Second Draft - May 9, 1983

MIMEOGRAPH SERIES #1628

FASCINATION OF STATISTICS BOOK

ESTIMATING THE SIZE OF WILDLIFE POPULATIONS USING CAPTURE TECHNIQUES

Kenneth H. Pollock
Cooperative Wildlife Statistics Project
Department of Statistics
North Carolina State University
Raleigh, North Carolina 27650 U.S.A.

Key Words: Ecology, Sampling, Estimation, Mark-Recapture Method, Removal Method, Catch-Effort Sampling.

How many trout are there in a stream? How many raccoons in a game reserve? In this article I discuss two approaches to addressing these types of difficult questions which are of such importance to fisheries and wildlife ecologists.

The first approach is the mark and recapture technique which has been widely used for a variety of species. A sample of animals are captured and marked before being released back into the population. Later another sample is taken and the proportion of marked animals noted.

In some situations biologists find it preferable to use a second approach which is based on permanent removal of the animals. This removal technique is also widely used especially in fisheries where the biologists may have to depend on samples collected by commercial fishermen.

To the statistician the problems posed involve estimation. The total population size (denoted by N) is an unknown parameter. Based on a statistical model I derive an estimator of N using the data from the mark and recapture or removal experiments. Each model is discussed in detail with emphasis placed on the assumptions required to obtain a valid estimator. I also consider the important question of how many animals should be sampled to obtain reliable estimators. Each technique is illustrated by a detailed practical example.

THE MARK AND RECAPTURE EXPERIMENT

1. Derivation Of The Estimator

To begin with I consider the simplest mark and recapture experiment. A sample of n_1 animals is taken from a population of size N animals. The animals are permanently marked in some manner depending on the species involved. For example, many mammals are marked with metal ear tags, and birds with metal leg bands. There are however a myriad of other marking techniques. After marking the animals are returned to the population. Later another sample of n_2 animals is taken. These animals are examined for marks and the number of marked animals (m_2) is recorded.

Intuitively one might expect the ratio of marked to total animals in the second sample to reflect the same ratio in the whole population. This can be expressed mathematically by the equation

$$\frac{m_2}{n_2} \simeq \frac{n_1}{N} \quad . \qquad (1)$$

This equation can be solved for N to obtain

$$\hat{N} = \frac{n_1 n_2}{m_2} \qquad (2)$$

Let us illustrate these equations numerically. Suppose the biologist marks 100 animals in the first sample and later captures 50 of which 5 are marked. The marked to total ratio in the sample is 0.1 (5/50). For the same ratio to apply to the population the population must consist of 1,000 animals (100/1000 = 0.1). Formally using the equation (2) we have

$$\hat{N} = \frac{n_1 n_2}{m_2}$$

$$= \frac{100 \times 50}{5}$$

$$= 1{,}000$$

$\hat{N}$ is called a point estimator of N, the unknown population size. Statisticians are careful to emphasize the distinction between the estimator ($\hat{N}$) which can be enumerated from the data and the parameter (N) which is the unknown population size. N cannot be found exactly without counting the whole population which is impossible in practice!

One interesting mathematical property of the Petersen estimator needs mention. What happens if no marked animals are captured in the second sample ($m_2 = 0$)? We find $\hat{N}$ is infinite! This is just telling us what we already would expect intuitively. An experiment where no marked animals are captured gives us no useful information on population size.

2. Historical Perspective

The mark and recapture technique has a very long history with the basic principle dating back at least to P. S. Laplace in the late eighteenth century. He used it to estimate the population size of France (Seber (1982; p. 104)). As his number "marked" (n_1) he used the size of a register of births for the whole country. His second sample consisted of a number of parishes of known total size (n_2). For these parishes he also obtained the total number of births (m_2). If one assumes that the ratio of births to total population size can be approximated by the ratio of births to total size in his known parishes then equation (2) can be used.

The use of the method on animal populations began much later. Carl Petersen, a Danish fisheries scientist, had realized the potential of mark and recapture in a paper published in 1896. In 1930 Frederick Lincoln, a United States Fish and Wildlife Service biologist, began to apply the method to waterfowl populations in North America. The estimator given in equation (2) is commonly called the Petersen estimator or Lincoln Index after these two men. It should be emphasized that index, however, is a misnomer because an index is usually defined to be some estimate of relative abundance. This estimator is for the total population size which is absolute abundance.

In an interesting new monograph by White et al. (1982) a brief but very informative history of mark and recapture techniques is given. They have an interesting series of photographs of some of the important contributors to the statistical theory of these methods.

3. Model Assumptions

To obtain the estimator of N given in equation (2) we have built a statistical model of the mark and recapture procedure. This model depends on several critical assumptions which may not always be satisfied in real animal populations.

1. Closure. It is assumed that the population is closed during the period of the study. That is there are no additions to the population due to birth or immigration. Also, there are no deletions from the population due to death or emigration.

2. Equal Catchability. It is assumed that for each sample every animal has the same probability of being captured. However, the probability may differ for the two samples.

3. Marks Permanent. It is assumed that marks are not lost before the second sample is taken.

The assumption of closure is often a reasonable one provided one takes a short time period between the mark and recapture samples so that births and deaths can be ignored. Migration may be a problem in some studies where there is no well defined natural boundary for the population.

The assumption of equal catchability is critical and is often violated in real populations. There are two different kinds of departures from this assumption. There may be inherent heterogeneity in capture probabilities for animals due to many factors like age and sex. If the same animals tend to be easy to capture in the first and second samples then the number marked in the second sample (m_2) becomes too large and hence the population size estimator ($\hat{N}$) is too small. A statistician calls this a negative bias on the estimator. The second departure is that marking may affect an animal's capture probability in the second sample. For example animals could become "trap shy" if their capture and marking was stressful lowering their capture probability compared to unmarked animals. On the other hand animals could become "trap happy" if their capture and marking was enjoyable. This often occurs if the traps contain food and results in marked animals having a higher probability of capture than unmarked animals. A "trap shy" response results in a positive bias on the population size estimates while a "trap happy" response results in a negative bias. Use of different capture techniques in the two samples often greatly reduces problems with heterogeneity and trap response of the capture probabilities.

Mark loss can be a serious problem especially in fisheries studies. It causes a positive bias on the population size estimator. Sometimes the biologist puts two marks on each animal and uses the information on how many animals retain one or two marks to estimate the rate of mark loss. The population size estimates can then be adjusted to eliminate the positive bias caused by the mark loss.

4. Practical Example.

In a recent study on raccoons in northern Florida by Conner (1982) a sample of 48 animals was captured using cage type live traps baited with fish heads. An unusual marking scheme was used on these animals. They were injected with a small amount of a radioactive cadmium isotope. The "recapture" sample involved an intensive search of the study area for scats (faeces). Those faeces which were "marked" (radioactive) were detected using a special scintillation counter.

Conner carried out his search for scats each week for five weeks. He decided to treat each week separately and obtain five different Petersen estimates so that he could consider the variation of the estimates. The data and resulting estimates for each week are now presented.

Week of Collection	Number Marked (n_1)	Number of Scats Collected (n_2)	Number of Marked Scats (m_2)	Petersen Estimate ($\hat{N}$)
1	48	71	31	109.9
2	48	22	11	96.0
3	48	74	35	101.5
4	48	28	9	149.3
5	48	35	19	88.4

One approach to proceeding is to assume that the five different estimates of N, the population size, are independent and have an approximate Normal (Bell-Shaped) distribution. The mean of the five values is 109 animals which is our "best" point estimate of the number of raccoons in the population. It is also informative to calculate a 95% confidence interval estimate of N based on the mean of the five estimates. The theory is given in any good statistics text (Snedecor and Cochran (1980; p. 56)) and will not be presented here. For this example the confidence interval ranges from 79 to 139 animals. Let us summarize our results in the following table.

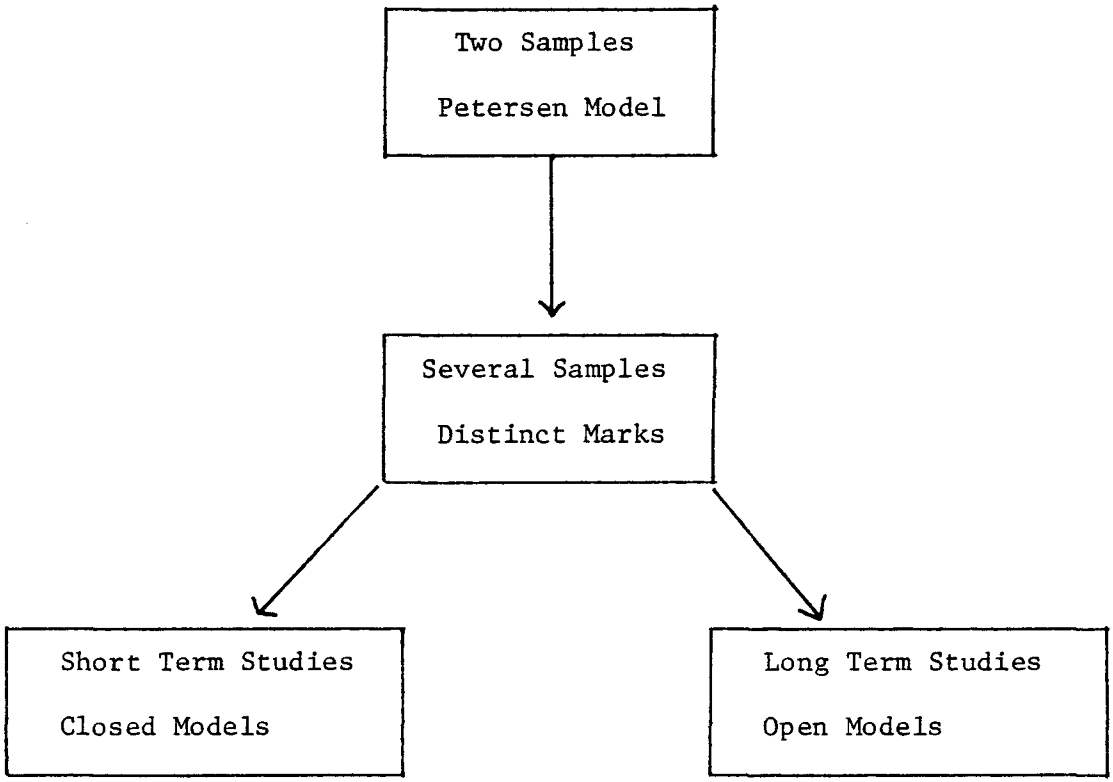
Point Estimate of Population Size	95% Confidence Interval Estimate of Population Size
109 animals	79 - 139 animals

The 95% confidence interval gives us an idea of the precision of our experiment. Here with approximately half of the animals in the population (48 out of 109) marked and the scat collection lasting five weeks the confidence interval had width 60 (139-79). To increase the precision of his study and reduce the confidence interval width the biologist has two options. He can increase the number of marked animals in his population or he can continue the scat collection over more weeks. I calculated the approximate confidence interval width if he had the same number of marked animals but had carried out scat collection for 10 weeks at the same sampling intensity. I found that the confidence interval width would have been 34, which is around half the current width. Of course this much more precise study would have cost a lot more to run. You do not get something for nothing in scientific work!

I believe that this is a sound study in terms of the assumptions behind the model. The biologist felt the population was approximately closed over the period of the experiment. He also felt that there was no appreciable chance of not detecting radioactivity in the scats. The key assumption to consider here (as it is in most studies) is that of equal catchability of the animals. Here there is no danger of obtaining "trap shy" or "trap happy" animals because two completely different methods of sampling are used. These are trapping in the first sample and scat collection in the second. Also, because of the two different methods, I believe it is unlikely that animals capture probabilities in the traps are related to their scat collection probabilities. In other words I do not believe that there will be any negative bias on the population size estimate caused by animals easy to capture in the traps also by having their scats collected at a higher rate.

5. Extensions

I have barely scratched the surface of mark and recapture methodology. Often biologists find it necessary to mark and recapture animals over many samples. In these studies each animal has its own distinct mark (often a numbered tag) and a comprehensive capture history is built up for every animal captured at least once. The scope of mark and recapture methodology can best be summarized in a diagram.



The closed population models can allow for unequal catchability of animals. The open population models also allow for the estimation of birth and death parameters which are very important to biologists.

There is a very large literature on mark and recapture methods. The new monograph by White et al. (1982) is an excellent introduction although it concentrates on closed population models. There are two good recent reviews written for biologists (Nichols et al. (1981), Pollock (1981)). Also an excellent, if rather technical book, is that by Seber (1982).

THE REMOVAL EXPERIMENT

1. Derivation Of The Estimator

Here I consider the simple two sample removal experiment which is closely related to the mark and recapture experiment considered in the previous section. A sample of n_1 animals is captured from a population of size N animals.

Now instead of being marked and returned the animals are permanently removed. In fisheries sometimes they may be part of a commercial catch. For some species live traps may be used and the animals are taken to another site and released. Later another sample of n_2 animals is removed using the same amount of effort as for the first sample.

Given that the same amount of effort is expended for each sample one might expect the proportion of the population removed to be constant for the two samples. This simple but powerful idea forms the basis for the estimator of N. Mathematically the result is that n_1 removals out of a population of size N should be equivalent to n_2 removals out of the reduced population of size $(N-n_1)$. This can be expressed by the formulae

$$\frac{n_1}{N} \approx \frac{n_2}{(N-n_1)} \quad . \tag{3}$$

Solving this equation gives us an estimator of N which is

$$\hat{N} = \frac{n_1^{\,2}}{(n_1-n_2)} \quad . \tag{4}$$

Notice that this estimator has some interesting mathematical properties. If n_1 equals n_2 then the estimate suggests to us that there are an infinite number of animals in our population! If n_1 is less than n_2 then the estimate suggests to us that there are a negative number of animals in our population! If the amount of effort is low then by chance n_1 could be less than or equal to n_2. These nonsensical results are really telling us that we do not have enough information to estimate the population size from such an experiment.

2. Model Assumptions

We have built a statistical model for the removal procedure which is very similar to the mark and recapture model. Again the model depends on several critical assumptions which may not be valid in real removal studies.

1. Closure. It is assumed that the population is closed during the period of the study.

2. Equal Catchability. It is assumed that for each sample every animal has the same probability of being captured.

3. Constant Capture Rate. It is assumed that the amount of effort used for each sample is the same and that this means that all animals have the same probability of capture in sample 1 as in sample 2.

The assumption of closure may often be a reasonable one provided there is a short period between the two samples to prevent migration. For example, if food is in short supply then after animals are removed in the first sample you may have animals moving in from surrounding areas.

Again the assumption of equal capture probabilities is crucial. If there is a tendency for the same animals to be hard (or easy) to capture in both samples the population size estimator $\hat{N}$ will have a negative bias.

The assumption that all animals have the same probability of capture in both samples is also crucial. The biologist can help insure this by using equal effort in both samples. He or she should also attempt to conduct the two samples under identical weather conditions. Obviously, however, this may be impractical in some cases.

3. Practical Example

Sometimes fisheries biologists block off a small section of stream and carry out a removal experiment on the section. One capture technique called electrofishing is to pass an electric current through the water. Hand nets are then used to remove all the stunned fish which have come to the surface.

Seber and Le Cren (1967) have presented data for such an experiment where 79 trout were removed in the first sample and 28 trout removed in the second sample. Use of equation (4) then gives us a point estimate of 122 trout for the population size. A 95% confidence interval of 104 to 140 trout is also given but this theory will not be developed here.

Population Size Estimate

$$\hat{N} = \frac{n_1^2}{(n_1-n_2)}$$

$$= \frac{79^2}{(79-28)}$$

$$= 122 \text{ trout}$$

95% Confidence Interval Estimate

122 ± 18

104 to 140 trout.

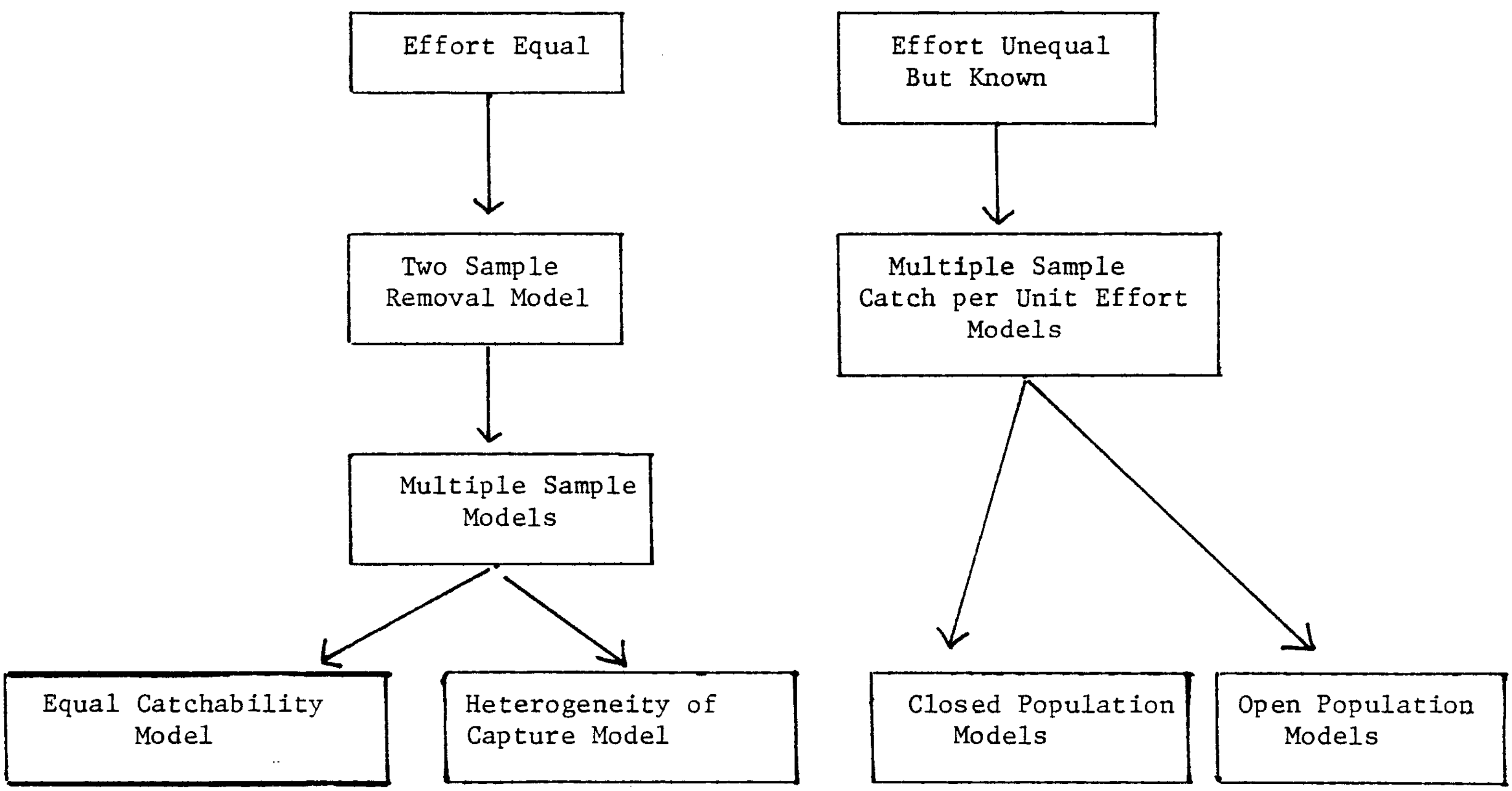
I believe that this is a valid estimate of the population size. Closure is guaranteed by there being a short period between samples and by having the stream blocked off at both ends to prevent fish escaping. The use of equal effort and the two periods close together suggests that the probability of capture should be constant over the two samples.

This is a very precise estimate for a biological study. The reason for this is obvious. The population size estimate is 122 trout and over both samples 107 trout (79 + 28) were actually removed. This degree of removal is often not feasible in practice.

4. Extensions

Often one cannot hope for the high rate of removal of the electrofishing example. Sometimes biologists extend the removals over several samples to compensate for this. It is possible to allow for some degree of heterogeneity of capture probabilities between animals in these multiple sample removal models.

Another extension relates especially to commercial fisheries. Often the effort in different samples is unequal but there is some measure of the relative effort. This leads to a whole class of what are called Catch per Unit Effort Models. There are versions for both closed and open populations. Again a summary diagram may be helpful.



There is also a very large literature on removal and catch per unit effort models. An excellent reference for the removal models is the new monograph by White et al. (1982). Important references for the catch per unit effort models are Ricker (1975) which is written for fisheries biologists and the more technical book by Seber (1982).

DISCUSSION

In this paper there have only been four equations presented. This was very deliberate on my part. I wanted to emphasize that the important concepts of statistics are easily understood because they involve intuition and commonsense. Mathematics should be viewed as a tool the statistician uses. Unfortunately, all too often the layman becomes bogged down in the mathematical details of statistics and "cannot see the forest for the trees".

The aim of this paper has been to show an interesting subset of statistical techniques that ecologists use in their work. For these methods to be used effectively it is necessary for the statistician and ecologist to work closely together. Therefore I feel it is essential for the statistician to have some knowledge of ecology and for the ecologist to have a sound training in the important concepts of statistics.

REFERENCES

Conner, M. C. (1982). Determination of bobcat (Lynx rufus) and raccoon (Procyon lotor) population abundance by radioisotope tagging. M.S. Thesis, University of Florida.

Nichols, J. D., Noon, B. R., Stokes, S. L., and Hines, J. E. (1981). Remarks on the use of mark-recapture methodology in estimating avian population size. In J. M. Scott and C. J. Ralph, eds. Estimating numbers of terrestrial birds, Studies in Avian Biology, 6, 121-136.

Pollock, K. H. (1981). Capture-recapture models: A review of current methods, assumptions and experimental design. In J. M. Scott and C. J. Ralph, eds. Estimating numbers of terrestrial birds, Studies in Avian Biology, 6, 426-435.

Ricker, W. E. (1975). Computation and interpretation of biological statistics of fish populations. Information Canada Bulletin 191, Ottawa, Canada.

Seber, G. A. F. (1982). The estimation of animal abundance and related parameters. 2nd Edition, Griffin, London, England.

Seber, G. A. F. and Le Cren, E. D. (1967). Estimating population parameters from catches large relative to the population. Journal of Animal Ecology, 36 631-643.

Snedecor, G. W. and Cochran, W. G. (1980). Statistical Methods. 7th Edition, Iowa State University Press, Ames, Iowa, U.S.A.

White, G. C., Anderson, D. R., Burnham, K. P., Otis, D. L. (1982). Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, Los Alamos, New Mexico, U.S.A.