ESTIMATION OF SURVIVAL RATES FROM A TAG-RECAPTURE

STUDY WITH TAG LOSS

BU-846-M by July, 1984

Walter K. Kremers

Biometrics Unit, Cornell, Ithaca, New York

ABSTRACT

A probability model is described for a tag-recapture study where birds are released annually and birds may be recovered and identified by a permanent band or resighted and identified by a non-permanent collar. Survival, capture, and sighting probabilities are assumed to depend on the year. Collar retention probabilities are assumed to depend on only the age of the collar. Closed form maximum likelihood estimates do not exist for the general model but numerical solutions are easily obtained by the EM algorithm.

INTRODUCTION

Survival rates can be estimated through the use of many different tag-recapture models. If survival rates are of primary interest apart from population size a series of models developed by Brownie et.al.(1978), and Seber(1971) are applicable. These works are based on recoveries, that is, on reobservations where the bird (animal) is caught and permanently removed from the population. Brownie's work is also extended to resightings or nonrecovery observations, where capture probabilities are replaced by probabilities of observation without reobservation thereafter.

Generally for tag-recapture studies retention of indentifying bands (or marks) ia assumed to be permanent or that retention rates are incorporated into survival rates. Here we consider a model similar to that of Brownie and Seber but based on two types of observations and two types of marks. One mark, an identifying "band", is assumed to be permanent and the second, an identifying "collar", is not. The first type of observation is a recovery from which we observe the band and if retained, the collar. The second type of observation is a resighting of the bird by the collar. Heuristically we estimate collar loss from those birds which are recovered and then estimate survival from birds resighted by accounting for collar loss, and by birds recovered.

Following the convention of Brownie et.al. we ignore intermediate observations and consider likelihood functions for recaptures and resightings using only the "final" observation. By considering only the final observation and ignoring intermediate observations for each bird the likelihood becomes markedly simpler though with this simplification is a loss of information. Closed form Maximum Likelihood Estimates (MLE's) do not exist for the model proposed here but numerical MLE's are easily obtained by the use of the EM algorithm (Dempster,*et.al.*,1977) or other numerical procedures. Variances and covariances must be estimated from estimates of the information matrix.

MODEL & NOTATION

Birds are released once a year, and observed for ℓ years. The probability of survival from the time of release in one year to the time of release in the next year is assumed to be the same for all birds alive at beginning of this period, and is allowed to depend on the year. If an animal survives from one year to the next we assume the probability of collar retention to depend on age of the collar, but not on the year. We also assume that for every year, each bird is recovered with the same probability, and each collared bird is sighted but not recovered with the same probability. Collar loss and mortality between the time of reobservation and most recent release are assumed negligible.

With notation similar to that of Brownie et.al. we consider the following random variables.

Xc1(i,j)= number of birds released in the i'th year, recovered in the j'th year <u>with</u> collars.

Xc0(i,j)= number of birds released in the i,th year, recovered in the j'th year <u>without</u> collars.

Xs1(i,j) = number of birds released in the i'th year, sighted (with collars) in the j'th year and not observed thereafter.

$X(i,j) = Xc1(i,j) + Xc0(i,j) + Xs1(i,j)$

$Cc(j) = \sum_{i=1}^{j} Xc1(i,j) + Xc0(i,j)$

$Cs1(j) = \sum_{i=1}^{j} Xs1(i,j)$

$R(i) = \sum_{m=i}^{\ell} X(m,j)$

$$T(i) = \sum_{m\le i,\, n\ge i+1} X(m,n)$$

$$D(i) = \sum_{m,n:n-m\ge i} (Xc1(m,n) + Xs1(m,n))$$

$$Ec0(i) = \sum_{m=1}^{\ell-i} Xc0(m,m+i),$$

As in, and in addition to, the notation of Brownie et.al. we define the following parameters.

S_i= probability that a bird survives until the (i+1)'th release given it has survived until the i'th release.

P_i= probability that a collar is retained i years given the bird bearing the collar has retained its collar and survived i-1 years since its release.

f_i= probability that a bird is recovered in the i'th year given it is alive at the time of the i'th release.

g_i= probability that a bird is resighted but not recovered in the i'th year given it is alive and has its collar at the time of the i'th release.

For this model we assume that the sighting of a bird in year j without its recovery in year j is not related to the survival of the bird during year j, except for the fact that the bird was not recovered. That is, because survival rates are the same for sighted and nonsighted birds,

P[Survives year j|sighted but not captured]

=P[Survives year j|not captured]

=P[Survives year j and is not captured]/P[not captured]

=P[Survives]/P[not captured]

$=S_j/(1-f_j)$.

Similarly the probability of a bird being seen in year j given it is not recovered in year j and has a collar is

P[seen in year j |alive,collar present & not recovered]$=g_j/(1-f_j)$

Let $\chi_{i,j}$ be the probability of a bird not being observed after the j'th year given the bird is sighted but not recovered in the j'th year and was released in the i'th year. Because the probability of retention of the collar is dependent on the age of the collar $\chi_{i,j}$ is not a function of j alone. In terms of the capture, sighting, survival, and retention probabilities

$$\chi_{i,j}=1-[S_j/((1-f_j)]$$
$$x\sum_{m=j+1}^{\ell}\{[\Pi_{n=j+1}^{m-1}S_n(1-g_n-f_n)/(1-f_n)]x[f_m+g_m\Pi_{n=j-i+1}^{m-i}P_n]\},$$

where $\Pi_{n=j+1}^{j}a_n$ is understood to be equal to one.

Similarly let $(1-\lambda_i)$ be the probability of a bird released in the i'th year never being observed. Then

$$(1-\lambda_i)=1-\sum_{m=i}^{\ell}\{[\Pi_{n=i}^{m-1}S_n(1-f_n-g_n)/(1-f_n)]x[f_m+g_m\Pi_{n=1}^{m-1}P_n]\}$$

Each bird released may be last observed in only one year, or not observed at all. The probability that a bird released in the i'th year is last sighted in the j'th year, is the probability that a bird survives and retains the collar until the j'th release, is sighted in the j'th year, and is not observed thereafter. This probability is

$$S_i..S_{j-1}P_1..P_{j-i}g_j\chi_{i,j}.$$

THE LIKELIHOOD

Let C be the constant

$$\Pi_{i=1}^{\ell}\begin{pmatrix} N(i) \\ Xc1(i,i),\ldots,Xc1(i,\ell),Xc0(i,i+1),\ldots,Xc0(i,\ell),Xs1(i,i),\ldots,Xs1(i,\ell)\end{pmatrix}$$

where

$$\begin{pmatrix} A \\ a_1,a_2,\ldots,a_n\end{pmatrix}=A!/(a_1!\ a_2!\ldots a_n![A-\sum_{i=1}^{n}a_i]!)$$

Since a bird may be last observed in only one year the probability distribution function or likelihood is that of a multinomial distribution and after some simplification, is found to be

$$L=C\times\{\Pi_{i=1}^{\ell} S_i^{T(i)}\ P_i^{D(i)}\ (1-P_1..P_{i-1})^{EcO(i)}\ f_i^{Cc(i)}\ g_i^{Csl(i)}$$
$$(1-\lambda_i)^{N(i)-R(i)}\}$$
$$\times\{\Pi_{i,j=1}^{\ell} \chi_{i,j}^{Xsl(i,j)}\}$$

For the analogous model by Brownie, where survival and recovery rates are dependent on year, the dimensionality of the sufficient statistic is the same as the parameter space and closed form solutions are derived. Here the dimensionality of the sufficient statistic is greater than the parameter space and closed form MLE's are not known (to the author).

To obtain MLE's we consider the E-M algorithm. Let "ghost" sightings be sightings of those birds which we have released but have lost their collars. Because these birds have lost their collars we are unaware of the resighting. However if estimates of the parameters are given we may calculate the expected number of ghost resightings and include this in the likelihood. When this is done we may obtain closed form "MLE's" from this adjusted likelihood. The new estimates are then used to estimate the expected number of ghost resightings and the procedure continues iteratively until a maximum is achieved.

Specifically the procedure is as follows.

Let Ψ_i=the probability that a bird sighted in the j'th year is not reobserved and neither is its ghost

$$=1-[S_j/((1-f_j))]\times\sum_{m=j+1}^{\ell}\{[\Pi_{n=j+1}^{m-1} S_n(1-g_n-f_n)/(1-f_n)]\times\ [f_m+g_m]\}$$

Let G_i = probability that a bird (or its ghost) is resighted but not recovered in the i'th year <u>and</u> not reobserved thereafter = $g_i\ \Psi_i$.

Let $(1-\lambda\lambda_i)$= probability that a bird released in the i'th year is never observed and neither is its ghost

$$= 1-\sum_{m=i}^{\ell}\{[\Pi_{n=i}^{m-1} S_n(1-f_n-g_n)/(1-f_n)] \times [f_m+g_m]\}$$

Execute the expectation procedure by letting $Xs0(i,j)$ be the expected number of birds last resighted as ghosts in the j'th year of those released in the i'th year

$$= \sum_{k=i+1}^{j-1} \{ Xs1(i,j) [\Pi_{m=k}^{j-1} S_m]$$
$$[\sum_{m=1}^{j-k} (1-P_{k-i+m}) \Pi_{n=1}^{m-1} P_{k-i+n} (1-g_{k+n}-f_{k+n})/(1-f_{k+n})] G_j/\chi_{i,k} \}$$
$$+ (N_i - R_i) [\Pi_{m=i}^{j-1} S_m$$
$$[\sum_{m=1}^{j-i} (1-P_m) \Pi_{n=1}^{m-1} P_n (1-g_{i+n}-f_{i+n})/(1-f_{i+n})] G_j/(1-\lambda_i)$$

Adjust the $Xs1(i,j)$ for the resightings of ghosts by multiplying by the probability of no reobservation by the collar or of the ghost given there is no reobservation by the collar. That is let

$$Xs1^*(i,j)= (\Psi_j/\chi_{i,j})Xs1(i,j).$$

Redefine all random variables in terms of the $Xs1^*(i,j)$, and in addition define

$$Cs(j) = Cs1(j) + \sum_{i=1}^{j} Xs0(i,j)$$

$$RR(i) = R(i) + \sum_{j=i}^{\ell} Xs0(i,j)$$

$$TT(i) = T(i) + \sum_{m\le i,\, n\ge i+1} Xs0(m,n)$$

$$Esc0(i) = Es0(i) + \sum_{m=1}^{\ell-i} Xs0(m,m+i)$$

Execute the maximization procedure by observing that the "likelihood" is proportional to

$$\Pi_{i=1}^{\ell} \{S_i^{TT(i)} P_i^{D(i)} (1-P_1\ldots P_{i-1})^{Esc0(i)} f_i^{Cc(i)} G_i^{Cs(i)}$$
$$(1-\lambda\lambda_i)^{N(i)-RR(i)}\}$$

and "solve for the MLE's" which are

$$S(i) = (RR(i)/N(i)) ((TT(i)-Cc(i)-Cs(i))/TT(i)) (N(i+1)/RR(i+1))$$

$$P(1) = (D(1)-D(2))/(D(1)-D(2)+Ecs0(1))$$

P(i) = (D(i)-D(i+1))/[(D(i)-D(i+1)-EscO(i))P(1)...P(i-1)]

f(i) = (RR(i)/N(i)) (Cc(i)/TT(i))

G(i) = (RR(i)/N(i)) (Cs(i)/TT(i))

Return to the expectation procedure and iterate until a stable set of parameter estimates are obtained.

NUMERICAL EXAMPLE

The data in Table I are contrived so that the MLE's are $S_1=S_2=0.8$, $S_3=0.7$, $P_1=0.8$, $P_2=P_3=0.9$, $g_1=g_2=0.6$, $g_3=g_4=0.5$ and $f_1=..=f_4=0.02$. $N_1=..=N_4=1000$. For a variety af starting values convergence is reached to within two significant digits of the (correct) MLE's within fifty iterations.

DISCUSSION

We have considered the model where all birds have the same survival and reobservation probabilities. These assumptions can be relaxed just as in the work of Brownie et.al. If collar retention probabilities are dependent on only the age of the collar estimates of parameters may be estimated using the EM algorithm as we have done here and models may be compared using the likelihood-ratio test. Future work might consider models where collar retention probabilities are not determined by the age of the collar.

A shortcoming of this model is that collar loss and mortality between the time of reobservation and most resent release are assumed negligible. This assumption may be reasonable for birds tagged and released in September and hunted or sighted in October but is not reasonable if the birds are released instead in March. Natural extensions of this this model would be to relax this assumption.

In practices where resighting rates are high and recovery rates are low most of the information for survival rates will be from resightings apart from a constant of proportionality determined by the collar retention probabilities. Without recovery information the model is overparamiterized as S_i and P_j always occur together. If all S_i are multiplied by a constant and all P_j divided by the same constant the likelihood is unchanged. However apart from this constant parameters may be estimated. In particular the S_i/S_j are estimable and in applications where trends in survival probabilities are of primary interest these trends may be estimated from resighting data alone without knowing collar retention probabilities

REFERENCES

Brownie, C.,D.R.Anderson,K.P.Burnham,and D.S.Robson (1978). *Statistical Inference From Band Recovery Data-A Handbook.* U.S. Fish and Wildlife Service, Resource Publication #131.

Dempster,A.p.,N.M.Laird,and D.B.Rubin (1977). Maximimum likelihood from incomplete data via the EM algorithm. *J.R.Statist.Soc.*B **39**:1-38.

Seber, G.A.F. (1970). Estimating time-specific survival and reporting rates for adult birds from band returns. *Biometrika* **57**: 313-318.

Table I

Contrived data for a tag-recapture study with $\ell=4$

year of release	year of observation 1	2	3	4
		Xc1(i,j)		
1	20	12.8	9.22	5.8
2		20	12.8	8.06
3			20	11.2
4				20
		Xc0(i,j)		
1	-	3.2	3.58	3.15
2		-	3.2	3.14
3			-	2.8
4				-
		Xs1(i,j)		
1	279	191	153	145
2		330	212	202
3			350	280
4				500