NOAA Technical Memorandum NMFS F/NWC-15

Estimation of a Decreasing Population Size Over Time

by
Russell F. Kappenman

September 1981

NOAA FORM 25–13 (1–78)	**BIBLIOGRAPHIC DATA SHEET**		**U. S. DEPARTMENT OF COMMERCE** NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

16. ABSTRACT

A technique for estimating the size of a population of animals at each of several equally spaced points in time is presented. It is assumed that there are no additions to the population over time, but loss of individuals through natural mortality, disease, predation, or cannibalism does occur. The technique is based upon the process of sampling and tagging individuals from the population periodically. (Author)

17. KEY WORDS AND DOCUMENT ANALYSIS

17A. DESCRIPTORS

*Aquaculture, *Sampling, Population, Marking

17B. IDENTIFIERS/OPEN-ENDED TERMS

Population size

17C. COSATI FIELD/GROUP

6F, 8A

18. AVAILABILITY STATEMENT	19. SECURITY CLASS (This report) UNCLASSIFIED	21. NO. OF PAGES 11 p.
Released for distribution:	20. SECURITY CLASS (This report) UNCLASSIFIED	22. PRICE

NOAA FORM 25–13   (1–78)   SUPERSEDES ALL PREVIOUS EDITIONS.

ESTIMATION OF A DECREASING POPULATION SIZE OVER TIME

by

Russell F. Kappenman

Fisheries Data and Management Systems Division
Northwest and Alaska Fisheries Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
2725 Montlake Boulevard East
Seattle, Washington 98112

September 1981

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

A technique for estimating the size of a population of animals at each of several equally spaced points in time is presented. It is assumed that there are no additions to the population over time, but loss of individuals through natural mortality, disease, predation, or cannibalism does occur. The technique is based upon the process of sampling and tagging individuals from the population periodically.

CONTENTS

INTRODUCTION

The culture of aquatic animals in enclosed areas is being practiced for some species and investigated for others in this country. Populations of fish or crustacea are put in enclosures, such as raceways and ponds, which allow neither migration nor immigration. The animals are raised until they are ready for consumption or introduction to a new environment.

Typically, the populations are enormously large (thousands or tens of thousands of animals are often involved) and it is important to monitor population size over time. For many culturings, the population size cannot increase' during confinement but decreases because of natural mortality, disease, predation, or cannibalism. Actual counting of the number of animals present in a given enclosure at a given point in time is often impractical or impossible. Thus a scheme for estimating, with given precision, the number of animals present in a large population at various points in time is of interest.

The purpose of this paper is to present one such scheme. We assume that 1) the initial population size (i.e., the number of animals put in the enclosure) is known; 2) estimates of population size at the end of several consecutive equal time intervals are desired; 3) the animals in the population may be tagged in such a way that if one is drawn from the population at any time during the course of the culture, whether or not it has been tagged previously can be determined; 4) sampling animals from and tagging and returning animals to the population can be performed at those points in time for which population size is to be estimated; 5) the population size can only decrease over time; with loss probability constant from one time interval to the next; and 6) loss probability is the same for tagged and untagged animals. We do not assume that a tag placed on an animal at one point in time can be distinguished from a tag placed at

another point in time. In the applications for which this procedure was developed only one type of tag has been devised.

In the next section, a tagging and sampling procedure for estimating population size over time is described and an estimator of population size at any given point in time is derived, under the assumptions just listed. In the last section, we address the problem of determining sample sizes and tagging numbers.

THE ESTIMATION PROCEDURE

Suppose that a population of N animals, where N is known, is stocked in an enclosure. We want to estimate, at each of several equally spaced time points, the population size. There are no additions to the population, but the size of it can decrease over time. Assume that the survival proportion, S, is constant from one time interval to the next.

In order to estimate the population size at the various time points, we proceed as follows. At stocking, t_o of the N animals are tagged. Then at the end of the i-th interval of time, select at random n_i animals, count the number, x_i, of these which have tags, tag t_i $(\leq n_i - x_i)$ untagged animals, and return all sampled animals to the population.

Now consider estimation of the population size at the m-th time point. The population size and the number of tags present in the population at the i-th point in time, for i = 1, 2, ... m, are, respectively,

$$N_i(S) = NS^1$$

and

$$X_i(S) = \sum_{j=0}^{i-1} t_j S^{i-j}.$$

The distribution of x_i is hypergeometric with parameters $N_i(S)$, $X_i(S)$, and n_1, for $i=1, \ldots . m$. The logarithm of the product of the density functions of x_1' X_2' $\ldots$, x_{m-1}, and x_m is the sum of a quantity, which is not a function of S, and

$$L(S) = \sum_{i=1}^{m} \{ h[N_i(S)-X_i(S)] + h[X_i(S)] + h[N_i(S)-n_i]$$

$$-h[N_i(S)-X_i(S)-n_i+x_i]-h[X_i(S)-x_i]-h[N_i(S)]\}$$

where

$$\mathbf{h[yl} = \mathbf{ln\ y} ! = \ln r(y+1),$$

Note that for large z, one might use the approximation

$$\ln\Gamma(z) = (z-\tfrac{1}{2})\ln z-z+\tfrac{1}{2}\ln(2\pi) + (1/12z) - (1/360z^3)$$

$$+ (1/1260z^5) - (1/1680z^7) + (1/1188z^9).$$

See Abramowitz and Stegun (1964).

A possible estimate of S is the value of S which maximizes L(S). This value may be found by using a simple computer search and, depending on how the t_i's are determined, is either a maximum likelihood or pseudomaximum **likelihood** estimate for S. The corresponding estimate of population size at the m-th time point is N times the m-th power of the estimate of S.

An an example, suppose N = 100,000, m = 6, t_o = 3500, n_i = 3500, for
i = 1, ... 6, x_1 = 154, x_2 = 240, x_3 = 374, x_4 7 531, x_5 = 665, x_6 = 821,
and t_i = n_i - xi, for i = 1, ... 5. The following table gives the values
of $L(S)$ for a range of values of S:

S	0.90	0.91	0.92	0.93	0.94
$L(S)$	-7802.51	-7799.62	-7798.55	-7799.12,	-7801.15

It is clear that the pseudomaximum likelihood estimate of S is 0.92 and an
estimate of the population size at the end of the sixth time interval is
$100,000(0.92)^6$ = 60,635.

If $x_i/X_i(S)$, $[n_i-x_i]/[N_i(s)-x_i(s)]$, and $n_i/N_i(S)$ are small (say, less than
0.1), then the distribution of x_i **may** be approximated by the binomial distribution
whose parameters are n_i and

$$g_i(S) = X_i(S)/N_i(S)$$
$$= [\sum_{j=o}^{i-1} (t_j/S^j)]/N.$$

See Parzen (1960). When these quantities are small for i=1, 2, ... m, the
value of S which maximizes $L(S)$ may be adequately approximated by the value
of S which maximizes

$$L_1(S) = \sum_{i=1}^{m} x_i \ln g_i(S) + (n_i-x_i)\ln[1-g_i(S)].$$

Again a simple search procedure can be used to find this value or one can
apply a numerical procedure, such as the Newton-Raphson one, to solve the
equation

$$\frac{dL_1(S)}{dS} = \sum_{i=1}^{m} \frac{g_i'(S)[x_i-n\, g_i(S)]}{g_i(S)[1-g_i(S)]} = o$$

for S. Here

$$g_i'(S) = -[\sum_{j=1}^{i-1} (j\ t_j/S^{j+1})]/N,$$

for i=1, ... m.

DETERMINATION OF TAGGING NUMBERS AND SAMPLE SIZES

This 'work was motivated by the following posed problem. One hundred thousand shrimp were to be raised for 6 months in a closed raceway. Estimates of population size at monthly intervals were desired. The investigators wanted an estimate, for population size at the end of 6 months, which was within 10% of the actual population size at that time, with about 95% confidence;

The investigators had the capability of inserting a magnetic wire tag internally in each shrimp and determining whether or not a given shrimp was tagged. They were unable to tag shrimp in more than one way. So a tag inserted at one point in time could not be distinguished from a tag inserted at another point in time.

The design of the investigation required guidelines for the number of taggings and the sample sizes. It was decided that if the required sample sizes were not too large, all untagged shrimp in each sample would be tagged before the sampled animals were returned to the population.

If it is possible to tag all untagged animals in each sample', then a sampling and tagging plan only requires values for t_o, the number of animals tagged at stocking, and the n_i's, the sample sizes for the time, points. The values of these needed to insure that the estimate of population size at the m-th time point has a certain precision may be obtained by simulation if one has at least a rough idea of what proportion of the original population will survive to the m-th time point.

Let S be the m-th root of this proportion. For large populations and fixed t_o, n_1, n_2, ... n_{m-1}, and n_m, the distribution of x_i is approximately binomial with parameters n_i and $g_i,(S)$. A simulated value for x_i can be found by counting the number of independent uniform (0, 1) random variates, out of n_i, that do not exceed $g_i(S)$, for i =1, ... m, The uniform random variates can be generated with a computer. Once the m x_i 's are generated, one can compute an estimate of S and an estimate of the population size at the m-th time point. The latter estimate can then be compared to the actual population size at the m-th time point to determine whether or not it is "close enough." For fixed t_o, n_1,,n_2, ... $n_{m\ 1}$, and nm, the whole process may be repeated over and over, a large number of times. When this is done, one is able to estimate the percentage of the time that the estimate of population size is close enough to actual population -size, for the fixed t_o, n_1, n_2, ;.., $n_{m\ 1}$, and n_m Trial and error of these fixed values enables one to arrive at values which insure that the estimate based on them has the desired precision.

An an example, for the posed problem, $N = 100,000$ and the investigators expected about 80% of the population to survive for 6 months. Thus we set $S = (0.80)^{1/6.} = 0.96$. It was found that when $t_o = n_1 = n_2 = \ldots = n_6 = 3500$, the estimate of population size at the end of 6 months was within 10% of the actual population size (80,000) 938 times out of 1000 simulations. Thus we are roughly 94% confident that if 100,000 animals are stocked, if 80,000 animals survive for 6 months, if 3,500 animals are tagged at stocking, if 3,500 animals are sampled monthly, if all untagged animals in each sample are tagged, if all sampled animals are returned to the population, and if the estimation procedure of the preceding section is used, then our estimate of population size is within 10% of the actual population size at the end of 6 months.

LITERATURE CITED

Abramowitz, M., and I. Stegun.

    1964. Handbook of mathematical functions with formulas, graphs, and mathematical tables. U.S. Dep. Commer., Natl. Bur. Stand., App. Math. Ser. 55, 1,046 p.

Parzen, E.

    1960. Modern probability theory and its applications. John Wiley & Sons, Inc., New York, 464 p.