LOGLINEAR MODELS FOR CAPTURE-RECAPTURE EXPERIMENTS ON OPEN POPULATIONS

R. M. Cormack

In The Mathematical Theory of the Dynamics of Biological Populations ed. R. W. Hiorns and D. Cooke. Academic Press, London, 1981.

Abstract

A general loglinear model is developed for capture-recapture experiments on open populations. The appropriate parameterization is that used by the GLIM computer package. Models appropriate to closed populations, death only, birth only, birth and death, all with or without trap dependence, and with or without constant effort are defined by specific subsets of the GLIM parameters. This allows easy exploration of capture-recapture data sets to identify the most parsimonious model.

BU-711-M in the Biometrics Unit Mimeo Series.

LOGLINEAR MODELS FOR CAPTURE-RECAPTURE EXPERIMENTS ON OPEN POPULATIONS

R. M. Cormack

Department of Statistics

University of St. Andrews

To appear in

The Mathematical Theory of the Dynamics of Biological Populations, ed. R. W. Hiorns and D. Cooke.
Academic Press, London 1981

1. Introduction

Capture-recapture is one of the methods by which the biologist attempts to obtain numerical estimates for the basic parameters - population size, birth and immigration, death and emigration - of an animal population in the wild. Repeated samples are taken from the population, some at least of the sampled individuals being given an identifiable mark and returned to the population. The record of recaptured marks and captures of unmarked individuals then provides a set of statistics from which, under certain assumptions, information on population parameters can be deduced.

Initially the population was assumed closed and subject to two samples, a marking sample and a recapture sample. If marked and unmarked animals are equally at risk to capture in the second sample, the simple Petersen estimate of population size is obtained by equating the unknown proportion of unmarked individuals caught in the second sample to the known proportion of marked individuals caught. More precise estimates are obtained from longer chains of samples, the Schnabel census, but by then the dangers of postulating a closed population are too great to be ignored. The longer the time between samples the more likely is the requirement of random mixing of the population to be met, but the less likely is the postulated closure of the population.

Early work attempted to correct for such dynamics, but this quickly changed to a positive effort to use the data from an extended capture-recapture experiment to estimate the parameters of the dynamic population. These culminated in the model of Jolly (1965) and Seber (1965) which yields estimates of birth parameters, death rates and population size of a population. Certain assumptions need to be satisfied, and these remain the bugbear of most practical studies.

There are three principal assumptions:

i) that individuals who have been captured retain on average the same parameters of survival and liability to capture as those who have not been captured;

ii) that all individuals are homogeneous in their behaviour as far as it affects their interaction with the sampling process;

iii) that no individuals emigrate from the population during one or more sampling periods and then return.

The third difficulty remains. Considerable progress has been made with the first two, trap dependence and behavioural heterogeneity, for the study of closed populations (Otis et al. 1978).

There are further technical difficulties with the Jolly-Seber model. A minor problem is that impossible estimates, negative birth or death rates, often occur. A major difficulty is the lack of parsimony in the form of the model which is usually applied. Estimates are readily available only if different survival parameters φ_i, different birth parameters B_i, and different catchability parameters p_i are postulated for each sampling period. Lack of parsimony increases the variability of parameter estimates. Again there has been recent progress, notably by Jolly (1979).

The conduct of such an experiment has been modelled by statistical distributions in many different ways and the history of these has been extensively documented (Cormack (1968, 1979); Seber (1973)). The subject provides an excellent example of the biological insight to be gained from different mathematical or statistical models of the same situation.

Many of these statistical methods, including the Jolly-Seber method, utilize only information which is provided by a batch mark; that is, one knows for each individual in the i^{th} sample only when that individual was last caught. If we are to try to model individual behaviour, we cannot expect to progress very far unless we know the complete capture history of each individual throughout the experiment. We assume henceforth that such records are available.

For such capture-recapture data we shall develop a sequence of models representing a closed population, birth, death, trap dependence, with variable or constant sampling effort, and show how the GLIM computer package can readily be used to select the model from among combinations of these factors most appropriate for the data set.

2. Loglinear Models for Contingency Tables

The modern approach to the analysis of data on counts expresses the logarithm of the expected value of every count as a linear function of a set of parameters. The general theory and methodology is described by Bishop, Fienberg and Holland (1975). Consider a 2 × 2 contingency table with

Observations		Probabilities		Expectations	
n_{11}	n_{12}	π_{11}	π_{12}	m_{11}	m_{12}
n_{21}	n_{22}	π_{21}	π_{22}	m_{21}	m_{22}

where $\Sigma\Sigma n_{ab} = N$, $\Sigma\Sigma \pi_{ab} = 1$, $m_{ab} = N\pi_{ab}$. We model the expectations as:

$$\log m_{11} = u + u_1 + u_2 + u_{12}$$

$$\log m_{12} = u + u_1 - u_2 - u_{12}$$

$$\log m_{21} = u - u_1 + u_2 - u_{12}$$

$$\log m_{22} = u - u_1 - u_2 + u_{12} \qquad (2.1)$$

If the four counts are from unrelated distributions, this is purely a reparameterization, the 4 m-parameters being replaced by the 4 u-parameters. Any hypothesis about the m_{ab} generates a corresponding hypothesis about the u. Thus, for example, the usual contingency table hypothesis of independence of the row and column categorizations: $\pi_{11}\pi_{22} = \pi_{12}\pi_{21}$ corresponds to the hypothesis: $u_{12} = 0$. The parameter u_{12} represents the interaction between the two categorizations.

Denote by $m_{abc\cdots s}$ the expected counts in a 2^s table, the suffices a,b,$\cdots$,s taking values 1 or 2 for the two alternatives in each of the s categorizations. In our application the s categorizations are the s samples, $m_{abc\cdots s}$ being the expected number of individuals with capture history abc$\cdots$s, a = 1 if these individuals were caught in the first sample, a = 2 if they were not caught. The complete, fully saturated, model reparameterizes log $(m_{abc\cdots s})$ as a sum or difference of:

an overall mean	u	
main effects	u_i	for the i^{th} sample
two-factor interactions	u_{ij}	between the i^{th} and j^{th} samples
three-factor interactions	u_{ijk}	etc.

We may write $\underset{\sim}{m}$, the 2^s-vector of $m_{abc\cdots s}$, in terms of $\underset{\sim}{u}$, the 2^s-vector of all the u's, as

$$\log \underset{\sim}{m} = A\underset{\sim}{u} \qquad (2.2)$$

where A is a $2^s \times 2^s$ orthogonal matrix with elements ± 1. The notation, and to a certain extent the interpretation, is the same as that for the analysis of variance of a 2^s factorial experiment. We adopt the convention of writing the elements of $\underset{\sim}{\ell} = \log \underset{\sim}{m}$ and of $\underset{\sim}{u}$ in the standard order of treatments and corresponding contrasts for the factorial experiment, as exemplified for the 3-sample experiment by:

$$\begin{bmatrix} \ell_{111} \\ \ell_{211} \\ \ell_{121} \\ \ell_{221} \\ \ell_{112} \\ \ell_{212} \\ \ell_{122} \\ \ell_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} u \\ u_1 \\ u_2 \\ u_{12} \\ u_3 \\ u_{13} \\ u_{23} \\ u_{123} \end{bmatrix} \qquad (2.3)$$

Fienberg (1972) presented the capture-recapture of a closed population in these terms. The observation $n_{22\cdot\cdot 2}$, the number of individuals not seen in any of the samples, is unknown, and it is the object of the study of a closed population to estimate this. For such an incomplete table a fully saturated model must have (2^s-1) parameters. We need to assume that (at least) the highest order interaction is identically zero to achieve identifiability for the other parameters. In the simple 2-sample case the samples must be assumed independent to permit estimation of N, just as, dually, knowledge of N in the contingency table permits

testing of the hypothesis of independence. In the s-sample case the usual model of independent samples is obtained by setting all the interactions $u_{ij}, u_{ijk}, \cdots$ to be zero. The advantage of the formulation (2.3) is in the exploration of models in which some of the interaction terms are non-zero. Interaction between sampling periods, in the form of trap-happiness or trap-shyness, is one phenomenon likely to invalidate seriously the classical closed-population estimates. Trap dependence lasting for one sampling period can be modelled by including in the model the neighbouring two-factor interaction terms $u_{i,i+1}$. This intuitively reasonable result can be justified rigorously by constructing appropriate matrices similar to those given in Sections 3, 4 and 5. Heterogeneity between individuals caused by different behaviour leads to an unpatterned set of interactions. Different models can be explored and the simplest acceptable model identified. This model is then extended to the missing cell of unseen animals, and an estimate $\hat{n}_{22..2}$ obtained from the chosen model.

The loglinear formulation provides a natural characterization of dependences between different samples in a way which other formulations of capture-recapture experiments do not provide. Distributionally the N individuals in the population are thought of as being multinomially allocated to the 2^s cells of different capture histories with probabilities $m_{ab..s}/N$. The (2^s-1) observed histories are also multinomial, index $(N-n_{22..2})$. The details of the sampling scheme in a capture-recapture experiment are usually not fixed in advance, since neither fixed sample size nor fixed effort are usually adhered to; but the multinomial should provide a good working distribution since statistically it can arise from the product of binomial distributions or of independent Poisson distributions conditioned on the observed sample size.

Different sets of non-zero interactions in the model (2.3) lead to different forms of estimator for the unknown population size. Some yield estimators in closed form, others require iterative solution. The details for hierarchical interaction models, in which an interaction can be non-zero only if all lower order interactions formed by a subset of the symbols are non-zero, are given in Bishop, Fienberg and Holland (1975).

3. Populations with Death

With an open population, trap dependence will still be manifest simply in the interactions between successive samples. The question is whether the natural dynamics of a population — birth, death and migration — result in any characteristic pattern of interaction. We can illustrate that they do by examining the detail of the 3-sample experiment with death occurring (Cormack, 1979). We use the standard notation (referring to these as the CR parameters):

φ_i : Probability (animal alive at (i+1) | alive at i)

p_i : Probability (animal caught at i | alive at i)

χ_i : Probability (animal not seen after i | alive at i)

noting that the set of χ_i is functionally dependent on the p_i, φ_i:

$$\chi_i = 1 - \varphi_i + \varphi_i(1-p_{i+1})\chi_{i+1} \quad . \tag{3.1}$$

The N animals in the population are multinomially distributed into the $8 = 2^3$ cells with expectations

$$
\begin{aligned}
m_{111} &= Np_1\varphi_1 p_2(1-\chi_2) \\
m_{211} &= N(1-p_1)\varphi_1 p_2(1-\chi_2) \\
m_{121} &= Np_1\varphi_1(1-p_2)(1-\chi_2) \\
m_{221} &= N(1-p_1)\varphi_1(1-p_2)(1-\chi_2) \\
m_{112} &= Np_1\varphi_1 p_2\chi_2 \\
m_{212} &= N(1-p_1)\varphi_1 p_2\chi_2 \\
m_{122} &= Np_1\chi_1 \\
m_{222} &= N(1-p_1)\chi_1 \quad .
\end{aligned} \tag{3.2}
$$

Since all expectations are products of parametric functions, a linear model exists for $\underset{\sim}{\ell} = \log \underset{\sim}{m}$ in terms of $\underset{\sim}{\beta}$, the vector of parameters:

$(\log N, \log p_1, \log (1-p_1), \log \varphi_1, \log p_2, \log (1-p_2), \log (1-\chi_2), \log \chi_2, \log \chi_1)'$

as $\underset{\sim}{\ell} = D\underset{\sim}{\beta}$ where D is a matrix with elements 0 or 1, readily obtainable from the expectations above. We can transfer from $\underset{\sim}{\ell}$ to the loglinear parameterization $\underset{\sim}{u}$, since $\underset{\sim}{\ell} = A\underset{\sim}{u}$ and $A^{-1} = \frac{1}{2^s} A$ imply

$$
\underset{\sim}{u} = \frac{1}{2^s} A\underset{\sim}{\ell} = \frac{1}{2^s} AD\underset{\sim}{\beta} = T\underset{\sim}{\beta} \; . \tag{3.3}
$$

The full relationship for a 3-sample experiment, which reveals the interactions which death imposes on the model, is:

$$\begin{bmatrix} u \\ u_1 \\ u_2 \\ u_{12} \\ u_3 \\ u_{13} \\ u_{23} \\ u_{123} \end{bmatrix} = \frac{1}{2^3} \begin{bmatrix} 8 & 4 & 4 & 6 & 4 & 2 & 4 & 2 & 2 \\ 0 & 4 & -4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 4 & -2 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 2 & 4 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & -2 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \log N \\ \log p_1 \\ \log (1-p_1) \\ \log \varphi_1 \\ \log p_2 \\ \log (1-p_2) \\ \log (1-\chi_2) \\ \log \chi_2 \\ \log \chi_1 \end{bmatrix} \qquad (3.4)$$

In this parameterization the only interaction induced by death is u_{23}. The interactions u_{12}, u_{13}, u_{123} remain identically zero. Five loglinear parameters correspond to five estimable CR parameters N, p_1, p_2, φ_1, χ_2. Unfortunately, with a longer chain of samples this correspondence fails. Only those interactions including the first sample are identically zero in this parameterization, and the number of loglinear parameters required in the model becomes greater than the number of CR parameters.

4. The GLIM Parameterization

GLIM is a program, developed by the Working Party on Statistical Computing of the Royal Statistical Society, which provides a framework for fitting generalized linear models to data. It is widely available throughout the world. The underlying theory is given by Nelder and Wedderburn (1972) and the operation is described in the NAG Manual (1978). The distributional assumptions for

contingency tables are that the counts in all cells have independent Poisson distributions whose means are exp $(\underset{\sim}{\beta}\cdot\underset{\sim}{x})$ for some known $\underset{\sim}{x}$ and unknown $\underset{\sim}{\beta}$, a formulation which accords perfectly with the capture-recapture models above.

For a 2^s factorial structure the GLIM parameterization differs from the log-linear vector $\underset{\sim}{u}$ displayed in (2.3). The representation for the 2^3 experiment is:

$$\begin{bmatrix} \ell_{111} \\ \ell_{211} \\ \ell_{121} \\ \ell_{221} \\ \ell_{112} \\ \ell_{212} \\ \ell_{122} \\ \ell_{222} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} GM \\ A \\ B \\ AB \\ C \\ AC \\ BC \\ ABC \end{bmatrix} \qquad (4.1)$$

of which the inverse is:

$$\begin{bmatrix} GM \\ A \\ B \\ AB \\ C \\ AC \\ BC \\ ABC \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \ell_{111} \\ \ell_{211} \\ \ell_{121} \\ \ell_{221} \\ \ell_{112} \\ \ell_{212} \\ \ell_{122} \\ \ell_{222} \end{bmatrix} . \qquad (4.2)$$

In this parameterization "main effects" and "interactions" are not evaluated from all observations, but only from those in which every other factor is at level 1. Since level 1 in our notation represents a set of animals which are seen in that sample and are therefore known, this form is particularly advantageous for modelling capture-recapture experiments. The unobserved $n_{22\cdots2}$ is the only observation whose expectation depends on the highest order interaction AB··S. Thus, analysis of the observed counts is unaffected by the value of AB··S and can be carried out even when, as in the case of birth and death, this interaction is known to be non-zero and non-estimable.

Multiplication of the appropriate matrices to obtain the vector of GLIM parameters in terms of the vector of CR parameters $\underset{\sim}{\beta}$ shows that the only interactions induced by the occurrence of death are those with consecutive factors ending in the final one: BCD···S, CD···S, ···, RS. The correspondence between GLIM interaction parameters and CR survival parameters is then 1 to 1 for any number of sample periods – (2s-1) parameters of each type. In the loglinear formulation (3.4) extended to a longer series of samples the extra constraints take the form of equality of sets of non-zero interactions. The form of the transforming matrix is also much simpler than in the loglinear parameterization, its elements all being ± 1 or 0 (see Table 5.2).

5. Birth and Death

The usual description of birth or immigration in CR studies is as an unknown, fixed or random, addition to the population prior to each sampling period. Thus Seber (1973) defines the unknown random variable B_i as the number of new animals joining the population in the interval from time t_i to time t_{i+1} which are still

alive and in the population at time t_{i+1}. To incorporate birth into the log-linear model we need to represent it by a multiplicative parameter. To this end we define parameters:

$$\psi_1 = 1 + B_1/N(1-p_1)\varphi_1$$

$$\psi_2 = 1 + B_2/N(1-p_1)\varphi_1\psi_1(1-p_2)\varphi_2 \qquad (5.1)$$

etc.

such that the population of unmarked individuals immediately before the $(i+1)^{th}$ sample is increased by a factor ψ_i. We may interpret $1/\psi_i$ as the probability that an unmarked individual alive at the i^{th} sample was also alive in the population at the preceding sample.

With this description of birth, and N remaining as the population size at the start of the experiment, the expectations of the numbers of animals having the 16 different capture histories in a 4-sample CR experiment are given in Table 5.1.

The parameter ξ_1 is a thoroughly untidy function of the φ, p and ψ, best defined by the recurrence

$$\xi_i = 1 - \varphi_i + \varphi_i\psi_i(1-p_{i+1})\xi_{i+1} \qquad (5.2)$$

with boundary condition $\xi_s = 1$. Since only m_{2222} (and analogously in longer experiments only $m_{22..2}$) depends on ξ_1, and since the GLIM model does not represent m_{2222}, it need not concern us further. The CR parameters appear in 16 different forms to represent the 15 observable classes. The transformation is shown in Table 5.2. This reveals that the GLIM interactions

Table 5.1. Expectations of numbers of animals with different capture histories in a 4-sample CR experiment with birth and death

$$m_{1111} = Np_1\varphi_1p_2\varphi_2p_3(1-\chi_3)$$

$$m_{2111} = N(1-p_1)\varphi_1\psi_1p_2\varphi_2p_3(1-\chi_3)$$

$$m_{1211} = Np_1\varphi_1(1-p_2)\varphi_2p_3(1-\chi_3)$$

$$m_{2211} = N(1-p_1)\varphi_1\psi_1(1-p_2)\varphi_2\psi_2p_3(1-\chi_3)$$

$$m_{1121} = Np_1\varphi_1p_2\varphi_2(1-p_3)(1-\chi_3)$$

$$m_{2121} = N(1-p_1)\varphi_1\psi_1p_2\varphi_2(1-p_3)(1-\chi_3)$$

$$m_{1221} = Np_1\varphi_1(1-p_2)\varphi_2(1-p_3)(1-\chi_3)$$

$$m_{2221} = N(1-p_1)\varphi_1\psi_1(1-p_2)\varphi_2\psi_2(1-p_3)(1-\chi_3)\psi_3$$

$$m_{1112} = Np_1\varphi_1p_2\varphi_2p_3\chi_3$$

$$m_{2112} = N(1-p_1)\varphi_1\psi_1p_2\varphi_2p_3\chi_3$$

$$m_{1212} = Np_1\varphi_1(1-p_2)\varphi_2p_3\chi_3$$

$$m_{2212} = N(1-p_1)\varphi_1\psi_1(1-p_2)\varphi_2\psi_2p_3\chi_3$$

$$m_{1122} = Np_1\varphi_1p_2\chi_2$$

$$m_{2122} = N(1-p_1)\varphi_1\psi_1p_2\chi_2$$

$$m_{1222} = Np_1\chi_1$$

$$m_{2222} = N(1-p_1)\xi_1$$

Table 5.2 Transformation $\underset{\sim}{g} = T\underset{\sim}{\beta}$ for 4-sample experiment with birth and death

$$
\begin{bmatrix} GM \\ A \\ B \\ AB \\ C \\ AC \\ BC \\ ABC \\ D \\ AD \\ BD \\ ABD \\ CD \\ ACD \\ BCD \end{bmatrix} =
\begin{bmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} \log N \\ \log p_1 \\ \log (1-p_1) \\ \log \varphi_1 \\ \log \chi_1 \\ \log \psi_1 \\ \log p_2 \\ \log (1-p_2) \\ \log \varphi_2 \\ \log \chi_2 \\ \log \psi_2 \\ \log p_3 \\ \log (1-p_3) \\ \log (1-\chi_3) \\ \log \chi_3 \\ \log \psi_3 \end{bmatrix}
$$

AC, BC, AD, BD, ABD and ACD

are identically zero. In general, birth induces, in a way exactly symmetrical to death, only those interactions with consecutive factors starting from the first: AB, ABC, ···, ABC··S. The remaining 9 effects stand in 1 to 1 correspondence with 9 estimable CR parametric functions as shown in Table 5.3.

Table 5.3. Equivalence of GLIM and CR parameters

GM	=	$\log [Np_1\varphi_1 p_2\varphi_2 p_3(1-\chi_3)]$
A	=	$\log [(1-p_1)\psi_1/p_1]$
B	=	$\log [(1-p_2)/p_2]$
AB	=	$\log \psi_2$
C	=	$\log [(1-p_3)/p_3]$
ABC	=	$\log \psi_3$
D	=	$\log [\chi_3/(1-\chi_3)]$
CD	=	$\log [\chi_2/\varphi_2(1-p_3)\chi_3]$
BCD	=	$\log [\chi_1/\varphi_1(1-p_2)\chi_2]$

If this model is selected as appropriate these equations, with estimators replacing parameters, yield the estimates of these parametric functions. Of course not all the basic parameters are expressible in terms of these functions: for example, p_1 and ψ_1 are not separately estimable. Nor is N, the initial population

size. But the population size N_2 at the time of the second sample can be expressed as $Np_1\varphi_1[1 + (1-p_1)\psi_1/p]$, which is estimable.

The structure of the estimating relationships is important for the choice of model. Models with death only and birth only are sub-models of that shown in Table 5.3. With no birth $\psi_i = 1$ for all i so that the interactions AB and ABC are identically zero, and the main effect A takes on the same form as the main effects B and C, representing the log odds against a live individual being captured in the corresponding sample. The death parameters look, but are not, slightly more complicated: $\varphi_i(1-p_{i+1})\chi_{i+1}/\chi_i$ is the conditional probability that an individual alive at i but not seen thereafter survives at least until (i+1) – the precise dual of the birth parameter $1/\psi_i$. With no death, $\varphi_i = 1$ and $(1-p_{i+1})\chi_{i+1}/\chi_i = 1$, so that the interactions CD and BCD are identically zero, and the main effect D becomes $\log [(1-p_4)/p_4]$.

Direct justification of the multinomial model used by Fienberg (1972) for the closed population is no longer possible, since the multiplicative birth parameters ψ_i are not probabilities (but are the inverses of probabilities). However, it seems a priori reasonable to postulate that each observed $n_{ab\cdot\cdot s}$ is an independent observation from a Poisson distribution with mean $m_{ab\cdot\cdot s}$. This distributional assumption can be tested by the analysis. Under this assumption the estimates derived above for the parametric functions are maximum likelihood estimates, and hence estimates of the basic CR parameters such as φ_i, p_i, N_i (where estimable) calculated from these functions are also maximum likelihood estimates. Details of the estimates will be published elsewhere, but it is worth noting here that, for an open population with different birth, death and capture probabilities for each sample period, the estimates obtained are the usual Jolly-Seber ones.

6. Choice of Model

When a capture-recapture study is to be carried out, the experimenter may attempt by careful attention to the detail of the design to restrict the type of change and behavioural interaction to which the population is subject during the study. He knows that, if trap dependence and heterogeneity can be eliminated but birth and death and/or migration are occurring, the Jolly-Seber estimates are valid. If he can restrict the study to a closed population, his estimate of population size will have a much smaller error than that obtained from an open population, but the estimate may be severely biased if the closure and other assumptions are unjustified. Within the assumption of closed populations and the hypergeometric sampling model a family of tests of different assumptions has been provided by Otis et al. (1978). Tests of closure have been given by Pollock et al. (1974). A general goodness-of-fit of the Jolly-Seber model is discussed in Seber (1973).

The loglinear model provides a unification of a number of previously disparate models, though by no means all previously considered models. The analysis of any model by the GLIM package yields a deviance, measuring the lack of fit of the model, which for Poisson variables has asymptotically a χ^2 distribution. Study of these deviances may allow identification of the best fitting, most parsimonious model. The importance of parsimony in reducing estimate variability in loglinear models is stressed by Bishop, Fienberg and Holland (1975).

The interaction patterns induced by birth and death were developed in the previous section. Constant sampling intensity will result in the equality of certain of the main effects. Trap dependence which lasts for one sampling period only will be reflected in the need for all interactions of the type R,R+1: it

should be noted, however, that AB is also induced by birth and S-1,S by death. Heterogeneity, in the form of varying behaviour of groups or of individuals with respect to the population dynamics or interaction with the sampling scheme, induces all interactions, and thus will yield an analysis in which no reduced (unsaturated) model is found to fit: typically when individuals have different capture probabilities all simpler models are found to have equally bad fits.

The complete range of patterns for a 4-sample experiment is given in Table 6.1. (Interpretation of the constant effort models accompanies Table 7.1.)

One further advantage of this analysis is its ability to force out-of-range estimates back within the limits of biological reality, at least when trap dependence is absent. Frequently a Jolly-Seber analysis yields survival estimates $\hat{\varphi}_i > 1$ or birth estimates $\hat{B}_i < 0$, a problem of analysis for which a solution has recently been proposed by Buckland (1980). Within the GLIM analysis such out-of-range estimates will show up as negative estimates for the corresponding birth or death interactions, since these represent the logarithms of the inverse of a conditional probability of birth or death. Thus, if trap dependence is not found to be necessary, no birth or death interaction terms should be negative in the final chosen model, and models with such interactions set to zero should be explored for the final choice.

7. The GLIM Program

If variables are declared as factors within GLIM, interactions must be hierarchical in the sense that if an interaction PQ is zero, all higher order interactions incuding PQ must be zero. Since birth and death induce patterns of interactions defying this rule the factor levels must be converted to vectors

Table 6.1. GLIM models for 4-sample experiment

Closed	A + B + C + D
Birth only	A + B + C + D + AB + ABC
Death only	A + B + C + D + CD + BCD
Birth and death	A + B + C + D + AB + CD + ABC + BCD
Trap dependence	A + B + C + D + AB + BC + CD
Birth with dependence	A + B + C + D + AB + BC + CD + ABC
Death with dependence	A + B + C + D + AB + BC + CD + BCD
Birth, death, dependence	A + B + C + D + AB + BC + CD + ABC + BCD
Constant effort: closed	GM
Constant effort with birth	As birth with B = C = D i.e., A + PB + AB + ABC
Constant effort with death	As death with A = B = C i.e., PD + D + CD + BCD
Constant effort with birth and death	As birth and death with B = C i.e., A + PBD + D + AB + CD + ABC + BCD

of independent regression variables. These vectors are the columns of $2^s \times 2^s$ matrices in the pattern of (4.1). The complete program allowing exploration of all the models described for a 4-sample experiment is given in Table 7.1. The three factors with labels starting with P are for models with constant effort: in GLIM one forces two parameters to be identical by constructing the sum of the corresponding explanatory variables and fitting that instead of the separate vectors. We read in a fictitious value for $n_{22..2}$ and then give it zero weight W in the analysis.

Table 7.1. GLIM program for analysis of 4-sample experiment

```
$UNITS 16                $DATA N              $READ
13 2 1 3 4 3 2 4 4 0 1 1 0 1 1 100
$CAL A=%GL(2,1)-1 : B=%GL(2,2)-1 :
     C=%GL(2,4)-1 : D=%GL(2,8)-1 :
    AB=A*B : BC=B*C : CD=C* D:
   ABC=AB*C : BCD=BC*D : W=1
   PBD=B+C : PB=PBD+D : PD=PBD+A
$EDIT 16 W 0  $YVAR N  $ERR P  $WEI W
$FIT   a selection of interesting models from Table 6.1
```

The data used in the above example were extracted from a long-term study of the fulmar petrel Fulmarus glacialis (see, for example, Dunnet and Ollason (1978), Cormack (1973)). They represent records from 1955-58 of all birds sighted during that period. Death is occurring, although the survival rate is high; "birth" occurs, since new marked birds were added to the population in each year; trap dependence will be limited since most birds are sighted without being caught or handled; heterogeneity is expected since the two sexes have different behaviour

patterns which affect their probabilities of being observed by the experimenters. This pattern is revealed by the deviances from selected models given in Table 7.2, none of which give an acceptable fit.

Table 7.2. Selected deviances from analysis of all birds

Closed	$\chi^2_{10} = 35.5$
Birth only	$\chi^2_8 = 20.5$
Birth and death	$\chi^2_6 = 13.3$
Birth, death and dependence	$\chi^2_5 = 11.8$

By contrast, when males alone were analyzed and the new birds added during these years excluded by restricting attention only to birds marked prior to 1955, so that death should be the only effective parameter, the deviances are given in Table 7.3.

Table 7.3. Selected deviances from analysis of males with no births

Closed	$\chi^2_{10} = 18.1$
Death only	$\chi^2_8 = 10.4$
Birth and death	$\chi^2_6 = 8.8$
Birth, death and dependence	$\chi^2_5 = 8.0$
Death with constant effort	$\chi^2_{10} = 12.4$

The final model is a parsimonious, well-fitting model which accords with the effort expended by the experimenter.

8. Conclusions

Loglinear models form a natural way of representing dependence between different sampling occasions of a capture-recapture study. Developed hitherto only for closed populations, they can be applied with the appropriate parameterization to dynamic populations, different patterns of interaction revealing different aspects of the population dynamics and the sampling behaviour. They are readily analyzed by GLIM, a widely available computer package, at least for a limited number of sampling periods. Inappropriate models can be immediately identified and evidence obtained for the most appropriate parsimonious model.

9. Acknowledgements

This work was completed during an extended visit to the Biometrics Unit of Cornell University, financed by a grant from NATO Scientific Affairs Division.

Discussions with colleagues C. D. Sinclair and I. F. West at St. Andrews and D. S. Robson at Cornell helped in the development of these models.

References

Baker, R. J. and Nelder, J. A. (1978). The GLIM System Manual – Release 3. Numerical Algorithms Group, Oxford.

Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975). Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, Massachusetts.

Buckland, S. T. (1980). A modified analysis of the Jolly-Seber capture-recapture model. Biometrics, 36, 000-000.

Cormack, R. M. (1968). The statistics of capture-recapture. Oceanography and Marine Biology Annual Review, 6, 455-506.

Cormack, R. M. (1973). Commonsense estimates from capture-recapture studies. In The Mathematical Theory of the Dynamics of Biological Populations, M. S. Bartlett and R. W. Hiorns, eds. Academic Press, London. 225-235.

Cormack, R. M. (1979). Models for capture-recapture. In Sampling Biological Populations, R. M. Cormack, G. P. Patil and D. S. Robson, eds. International Co-operative Publishing House, Fairland, Maryland. 217-255.

Dunnet, G. M. and Ollason, J. C. (1978). The estimation of survival rate in the fulmar Fulmarus glacialis. J. Anim. Ecol., 47, 507-520.

Fienberg, S. E. (1972). The multiple recapture census for closed populations and incomplete 2^k contingency tables. Biometrika, 59, 591-603.

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 models exemplified by a constant survival rate model. In Sampling Biological Populations, International Co-operative Publishing House, Fairland, Maryland. 277-282.

Nelder, J. A. and Wedderburn, R. W. M. Generalised linear models. J. Roy. Statist. Soc. A, 135, 370-384.

Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. Wildlife Monograph No. 62.

Pollock, K. H., Solomon, D. L. and Robson, D. S. (1974). Tests for mortality and recruitment in a K-sample tag-recapture experiment. Biometrics, 30, 77-87.

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. Griffin, London.