Note: This is a semi-final version of the paper cited below. There are additions to the tables and an appendix that did not appear in the final paper, but otherwise the text is close to that published. The Tables are gathered at the end of the paper.

A general methodology for the analysis of capture-recapture experiments in open populations

Carl James Schwarz

Department of Mathematics and Statistics, Simon Fraser University,

Burnaby, British Columbia   V5A 1S6, Canada

and

A. Neil Arnason

Department of Computer Science, University of Manitoba,

Winnipeg, Manitoba R3T 2N2, Canada

SUMMARY

We trace the development of a likelihood function representation for the open-population capture-recapture (Jolly-Seber) experiment. We find that the modelling of the birth process in the general model is not consistent with the reduced death-only model and that all formulations to date lead to difficulties in imposing constraints upon the parameters of the birth process. We propose a generalisation to the usual Jolly-Seber representation that models births using a multinomial distribution from a super-population. We show how this leads to simplifications in the numerical optimization of the likelihood and how constraints upon the parameters of the model can now be easily imposed. We show how covariate models using auxiliary variables such as sampling effort or weather conditions to explain capture or survival rates can also be easily added. We also show how this model can be generalised to more than one group of animals. Finally a numerical example is provided which fits a class of models where the capture probabilities, survival probabilities and birth probabilities can each vary over time or among groups or both. This permits sequential model fitting within a comprehensive model framework; an approach akin to that of Lebreton et al (*Ecological Monographs*, **62**, 67-118).

1. Introduction

The likelihood function is one of the cornerstones of modern statistical theory because it provides an explicit link between the observable data and the unknown parameters thought to be important to the process that produced the data. It is then somewhat surprising that the likelihood function for open-population capture-recapture experiments is ill-defined. In particular, the modeling of the birth process in the usual Jolly-Seber model is not consistent with the reduced (death-only) model in which births are constrained to be absent and the formulations used to date lead to difficulties in imposing constraints upon the parameters of the birth process.

In this paper we propose a generalisation to the usual Jolly-Seber model that is analogous to that of Crosbie and Manly (1985). We show how this leads to simplifications in the numerical optimization of the likelihood and how constraints upon the parameters of the model can be easily imposed. We also show how this model can be generalised to more than one group of animals. Finally a numerical example is provided which fits a class of models akin to those proposed by LeBreton et al. (1992).

2. Notation

As is usual with Jolly-Seber models, "birth" refers to any mechanism by which new animals are added at unknown times to the catchable population (by immigration, recruitment, etc.). Similarly, "death" refers to all processes that permanently remove animals from the catchable population (emigration, death, inactivity, etc.). Births at known sample times (e.g. by deliberate addition of marked animals) are called injections and deaths at known sample times are called losses on capture.

Statistics:

n_i number of animals captured at sample time i, i=1, …, k.

      $n_i = m_i + u_i$.

m_i number of animals captured at sample time i that were previously marked.

u_i number of animals captured at sample time i that are unmarked.

l_i number of animals lost on capture at time i.

R_i number of animalsthat are released after the ith sample. R_i need not equal n_i if losses on capture or injections of new animals occur at sample time i.

r_i number of R_i animals released at sample time i that are recaptured at one or more future sample times.

z_i number of animals captured before time i, not captured at time i, and captured after time i.

Fundamental Parameters:

k   number of sample times.

p_i probability of capture at sample time i, i=1,…,k.

ϕ_i probability of an animal surviving between sample time i and sample time i+1 given it was alive at sample time i, i=1, …, k-1.

B_i number of animals that enter after sample time i and survive to sample time i+1, i=0, …, k-1. The B_i are referred to as the net births. B_0 is defined as the number of animals alive just prior to the first sample time.

N   total number of animals that enter the system and survive until the next sample time.

   $N = B_0 + B_1 + … + B_{k-1}$.

β_i fraction of the total net births that enter the system between sample times i and i+1, i=0, …, k-1. We refer to these as the entry probabilities. $\beta_i = B_i / N$.

v_i probability that an animals captured at time i will not be released, i=1, …, k.

Functions of parameters:

λ_i probability that an animal is seen again after sample time i, i=1, …, k.

   $\lambda_i = \phi_i p_{i+1} + \phi_i (1-p_{i+1}) \lambda_{i+1}$, i=1,…,k-1; $\lambda_k = 0$;

τ_i conditional probability that an animal is seen at sample time i given that it was seen at or after sample time i, i=1, …, k.

   $\tau_i = p_i / (p_i + (1-p_i)\lambda_i)$, i=1,…,k.

ψ_i probability that an animal enters the population and is not seen before time i, i=1,…, k-1.

   $\psi_1 = \beta_0$, $\psi_{i+1} = \psi_i(1-p_i)\phi_i + \beta_i$.

N_i population size at time i

   $N_1 = B_0$, $N_{i+1} = (N_i - n_i + R_i)\phi_i + B_i$

U_i number of unmarked animals in the population at time i.

        $U_1=0$; $U_{i+1}=U_i(1-p_i)\phi_i + B_i$

B_i^* gross number of animals that enter between sampling occasion i and i+1. These include animals that enter and die before the next sampling occasion.

3. Development

3.1 Development for a single group of animals

As shown by Darroch (1959), Jolly (1965), and Seber (1965), the likelihood function for a capture-recapture experiment can be partitioned into three components:

     L = $L_1 \times L_2 \times L_3$

       =   P(first capture | $\{p_i\}$, $\{\phi_i\}$, $\{B_i\}$) ×

             P(losses on capture | $\{v_i\}$) ×

             P(recapture | $\{p_i\}$, $\{\phi_i\}$).

The middle component is normally uninformative about the capture rates and survival rates and is usually modelledby a binomial distribution:

$$L_2 = \text{P(losses on capture)} = \prod_{i=1}^{k} P(l_i | \{n_i\}, \{v_i\}) = \prod_{i=1}^{k} \binom{n_i}{l_i}(1-v_i)^{l_i}(v_i)^{n_i-l_i}.$$

The third component can be factored into a series of conditionally independent binomial terms (Burnham, 1990):

$$L_3 = \text{P(recaptures)} = \prod_{i=1}^{k-1} P(r_i|R_i) \prod_{i=2}^{k-1} P(m_i|m_i+z_i)$$

$$= \prod_{i=1}^{k-1} \binom{R_i}{r_i}(\lambda_i)^{r_i}(1-\lambda_i)^{R_i-r_i} \prod_{i=2}^{k-1} \binom{m_i+z_i}{m_i}(\tau_i)^{m_i}(1-\tau_i)^{z_i}$$

The difficulty arises in modeling the first component. Darroch (1959) treated the $\{B_i\}$ as fixed constants, derived the generating function for the likelihood and noted that the actual likelihood was intractable because of the presence of up to (k-1) dimensional sums of probabilities, i.e.:

$$L_1 \;=\; \text{P(first captures)}$$

$$=\; \text{P}(u_1|B_0)\,\text{P}(u_2|B_0, B_1, u_1) \ldots \text{P}(u_k|B_0, B_1, \ldots, B_{k-1}, u_1, u_2, \ldots, u_{k-1})$$

$$=\; \binom{B_0}{u_1}(p_1)^{u_1}(1-p_1)^{B_0-u_1} \times$$

$$\sum_{d_1=0}^{B_0-u_1}\binom{B_0-u_1}{d_1}(1-\phi_1)^{d_1}(\phi_1)^{B_0-u_1-d_1}\binom{B_0-u_1-d_1+B_1}{u_2}(p_2)^{u_2}(1-p_2)^{B_0-u_1-d_1+B_1-u_2} \times$$

$$\vdots$$

where d_i is the number of animals that die between sampling times i and i+1. Darroch (1959) ruled out obtaining MLEs from this equation, but obtained estimates by the method of moments using the fact that the MLEs were equivalent to the moment estimators in the cases of no birthor of no death.

An alternate approach used by both Jolly (1965) and Seber (1965) and summarized by Seber (1982) assumes that the U_i are fixed parameters and B_i are *defined* as $B_i = U_{i+1} - \phi_i(U_i - u_i)$. The first component of the likelihood then collapses to:

$$L_1' = \prod_{i=1}^{k}\binom{U_i}{u_i}(p_i)^{u_i}(1-p_i)^{U_i-u_i} \tag{1}$$

with the simple solution $\hat{U}_i = u_i/\hat{p}_i$. Estimates for B_i are then obtained using the defining relationship between B_i and U_i.

There are several problems with this simplification. First birthsdo not explicitly appear in the likelihood and so it is difficult to impose constraints upon the B_i. For example, how are constraints that the births are known to be zero between certain sampling times imposed? Second, it often happens that estimates of B_i are negative. How can the likelihood be numerically maximized keeping all B_i non-negative? Third, the general Jolly-Seber model can be simplified by assuming no birth or no death or both over all sample times but the above likelihood does not reduce to the usual likelihood often used for these simpler cases.

Cormack (1989) modelled births in a log-linear context by defining his parameter ψ_i such that the number of unmarked individuals in the population at the time of the $(i+1)^{st}$ sample is ψ_i times the number of unmarked animals surviving from the ith sample, i.e., $U_{i+1} = U_i\,(1-p_i)\phi_i\psi_i = U_i(1-p_i)\phi_i + B_i$. Cormack (1989) found a direct correspondence between the GLIM parameters used to fit the log-linear models, and the model parameters

representing births, survival and capture rates. The advantage of his formulation is that out-of-range estimates (e.g. estimates of births < 0) are easily handled by constraining the ψ_i to be non-negative. [Note that Cormack's ψ_i is not the same parameter as our ψ_i used later in this paper.]

However Cormack (1989) also identified several disadvantages of his procedure. Because his ψ_i only indirectly estimates B_i, models with constraints upon the B_i are difficult to implement. For example, equal B_i across sampling intervals has no simple equivalence in terms of his ψ_i. Also it is difficult to derive estimates of the standard errors of the estimates (but see Cormack, 1993).

Burnham (1991) derived an expression for L_1 by organizing the first captures of new recruits in an upper triangular array (his Table 1… but see Table Supp.1 at end of this paper) with elements b_{ij}. representing the number of animals from B_i first captured in time j. Only the $u_1, …, u_k$ where $u_j = b_{\bullet j}$, are observable. By conditioning upon the total number of recoveries ($b_{i\bullet}$), he was able to show that

$$b_{0\bullet} \sim Binomial\left(B_0 ; p_1 + (1-p_1)\lambda_1\right)$$

$$u_i \sim Binomial\left(b_{0\bullet} + \cdots + b_{i-1,\bullet} - u_1 - \ldots - u_{i-1}; \tau_i\right) \quad i = 1,…,k-1$$

and consequently derived

$$L_1'' = \binom{B_0}{b_{0\bullet}} \left(p_1 + (1-p_1)\lambda_1\right)^{b_{0\bullet}} \left(1 - p_1 - (1-p_1)\lambda_1\right)^{B_0 - b_{0\bullet}} \times$$

$$\prod_{i=1}^{k-1} \binom{b_{0\bullet} + \cdots + b_{i-1,\bullet} - u_1 - \cdots - u_{i-1}}{u_i} (\tau_i)^{u_i} (1 - \tau_i)^{b_{0\bullet} + \cdots + b_{i-1,\bullet} - u_1 - \cdots - u_i}$$

The parameters of the model are now $B_0, b_{0\bullet}, …, b_{k-1,\bullet}$ and Burnham (1991) showed that this likelihood gave rise to the usual results.

This representation is not entirely satisfactory. First, the parameter set $B_0, b_{0\bullet}, …, b_{k-1,\bullet}$ is a mixture of parameters of interest (B_0) and unobservable random variables ($b_{0\bullet}, …, b_{k-1,\bullet}$). Second, the likelihood function involves the parameters in factorial terms which make numerical optimization difficult. Third Burnham (1991) states that $B_1 = … = B_{k-1} = 0$ is equivalent to $b_{1\bullet} = … = b_{k-1,\bullet} = 0$. However the converse does not hold - the $b_{i\bullet}$ are the number of recaptures of B_i which could be zero for non-zero $B_{i\bullet}$.

We propose a representation that is more natural than previous representations and which also has a number of advantages for numerical optimization of the likelihood. Following the development outlined in Crosbie and Manly (1985), we reparameterize $B_0,\ldots,B_{k-1}$ by N (the total number of unique animals available for capture) and $\beta_{0, \ldots, \beta_{k-1}}$, the fraction of N that enter between sampling occasion i and i+1 and survive to the next sampling occasion. In other words, we hypothesize a super-population of N animals that enter the experiment according to a multinomial distribution:

$$B_1, \ldots, B_{k-1} \sim Multinomial(\text{N}; \beta_1, \ldots, \beta_{k-1}). \quad (B_0 = N - \sum_{i=1}^{k-1} B_i)$$

It is now straightforward to show that

$$u_1, \ldots, u_k \sim Multinomial(\text{N}; \psi_1 p_1, \psi_2 p_2, \ldots, \psi_k p_k)$$

where $\psi_1 = \beta_0$ and $\psi_{i+1} = \psi_i(1-p_i)\phi_i + \beta_i$. Now the first term of the likelihood can be written as:

$$L_1''' = \binom{N}{u_1, u_2, \ldots, u_k}\left(1 - \sum_{i=1}^{k} \psi_i p_i\right)^{N-u_\bullet} \prod_{i=1}^{k}\left(\psi_i p_i\right)^{u_i}$$

where the parameters are N, $\beta_0, \ldots, \beta_{k-1}$ subject to $\beta_0 + \beta_1 + \ldots + \beta_{k-1} = 1$.

By conditioning upon the total number of unmarked animals observed ($u_\bullet$), the first component of the likelihood can be further factored as:

$$L_1''' = L_{1a}''' \times L_{1b}'''$$

$$= \binom{N}{u_\bullet}\left(\sum_{i=1}^{k} \psi_i p_i\right)^{u_\bullet}\left(1 - \sum_{i=1}^{k} \psi_i p_i\right)^{N-u_\bullet} \times \binom{u_\bullet}{u_1, u_2, \ldots, u_k}\prod_{i=1}^{k}\left(\frac{\psi_i p_i}{\sum_{i=1}^{k} \psi_i p_i}\right)^{u_i}$$

The full likelihood can now be written as

$$L = L_{1a}'''(N, \{\beta_i\}, \{p_i\}, \{\phi_i\}) \times L_{1b}''(\{\beta_i\}, \{p_i\}, \{\phi_i\}) \times L_2(\{v_i\}) \times L_3(\{p_i\}, \{\phi_i\}).$$

which can be expressed as the product of the following terms:

$P(u_{\cdot}|N)$	$\times$		$Binomial(N;\sum_{i=1}^{k}\psi_i p_i)$	$\times$
$P(\{u_i;i=1,\ldots,s\}|u_{\cdot})$	$\times$		$Multinomial(u_{\cdot};\{\psi_i p_i\big/\sum_{i=1}^{k}\psi_i p_i\,;i=1,\ldots,s\})$	$\times$
$P(R_1|n_1)$	$\times$		$Binomial(n_1;\upsilon_1)$	$\times$
$P(r_1|R_1)$	$\times$		$Binomial(R_1;\lambda_1)$	$\times$
$P(m_2|T_2)$	$\times$		$Binomial(T_2;\tau_2)$	$\times$
$P(R_2|n_1)$	$\times$	**OR**	$Binomial(n_2;\upsilon_2)$	$\times$
$P(r_2|R_2)$	$\times$		$Binomial(R_2;\lambda_2)$	$\times$
$P(m_3|T_3)$	$\times$		$Binomial(T_3;\tau_3)$	$\times$
$\vdots$			$\vdots$	
$P(m_{k-1}|T_{k-1})$	$\times$		$Binomial(T_{k-1};\tau_{k-1})$	$\times$
$P(R_{k-1}|n_{k-1})$	$\times$		$Binomial(n_{k-1};\upsilon_{k-1})$	$\times$
$P(r_{k-1}|R_{k-1})$			$Binomial(R_{k-1};\lambda_{k-1})$	

subject to the constraint that $\Sigma\beta_i=1$.

This formulation leads to the usual estimates of $\{\phi_i\}$ and $\{p_i\}$. Crosby and Manly (1985) showed that it also leads to the usual estimates for $\{B_i\}$ and $\{N_i\}$. Also of interest, particularly to fisheries managers, are the estimates of *gross births*. If one assumes a uniform entry of new animals in the sampling interval and a uniform mortality during the interval, then $\hat{B}_i^{*} = \hat{B}_i \dfrac{\log(\hat{\phi}_i)}{\hat{\phi}_i - 1}$ and the estimated total number of gross births is

$\hat{N}^{*} = \sum_{i=1}^{k=1}\hat{B}_i^{*}$ (Schwarz et al., 1993). The later term is of interest when estimating a salmon spawning population as it estimates the total escapement from the fishery that return to spawn.

This formulation also leads to the same asymptotic variances for $\{\hat{\phi}_i\}$ and $\{\hat{p}_i\}$ as given by Pollock et al (1990). The asymptotic variances of the $\hat{B}_i$ have an extra source of variation (because the B_i are now assumed to be random variables). Usually this is small relative to sampling error if B_i were fixed, and can be removed using results for conditional variation:

$$E\left[V\left(\hat{B}_i|B_i\right)\right]= V\left[\hat{B}_i|N\right]- V\left[E\left(\hat{B}_i|B_i\right)\right]$$

$$\approx V\left[\hat{B}_i|N\right]- V\left[B_i|N\right]$$

$$\approx V\left[\hat{B}_i|N\right]- \frac{B_i(N-B_i)}{N}$$

The case of no births is obtained by setting $\beta_0=1$, $\beta_1=0$, $\beta_2=0$, …, $\beta_{k-1}=0$. The likelihood then simplifies to that of the death-only model in a parallel fashion to that described by Burnham (1990, p.48-51). The same estimators for the death-only model from this formulation are obtained as shown in Pollock et al (1990, Section 5.3.1).

The case of no deaths is obtained by setting $\phi_1=\phi_2= … =\phi_{k-1}=1$. Again the likelihood simplifies to that of the birth-only model in a parallel fashion as described by Burnham (1990, pp. 54-58).

This formulation can also be used for more biologically interesting models whose estimates must be found numerically (see next section). Models where births are known to be zero between two sampling points correspond to setting the appropriate $\beta_i=0$. Similarly, models where the number of births are thought to be functions of external covariates can also be formulated in terms of the β_i. Neither of these models could be previously imposed upon the sampling experiment using (1).

Likelihood ratio and AIC criteria can be used for model selection. The number of parameters is found as the 3k minus the number of imposed constraints upon the model. The full model actually has 3k parameters, but three constraints need to be imposed: $\Sigma\beta_i=1$ and two constraints because of non-identifiability, namely $p_1=p_k=1$, leading to 3k-3 parameters that can be estimated. Schwarz et al (1993) give a list of the identifiable parameters. Problems usually arise with identifiability issues if {p} are allowed to vary over time. They can also arise in sparse data, e.g., if $z_i=0$, then some parameters are no longer identifiable.

3.2. Generalisations to More Than One Group

The above theory is readily generalised to more than one group of animals (e.g., males vs. females). Now the parameters have an additional subscript to represent the group membership, i.e., $\{p_{gi}\},\{\phi_{gi}\},\{\beta_{gi}\}$, and $\{N_g\}$.

Now an entire new class of models can be fit to open population experiments akin to those of LeBreton et al. (1992) who examined the case of modelling survival and capture rates across and within groups. A similar notation can be used to specify models, e.g., the model $\{p_g, \phi_t, \beta_{g*t}\}$ refers to a model where the capture probabilities vary among groups but are constant over time within each group; the survival rates vary over time intervals but are the same among groups; and the entry probabilites vary over both time and among groups. We can also fit models with survival equal per unit time, denoted by $\phi_{\Delta t}$ or $\phi_{g*\Delta t}$.

As with a single group, a difficulty with multiple groups determining the identifiable parameters and imposing suitable constraints. For example, the model $\{p_{g*t}, \phi_t, \beta_{g*t}\}$ would seem to have Gk+(k-1)+Gk parameters for the capture probabilities, survival probabilites, entry probabilities, and population size where G is the number of groups. However, the parameters ϕ_{k-1} and $p_{g,k-1}$ are not separately idenfiable; nor are the parameters $\beta_{g,0}$ and $p_{g,1}$. As before we need to impose arbitrary constraints, e.g., $p_{g,1}=1$ and $p_{g,k}=1$ for g=1,…,G. This identifiability issue is usually only a problem in models that include $\{p_{g*t}\}$ or $\{p_t\}$ terms or in the case of sparse data.

4. Numerical Implementation

This form of the likelihood suggests a very convenient way of obtaining estimates of the parameters. First, L_2 and L_3 are maximized to obtain estimates of the recapture and survival rates. Because L_3 can be written as the product of independent binomial distributions, the maximization can be done using any of the robust numerical algorithms for least squares (Green, 1984; Burnham, 1989). After the estimates are obtained, these are used to maximize $L_{1b}^{''}$ to obtain estimates of $\beta_0, …, \beta_{k-1}$. Because   is a multinomial, it too can be rewritten as a product of binomial distributions and the same numerical algorithms used. Lastly, the estimates from the previous two steps are used to estimate N, the $\{B_i\}$, the $\{N_i\}$ and other derived parameters.

Sanathanan (1977) showed that the conditional MLEs derived in this way are asymptotically equivalent to the unconditional MLEs. In fact, in the full Jolly-Seber model, the only recoverable information on capture and survival rates is found in the recoveries of previously marked animals, the unmarked animals contain information only

about births and the conditional MLEs are identical to the unconditional MLEs (see Appendix A). This is not true for simpler models where constraints are placed upon the parameters. In these cases, there is information about the capture and survival probabilities in the unmarked animals.

Consequently, we do not use the above approach, but rather maximize the product $L''' = L'''_{1b} \times L_2 \times L_3$ by successively maximizing over subspaces of the parameters until convergence is obtained. This is the method of cyclic fixing which is known to be globally convergent under general conditions (Jensen et al., 1991). We apply the method as follows we start with some initial estimates; then L''' is maximized with respect to $\{\hat{p}_i\}$ keeping $\{\hat{\phi}_i\}$ and $\{\hat{\beta}_i\}$ fixed; then maximized with respect to $\{\hat{\phi}_i\}$ keeping $\{\hat{p}_i\}$ and $\{\hat{\beta}_i\}$ fixed; and then maximized with respect to $\{\hat{\beta}_i\}$ keeping $\{\hat{p}_i\}$ and $\{\hat{\phi}_i\}$ fixed. This cycle is repeated until the change in the likelihood is small. Finally, we perform one further step where the likelihood is maximized with respect to all of the parameters. This procedure quickly converges because each cycle is a series of k-dimensional minimization problems rather than a 3k-dimensional problem. The final estimates of $\{\hat{p}_i\}$, $\{\hat{\phi}_i\}$ and $\{\hat{\beta}_i\}$ are used to estimate N using L'''_{1a} . Finally, the derived parameters $\{\hat{B}_i\}$, $\{\hat{N}_i\}$, etc., are found.

The joint variance-covariance matrix for all the parameters in the likelihood is constructed as outlined by Sananathan (1977). The delta-method is used to obtain estimates of the variances and covariances for functions of these parameters. The variances of $\hat{B}_i$ and other derived parameters can be corrected for the additional variability caused by the assumption that B_i are random variables as outlined earlier.

To keep all estimates within the parameter space, we parameterize the $\{p_i\}$, $\{\phi_i\}$, and $\{\beta_i\}$ in terms of their logits. Now values of $-\infty$ to $+\infty$ on the logit scale translate to the range [0,1] and it is quite impossible to obtain inadmissible estimates of capture, survival, or births.

All models must impose the constraint that $\Sigma\beta_i=1$. Other constraints may be imposed for biological realism (e.g., $\beta_i=0$ for some i), to overcome identifiability problems (e.g., $p_1=1$ and $p_k=1$ in the full model), or to increase precision with sparse data

(e.g., $p_1 = p_2 = \ldots = p_k$). Models with linear covariates for parameters are reformulated in terms of the fundamental parameters. For example, suppose we wish to fit a covariate model $p_i = \mathbf{x}_i' \gamma$ where $\mathbf{x_i}$ and γ are s×1 vectors. An s×s system of equations can be set up using the first s parameters:

$$
\begin{aligned}
p_1 &= \mathbf{x}_1' \gamma \\
p_2 &= \mathbf{x}_2' \gamma \\
&\vdots \\
p_s &= \mathbf{x}_s' \gamma
\end{aligned}
\quad \equiv \quad
\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_s \end{bmatrix} =
\begin{bmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_s \end{bmatrix} \gamma
\quad \equiv \quad
\mathbf{p_{1:s}} = \mathbf{X}\gamma
$$

Now the γ can be solved for in terms of the first s values of $\{p_i\}$, $\gamma = X^{-1}\mathbf{p_{1:s}}$. The first s member of $\{p_i\}$ can be let free to vary, and constraints upon the remaining k-s parameters are modeled using:

$$
\begin{aligned}
p_{s+1} &= \mathbf{x}_{s+1}' \gamma = \mathbf{x}_{s+1}' \mathbf{X}^{-1} \mathbf{p_{1:s}} \\
p_{s+2} &= \mathbf{x}_{s+2}' \gamma = \mathbf{x}_{s+2}' \mathbf{X}^{-1} \mathbf{p_{1:s}} \\
&\vdots
\end{aligned}
$$

At the end of the iterative process, estimates of $\{p_i\}$ satisfying the constraints are returned automatically along with their estimated variance-covariance matrix $\hat{\Sigma}_\mathbf{p}$. From this, estimates of γ are obtained as $\hat{\gamma} = X^{-1}\hat{\mathbf{p}}_{1:s}$ and the estimated variances of the γ can be obtained using a Taylor-series expansion as $\hat{\Sigma}_\gamma = X^{-1}\hat{\Sigma}_\mathbf{p}(X^{-1})'$.

Constraints are imposed in the numerical optimization by using the methods of Lagrange multipliers as outlined by Aitchison and Silvey (1958) or Henk Don (1985). For example, let

θ represent the parameters over which the likelihood is being maximized,

$\hat{\theta}_m$ be the estimates at the mth iteration,

$\mathbf{G(\theta)=0}$ be the set of constraints to be applied,

$\mathbf{g}(\theta) = \dfrac{\partial \mathbf{G}(\theta)}{\partial \theta}$ be the partials of the constraints with respect to the parameters,

$\mathbf{S(\theta)}$ be the score functions for the unrestricted model, and

$\mathbf{I(\theta)}$ be the information matrix for the unrestricted model.

Then the values of the estimates on the (m+1)[st] iteration are computed as:

$$\begin{bmatrix} \hat{\theta}_{m+1} \\ \bullet \end{bmatrix} = \begin{bmatrix} \hat{\theta}_{m} \\ \bullet \end{bmatrix} + \begin{bmatrix} I(\hat{\theta}) & -g(\hat{\theta}) \\ -g(\hat{\theta})^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} S(\hat{\theta}) \\ G(\hat{\theta}) \end{bmatrix}$$

where $\bullet$ indicates elements not of interest.

At the final iteration, this procedure will automatically return a covariance matrix for the estimates that accounts for the constraints. For example, if a parameter is constrained to a particular value, the estimated variance of the estimate and its covariances with other estimates will be zero. If two or more parameters are constrained to be equal, their variances will be identical and the sampling correlation will equal 1. An important consequence of this is that, when fitting a reduced model resulting from constraints on the full model, it is not necessary to re-write the fitting routines. Rather, one need only supply the appropriate constraints, $\mathbf{G}(\mathbf{\theta})$, and partials, $\mathbf{g}(\mathbf{\theta})$, and one general algorithm will produce the restricted model estimates along with their variances and covariances. One disadvantage is that it is impossible to distinguish between parameters that are constrained to be zero because there is real knowledge that the parameter is zero, or parameters that are constrained to be zero because the unconstrained estimates are negative. In both cases the estimated variance is zero, but clearly in the latter situation, there is still some imprecision in the estimate that has not been accounted for in the estimated variance.

The major advantage of this method of imposing constraints over the design matrix methods is that arbitrary non-linear constraints can be imposed just as easily as linear constraints. This is a particular advantage when time intervals between sampling occasions (Δ_i) are unequal when constraints such as equal survival per unit time ($\phi_i^{1/\Delta_i} = \phi_{i'}^{1/\Delta_{i'}}$) may be of interest. As well, covariate models on any scale (e.g. logit(p_i) = f(covariates)) can also be readily imposed. The estimates of the covariate coefficients (γ) and their variance covariance generalise to this case as follows: if T(p) is any differentiable, bijective mapping of the parameter p, then define $\mathbf{t_{1:s}}$ as the vector whose i^{th} element is T(p_i) evaluated at $\hat{p}_i$ and define $\hat{\Sigma}_\mathbf{T}$ as the s×s matrix whose i,jth element is $t_i' t_j' \sigma_{ij}$ where t_i' is the derivative of T(p_i) with respect to p_i evaluated at $\hat{p}_i$ and σ_{ij} is the i,jth element of $\hat{\Sigma}_\mathbf{p}$; then $\hat{\gamma} = X^{-1}\mathbf{t_{1:s}}$ and $\hat{\Sigma}_\gamma = X^{-1}\hat{\Sigma}_\mathbf{T}(X^{-1})'$.

The above theory is readily extended to multiple groups, and the same method of imposing constraints can be used to allow different groups to share common parameters, e.g., $\phi_{gi}=\phi_{g'i}$ for g'≠g but for all i specifies that the groups have common survival probabilities that vary over time.

5 Software

A stand-alone FORTRAN77 program that implements the above procedures for both single and multiple groups is available from the first author. Several common models are pre-programmed. A general interface for equality constraints over and above the pre-programmed models is also available. Even more complex models can be fit if the user re-codes the subroutines that return $G(\theta)$ and $g(\theta)$ values so they return values appropriate to the user's constrained model. The program is very general and powerful, but does not integrate data management with analysis.

Many users would find it difficult to program the required subroutines and are willing to sacrifice complete generality for ease of use. POPAN-3 (Arnason and Schwarz, 1987) is a comprehensive computer system for the management and analysis of capture-recapture experiments from open-populations. A new release, POPAN-4 (Arnason and Schwarz, 1995) includes a UFIT procedure that allows easy specification of most common constraints for single groups. This new release has the general abilities to do for the Jolly-Seber open population model what SURGE has done for the Cormack-Jolly-Seber release-recapture experiments where only recovery and survival probabilities are modeled. For a a more detailed description of POPAN-4 and for obtaining POPAN-4, please contact the second author. We expect to make versions available for Unix (SunOS and Linux) and the PC (Windows or OS/2) machines via anonymous FTP. POPAN-5, allowing for multiple groups, is currently under development and should be available by the fall of 1995.

6. Example

Schwarz et al. (1993) investigated the use of capture-recapture methods to estimate the number of salmon returning to spawn to the Chase River in British Columbia. Weekly electrofishing trips were made; unmarked fish were tagged with individually numbered tags; marked fish had their tag number recorded. The fish were stratified into adult males,

adult females, and precocious males (called "jacks") who have returned to spawn a year earlier than usual. Because of low sample sizes in weeks 1 and 9, these weeks were pooled with weeks 2 and 8 respectively. The summary statistics for the adult males and the jacks are shown in Table 1. Note that because weeks 1 and 2, and weeks 8 and 9 were pooled, the recorded time periods are 1.5 and 8.5 respectively.

6.1 Inadmissible Estimates

We first demonstrate how our methodology keeps all estimates within their respective parameter spaces. Table 2 compares the estimates formed using the closed form estimators (bias corrected) for the adult males as reported by POPAN or other programs (e.g. program Jolly, Pollock et al,. 1990 and those found using our methodology for the adult males and the jacks.

As noted earlier, not all parameters are identifiable in the full model. We have imposed the constraints that p_1 and p_{10} are equal to 1. Consequently, other estimates must be interpreted carefully because they will estimate the confounded parameters noted earlier.

Estimates of births from the closed form estimates that are negative, though unbiased, are certainly not useful. By constraining all estimates of births to be admissible, more defensible estimates are obtained and the precision is greatly increased - not only for the estimates of births but for population sizes as well. The total escapement (roughly the sum of the births) is about the same in both cases - note that the negative estimates of birth are used as is when totaling the births. Also note that the estimated standard error for estimates occurring at the boundary space is reported as zero even though there clearly is some uncertainly in the estimates. We suspect that profile methods would give a range of uncertainty in these cases. Schwarz et al. (1993) found that even though the estimated standard error is reported as zero in these cases, the average standard error over multiple simulations of this dataset gave good estimates of imprecision in the estimates.

6.2 Model selection

If the estimates for the adults and jacks are examined (Table 2), it appears that:

- the capture probabilites may be constant over time but differ among the groups;
- the survival probabilites may vary over time, but may be equal across the groups;
- the entry probabilites certainly change over time, but may be equal across groups.

Consequently, we fit a number of models searching for a parsimonious description of this experiment. The results are shown in Table 3. The notation used is similar to that employed by LeBreton et al. (1992), e.g., ϕ_{g*t} implies that survival probabilities are allowed to vary over groups and over time periods. Both AIC and likelihood ratio tests (not shown) lead to the same final model, namely, $\{p_g, \phi_t, \beta_{g*t}\}$. The final estimates are shown in Table 4.

7. Summary

By adopting a slightly more complicated model for a capture-recapture experiment (i.e, modelling births using a multinomial distribution from a super-population), it is now possible to subsume all simpler models in the Jolly-Seber framework into one model with appropriate constraints. This has the advantage that a single, general methodology can be used and that estimation and hypothesis testing can be performed using standard methods (i.e., maximum likelihood estimation and likelihood ratio tests). As well, biologists can now concentrate on fitting models that are biologically reasonable, rather than being constrained to fit a smaller class of models that have been "preprogrammed" into a computer package.

The modelling structure introduced by Lebreton et al. (1992) for mark-recovery experiments (where estimation of abundance is not done) has been quickly adopted by biologists. For example, the recent EURING94 conference held at the Patuxent Wildlife Research Centre in September 1994 had over 20 papers that used this methodology. The results in this paper permit application of the same sort of methods to mark-recapture experiments. As we showed in Section 6, it is now possible to fit a wide class of models under a general framework for model selection and testing.

REFERENCES

Aitchison, J. and Silvey, S.D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics* , **29**, 813-829.

Arnason, A.N. and Schwarz, C.J. (1987). POPAN-3: Extended analysis and testing features for POPAN-2. Box 272, St. Norbert, Manitoba. Charles Babbage Research Centre.

Arnason, A.N. and Schwarz, C.J. (1995). POPAN-4. Enhancements to a system for the analysis of mark-recapture data from open populations. *Journal of Applied Statistics*, **22**, 785-800.

Burnham, K.P. (1989). Numerical survival rate estimation for capture-recapture models using SAS Proc Nlin. Page 416-435 in Estimation and Analysis of Insect Populations. McDonald, L, Manly, B., Lockwood, J. and Logan, J. (eds). Springer-Verlag, New York, 492 pp.

Burnham, K.P. (1991). On a unified theory for release-resampling of animal populations. In Proceeding of 1990 Taipai Symposium in Statistics. M.T. Chao and P.E. Cheng (editors), p.11-36. Institute of Statistical Science, Academia Sinica.

Burnham, K. P. Anderson, D. R.,White, G. C., Brownie, C., and Pollock, K.H. (1987). Design and analysis methods for fish survival experiments based on release-recapture. *American Fisheries Society Monograph* **5**.

Cormack, R. (1993). Variance of mark-recapture estimates. *Biometrics*, **49**, 1188-1193.

Cormack, R. M. (1989). Log-linear models for capture-recapture. *Biometrics*, **45**, 395-413.

Crosbie, S.F. and Manly, B.F. (1985). Parsimonious modeling of capture-mark-recapture studies. *Biometrics* , **41**, 385-398.

Darroch, J. N. (1959). The multiple recapture census. II. Estimation when there is immigration or death. *Biometrika*, **46**, 336-359.

Green, P.J. (1984). Iteratively reweighted least squares for maximum likelihood estimation and some robust and resistant alternatives. *Journal or the Royal Statistical Association B*, **46**, 149-192.

Henk Don, F. J. (1985). The use of generalized inverses in restricted maximum likelihood. *Linear Algebra and its Applications*, **70**, 225-240.

Jensen, S. T., Johansen, S., and Lauritzen, S. L. (1991). Globally convergent algorithm for maximizing a likelihood function. *Biometrika*, **78**, 867-877.

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

Lebreton, J.-D., Burnham, K.P., Clobert, J., and Anderson, D.R. (1992). Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs*, **62**, 67-118.

Pollock, K. H. (1981). Capture-recapture models allowing for age-dependent survival and capture rates. *Biometrics*, **37**, 521-529.

Pollock, K. H., Nichols, J. D., Brownie, C., and Hines, J. E. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs **107**

Sanathanan, L.P. (1977). Estimating the size of a truncated sample. *Journal of the American Statistical Association*, **72**, 669-672.

Schwarz, C.J., Bailey, R.E., Irvine, J.E., and Dalziel, F.C. (1993). Estimating salmon spawning escapement using capture-recapture methods. *Canadian Journal of Fisheries and Aquatic Sciences*, **50**, 1181-1191.

Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika*, **52**, 249-259.

Seber, G. A. F. (982). The estimation of animal abundance and related parameters, 2nd edition. Griffin:London, 654 pp.

APPENDIX A

Showing that the condition MLE are equivalent to the unconditional MLE in the full Jolly-Seber Model

The likelihood can be written as (ignoring term involving v_i):

$$L_1''' = L_{1a}''' \times L_{1b}''' \times L_3$$

$$= \binom{N}{u_.} \left(\sum_{i=1}^{k} \psi_i p_i \right)^{u_.} \left(1 - \sum_{i=1}^{k} \psi_i p_i \right)^{N - u_.} \times \binom{u_.}{u_1, u_2, \ldots, u_k} \prod_{i=1}^{k} \left(\frac{\psi_i p_i}{\sum_{i=1}^{k} \psi_i p_i} \right)^{u_i} \times L_3(\{\phi_i\}, \{p_i\})$$

The MLE of integer valued parameters can be found by equating the first backwards difference of the log-likelihood to zero, i.e.

$$\frac{\Delta \log L}{\Delta N} = 0 \quad \Rightarrow \quad \frac{N}{N - u_.}(1 - \sum_{i=1}^{k} \psi_i p_i) = 1 \quad \Rightarrow \quad \hat{N} = \frac{u_.}{\sum_{i=1}^{k} \psi_i p_i} \tag{1}$$

Now let θ_i be one of the remaining parameters $\theta = \{\{\beta_i\}, \{\phi_i\}, \{p_i\}\}$. The score equation for θ_i is

$$\frac{\partial \log L}{\partial \theta_i} = 0 = \frac{u_.}{\sum_{i=1}^{k} \psi_i p_i} \frac{\partial \left(\sum_{i=1}^{k} \psi_i p_i \right)}{\partial \theta_i} - \frac{N - u_.}{1 - \sum_{i=1}^{k} \psi_i p_i)} \frac{\partial \left(\sum_{i=1}^{k} \psi_i p_i \right)}{\partial \theta_i} + \frac{\partial (L_{1b}''' L_3)}{\partial \theta_i} \tag{2}$$

Now if the solutions to (1) is substituted into (2), the first two term vanish and

$$\frac{\partial \log L}{\partial \theta_i} = 0 = \frac{\partial L_3}{\partial \theta_i}.$$

This implies that there is no information in the first term of the likelihood about any parameter other than N, and the first term can be ignored when obtaining MLEs of the remaining parameters.

In a similar fashion, the score function for $\{\psi_i\}$ is:

$$\frac{\partial \log L}{\partial \psi_i} = 0 = \frac{u_i}{\psi_i} - \frac{N - u_.}{1 - \sum_{i=1}^{k} \psi_i p_i} p_i \quad \Rightarrow \frac{u_i}{\psi_i p_i} - \frac{N - u_.}{1 - \sum_{i=1}^{k} \psi_i p_i} \tag{4}$$

Now let θ_i be an element of the remaining parameters $\theta = \{\{\phi_i\}, \{p_i\}\}$. The score function for θ_i is:

$$\frac{\partial \log L}{\partial \theta_i} = 0 = \sum_{i=1}^{k} \frac{u_i}{\psi_i p_i} \frac{\partial(\psi_i p_i)}{\partial \theta_i} - \frac{N - u_.}{1 - \sum_{i=1}^{k}(\psi_i p_i)} \frac{\partial\left(\sum_{i=1}^{k} \psi_i p_i\right)}{\partial \theta_i} + \frac{\partial L_3}{\partial \theta_i} \qquad (5)$$

When (4) is substituted into (5), the first two terms again vanish and

$$\frac{\partial \log L}{\partial \theta_i} = \frac{\partial L_3}{\partial \theta_i}.$$

This implies that the first two components of the likelihood contain no information about the capture and survival rates.

Intuitively, only recaptures of marked animals give information about the capture and survival rates; the relative proportions of unmarked animals captured gives information about the relative number of new recruits and the total number of unmarked animals give information about the "total" population size.

When constraints are applied to the model, the above development does not hold, and captures of unmarked animals does supply information about the capture and survival rates. In these cases, the conditional likelihood estimates will differ from the MLEs.

Table 1 (Supp)

Representation of first recaptures (b_{ij}) from recruits that arrive just before capture time i

(B_{i-1}). [Taken from Burnham, 1990].

Sample Time	New Recruits	1	2	3	…	k	Total Recaptures
1	B_0	b_{01}	b_{02}	b_{03}	…	b_{0k}	$b_{0\bullet}$
2	B_1		b_{12}	b_{13}	…	b_{1k}	$b_{1\bullet}$
3	B_2			b_{23}	…	b_{2k}	$b_{2\bullet}$
.	.					.	
.	.					.	
.	.					.	
k	B_{k-1}					$b_{k-1,k}$	$b_{k-1,\bullet}$
Total		u_1	u_2	u_3		u_k	

Table 2 Supp

Examples of the syntax of POPAN-4 in the UFIT paragraph for imposing constraints

Example of Constraint Type	POPAN-4 specification
Constant Constraints	
$\beta_2=0$, $\beta_4=0$, $\phi_3=1$	BPCONST=(P2-0)(P4-0); SPCONST=(P3-1); [a]
Equality Constraints	
$p_1=p_2=p_3$, $p_4=p_5=p_6$, $\phi_1=\phi_2=\phi_3$	CPCONST=(P1:2-P3)(P4:5-P6); SPCONST=(P1:2-P3); [b]
Mixed constraints	
$p_1=p_2=1$, $p_3=p_4$	CPCONST=(P1:P2-1)(P3-P4);
Equality Constraints per unit time	
$\phi_1^{1/\Delta_1} = \phi_2^{1/\Delta_2}$	SPCONST=(P1-P2); ADJUST=YES; [c]
Covariate constraints [d]	
$p_i=\gamma_0+\gamma_1 X_{1i}+\gamma_2 X_{2i}$	CPCONST=P-(C0,C1,C2);
$p_i=\gamma_0+\gamma_1 X_{1i}+\gamma_2 X_{1i}^2$	CPCONST=P-(C0,C1,C11) ; [e]
$\mathrm{logit}(p_i)=\gamma_0+\gamma_1 X_{1i}$	CPCONST=LOGITP-(C0,C1);
$\phi_1^{1/\Delta_1}=\gamma_0+\gamma_1 X_{1i}$	SPCONST=P-(C0,C1); ADJUST=YES;

[a] CPCONST (Capture Probability CONSTraint), SPCONST (Survival Probability CONSTraint), and BPCONST (Birth Proportion CONSTraint) may be used interchangeably as required.

[b] The phrase time1:time2 specifies a range of sample times.

[c] The keyword ADJUST=YES imposes constraints on a per-unit-time basis. BIRTHS=GROSS (not shown) imposes constraints on the gross births.

[d] Covariate constraints also involve specification of the covariate vectors of length k; the keyword C1=(X_{11}, X_{12}, ..., X_{1k}); is used to specify the covariate 1 values (X) for sample times 1 to k.

[e] Only 9 covariates (C1,…,C9 in addition to the constant C0) are allowed in POPAN-4; the notation C11 unambiguously implies C1^2

Table 1

Summary statistics for salmon escapement survey

Week	Number of captures	Number of marks	Released after marking	Subsequently recaptured	Seen before i, not at i, and after i
	n_i	m_i	R_i	r_i	z_i
Adult Males					
1.5	85	0	85	28	0
3	35	12	34	19	16
4	97	14	72	31	21
5	84	25	78	34	27
6	67	39	56	14	22
7	51	28	37	5	8
8.5	39	6	29	7	7
10	18	14	0	0	0
Jacks					
1.5	67	0	62	21	0
3	28	9	25	7	12
4	46	6	44	9	13
5	47	12	45	5	10
6	25	9	24	3	6
7	16	6	12	1	3
8.5	7	1	5	1	3
10	7	4	0	0	0

Table 2

Comparing the estimates from the closed form equations and from our methodology

Week	$\hat{P}_i$	$se(\hat{P}_i)$	$\hat{\phi}_i$	$se(\hat{\phi}_i)$	$\hat{\beta}_i$	$se(\hat{\beta}_i)$	$\hat{B}_i$	$se(\hat{B}_i|B_i)$	$\hat{N}_i$	$se(\hat{N}_i|N_i)$	$\hat{B}_i^*$	$se(\hat{B}_i^*)$
Adult estimates using closed form MLEs												
0												
1.5	ne	-	0.478	0.093			ne	-	ne	-	ne	-
3	0.295	0.086	1.002	0.196			317.2	121.3	118.5	29.9	316.8	102.4
4	0.223	0.064	0.720	0.117			-3.0	88.0	434.9	119.1	-3.5	100.0
5	0.288	0.063	0.908	0.202			-41.5	45.7	292.1	57.7	-43.5	46.9
6	0.307	0.075	0.606	0.244			33.4	26.1	218.2	48.5	42.4	29.1
7	0.321	0.126	0.364	0.184			174.8	108.9	158.8	59.5	277.8	145.9
8.5	0.171	0.092	0.241§	0.079			ne	-	227.5	117.3	ne	-
10	ne	-							ne	-		
Adult estimates (using our methodology)												
0					0.164¶	0.020	85.0¶	0.0			85.0¶	8.4
1.5	1.000#	0.000	0.478	0.093	0.151[a]	0.054	77.9[a]	26.4	85.0¶	0.0	110.1[a]	36.6
3	0.295	0.086	0.978	0.181	0.536	0.070	277.3	43.7	118.5	29.9	280.4	43.0
4	0.247	0.037	0.692	0.114	0.000	0.000	0.0	0.0	392.3	47.4	0.0	0.0
5	0.331	0.050	0.933	0.211	0.000	0.000	0.0	0.0	254.1	30.7	0.0	0.0
6	0.289	0.071	0.606	0.247	0.049	0.045	25.3	23.1	231.5	51.8	32.2	28.0
7	0.321	0.126	0.229	0.106	0.100	0.030	51.5	13.9	158.8	59.5	98.5	29.4
8.5	0.461	0.118	0.241§	0.079	0.000&	0.000	0.0&	0.0	84.6	19.2	0.0&	0.0
10	1.000#	0.000							18.0	0.6		
					$\hat{N}=$ 517.0$	38.2				$\hat{N}^*=$ 606.2$	47.7	

Table 2 (*continued*)

Week	$\hat{p}_i$	$se(\hat{p}_i)$	$\hat{\phi}_i$	$se(\hat{\phi}_i)$	$\hat{\beta}_i$	$se(\hat{\beta}_i)$	$\hat{B}_i$	$se(\hat{B}_i\|B_i)$	$\hat{N}_i$	$se(\hat{N}_i\|N_i)$	$\hat{B}_i^*$	$se(\hat{B}_i^*)$
					Jack estimates (using our methodology)							
0	1.000#	0.000			0.142¶	0.030	67.0¶	0.0			67.0¶	7.6
1.5	0.177	0.071	0.820	0.221	0.228[a]	0.099	107.3[a]	48.5	67.0¶	0.0	118.3[a]	47.5
3	0.108	0.028	1.000	0.000	0.576	0.125	270.9	82.4	158.1	57.3	270.9	83.1
4	0.121	0.054	0.919	0.418	0.000	0.000	0.0	0.0	426.0	93.2	0.0	0.0
5	0.147	0.091	0.437	0.288	0.000	0.000	0.0	0.0	389.6	165.4	0.0	0.0
6	0.143	0.146	0.583	0.612	0.029	0.090	13.7	42.5	169.6	100.3	17.7	52.6
7	0.189	0.162	0.237	0.289	0.024	0.035	11.4	16.0	112.0	111.3	21.5	29.2
8.5			0.200§	0.179	0.000&	0.000	0.0&	0.0	37.0	29.0	0.0&	0.0
10	1.000#	0.000							7.0	0.3		
					$\hat{N}=$ 470.3$	82.6					$\hat{N}^*=$ 495.4$	78.4

ne Not estimated by most program because of identifiability problems

\# Imposed restriction because of identifiability problems.

§ Confounded parameter. Estimates the product $\phi_{8.5}p_{10}$.

¶ Confounded parameter. Estimates the products $\beta_0 p_1$, or $N\beta_0 p_1 = n_1$

[a] Confounded parameter. Estimates $\beta_{1.5}+\beta_0(1-p_{1.5})\phi_{1.5}$, $N(\beta_{1.5}+\beta_0(1-p_{1.5})\phi_{1.5})$, and $N(\beta_{1.5}+\beta_0(1-p_{1.5})\phi_{1.5})\log(\phi_{1.5})/(1-\phi_{1.5})$

& Confounded parameter. Estimates $\beta_{8.5}/\phi_{8.5}$, $N\beta_{8.5}/\phi_{8.5}$, and $N\beta_{8.5}/\phi_{8.5}\log(\phi_{8.5}p_{10})/(1-\phi_{8.5}p_{10})$ respectively.

$ Confounded parameter. In the full model, this estimates the total number of identifiable net or gross births respectively.

Table 4

Estimates from model {p$_g$, φ$_t$, β$_{g*t}$}

Week	$\hat{p}_i$	$se(\hat{p}_i)$	$\hat{\phi}_i$	$se(\hat{\phi}_i)$	$\hat{\beta}_i$	$se(\hat{\beta}_i)$	$\hat{B}_i$	$se(\hat{B}_i|B_i)$	$\hat{N}_i$	$se(\hat{N}_i|N_i)$	$\hat{B}_i^*$	$se(\hat{B}_i^*)$
Adult estimates												
0					0.422	0.038	253.3	29.0			253.3	31.4
1.5	0.323	0.031	0.472	0.052	0.000	0.000	0.0	0.0	253.3	29.0	0.0	0.0
3	0.323	0.031	0.927	0.116	0.343	0.048	206.1	29.3	119.6	14.6	214.0	36.1
4	0.323	0.031	0.785	0.094	0.070	0.051	42.2	29.5	316.1	35.5	47.5	35.4
5	0.323	0.031	0.748	0.078	0.000	0.000	0.0	0.0	270.8	25.7	0.0	0.0
6	0.323	0.031	0.623	0.096	0.055	0.027	33.0	15.3	198.1	22.4	41.4	21.5
7	0.323	0.031	0.308	0.081	0.110	0.024	66.1	13.1	149.6	18.9	112.4	30.1
8.5	0.323	0.031	0.614	0.136	0.000	0.000	0.0	0.0	107.8	15.2	0.0	0.0
10	0.323	0.031							60.1	11.5		
					$\hat{N}=$ 600.7	40.4					$\hat{N}^*=$ 668.7	46.5
Jack estimates												
0					0.681	0.064	388.9	70.3			388.9	71.2
1.5	0.158	0.025	0.472	0.052	0.000	0.000	0.0	0.0	388.9	70.3	0.0	0.0
3	0.158	0.025	0.927	0.116	0.231	0.077	132.1	44.8	181.3	34.8	137.2	50.1
4	0.158	0.025	0.785	0.094	0.052	0.074	29.6	41.7	297.4	55.7	33.3	48.2
5	0.158	0.025	0.748	0.078	0.000	0.000	0.0	0.0	261.6	38.2	0.0	0.0
6	0.158	0.025	0.623	0.096	0.000	0.000	0.0	0.0	194.2	30.2	0.0	0.0
7	0.158	0.025	0.308	0.081	0.036	0.023	20.6	12.7	120.4	22.3	35.1	24.0
8.5	0.158	0.025	0.614	0.136	0.000	0.000	0.0	0.0	56.5	13.6	0.0	0.0
10	0.158	0.025							33.5	8.7		
					$\hat{N}=$ 571.3	75.5					$\hat{N}^*=$ 594.5	77.5

Table 3

Summary of models examined for escapement example.

Model			log-likelihood	Identifiable parameters	AIC
$\{p_{g*t}$	ϕ_{g*t}	$\beta_{g*t}\}$	-1020.8	42	2125.6
$\{p_t$	ϕ_{g*t}	$\beta_{g*t}\}$	-1027.0	36	2126.0
$\{p_g$	ϕ_{g*t}	$\beta_{g*t}\}$	-1027.1	32	2118.2
$\{p_g$	$\phi_{g*\Delta\tau}$	$\beta_{g*t}\}$	-1041.1	20	2122.2
$\{p_g$	$\phi_{\Delta t}$	$\beta_{g*t}\}$	-1041.2	19	2120.4
$\{p_{g*t}$	$\phi_{g*\Delta t}$	$\beta_{g*t}\}$	-1028.3	32	2120.6
$\{p_{g*t}$	ϕ_t	$\beta_{g*t}\}$	-1023.3	36	2118.6
$\{p_g$	ϕ_{g*t}	$\beta_t\}$	-1033.7	25	2117.4
$\{p_g$	ϕ_t	$\beta_t\}$	-1043.6	18	2123.2
$\{p_g$	ϕ_t	$\beta_{g*t}\}$	-1031.3	25	2112.6

The notation used is similar to LeBreton et al. (1992) along with the symbol Δt to indicate survival rates are equal per unit time.