A Note on Estimation Methodology in
Contemporary Small Mammal Capture-
Recapture Studies

James D. Nichols and Kenneth H. Pollock

Inst. of Statistics Mimeo Series #1360

A NOTE ON ESTIMATION METHODOLOGY IN CONTEMPORARY SMALL MAMMAL CAPTURE-RECAPTURE STUDIES[1]

James D. Nichols[2] and Kenneth H. Pollock[3]

Enumeration estimators, which are also known as "minimum number known alive" and "calendar-of-catches" estimators (e.g. see Newson and Chitty 1962, Krebs 1966, Krebs et al. 1969), currently dominate the literature of small mammal capture-recapture studies. Although it is widely recognized that these estimators are negatively biased, they still seem to be preferred by mammalogists to estimators based on probabilistic models such as the Jolly-Seber model (Jolly 1965, Seber 1965). It is our belief that estimators based on the Jolly-Seber and related models are preferable to enumeration estimators and should thus be adopted in small mammal capture-recapture studies.

In this note we first give a brief description of the Jolly-Seber model and then write down the expected values of enumeration survival rate and population size estimators in terms of the parameters of the Jolly-Seber model. We show that these estimators actually estimate complicated functions of the Jolly-Seber model parameters. We then list frequently used reasons for choosing enumeration over Jolly-Seber estimators and argue that Jolly-Seber estimators are still to be preferred.

The Jolly-Seber Model

The most important open population model is the Jolly-Seber model, independently derived by Jolly (1965) and Seber (1965). The best reference for detailed study of this model is Seber (1973:196). This model makes the following assumptions:

(i)   Every animal in the population (marked or unmarked) has the same

probability (p_i) of being caught in the ith sample (i = 1,..., K), given

that it is alive and in the population when the sample is taken.

(ii) Every animal has the same probability (ϕ_i) of surviving from the ith

to the (i + 1)th sample, given that it is alive and in the population

immediately after the ith release (i = 1,..., K-1).

(iii) Marked animals do not lose their marks and all marks are reported on

recovery.

(iv) The actual time spent sampling occupies a short period.

Here an intuitive discussion of parameter estimation will be given.

Imagine to begin with that M_i, the number of marked animals in the population

just before the ith sample, is known for all values i = 2,..., K (there are no

marked animals at the time of the first sample so that M_1 = 0).

Obviously an intuitive estimator of N_i, the population size at time i,

is the Petersen estimator. If the model assumptions are valid, then

$$\frac{m_i}{n_i} \approx \frac{M_i}{N_i} \quad,$$

which gives

$$\hat{N}_i = \frac{n_i M_i}{m_i} \quad, \tag{1}$$

where m_i and n_i are the marked and total numbers of animals captured in the

ith sample, respectively.

An estimator of the survival rate from sample i to sample (i+1) is

M_{i+1}, which is the total number of marked animals in the population just

before the (i+1)th sample, divided by the total number of marked animals

in the population immediately after sample i, which is $M_i - m_i + R_i$.

Note that R_i is the number of the n_i animals captured that are released.

$$\hat{\phi}_i = \frac{M_{i+1}}{(M_i - m_i + R_i)} \qquad \qquad (2)$$

An intuitive estimator of the recruitment in time interval i to (i+1) is

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\phi}_i (\hat{N}_i - n_i + R_i). \qquad (3)$$

This is simply the estimated difference between the population size at time (i+1) (which is N_{i+1}) and the expected number of survivors from time i to time (i+1) (which is $\phi_i(N_i - n_i + R_i)$).

To complete this intuitive outline we need an estimator of the M_i because they are obviously unknown in an open population. This can be obtained by equating the two ratios

$$\frac{Z_i}{M_i - m_i} \simeq \frac{r_i}{R_i} \qquad ,$$

which are the future recovery rates of the two distinct groups of marked animals:

(i) $(M_i - m_i)$ are the marked animals <u>not seen at i</u>

and (ii) R_i are the animals seen at i and then released for possible recapture. Note that Z_i and r_i are the members of $(M_i - m_i)$ and R_i, respectively, which are captured again at least once. The estimator of M_i is thus given by

$$\hat{M}_i = m_i + \frac{R_i Z_i}{r_i} \qquad \qquad (4)$$

and is defined only for i = 2,..., K-1. It follows that $\hat{N}_i$ in (1) is defined for i = 2,...,K-1; $\hat{\phi}_i$ in (2) for i = 1,..., K-2; and $\hat{B}_i$ in (3) for i = 2,..., K-2.

A summary of these parameter estimators and their approximate large sample variances is given in Seber (1973:205), which should also be consulted for some detailed examples. A very important new development is the work of

Jolly (1979, 1981). He restricts the large number of parameters by assuming a constant survival rate and/or a constant capture rate <u>over the whole study</u>. If these assumptions are realistic, as they often are, then he shows that there can be large gains in precision of estimators. Crosbie (1979) has also considered similar models.

Enumeration Estimators

The enumeration estimator for survival rate is simply:

$$\hat{\phi}_i^E = r_i/R_i, \tag{5}$$

where $\hat{\phi}_i^E$ denotes the enumeration survival rate estimator, and r_i and R_i are as previously defined. The expected value of this estimator can be written as:

$$E(\hat{\phi}_i^E) \simeq 1-\chi_i = \phi_i(1-q_{i+1}\chi_{i+1}) \tag{6}$$

where χ_i is the probability of never capturing an animal after time i given that the animal is alive and in the population at the time of release in period i (note that $\chi_K = 1$, where K is the last sampling period), and $q_{i+1} = 1-p_{i+1}$. χ_{i+1} can be expressed in terms of ϕ_{i+1}, q_{i+1}, and χ_{i+2}, and subsequent χ_j can also be expressed in terms of survival and capture probabilities. The estimator, $\hat{\phi}_i^E$, is thus seen to estimate a complicated function involving every survival and capture probability subsequent to period i.

The enumeration estimator for population size at time i is given by:

$$\hat{N}_i^E = n_i + Z_i . \tag{7}$$

The expected value of this estimator can be written as:

$$E(\hat{N}_i^E) = N_i p_i + M_i q_i (1-\chi_i) . \tag{8}$$

Because of the χ_i term this expected value is seen to involve all of the p_j and ϕ_j occurring subsequent to period i. In addition, M_i is itself a

function of all p_j and ϕ_j occurring before time i. Therefore, $\hat{N}_i^E$ also estimates a complicated function involving a large number of parameters.

Comparison of Methods

The most frequently-provided reason for preferring enumeration to Jolly-Seber estimators is that capture probabilities, p_i, often are not the same for all individuals in the population (or in the stratum of interest) for a given time period, and that this assumption is required by the Jolly-Seber model. Conceptually, we can think of unequal capture probabilities as resulting either from heterogeneity or trap response (e.g. see Pollock 1981a). Heterogeneity refers to the situation in which capture probability is a property of the individual and varies from one animal to another in the population (for small mammal examples, see review of Smith et al. 1975). Heterogeneity results in negatively-biased Jolly-Seber estimates of population size and survival rate, although the survival rate estimates are much less affected (see Cormack 1972, Carothers 1973, Gilbert 1973). Trap response refers to the situation in which an animal's capture probability depends on his previous capture history (e.g. whether or not it has been caught before). Many small mammals are thought to exhibit greater capture probabilities after their initial capture (see review in Tanaka 1980). This type of response results in negatively biased Jolly-Seber estimates of population size, whereas a decrease in capture probability after initial capture can result in a positive bias.

Although heterogeneity of capture probability and trap response do result in biases in Jolly-Seber estimates, we do not believe that this should result in the preference of enumeration estimators. Enumeration estimators are always biased whenever capture probabilities are less than 1. In addition,

the situation (high average capture probability) for which the bias of

enumeration estimators is said to be acceptably small (see Hilborn et al.

1976) is the same situation for which the Jolly-Seber estimators are known

to be robust to heterogeneity of capture probability (Carothers 1973,

Gilbert 1973). We noted that the bias of the Jolly-Seber estimates of

ϕ_i and N_i is negative in the case of heterogeneity. The bias of the

estimate of N_i is also negative in the case of a "trap-happy" response to

initial capture. However, in the case of negative bias, we can show that

the absolute value of the bias of the Jolly-Seber estimates is either equal

to or smaller than the absolute value of the bias of the enumeration estimates

(see Appendix). Therefore, the Jolly-Seber estimates perform better than

the enumeration estimates even in the case of heterogeneity and trap-happy

response, the 2 sources of unequal capture probability most likely to

occur in small mammal studies. Finally, we note that trap response has

been explicitly incorporated in the general probabilistic models of Robson

(1969) and Pollock (1975). Despite the claim that trap response occurs commonly

in small mammal capture-recapture studies, we are not aware of any uses of these

general models in such studies.

It has also been claimed that the inability of the Jolly-Seber model to

handle more than one age class severely limits its utility for small mammal

studies. However, a model permitting different capture and survival

probabilities for different age classes has recently been developed (Pollock

1981b, Stokes in prep.).

It is sometimes argued that the actual values of estimates of demographic

parameters are not as important as comparisons of such estimates (e.g. between

populations, time periods, etc.). Therefore, if the enumeration estimators

provided relative estimates that could be used in such comparisons, then an

argument could be made for their use. However, if we wish to use comparisons of different estimates to draw inferences about differences between the underlying parameters, then we must be willing to assume that the biases are equal for the estimates being compared. A simple example will illustrate that this is never likely to be the case. Suppose we have two populations each with perfect survival ($\phi_i = 1$) and each is sampled 3 times. In the first population the probability of capture in each sample is $p_i = 0.9$ whereas in the second population the probability of capture is 0.8. Using equation (6) we can calculate the expected values of the survival rate estimates which are:

Population 1	Population 2
$E(\hat{\phi}_1^E) = 0.99$	$E(\hat{\phi}_1^E) = 0.96$
$E(\hat{\phi}_2^E) = 0.90$	$E(\hat{\phi}_2^E) = 0.80$

Admittedly this is a very small experiment but it does illustrate the fact that the bias can differ markedly between time periods as well as between populations. The data in Newson and Chitty (1962:734) also illustrate this. In all cases there is a marked drop in the enumeration survival estimate in the final period. This could be due to a real drop in survival, but it is also consistent with our example. The negative bias of survival estimates is always much larger at the end of the study. Temporal changes in capture probability can also cause problems in comparing enumeration estimates. We have seen examples in the literature in which enumeration survival estimates are lowest during the period of the year in which capture probabilities are lowest. Although such results are generally interpreted as evidence of low survival rates during the period of interest, they would also be expected to occur simply as a result of low capture probabilities.

Finally the argument that the Jolly-Seber estimates are difficult to compute is not valid. A very powerful, flexible program package called POPAN-2 (Arnason and Baniuk 1978) is now available. The data required for the program is each individual's complete capture history together with any attribute data such as age or sex. There is a strong data manipulation capability so that, for example, it is easy to obtain analyses stratified by age or sex or any other coded attribute data.

Conclusion

In summary, we strongly recommend the use of probabilistic models for the analysis of small mammal capture-recapture data. In addition to the standard model of Jolly (1965) and Seber (1965), the age-specific model of Pollock (1981b) and Stokes (in prep.) should prove useful, as might the trap-response models of Robson (1969) and Pollock (1975) and the modified Jolly-Seber models of Jolly (1979, 1981) which restrict the number of parameters. Because of the heterogeneity and trap response believed to operate in small mammal populations, perhaps a sensible design for capture-recapture studies is that described by Pollock (1981c). In this design population size is estimated using capture-recapture data within a sampling period (using the closed population models of Otis et al. 1978 which allow unequal catchability) and survival rates are estimated from recaptures occurring between sampling periods (using the Jolly-Seber model). We conclude that estimates based on probabilistic models should generally be superior to those based on enumeration techniques, and we hope that small mammal population ecologists will begin to take advantage of these models.

Literature Cited

Arnason, A.N. and L. Baniuk. 1978. POPAN-2, A data maintenance and analysis system for recapture data. Charles Babbage Research Centre, St. Pierre, Manitoba, Canada.

Carothers, A. D. 1973. The effects of unequal catchability on Jolly-Seber estimates. Biometrics 29:79-100.

Cormack, R. M. 1972. The logic of capture-recapture estimates. Biometrics 28:337-343.

Crosbie, S. F. 1979. The mathematical modelling of capture-mark-recapture experiments on open animal populations. PhD Thesis, University of Otago, Dunedin, New Zealand.

Gilbert, R. O. 1973. Approximations of the bias in the Jolly-Seber capture-recapture model. Biometrics 29:501-526.

Hilborn, R., J. A. Redfield, and C. J. Krebs. 1976. On the reliability of enumeration for mark and recapture census of voles. Canadian Journal of Zoology 54:1019-1024.

Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52:225-247.

Jolly, G. M. 1979. A unified approach to mark-recapture stochastic model, exemplified by a constant survival rate model. Pages 277-282 in R. M. Cormack, G. P. Patil, and D. S. Robson, editors. Sampling biological populations. Statistical Ecology Series, Vol. 5. International Cooperative Publishing House. Burtonsville, Md.

Jolly, G. M. 1981. Mark-recapture models with parameters constant in time. Biometrics 37. (in press).

Krebs, C. J. 1966. Demographic changes in fluctuating populations of Microtus californicus. Ecological Monographs 36:239-273.

Krebs, C. J., B. L. Keller, and R. H. Tamarin. 1969. _Microtus_ population

biology: demographic changes in fluctuating populations of _M. ochrogaster_

and _M. pennsylvanicus_ in southern Indiana. Ecology 50:587-607.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978.

Statistical inference from capture data on closed animal populations.

Wildlife Monograph No.62. 135 pp.

Newson, J., and D. Chitty. 1962. Haemoglobin levels, growth and survival

in two _Microtus_ populations. Ecology 43:733-738.

Pollock, K. H. 1975. A K-sample tag-recapture model allowing for unequal

survival and catchability. Biometrika 62:577-583.

Pollock, K. H. 1981a. Capture-recapture models: A review of current

methods, assumptions and experimental design. _In_ C. J. Ralph and

J. M. Scott, editors. Estimating the numbers of terrestrial birds.

Studies in Avian Biology 6. (in press).

Pollock, K. H. 1981b. Capture-recapture models allowing for age-dependent

survival and capture rates. Biometrics 37. (in press).

Pollock, K. H. 1981c. A capture-recapture sampling design robust to

unequal catchability. Journal of Wildlife Management. (in press).

Robson, D. S. 1969. Mark-recapture methods of population estimation

Pages 120-140 _in_ N. L. Johnson and H. Smith, editors. New developments

in survey sampling. John Wiley and Sons, Inc., New York.

Seber, G. A. F. 1965. A note on the multiple recapture census. Biometrika

52:249-259.

Seber, G. A. F. 1973. The estimation of animal abundance and related

parameters. Griffin, London.

Smith, M. H., R. H. Gardner, J. B. Gentry, D. W. Kaufman, and M. H. O'Farrell.
1975. Density estimation of small mammal populations. Pages 25-53 in
F. B. Golley, K. Petrusewicz, and L. Ryszkowski, editors. Small mammals:
their productivity and population dynamics. Cambridge University Press,
Cambridge.

Tanaka, R. 1980. Controversial problems in advanced research on estimating
population densities of small rodents. Researches on Population Ecology.
Supplement Number 2. 67 pp.

APPENDIX

Here we show that the Jolly-Seber estimates are always at least as
large as enumeration estimates, and this indicates that Jolly-Seber estimates
are at least as good as enumeration estimates in situations producing negative
bias.

Population Estimation

The enumeration estimator is given by (7)

$$\hat{N}_i^E = n_i + Z_i,$$

whereas the Jolly-Seber estimator is given by (1) and can be rewritten as

$$\hat{N}_i = n_i Z_i R_i / r_i m_i + n_i.$$

Since $R_i/r_i \geq 1$ and $n_i/m_i \geq 1$, we know that:

$$n_i Z_i R_i / r_i m_i \geq Z_i.$$

Therefore, $\hat{N}_i \geq \hat{N}_i^E$.

Survival Estimation

The enumeration estimator is given by (5)

$$\hat{\phi}_i^E = r_i/R_i \quad ,$$

whereas the corresponding Jolly-Seber estimator is given by (2) and can be

rewritten as

$$
\hat{\phi}_i \;=\; \left[\frac{Z_{i+1}\,R_{i+1}}{r_{i+1}\,(Z_i + r_i)} + \frac{m_{i+1}}{Z_i + r_i}\right]\left[\frac{r_i}{R_i}\right]
$$

We know that $R_{i+1}/r_{i+1} \geq 1$, and $(Z_{i+1} + m_{i+1})/(Z_i + r_i) = 1$, because all the marked animals known to be in the population at time $i + 1$ $(z_{i+1} + m_{i+1})$ is equivalent to those animals marked at or before time i that are subsequently recaptured $(Z_i + r_i)$. Hence, $\hat{\phi}_i \geq \hat{\phi}_i^{E}$.