

ESTIMATION OF ABSOLUTE AND RELATIVE SURVIVAL RATES FROM THE RECOVERIES OF DEAD ANIMALS

BRYAN F. J. MANLY

Biometrics Unit, University of Otago, P.O. Box 56, Dunedin

SUMMARY : Circumstances that give rise to samples of dead animals from natural populations are considered and five important particular situations are emphasized. In two of these situations it is possible to estimate the absolute mortality rates of animals in the natural populations concerned. In the other three situations the populations comprise two or more different types of animal and only the relative mortality rates of these can be estimated.

The most obvious examples of the first two situations come from bird banding experiments. Models for such experiments are therefore briefly reviewed. A Poisson model for samples of dead animals from a population with an unknown initial size is proposed and is shown to produce survival rate estimates that can be readily calculated on the assumption that the survival rate per unit time becomes constant for older animals. This model is of value since the estimation does not require iterative computer calculations whereas other models making essentially the same assumptions do require these.

The third and fourth situations that have been considered concern large populations with relatively small numbers of deaths. The relative mortality rates of the different types of animals in the populations can be estimated by comparing the proportions of dead animals of the different types with the corresponding proportions of live animals.

The final situation discussed occurs when animals have associated with them values for certain characters X_1, X_2, \dots, X_p and the relative mortality rates of animals with different X values is to be determined by comparing the distribution of the X 's for live and dead animals.

INTRODUCTION

Data obtained from records of dead individuals can be used to estimate absolute and relative survival rates for populations of animals living under natural conditions. The purpose of this paper is to review experimental procedures that give rise to counts of dead animals and to discuss appropriate methods for analysing data. To begin with some examples will be considered in order to illustrate the range of situations that will be covered.

Probably the best known situation occurs with bird banding. Each year for a number of years a group of birds is banded and released. A proportion of the bands are recovered from dead birds and providing that the bands are dated it is possible to build up a record of the recoveries made from birds in successive years after they were banded. Table 1 shows the results obtained in this way by Fordham and Cormack (1970) for Dominican gulls on Somes Island, New Zealand. On occasions the number of birds banded in each year is not certain because some birds are banded without this being recorded. This is the case with the British Heron for which North and Morgan (1979) report the data shown in Table 2.

A rather different type of experiment was

described by Sheppard (1951). He was studying various aspects of predation of the snail *Capea nemoralis* by the song thrush and, in particular, whether the birds are selective in their choice of different colours of the snail. To test this he collected 1358 *C. nemoralis* snails from several locations, marked them, and scattered them near some thrush "anvils" in Ten Acre Copse, Wytham Woods, on 26 April 1950. It is known that thrushes break snail shells on the stone "anvils" so that they can eat the soft parts. Sheppard therefore collected the remains of broken marked shells at various times after 26 April and compared their colour distribution

TABLE 1. *Band returns from dead Dominican gulls on Somes Island, New Zealand, according to Fordham and Cormack (1970).*

Year of banding	Number of banded	Year after banding in which recovery was made					
		1	2	3	4	5	6
1961	574	16	10	10	6	7	5
1962	728	20	12	4	5	5	
1963	710	23	9	5	2		
1964	561	21	9	8			

TABLE 2. Band returns from British Herons given by North and Morgan (1979) with the original source being the British Trust for Ornithology. All the Herons were banded as nestlings.

Year of Banding	Year after banding in which recovery was made																				
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th	16th	17th	18th	19th	20th	21st
1955	31	5	0	0	3	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
1956	14	5	5	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1957	27	10	1	3	3	1	0	0	0	0	0	0	0	0	1	1	0	0	0		
1958	13	2	2	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0			
1959	35	22	7	6	1	2	1	2	0	0	0	0	0	0	0	0	0				
1960	5	6	5	0	2	0	0	0	0	1	0	0	0	0	0	0					
1961	22	5	2	1	1	0	0	0	0	0	0	0	0	0	0						
1962	7	0	0	0	0	0	0	0	0	0	0	0	0	0							
1963	3	2	1	2	0	0	0	0	0	0	1	0	0								
1964	7	3	2	0	0	0	0	0	1	0	0	0	0								
1965	12	5	1	0	0	1	0	0	0	3	1										
1966	7	9	4	0	1	0	1	0	0	0											
1967	31	9	5	4	1	2	2	0	0												
1968	35	11	2	0	4	0	2	2													
1969	58	16	6	0	0	1	1														
1970	40	17	8	6	3	2															
1971	30	17	4	6	3																
1972	24	14	4	1																	
1973	32	5	3																		
1974	21	5																			

with the known distribution in the marked population. The results of his collections are shown in Table 3.

While carrying out his experiments in Ten Acre Copse, Sheppard also collected broken shells from some thrush "anvils" in Marley Wood. In this case there was not a known population of marked shells. However samples from the population being predated were taken on two occasions during the collection period and these give an estimate of the population proportions of different colours of snail. The results of the Marley Wood experiment are shown in Table 4.

Another example of data based upon counting dead individuals comes from Wong and Ward's (1972) experiment on predation of *Daphnia publicaria* by yellow perch fry. Wong and Ward compared the size distribution of *D. publicaria* in the stomachs of perch fry with the distribution in the plankton in West Blue Lake, Manitoba, on five occasions over the period 1 July to 25 August 1969. Any differences between the stomach and plankton distributions is presumably due to size selective predation. The experimental results are shown in Table 5 for one of the five sampling times.

All of these example have one important common feature: the dead animals recorded are only a sample of all the animals that die. It is this feature that makes the analysis of the various sets of data

somewhat complicated. The estimation of survival rates and other population parameters is usually relatively straightforward if the actual numbers of living and dead animals are known at several points

TABLE 3. Samples of broken *Cepaea nemoralis* shells from thrush "anvils" at Ten Acre Copse. It is known that the broken shells came from a population consisting of 747 yellow snails and 611 pink and brown snails on 26 April 1950. Data from Sheppard (1951).

Date	Day Number	Broken Shells Collected	
		Pink and Brown	Yellow
28 April	2	0	2
1 May	5	0	1
2 "	6	1	1
5 "	9	1	3
8 "	12	3	9
11 "	15	0	1
12 "	16	7	4
16 "	20	0	1
17 "	21	0	1
20 "	24	1	0
22 "	26	1	1
30 "	34	0	2
3 June	38	2	1
5 "	40	3	0
		19	27

TABLE 4. *Samples of broken Capaea nemoralis shells from thrush "anvils" in Marley Wood. The anvils were cleared of broken shells on 6 April 1950. Random samples taken from the population being predated contained 397 pink and brown shells and 137 yellow shells.*

TABLE 5. *Distribution of the lengths of Daphnia publicaria in plankton samples and in the stomachs of perch fry at West Blue Lake on 1 July 1969. This table was constructed from Figure 1 of Wong and Ward (1972) and might therefore contain some small errors.*

Length (mm)	Number in length class	
	Plankton	Stomach
0.5-0.7	20	59
0.7	22	84
0.9	20	154
1.1	18	138
1.3	26	44
1.5	24	10
1.7	22	5
1.9	24	0
2.1	26	0
2.3	16	0
2.5	11	0
2.7	7	0
2.9-3.1	1	0
Total	237	494
Mean length (mm)	1.56	1.03
Standard deviation	0.59	0.25

in time. Only experiments involving samples of dead animals are considered in this paper.

There is perhaps a need to specifically mention one situation that is not going to be considered. This occurs when bird band data are obtained from birds that are shot by hunters. The data then have exactly the same form as the examples shown in Tables 1 and 2. However, shooting clearly does not give a random sample of birds that die from all

causes. Therefore a statistical model for shooting returns will have different parameters from a model for returns from dead birds in general. Actually, it turns out that both these situations can be considered as special cases of the multi-sample single recapture census (Sebef, 1980). However this is an aspect of modelling that will not be discussed in the present paper,

The examples suggest that there are essentially five different types of model for counts of dead animals that need to be covered, as follows:

- (a) There is a population with a known size at time to which decreases through deaths at an appreciable rate from then on. The data available for analysis are counts d_1, d_2, \dots, d_s of dead animals, where d_i is an unknown proportion of the animals that died between time t_{i-1} and time t_i . The band returns from one year of banding of Dominican gulls (Table 1) is an example of this type of situation. The problem is to estimate survival rates and also the probability of recovering a dead animal.
- (b) This is like (a), except that the initial population size is unknown. The band returns from one year of banding of British Herons (Table 2) is an example of this type of situation. The problem is to estimate survival rates.
- (c) The population from which the deaths come is very large in comparison with the number of deaths during the experimental period. This means that the counts of dead animals do not change appreciably due to a reduction in the population size. The object of analysing the data is not to estimate the absolute survival rate, which is very high. Rather, the population consists of $K(>2)$ different types of animal and it is desired to estimate their relative mortality rates. To do this the proportions of different types of dead animals can be compared with the proportions in the population, which are known exactly. Sheppard's (1951) Ten Acre Copse experiment yielded this type of data (Table 3).
- (d) This is the same as (c) except that the proportions of different types of animal in the population are only known from random sampling of live individuals. Interest still centres on the estimation of relative survival rates. Sheppard's (1951) Marley Wood experiment yielded this type of data (Table 4).
- (e) The population from which the deaths come is very large compared to the number of deaths. Each individual in the population is characterised by the values that it possesses for the

variables X_1, X_2, \dots, X_p . The multivariate distribution of the X 's in the population is estimated from a random sample and this can be compared with the distribution in a random sample of dead individuals. The problem is to see how the survival varies for individuals with different values for the X 's. Wong and Ward's (1972) data on predation of *Daphnia publicaria* (Table 5) give an example of this type of situation with one X variable only (length).

In all these situations permanent emigration is taken into account as being equivalent to "death". However, none of the models that are discussed in this paper are appropriate when there is an appreciable amount of temporary emigration.

Situations (a) to (e) will now be considered in turn.

A REDUCING POPULATION WITH A KNOWN INITIAL SIZE

Suppose that at time t_0 there are A animals in a population, and that the probability of surviving the period from time t_{i-1} to time t_i is ϕ_i for all animals still alive at time t_{i-1} . Then the probability of an animal surviving until time t_{j-1} , dying in the interval t_{j-1} to t_j and its death being recorded, is

$$\alpha_j = \phi_1 \phi_2 \dots \phi_j (1 - \phi_j) \gamma_j \tag{1}$$

where γ_j is the probability of the death being recorded. It then follows that the probability of recording d_1 dead animals for the time period from t_0 to t_1 , d_2 dead animals for the time period from t_1 to t_2 , . . . , d_s dead animals for the time period from t_{s-1} to t_s is given by the multinomial distribution as

is the probability of a particular animal's death being recorded.

This multinomial distribution is the basis for many of the models that have been proposed for bird banding data with A denoting the number of birds banded in a particular year. There is then a

separate probability of the form of equation (2) for each year of banding and the probability of the full set of data (the likelihood function) is obtained by multiplying the yearly probabilities together. Maximum likelihood estimates of the various parameters of the model are then equal to the values that make this likelihood function as large as possible. (In general maximum likelihood estimates of parameters have good properties relative to ad-hoc estimates calculated in other ways. That is why it is desirable to find maximum likelihood estimates of the parameters of models.)

Explicit formulae for the maximum likelihood estimation of survival and recording probabilities are only available for a few specific situations with bird band data. Seber (1970) gives such formulae for the case when ϕ_j and γ_j are time-specific so that the probability of survival and the probability of recording a death are the same for all birds in any particular calendar period, irrespective of when they were banded. Seber (1971) has also found explicit estimates of survival probabilities on the assumption that these are age-specific and that the recovery probability remains constant over time. These survival estimates are maximum likelihood estimates conditional on the total number of bands recovered, R . Unfortunately they do not have good properties for observations taken over a long period of time and Seber suggests that in practice it may be better to follow Fordham and Cormack (1970) and assume that the survival probability becomes constant for birds over a certain age. Explicit estimates are then no longer available from data involving more than one year of banding.

The conditional maximum likelihood method used by Seber (1971) to obtain age-specific survival probabilities was first used by Haldane (1953, 1955) for the case where survival and recovery probabilities are constant over time.

Brownie *et al.* (1978) have produced a comprehensive handbook on the analysis of bird banding data, including details of computer programs that they have written for carrying out calculations. Allowance is made for animals of different ages to have different survival rates and also for survival and recovery rates to vary from year to year. Brownie *et al.* were mainly concerned with the analysis of band recoveries from birds shot by hunters. However, their computer programs can be used to analyse data coming from natural deaths.

A REDUCING POPULATION WITH UNKNOWN INITIAL SIZE

Consider now the same situation as in the previous

section, except that the initial population size is not known.

The simplest approach to adopt in this case involves assuming that the recording probabilities $\gamma_1, \gamma_2, \dots, \gamma_s$ of equation (1) are all equal, and noting that the probability of recording the numbers d_1, d_2, \dots, d_s of dead animals, conditional on the total number of them being R , is the multinomial form

$$P(d_1, d_2, \dots, d_s/R) = \frac{R!}{s} \beta_1^{d_1} \beta_2^{d_2} \dots \beta_s^{d_s} \xi^{-R} \prod_{j=1}^s d_j! \quad (3)$$

where

$$\beta_j = \phi_1 \phi_2 \dots \phi_{j-1} (1 - \phi_j),$$

and

$$\xi = \sum_{j=1}^s \beta_j$$

This probability does not involve the unknown total number of animals in the population at time t_0 , or the recording probability.

Haldane's (1953, 1955) models for bird band experiments were of this form for each year of banding. He multiplied the probabilities for the different years together to get the full likelihood function for all the data and maximised this with all the survival probabilities set equal to ϕ . This produces an equation for the conditional maximum likelihood estimate, $\hat{\phi}$, of a constant survival probability. Seber (1971) used the same approach for estimating age-specific survival rates. North and Morgan (1979) did the same when analysing their Heron data (Table 2) but they assumed that first year survival rates varied from year to year whilst the survival rate was constant from year to year for older birds. An interesting aspect of North and Morgan's work is the way that they have been able to relate first year survival to the severity of the winter.

One problem with bird banding models that assume constant survival for older birds is the lack of explicit formulae for maximum likelihood survival estimates. The result is that usually the estimation equations need to be solved numerically, possibly using an electronic computer. It therefore seems worth noting the fact that estimates based upon the assumption of a constant survival rate after an initial period of varying survival can easily be calculated for a single cohort of animals that reduces over time. In the case of bird-banding experiments the cohort would be the birds banded in one particular year.

To see this, suppose that the number of animals in a population at time $t_0 = 0$ is unknown but large relative to the total number of dead animals recovered at different times (say ten times as large, or more). Also suppose that samples of dead animals all relate to a

unit period of time (such as a year) so that d_j is the number of animals recorded as dying between time $t_{j-1} = j-1$ and time $t_j = j$. Then d_1, d_2, \dots, d_s will approximately be independent Poisson variates with the mean value of d_j given by

$$E_j = \phi_1 \phi_2 \dots \phi_{j-1} (1 - \phi_j) \gamma_j A,$$

or

$$E_j = \phi_1 \phi_2 \dots \phi_{j-1} (1 - \phi_j) B_j,$$

where $B_j = \gamma_j A$. The likelihood for the observed values will then be

$$P(d_1, d_2, \dots, d_s) = \prod_{j=1}^s \exp(-E_j) E_j^{d_j} d_j! \quad (4)$$

This type of Poisson model has also been used by Robson (1963) and Jolly (1979).

The particular situation that is of interest occurs when $\phi_{r+1} = \phi_{r+2} = \dots = \phi_s$ for some value r , with $B_j = B$ for all j . With $r = 2$ this corresponds to North and Morgan's (1979) assumptions.

If the usual procedure for obtaining maximum likelihood estimates of $\phi_0, \phi_2, \dots, \phi_r, \phi$ and B is followed then the estimates $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_r, \phi$ and B are found to satisfy the equations

$$T_r/R_r - 1/(1 - \hat{\phi}) + (s-r) \hat{\phi}^{s-r}/(1 - \hat{\phi}^{s-r}) = 0, \quad (5)$$

$$\hat{B} = R_0 + R_r \hat{\phi}^{s-r}/(1 - \hat{\phi}^{s-r}), \quad (6)$$

and

$$\hat{\phi}_i = (\hat{B} - R_0 + R_i) / (\hat{B} - R_0 + R_{i-1}), \quad (7)$$

$i = 1, 2, \dots, r$. Here

$$R_i = \sum_{j=i+1}^s d_j \text{ and } T_i = \sum_{j=i+1}^s (j-i-1)d_j \quad (8)$$

(See the Appendix 1 for more details of how these equations are obtained and also the derivation of equations (12) to (18) below).

If $s-r$ is large so that $\phi^{s-r} \approx 0$ then equation (5) reduces to

$$\phi = 1 - R_r/T_r \quad (9)$$

and it also follows that

$$B = R_0, \quad (10)$$

and

$$\hat{\phi}_i = R_i/R_{i-1} \quad (11)$$

If ϕ^{s-r} is not near zero then the solution for equation (5) can be read easily from Figure 1 and estimates of B and ϕ_i follow directly from equations (6) and (7). With $r=0$ equation (9) gives Lack's estimate of survival (Seber, 1973, page 247). Equation (5) can also be solved using a table given by Robson and Chapman (1961) and reproduced by Seber (1973, Appendix A6).

The variances and covariances of the estimators are approximately given by the following equations;

$$\text{Var}(\hat{B}) \approx B/(1 - \phi_1 \phi_2 \dots \phi^{s-r}) \quad (12)$$

covariance formulae may be rather inaccurate (see Appendix 1). Note that if $\phi_{s-r} \approx 0$, then covariances are all approximately zero.

Example

As an example of the above equations, consider the data on British Herons (Table 2). Table 6 shows the estimate obtained from these data, together with estimated standard errors, on the assumption that the survival rate became constant after two years of life. This is the assumption made by North and Morgan (1979). Actually, likelihood ratio goodness of fit tests (Brownie *et al.* 1978, p. 203) indicate that for birds banded in most years it is sufficient to assume a constant survival rate for birds over one year of age.

Equation (14) indicates that $\text{Var}(\hat{\phi}_i)$ is proportional to $1/(\phi_1\phi_2 \dots \phi_{i-1}B)$, which is approximately proportional to $1/R_{i-1}$. It is therefore appropriate to estimate the mean value of ϕ_i using a weighted mean of the estimate $\hat{\phi}_i$, weighted by R_{i-1} (Seber, 1973, p.6). In

FIGURE 1. The solution of the equation $T/R - l/(1 - \hat{\phi}) + n\hat{\phi}N/(1 - \hat{\phi}N) = 0$ can be read from this figure, for given values of R/T and N . For example if $R/T = 0.5$ and $N = 4$ then $\hat{\phi} \approx 0.67$. The solution of equation (5) of this paper can be found by putting $T/R = T_r/R_r$ and $N = s-r$.

TABLE 6. Survival probabilities estimated from the British Heron data shown in Table 2.

Year of Banding	$\hat{\phi}_1$	Std. err.	$\hat{\phi}_2$	Std. err.	$\hat{\phi}$	Std. err.
1955	0.26	0.00	0.55	0.15	0.79	0.08
1956	0.52	0.09	0.67	0.12	0.52	0.11
1957	0.43	0.07	0.50	0.11	0.79	0.06
1958	0.35	0.11	0.71	0.17	0.58	0.14
1959	0.54	0.06	0.46	0.08	0.60	0.07
1960	0.74	0.10	0.57	0.13	0.58	0.11
1961	0.29	0.08	0.44	0.17	0.43	0.19
1962	0.00	-	-	-	-	-
1963	0.67	0.16	0.67	0.19	0.71	0.08
1964	0.46	0.14	0.50	0.21	0.67	0.19
1965	0.56	0.14	-	-	0.88	0.04
1966	0.68	0.10	0.40	0.13	0.51	0.15
1967	0.43	0.07	0.62	0.10	0.63	0.09
1968	0.40	0.07	-	-	0.70	0.07
1969	0.30	0.05	0.34	0.10	0.52	0.14
1970	0.50	0.06	0.57	0.08	0.63	0.08
1971	0.67	0.05	0.72	0.06	0.89	0.02
1972	0.45	0.08	0.28	0.10	0.25	0.20
1973	0.29	0.08	0.61	0.11	-	-
1974	0.37	0.09	-	-	-	-

Notes: Equation (5) gives $\hat{\phi}_2 > 1$ for birds banded in 1965 and 1968. Only $\hat{\phi}_1$ and $\hat{\phi}$ have been separately estimated in these years. Lack of data precludes the estimation of some of the parameters for birds banded in 1962, 1973 and 1974. Standard errors (Std. err.) were calculated using equations (13) and (14) with unknown parameter values replaced by their estimates.

The terms multiplied by the factor ϕ_{s-r} represent corrections to take into account the fact that sampling has not continued until all animals are dead (i.e., until $\phi_{s-r} \approx 0$). These terms should be small compared to the other terms. If they are not then these variance and

particular the weighted mean of the $\hat{\phi}_2$ estimate in Table 6 is

$$\hat{\phi}_2 = \Sigma R_1 \hat{\phi}_2 / \Sigma R_1 \approx 0.527,$$

with standard error

$$SE(\hat{\phi}_2) = \{\Sigma R_1^2 \text{Var}(\hat{\phi}_2) / (\Sigma R_1)^2\}^{1/2} \approx 0.028.$$

This agrees closely with North and Morgan's full maximum likelihood estimate of the mean ϕ_2 value which is 0.532 with standard error 0.029.

Likewise equation (13) indicates that $\text{Var}(\hat{\phi})$ is proportional to $1/(\phi_1 \phi_2 \dots \phi_r B)$, which is approximately proportional to $1/R_r$. An appropriate estimate of the mean ϕ value for the Herons is therefore

$$\hat{\phi} = \Sigma R_2 \hat{\phi} / \Sigma R_2 = 0.654,$$

with standard error

$$SE(\hat{\phi}) = \{\Sigma R_2^2 \text{Var}(\hat{\phi}) / (\Sigma R_2)^2\}^{1/2} \approx 0.028,$$

which agrees moderately well with North and Morgan's estimate of 0.697 with standard error 0.027.

A SMALL NUMBER OF DEATHS WITH SEVERAL TYPES OF ANIMAL IN KNOWN PROPORTIONS

The situation that will be considered in this section occurs when the population from which dead animals come is very large in comparison with the number of deaths that occur. Therefore the counts of dead animals do not decline appreciably over time because of a reduction in the population size. This means that the methods of analysis that have been reviewed above cannot be applied to the data. Indeed, the main point in analysing data must be to estimate the relative mortality rates of the different types of animal in the population since it is not possible to estimate the absolute rates.

It will be assumed that the dead animals come from a large population comprising K distinct types with relative frequencies A_1, A_2, \dots, A_K . It will also be assumed that the data available consists of the counts of dead animals of the various types for s time periods, with d_{ij} being the number of type i dead animals observed for the time period t_{j-1} to t_j ($i = 1, 2, \dots, K$; $j = 1, 2, \dots, s$).

If the probability of survival from time t_{j-1} to time t_j for a type i animal alive at time t_{j-1} is

$$\phi_{ij} = \exp\{-\lambda_{ij}(t_j - t_{j-1})\}, \tag{19}$$

where λ_{ij} is small, then the probability of death will be

$$1 - \phi_{ij} \approx 1 - \{1 - \lambda_{ij}(t_j - t_{j-1})\} = \lambda_{ij}(t_j - t_{j-1}). \tag{20}$$

Therefore the expected number of type i dead animals recovered in this time interval is

$$E_{ij} \approx A_i \gamma_{ij} \lambda_{ij} (t_j - t_{j-1}) B, \tag{21}$$

where $A_i B$ is the total number of type i animals alive at time t_j and γ_j is the recovery rate for all dead animals for the time interval. For a large population it will be reasonable to assume that the actual number of type i animals recovered is a Poisson variate with mean E_{ij} . Providing that the number of deaths is small it will be reasonable to treat $A_i B$ as remaining constant over time.

Putting $\gamma_j = \exp(\alpha_j)$ and $\lambda_{ij} = \exp(\Pi_{ij})$, equation (21) can be written as

$$E_{ij} = \exp\{\alpha_j + \Pi_{ij} + \log_e\{A_i(t_j - t_{j-1})\} + \log_e B\}. \tag{22}$$

This is then a log-linear model of the type that can be estimated by the computer program GLIM (Neider, 1975; Manly, 1977a), and other similar programs. If one of these programs is used then various different assumptions can be made about the parameters in the model and the goodness of fit of these assumptions can be assessed using chi-squared tests.

A simpler approach than fitting a log-linear model involves noting from equation (21) that

$$(E_{ij} / A_i) / \sum_{r=1}^K (E_{rj} / A_r) \approx \lambda_{ij} / \sum_{r=1}^K \lambda_{rj} = \beta_{ij} \tag{23}$$

where β_{ij} is the relative death rate for type i animals for the time interval t_{j-1} to t_j . If there are an equal number of animals of each type then a proportion β_{ij} of all deaths are expected to be of type i . Equation (23) suggests that we estimate β_{ij} by

$$\beta_{ij} = (d_{ij} / A_i) / \sum_{r=1}^K (d_{rj} / A_r). \tag{24}$$

On the assumption that the d_{ij} 's are independent Poisson variates the biases, variances and covariances of the β_{ij} 's are then found by Taylor series approximation (Seber, 1973, p.7) to be as follows:

Example

As an example consider Sheppard's (1951) data on predation of *Capaea nemoralis* by thrushes (Table 3). In this case $K=2$, corresponding to two colour classes of snails, with $s=14$ samples of dead snails.

Equation (22) was fitted to the data using the program GLIM. (The OFFSET directive can be used to handle the known values for $\log_e \{A_j (t_j - t_{j-1})\}$.) Various models were tried, as follows:

- (a) It was assumed that the α_j values (the recovery rate parameters) and the Π_{ij} values (the mortality rate parameters) were constant throughout the experiment. This gives a very poor fit to the data ($\chi^2_{27} = 84.78, p < 0.001$).
- (b) It was assumed that the α_j values were not equal but the Π_{ij} values were. This gives a fair fit to the data ($\chi^2_{14} = 20.83, p < 0.01$).
- (c) It was assumed that the α_j values were not equal and the Π_{ij} values varied with the two colour classes of snail but were constant over time for snails in the same class. This gives about the same fit as model (b) ($\chi^2_{13} = 20.58, 0.1 > p > 0.05$).
- (d) Finally it was assumed that the α_j values were not equal and the Π_{ij} values changed linearly with time relative to the Π_{2j} values, so that there is a relationship of the form

$$\Pi_{1j} = A + Bt_{j-1} + \Pi_{2j}$$

where A and B are constants and t_{j-1} is time in days. This model gives an excellent fit to the data ($\chi^2_{12} = 14.31, p > 0.1$) and is a considerably better fit than the other models that were tried.

Model (d) seems appropriate for the data. It gives the expected number of pink and brown shells in the j th sample to be of the form

$E_{1j} = \exp[\alpha_j - 9.23 + 0.079t_{j-1} + \log \{A_1(t_j - t_{j-1})\}]$
while the expected number of yellow shells is

$$E_{2j} = \exp[\alpha_j - 7.78 + \log \{A_2(t_j - t_{j-1})\}].$$

Equation (23) then gives relative death rates of

$$\beta_{1j} = \exp(-1.36 + 0.079t_{j-1}) / \{1 + \exp(-1.36 + 0.079t_{j-1})\},$$

and $\beta_{2j} = 1 - \beta_{1j}$, for the two colour classes. This suggests that at the start of Sheppard's experiment ($t_{j-1} = 0$) the β_{1j} value was 0.20 while at the end of the experiment ($t_{j-1} = 40$) it was 0.86. In other words at the start of the experiment the death rate was four times higher for yellow snails than for pink and brown snails, while by the end of the experiment pink and brown snails had a death rate about six times higher than yellow snails.

An alternative to this computer analysis with the program GLIM would have been to estimate the β_{ij} values directly using equation (24), and then use multiple regression to relate changes in these values to time. However this would give a rather unsatisfactory analysis with this particular example because the small numbers of dead animals would make the

individual β_{ij} estimates very unreliable. (A referee has pointed out the small number of dead animals means that the chi-square goodness of fit values given by the GLIM analysis need to be regarded with some caution too. They are, strictly speaking, only valid for large expected frequencies.)

Sheppard's (1951) own treatment of his data involved a probit analysis of the proportion of pink and brown shells in the samples at different times. This is somewhat unsatisfactory because it does not make use of the known colour composition of the population of marked snails.

A SMALL NUMBER OF DEATHS WITH SEVERAL TYPES OF ANIMAL IN ESTIMATED PROPORTIONS

The models discussed in the previous section are easily modified for the case where the population relative frequencies A_1, A_2, \dots, A_K are estimated from a random sample of the live animals rather than being known exactly.

Equation (22) still holds and it can be written as

$$E_{ij} = \exp\{\alpha_j + \Pi_{ij} + \delta_i + \log_e(t_j - t_{j-1}) + \log_e(B)\} \quad (28)$$

where $\exp(\delta_i) = A_i$. Because A_i is not known, δ_i must be estimated along with the other parameters. If a random sample of live animals is also available then we can let a_i denote the number of type i animals in this, where a_i will have an expected value of the form

$$E_i^* = A_i \exp(\theta) = \exp(\delta_i + \theta), \quad (29)$$

where $\exp(\theta)$ is a constant that reflects the sample size.

Between them equations (28) and (29) provide a log-linear model that can be fitted to data using GLIM or a similar computer program.

Equation (24) can be modified to

$$\hat{B}_{ij} = (e_{ij} / a_i) / \sum_{r=1}^K (e_{ij} / a_r) \quad (30)$$

to allow for estimated population relative frequencies a_1, a_2, \dots, a_K for the different types of animal. Equations (25) to (27) then become

Example

Equations (28) and (29) were fitted to Sheppard's (1951) Marley Wood data (Table 4) using GLIM. The OFFSET directive in the program was used to account for the known values of $\log_e(t_j - t_{j-1})$. Without going into details it will merely be said that a model similar to the one that fitted with the Ten Acre Copse data was found to be appropriate. This model gives

$$\beta_{ij} = \exp(-0.94 + 0.0387t_{j-1}) / \{1 + \exp(-0.94 + 0.0387t_{j-1})\}$$

for the relative death rate of pink and brown snails and $\beta_{2j} = 1 - \beta_{1j}$ for the relative death rate of yellow snails, where t_{j-1} is the time in days from 6 April. At the start of the experiment ($t_{j-1} = 0$) this gives $\beta_{1j} = 0.28$ so that a pink and brown snail was rather less likely to be eaten than a yellow snail. At the end of the experiment ($t_{j-1} = 50$) $\beta_{1j} = 0.73$, so that the reverse was true. Thus, the Marley Woods results confirm the trend found at Ten Acre Copse of an initial advantage for pink and brown snails that changed to a disadvantage as time went on. Sheppard attributed this to changes in the colour of the background vegetation.

MORTALITY RELATED TO QUANTITATIVE VARIABLES

The final situation that will be considered occurs when each animal has associated with it values for certain characters X_1, X_2, \dots, X_p . A random sample from the living population is available and also a random sample of dead animals. The problem is to see how survival relates to the X variables. The living population is assumed to be very large.

For simplicity assume that there is only a single X variable, which has probability density function $f_0(x)$ in the live population at time zero. Assume also that the probability of an individual with $X = x$ surviving for a time t is

$$p_t(x) = \exp\{-\lambda(x)t\}, \tag{34}$$

Then the probability density function for the survivors at time t is

$$f_t(x) = A f_0(x) p_t(x) \tag{35}$$

and the probability density function for the non-survivors is

$$g_t(x) = B f_0(x) \{1 - p_t(x)\},$$

where A and B are simply constants which ensure that

(x) and $g_t(x)$ integrate to 1.

Now, if t is small then

$$p_t(x) \approx 1 - \lambda(x)t$$

so that $\lambda(x)$ is the death rate for individuals with $X = x$. Also in this case

$$g_t(x) \approx Bt f_0(x) \lambda(x) \tag{36}$$

so that $g_t(x)$ and $f_t(x)$ have a similar form: they are both equal to $f_0(x)$ multiplied by a function of x. This shows that if a method is available for estimating $p_t(x)$ by comparing a sample of survivors with a sample from the original population then this method can be applied using a sample of deaths in place of the sample of survivors and it will then give an estimate of $\lambda(x)$ instead of $p_t(x)$.

There are indeed methods available for estimating $p(x)$ (the "fitness function") by comparing a sample of survivors with a sample from the initial population (O'Donald, 1970; Cavalli-Sforza and Bodmer, 1972; Manly I 977b, 1981). One approach involves assuming that

$$p_t(x) = \exp\{(\alpha_0 + \sum_{i=1}^r \alpha_i x^i)t\}$$

for some value of r (Manly, 1981). If this is used with a sample of dead animals instead of a sample of survivors then it estimates

$$\lambda(x) = \exp(\alpha_0 + \sum_{i=1}^r \alpha_i x^i)$$

instead of $p_t(x)$. In practice α_0 cannot be estimated from sample data. However, this is not important for the comparison of $\lambda(x)$ with different values of x.

The method of estimation discussed by Manly (1981) for data from a non-normal distribution was applied to Wong and Ward's (1972) data for 1 July (Table 5), for which it estimated

Here x is the length of *Daphnia publicaria*, which has a mean of 1.56 mm and a standard deviation of 0.59 mm in the plankton sample. This function gives $\lambda(1.56) = 1$, corresponding to *D. publicaria* of average length. Relative to this *D. publicaria* with a length of only 0.5 mm have a mortality rate of 2.23, while those with the maximum length of 3.1 mm have a mortality rate of only about 6×10^{-14} .

If there are p X variables rather than just one, then equations (34) to (36) generalise to

$$p_t(x_1, x_2, \dots, x_p) = \exp\{-\lambda(x_1, x_2, \dots, x_p)t\}, \tag{37}$$

$$f_t(x_1, x_2, \dots, x_p) = A f_0(x_1, x_2, \dots, x_p) \times p_t(x_1, x_2, \dots, x_p)$$

(38)

and

$$g_t(x_1, x_2, \dots, x_p) = Bt f_0(x_1, x_2, \dots, x_p) \times \lambda(x_1, x_2, \dots, x_p), \tag{39}$$

respectively. The similarity between equations (38) and (39) shows that the multivariate methods for estimating $p_i(x_1, x_2, \dots, x_p)$ discussed by Manly (1981) can also be used to estimate $\lambda(x_1, x_2, \dots, x_p)$. As is the case of a single X variable, using a sample of dead animals in place of a sample of survivors results in an estimate of $\lambda(x_1, x_2, \dots, x_p)$ rather than $p_i(x_1, x_2, \dots, x_p)$.

REFERENCES

- BROWNIE, C.; ANDERSON, D. R.; BURNHAM, K. P.; ROBSON, D. S. 1978. *Statistical Inference from Band Recovery Data*. U.S. Department of the Interior, Fish and Wildlife Service Resource Publication 131, Washington D.C.
- CAVELLI-SFORZA, L. L.; BODMER, W. F. 1972. *The Genetics of Human Populations*. W. H. Freeman and Co., San Francisco.
- FORDHAM, R. A.; CORMACK, R. M. 1970. Mortality and population change of dominican gulls in Wellington, New Zealand. *Journal of Animal Ecology* 39: 13-27.
- FRASER, R. A.; DUNCAN, W. J.; COLLAR, A. R. 1963. *Elementary Matrices*. Cambridge University Press, London.
- HALDANE, J. B. S. 1953. Some animal life tables. *Journal of the Institute of Actuaries* 79: 83-9.
- HALDANE, J. B. S. 1955. The calculation of mortality rates from ringing data. *Proceedings of the 11th International Ornithological Congress, Basel, Supplementum* 3, pp. 454-8.
- JOLLY, G. M. 1979. A unified approach to mark-recapture stochastic models, exemplified by a constant survival rate model. In: Cormack, R. M.; Patil, G. P.; Robson, D. S. (Editors) *Sampling Biological Populations*. Statistical Ecology Series, Vol. 5, pp. 277-82. International Co-operative Publishing House, Maryland, USA
- MANLY, B. F. J. 1977a. Examples of the use of GLIM. *New Zealand Statistician* 12(1): 26-42.
- MANLY, B. F. J. 1977b. The estimation of the fitness function from several samples taken from a population. *Biometrical Journal* 19: 391-401.
- MANLY, B. F. J. 1981. The estimation of multivariate fitness function from several samples taken from a population. *Biometrical Journal* 28: 267-81.
- NELDER, J. A. 1975. *GLIM Manual*. Distributed by the Numerical Algorithms Group, 13 Banbury Road, Oxford, England.
- NORTH, P. M.; MORGAN, B. J. T. 1979. Modelling Heron survival using weather data. *Biometrics* 35: 667-81.
- O'DONALD, P. 1970. Change of fitness by selection for quantitative characters. *Theoretical Population Biology* 1: 219-32.
- ROBSON, D. S. 1963. Maximum likelihood estimation of a sequence of annual survival rates for a capture-recapture series. In: Champion, W. H. (Editor). *North Atlantic Fish Marking Symposium*. International Commission for Northwest Atlantic Fisheries Special Publication 4; 330-5.
- ROBSON, D. S.; CHAPMAN, D. G. 1961. Catch curves and mortality rates. *Transactions of the American Fisheries Society* 90: 181-9.
- SEBER, G. A. F. 1970. Estimating time-specific survival and reporting rates from adult bird-band returns. *Biometrika* 57: 313-8.
- SEBER, G. A. F. 1971. Estimating age-specific survival rates from bird-band returns when the reporting rate is constant. *Biometrika* 58: 491-7.
- SEBER, G. A. F. 1973. *Estimation of Animal Abundance and Related Parameters*. Griffin, London.
- SEBER, G. A. F. 1980. *Some Recent Advances in the Estimation of Animal Abundance*. Washington Sea Grant Technical Report WSG 80-1.
- SHEPPARD, P. M. 1951. Fluctuations in the selective value of certain phenotypes in the polymorphic land snail *Cepaea nemoralis* (L.). *Heredity* 5: 125-34.
- WONG, B.; WARD, F. J. 1972. Size selection of *Daphnia publicaria* by yellow perch (*Perca flavescens*) fry in West blue lake, Manitoba. *Journal of the Fisheries Research Board of Canada* 29: 1761-4.

APPENDIX 1

THE POISSON MODEL FOR EQUATIONS (4) TO (7) AND (12) TO (18)

The likelihood given by equation (4) is maximised with respect to parameters B, CPI, $\phi_1\phi_2, \dots, \phi_r$ and ϕ when the log-likelihood

$$\lambda(B, \phi_1, \phi_2, \dots, \phi_r, \phi) = \sum_{j=1}^s \{d_j \log E_j - E_j - \log(d_j!)\}$$

is maximised. This occurs when

$$\frac{\partial \lambda}{\partial B} = \frac{\partial \lambda}{\partial \phi r} = \dots = \frac{\partial \lambda}{\partial \phi r} = \frac{\partial \lambda}{\partial \phi} = 0. \quad (A1)$$

Equations (A1) produce the estimation equations (5) to (7) for the maximum likelihood estimators.

Let θ_p and θ_q denote two of the parameters B, $\phi_1, \phi_2, \dots, \phi_r, \phi$. Then the second derivative of the likelihood function with respect to these is of the form

The standard theory of maximum likelihood estimation (Seber, 1973, p.5) shows that the matrix of variances and covariances of parameter estimates

The matrix \underline{V} can be calculated by inverting $\underline{D}-\underline{H}$ numerically. However providing most animals are dead by the end of the experiment $w_r \phi^{s-r} \approx 0$ and \underline{V} is given by

$$\underline{V} \approx \underline{D}^{-1} + \underline{D}^{-1} \underline{H} \underline{D}^{-1} \quad (\text{A4})$$

(Fraser, Duncan and Collar, 1963, p. 120). This approximation produces equations (12) to (18) of this paper. If the elements of $\underline{D}^{-1} \underline{H} \underline{D}^{-1}$ are not small relative to the elements of \underline{D}^{-1} then \underline{V} should be evaluated as $(\underline{D} - \underline{H})^{-1}$ rather than by using the approximation (A4).

is approximately given by minus the inverse of the matrix whose general term is (A2). This gives