Net Size	Sets the volume of the net in liters (Small = 3, Medium = 11, Large = 25)
Mark	Marks one of the unmarked tadpoles in the bucket red
Unmark	Unmarks one of the marked tadpoles in the bucket
Release	Places all of the tadpoles in the bucket back in the pond
Hold	Places all of the tadpoles in the bucket in the holding pen
Unmark All	Unmarks all the tadpoles in the bucket, pond, and pen
Empty Pen	Places all of the tadpoles in the pen back in the pond
Dip Net	Samples the population

Table 2: Model reporters

Reporter	Description
Pond Volume	The volume of the pond in liters
Total Marked	The total number of marked tadpoles in the pond, bucket, & pen
Marked in Bucket	The number of marked tadpoles in the sample bucket
Total in Bucket	The number of tadpoles in the bucket (marked & unmarked)
Current Sample	The number of tadpoles in the bucket only
Total Caught	The number of tadpoles in the bucket and pen combined

Using the controls, you can generate data to calculate all of the estimates above. Direct sampling is done by dividing the number of tadpoles in a given scoop of the net by the proportion of the whole pond volume represented by that scoop. Sampling with removal is done by taking a sample, recording the number caught, and then sequestering the sample in the holding pen. Mark/recapture estimates involve capturing a sample, marking a known quantity of tadpoles, releasing them into the pond, and then recapturing a sampling.

Sample Questions

1. *How does sample size relative to population density affect the accuracy of the estimates? That is, how close is the mean estimate (given a set of parameters) to the actual value?*

2. *How does sample size relative to population density affect the precision of the estimates? That is, how much variation is there around the mean (for a given set of parameters)?*

3. *Something to consider is even if a larger sample is always better, in practice sampling effort is usually limited. How large is good enough to have confidence in the estimate?*

4. *Another consideration is accuracy vs. precision. For example, an estimate may consistently underestimate the population and would then be precisely inaccurate. How useful would this be compared to a method to a method that on average gives an accurate estimate?*