Chapter 2. Models for Birds Banded as Adults

2.1 Introduction

The Experimental Situation

This chapter discusses the experimental situation where a sample of adult birds are captured, banded, and released into the population at roughly the same time each year, for a number of successive years. The banded population should be a representative sample of the population of interest. As described in Chapter 1, the population is subjected to hunting each year and hunters are requested to report bands from birds they have shot to the Bird Banding Laboratory. Data collection, or the recording of numbers of bands returned, may continue for several years after the last release of banded birds. A "year" of the experiment or banding study is the period (of approximately 1 year) between successive releases of banded birds. The period of survival is from the time of banding in year i to the time of banding in year $i+1$. The period of survival is *not* the interval between hunting seasons.

Notation and Description of Data

We let k represent the number of successive years at the start of which a release of banded birds is made. Also we define

ℓ = the number of years during which recoveries are recorded, $\ell \geq k$,
$s = \ell - k$, the number of years beyond the year of the last release when recoveries are recorded, $s \geq 0$,
N_i = the number of adults banded and released at the start of the i^{th} year, $i = 1,2,\ldots,k$,
R_{ij} = the number of bands recovered in year j from the adults released in year i, $i = 1,\ldots,k, j = i,\ldots,\ell$.

The data collected can be displayed in a table as shown below.

Table 2.1. *Representation of the data for a 5-year banding study when three releases were made (i.e., when $k=3$, $\ell=5$, $s=2$).*

Year banded	Number banded	Year of recovery 1	2	3	4	5	Row totals
1	N_1	T_1 R_{11}	R_{12}	R_{13}	R_{14}	R_{15}	$R_1 = T_1$
2	N_2		T_2 R_{22}	R_{23}	R_{24}	R_{25}	R_2
3	N_3			T_3 R_{33}	T_4 R_{34}	R_{35}	R_3
Column totals		C_1	C_2	C_3	C_4	$C_5 = T_5$	

As indicated in Table 2.1 we let R_i and C_i represent the row and column totals, respectively, and define the indicated block totals by

$$T_1 = R_1 ,$$

$$T_i = R_i + T_{i-1} - C_{i-1} \qquad ,i = 2,\ldots,k ,$$

and if $\ell > k$,

$$T_{k+j} = T_{k+j-1} - C_{k+j-1} \qquad ,j = 1,\ldots,s .$$

Table 2.2 contains real banding and recovery data for such an experiment with calculation of the corresponding subtotals R_i, C_j, and T_j.

Table 2.2 *Banding and recovery data for male wood duck (Aix sponsa), banded preseason in a midwestern State.*

			Year of recovery					
i	Year banded	Number banded	1964 $j=1$	1965 2	1966 3	1967 4	1968 5	R_i
1	1964	1,603	127	44	37	40	17	265
2	1965	1,595		62	76	44	28	210
3	1966	1,157			82	61	24	167
			$C_j=127$	106	195	145	69	
			$T_j=265$	348	409	214	69	

Definition of Parameters

A banded adult alive in the population at the start of the i^{th} year of the experiment will survive the year, be killed by a hunter, or die from a cause unrelated to hunting during the year. We define the following annual rates for banded adults alive at the start of a given year of the experiment:

S = survival rate or probability of surviving the year
K = kill rate or probability of being killed by a hunter during the year
H = harvest rate or probability of being killed and retrieved by a hunter during the year
$1-S-K$ = nonhunting mortality rate.

Not all birds killed by hunters are retrieved; therefore, we have $H=cK$ where c represents a retrieval rate. When a hunter kills and retrieves a banded bird, he may or may not report the band to the Bird Banding Laboratory, so we also define

λ = band reporting rate = probability that a hunter will report the band given that he has killed and retrieved a banded bird.

These different outcomes are summarized schematically, for any year of the experiment, below.

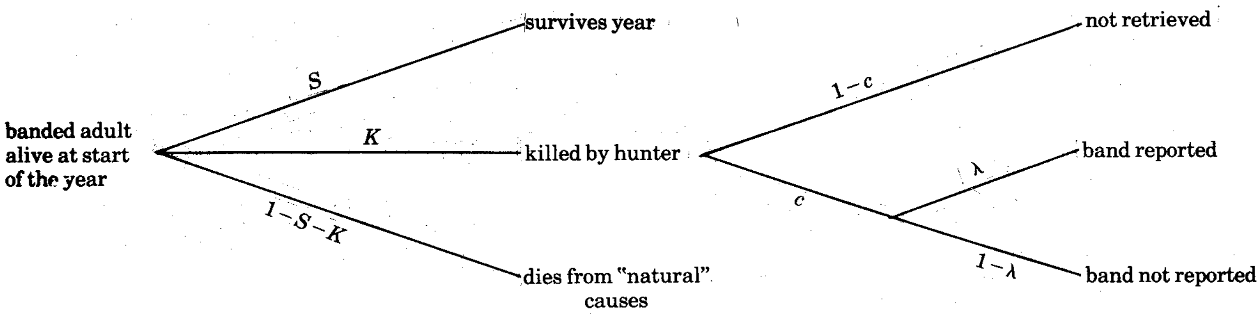

Fig. 2.1 Potential fates of a banded bird alive at the start of the year.

Note that the type of data collected supplies information directly about only those birds which are shot *and* reported. Thus, only the product λH is estimable but the component rates λ and H are not estimable without additional information such as the use of REWARD bands (Henny and Burnham 1976). Defining $f=\lambda H$ = band recovery rate (or the "reported exploitation rate"), we modify Fig. 2.1 as follows:

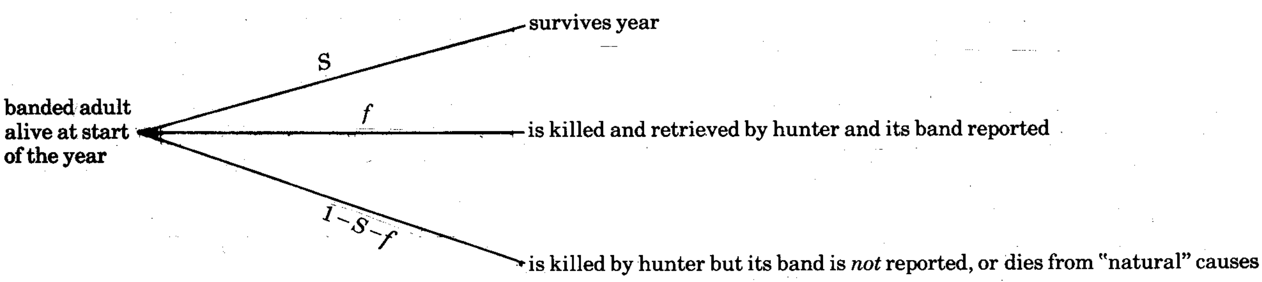

Fig. 2.2 Three critical fates of a banded bird alive at the start of the year.

Different assumptions about the variation of the parameters $f(=\lambda H)$, and S give rise to models of different degrees of complexity. All models in this chapter assume the rate parameters are age-independent (hence the restriction to data from banded adults only). In this chapter we describe the assumptions and biological significance of several such models, and the corresponding estimation schemes for each model. Tests to discriminate between competing models and goodness of fit tests are provided. Examples are based on output from a FORTRAN computer program we call ESTIMATE which is available for carrying out the numerical computations. This program is described in Chapter 6.

2.2 Model 1

One of the two most useful models of this chapter is that developed separately by Seber (1970) and by Robson and Youngs (1971), which we call Model 1. In addition to the assumption of age-independence referred to above, it is assumed that survival, hunting, and reporting rates are year-specific but independent of the year of banding. Thus, the parameters f and S are subscripted to indicate dependence on a specific year. For example, f_1 and S_1 are the recovery rate and survival rate, respectively, for year 1 of the banding study. In general, f_i and S_i are the corresponding rates for the i^{th} year of the study.

Model 1 is characterized as in Table 2.3 in terms of the expected or average numbers of band recoveries, expressed as functions of N_i, f_i, and S_i, as described in Chapter 1.

Table 2.3. *Expected numbers of band recoveries under Model 1 for a banding study with* $k=3$, $\ell=5$, *and* $s=2$.

Year banded	Number banded	Year of recovery 1	2	3	4	$\ell=5$
1	N_1	N_1f_1	$N_1S_1f_2$	$N_1S_1S_2f_3$	$N_1S_1S_2S_3f_4$	$N_1S_1S_2S_3S_4f_5$
2	N_2		N_2f_2	$N_2S_2f_3$	$N_2S_2S_3f_4$	$N_2S_2S_3S_4f_5$
$k=3$	N_3			N_3f_3	$N_3S_3f_4$	$N_3S_3S_4f_5$

Before describing the estimation of the parameters f_i and S_i for general k and ℓ, we note that in Table 2.3, where $k=3$, $\ell=5$, and $s=2$, the parameters S_3 and f_4 always occur together as the product S_3f_4. We find that the parameters S_3 and f_4 are not separately estimable, but the product S_3f_4 is estimable. In general, Under Model 1, the parameters $f_1,f_2,\ldots,f_k$, $S_1,S_2,\ldots,S_{k-1}$ are separately estimable, but if $s>0$, only products such as S_kf_{k+1}, $S_kS_{k+1}f_{k+2},\ldots$, $S_kS_{k+1}\cdots S_{k+s-1}f_{k+s}$ are also estimable, not the individual parameters S_{k+j-1}, and f_{k+j}, $j=1,\ldots,s$. Estimates of S_kf_{k+1},$S_kS_{k+1}f_{k+2}$, etc., are not of biological interest since, for example, $S_kS_{k+1}f_{k+2}$ represents the probability of surviving years k and $k+1$ and then being shot and reported in year $k+2$. Estimates of these products are necessary, however, to test goodness of fit of the model, as we explain later in this chapter.

Estimation of Parameters

The estimators developed by Seber (1970) and Robson and Youngs (1971) are based on the principle of Maximum Likelihood. ML estimators have many optimal properties for large sample sizes, but often have a small statistical bias. Seber (1962, 1965) and others have studied this bias in capture-recapture models (which are similar to the models used here) and found adjusted ML estimators that are essentially unbiased. We have made similar adjustments to the appropriate estimators in the models discussed here and in Chapter 3.

The ML estimator of the recovery rate f_i, in year i, is

$$\hat{f}_i = \frac{R_i}{N_i}\frac{C_i}{T_i} \qquad , i=1,\dots,k.$$

The bias-adjusted ML estimator of the survival rate S_i, in year i, is

$$\tilde{S}_i = \frac{R_i}{N_i}\left(\frac{T_i - C_i}{T_i}\right)\frac{N_{i+1}+1}{R_{i+1}+1} \qquad , i=1,\dots,k-1.$$

As stated in Section 1.4, we denote ML estimators by a "hat" (^) over the parameter symbol; e.g., $\hat{f}_i$ and $\hat{S}_i$. All bias-adjusted ML estimators are denoted by a tilde (~); e.g., $\tilde{S}_i$. The ML estimators of f_i under Model 1 are unbiased, but the $\hat{S}_i$ are biased. The estimator $\tilde{S}_i$ given above is essentially unbiased. For small sets, $\hat{f}_i$ and $\tilde{S}_i$ are easily evaluated as shown below using the data in Table 2.2:

$$\hat{f}_1 = \frac{R_1}{N_1}\frac{C_1}{T_1} = \frac{265\times 127}{1{,}603\times 265} = 0.0792 \text{ or } 7.92\%$$

$$\hat{f}_2 = \frac{R_2}{N_2}\frac{C_2}{T_2} = \frac{210\times 106}{1{,}595\times 348} = 0.0401 \text{ or } 4.01\%$$

$$\hat{f}_3 = \frac{R_3}{N_3}\frac{C_3}{T_3} = \frac{167\times 195}{1{,}157\times 409} = 0.0688 \text{ or } 6.88\%$$

$$\tilde{S}_1 = \frac{R_1\times(T_1-C_1)\times(N_2+1)}{N_1\times T_1\times(R_2+1)} = \frac{265\times(265-127)\times 1{,}596}{1{,}603\times 265\times 211} = 0.6512 \text{ or } 65.12\%,$$

$$\tilde{S}_2 = \frac{R_2\times(T_2-C_2)\times(N_3+1)}{N_2\times T_2\times(R_3+1)} = \frac{210\times(348-106)\times 1{,}158}{1{,}595\times 348\times 168} = 0.6311 \text{ or } 63.11\%\,.$$

The actual (unadjusted) ML estimator of S_i is

$$\hat{S}_i = \frac{R_i}{N_i}\left(\frac{T_i - C_i}{T_i}\right)\frac{N_{i+1}}{R_{i+1}} \qquad , i=1,\dots,k-1\,.$$

For example, from Table 2.2,

$$\hat{S}_1 = \frac{R_1}{N_1}\left(\frac{T_1 - C_1}{T_1}\right)\frac{N_2}{R_2} = \frac{265}{1{,}603}\;\frac{(265-127)}{265}\;\frac{1{,}595}{210} = 0.6539 \text{ or } 65.39\%\,.$$

The adjustment of the ML estimator of S_i involves merely the addition of 1 in two terms of the formula. The bias-adjusted estimates $\tilde{S}$ are always slightly smaller than the ML estimates $\hat{S}$.

Estimated average recovery and survival rates are easily computed.

$$\hat{\bar{f}} = \frac{1}{k-1}\sum_{i=1}^{k-1}\hat{f}_i, \qquad \text{and} \qquad \tilde{\bar{S}} = \frac{1}{k-1}\sum_{i=1}^{k-1}\tilde{S}_i\,.$$

These are simple arithmetic means; the geometric mean survival has been recommended in the literature, but is not used here (see Section 8.5). The geometric mean of the survival estimates is not a consistent estimator of the true arithmetic average survival rate $\bar{S} = (\sum_{i=1}^{k-1} S_i)/(k-1)$. Also we note that $\hat{f}_k$ is not included in $\hat{\bar{f}}$ by program ESTIMATE; rather $\hat{\bar{f}}$ is computed as shown above. The rationale came from a desire to compare variations in average recovery and survival rates. Because there is no estimate of S_k, comparability is enhanced by not including $\hat{f}_k$ in computing $\hat{\bar{f}}$.

As stated above, if $\ell > k$, estimates of $S_k f_{k+1}, S_k S_{k+1} f_{k+2}$, etc., are probably not of great interest. However, they must be computed to calculate the goodness of fit test described below, so the appropriate estimators are defined here,

$$\widehat{S_k \cdots S_{k+j-1} f_{k+j}} = \frac{R_k}{N_k} \frac{C_{k+j}}{T_k} \qquad ,j = 1, \ldots, s.$$

These estimates are printed in the output of program ESTIMATE as GAMMA (following Seber's original notation).

Sampling Variances, Standard Errors, and Confidence Intervals

Confidence intervals for the parameter estimates are very important since they provide an indication of the precision of the estimate. Approximate confidence intervals for $\hat{f}_i$, $\hat{S}_i$, and $\tilde{S}_i$ are obtained as follows:

Let $\text{var}(\hat{f}_i)$ be the estimator of the sampling variance of $\hat{f}_i$, then

$$\text{var}(\hat{f}_i) = (\hat{f}_i)^2 \left[\frac{1}{R_i} - \frac{1}{N_i} + \frac{1}{C_i} - \frac{1}{T_i}\right] \qquad ,i = 1, \ldots, k.$$

Then $\text{se}(\hat{f}_i) = \sqrt{\text{var}(\hat{f}_i)}$ is the estimator of the standard error of $\hat{f}_i$, and an approximate 95% confidence interval on the recovery rate f_i is $(\hat{f}_i - 1.96 \times \text{se}(\hat{f}_i), \hat{f}_i + 1.96 \times \text{se}(\hat{f}_i))$.

For example, if we use the data of Table 2.2 and the estimate $\hat{f}_2$ calculated above,

$$\text{var}(\hat{f}_2) = (0.0401)^2 \left[\frac{1}{210} - \frac{1}{1{,}595} + \frac{1}{106} - \frac{1}{348}\right] = 0.00001720,$$

$$\text{se}(\hat{f}_2) = \sqrt{0.00001720} = 0.00415,$$

$$1.96\ \text{se}(\hat{f}_2) = 1.96 \times 0.00415 = 0.0081,$$

and the approximate 95% confidence interval for f_2 is $(0.0401 - 0.0081, 0.0401 + 0.0081)$, which is $(0.0320, 0.0482)$, or in terms of percentage recovery (3.20,4.82).

Approximate confidence intervals are obtained in the same way for S_i based on the sampling variance estimators,

$$\text{var}(\tilde{S}_i) = (\tilde{S}_i)^2 \left[\frac{1}{R_i} - \frac{1}{N_i} + \frac{1}{R_{i+1}} - \frac{1}{N_{i+1}} + \frac{1}{T_{i+1} - R_{i+1}} - \frac{1}{T_i}\right] \qquad ,i = 1, \ldots, k-1.$$

The limits of an approximate 95% confidence interval on S_i are given by $\tilde{S}_i \pm 1.96\ \text{se}(\tilde{S}_i)$, where

$$\text{se}(\tilde{S}_i) = \sqrt{\text{var}(\tilde{S}_i)} \qquad ,i = 1, \ldots, k-1.$$

Program ESTIMATE computes confidence intervals for f_i and S_i in this way.

Sampling Covariances and Correlations

Estimates of annual recovery and survival rates are derived from the same information and therefore we might suspect that there are sampling correlations between these estimators. The program ESTIMATE computes estimates of the sampling covariances and correlations between the estimators, using formulae given in Robson and Youngs (1971). These correlations are estimates of the linear relationship between the estimators. The correlations may be substantial in some cases, e.g., $\text{corr}(\hat{\tilde{S}}_i,\hat{f}_{i+1})$ and $\text{corr}(\hat{\tilde{\bar{S}}},\hat{\bar{f}})$. These high correlations are important because they indicate that apparent relationships between parameter estimates often cannot be interpreted as evidence of a similar relationship between the true parameters. This subject is discussed further in Section 8.4 (also see Anderson and Burnham 1976).

The covariance estimators under Model 1 are given below:

$$\text{cov}(\tilde{S}_i,\hat{f}_i)=\hat{f}_i\tilde{S}_i\left[\frac{1}{R_i}-\frac{1}{N_i}-\frac{1}{T_i}\right] \qquad ,i=1,\ldots,k-1,$$

$$\text{cov}(\hat{f}_i,\hat{f}_{i+j})=0 \qquad ,j>1,$$

$$\text{cov}(\hat{f}_{i+j},\tilde{S}_i)=\begin{cases}0 & ,j>1,\\ -\tilde{S}_i\hat{f}_{i+1}\left[\frac{1}{R_{i+1}}-\frac{1}{N_{i+1}}\right] & ,j=1,\end{cases}$$

$$\text{cov}(\tilde{S}_i,\tilde{S}_{i+j})=\begin{cases}0 & ,j>1,\\ -\tilde{S}_i\tilde{S}_{i+1}\left[\frac{1}{R_{i+1}}-\frac{1}{N_{i+1}}\right] & ,j=1.\end{cases}$$

Estimated sampling correlation coefficients are derived by definition,

$$\text{corr}(\hat{X},\hat{Y})=\frac{\text{cov}(\hat{X},\hat{Y})}{\text{se}(\hat{X})\ \text{se}(\hat{Y})}$$

where $\hat{X}$ and $\hat{Y}$ are any two estimators such as $\hat{f}_1$ and $\tilde{S}_1$. For example, the estimated sampling correlation between $\tilde{S}_1$ and $\tilde{S}_2$ is

$$\text{corr}(\tilde{S}_1,\tilde{S}_2)=\frac{\text{cov}(\tilde{S}_1,\tilde{S}_2)}{\text{se}(\tilde{S}_1)\ \text{se}(\tilde{S}_2)}.$$

The computation of estimates of sampling covariances and correlations is easy but somewhat tedious as we illustrate for $\hat{f}_2$ and $\tilde{S}_2$

$$\text{cov}(\hat{f}_2,\tilde{S}_2)=\hat{f}_2\tilde{S}_2\left[\frac{1}{R_2}-\frac{1}{N_2}-\frac{1}{T_2}\right]=0.0401\times0.6311\left[\frac{1}{210}-\frac{1}{1{,}595}-\frac{1}{348}\right]=0.0000319,$$

$$\text{corr}(\hat{f}_2,\tilde{S}_2)=\frac{\text{cov}(\hat{f}_2,\tilde{S}_2)}{\text{se}(\hat{f}_2)\ \text{se}(\tilde{S}_2)}=\frac{0.0000319}{0.00415\times0.0647}=0.1189.$$

In this case the sampling correlation is fairly low; however, we note that $\text{corr}(\hat{f}_2,\tilde{S}_1)=-0.385$.

The calculation of the estimated sampling correlation between the estimates of average recovery and survival rate is more tedious, but is of the same general form:

$$\text{corr}(\bar{\tilde{S}},\bar{\hat{f}}) = \frac{\text{cov}(\bar{\tilde{S}},\bar{\hat{f}})}{\text{se}(\bar{\tilde{S}})\ \text{se}(\bar{\hat{f}})}$$

where

$$\text{cov}(\bar{\tilde{S}},\bar{\hat{f}}) = \frac{1}{(k-1)^2}\left(\sum_{i=1}^{k-1} \text{cov}(\hat{f}_i,\tilde{S}_i) + \sum_{i=1}^{k-2} \text{cov}(\hat{f}_{i+1},\tilde{S}_i)\right),$$

$$\text{se}(\bar{\hat{f}}) = \frac{1}{(k-1)}\left(\sum_{i=1}^{k-1} \text{var}(\hat{f}_i)\right)^{1/2}$$

and

$$\text{se}\,(\bar{\tilde{S}}) = \frac{1}{(k-1)}\left(\sum_{i=1}^{k-1} \text{var}(\tilde{S}_i) + 2\sum_{i=1}^{k-2} \text{cov}(\tilde{S}_i,\tilde{S}_{i+1})\right)^{1/2}.$$

The estimated sampling correlation between the estimates of average recovery rate and the average survival rate is -0.57 for the wood duck data. This represents a substantial correlation.

Goodness of Fit Test

The general strategy we are recommending for the analysis of banding and recovery data involves three steps:

(1) Formulation of tentative assumptions;

(2) estimation of parameters, using the principle of Maximum Likelihood, from models explicitly based on these tentative assumptions; and

(3) testing of the models employed and the assumptions made, using the observed data.

The testing phase is important and *has received inadequate attention in quantitative ecology in the past.* The parameter estimates, and particularly their sampling variances and covariances, are dependent on the assumptions being made. The assumptions are explicitly built into the model. It is, therefore, crucial that the model is realistic and that it adequately describes the observed data. This section treats the use of a goodness of fit test to examine the adequacy of the model. Tests between specific models are discussed in Section 2.6.

The goodness of fit testing procedure is illustrated by a conventional chi-square test. First, we have our observed data (presented in Table 2.2). Second, we have formal expressions for the expected number of band recoveries under the tentative assumptions explicitly made under Model 1 (Table 2.3). The expected numbers involve the parameters S_i and f_i, as well as the numbers banded, N_i. However, we now have estimates of these parameters, and can compute the estimates of the expected number of recoveries. For example, the expected number of recoveries the second year (1965, $j=2$) from the birds banded the first year (1964, $i=1$) is $N_1S_1f_2$ (see Table 2.3). This is the expected value of R_{12} under Model 1. The ML estimate of this expectation is $N_1\hat{S}_1\hat{f}_2$ or $1603 \times 0.6539 \times 0.0401 = 42.0$. Such a procedure is carried out for each cell in the array of expected recoveries and the estimates are denoted as $\mathbf{E}_{ij}$. A single chi-square value is computed as

$$\frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} \text{ or } \frac{(O-E)^2}{E}.$$

The overall test is made by adding the chi-square values for each cell,

$$\chi^2 = \sum_i\sum_j \frac{(\text{O}_{ij} - \mathbf{E}_{ij})^2}{\mathbf{E}_{ij}}.$$

This intuitive procedure represents a formal statistical test of the fit of the model to the observed data. The null hypothesis being tested is that the data fit the model and its assumptions. If this null hypothesis is rejected, the parameter estimates may be very biased, and their sampling variance and covariance estimates inappropriate. Under Model 1, this test has $(k(k+1)/2) + (\ell - k)k - (k + \ell - 1)$ degrees of freedom.

The test statistic is approximately distributed as chi-square, with the above degrees of freedom, under the null hypothesis. In computing the test statistic, it may be necessary to pool cells with small (e.g., <2) expected values to justify the chi-square approximation. Each time cells $\mathbf{E}_{ij}$ are pooled, the corresponding data values (R_{ij}) must be combined also. Pooling is done between cells in the same row. One degree of freedom is lost for each cell combined. This procedure is performed by program ESTIMATE and the pooled observed data and pooled estimated expected values are printed instead of the original, unpooled results.

An Example

The wood duck data have been used to illustrate the details of the calculations under Model 1. A computer analysis of these data appears in Example 2.1. Many of the estimates have been computed in this section and can be recognized in the computer example. Note that estimates obtained with a hand calculator may not agree exactly with the computer output. The computer carries about 14 significant digits and, therefore, rounding errors are small indeed. This is particularly important in the calculation of variances, covariances, and correlations.

As shown in Example 2.1a, the computer output displays the model structure in terms of the cell frequencies. The banding and recovery data and various column, row, and block totals are also displayed. Estimates of recovery rates and their standard errors are printed along with the 95 % confidence intervals on the true recovery rates. The first-year recovery rates (R_{ii}/N_i) are also printed; these inefficient estimates of f_i have been commonly used in bird banding studies. We print them only to allow comparison and do not recommend their use.

Estimates of annual and average annual survival rates and associated statistics are displayed. In Example 2.1b, the average survival rate was estimated at 64.11 ± 3.66 %. The mean life span was estimated to be 2.25 years $\pm$ 0.29 year (see Section 2.7 for the method). By multiplying the mean life span by 0.69 an estimate of "half life" is obtained. In the example it was estimated that half of the banded birds were dead after approximately 1.55 years (2.25×0.69). (The derivation of the value of 0.69 is explained in Appendix A.)

Statistics related to the goodness of fit test are printed next (see Example 2.1b). Matrices of observed data, estimated expected values, and individual chi-square values are displayed. By summing all the elements in the matrix of chi-square values we can test the fit of the model to the data. In this case a chi-square value of 5.87 was obtained with 5 degrees of freedom. A chi-square value as large as 5.87 is not unusual $(P=0.32)$ and we conclude that Model 1 fits the data satisfactorily. On the other hand, a chi-square value of 25.0 with 5 degrees of freedom would be unusually large $(P<0.01)$ and would provide reason to regard Model 1 as inappropriate for the analysis of these data.

The computer analysis of the example data under Model 1 concludes with the estimates of the sampling covariances and correlations of the estimators.

2.3 Model 2

Another model that is very useful in estimating age-independent parameters from band recovery data is referred to here as Model 2. The assumptions of Model 2 are more restrictive than those of Model 1, in that the survival rate is assumed to be constant from year to year. Band reporting rates λ and harvest rates H (and hence recovery rates f) are assumed to be year-specific. As before, the recovery rate in year i is denoted by f_i; however, the constant annual survival rate is denoted simply by S, without a subscript. Model 2 is a special case of Model 1 where one assumes $S_1 = S_2 = \cdots = S_{\ell-1} \equiv S$.

Model 2 is represented in Table 2.4 in terms of the expected numbers of band recoveries, expressed as functions of N_i, f_i, and S. Information in Table 2.4 specifies the assumptions upon which the model is explicitly based. The key assumptions of Model 2 are that recovery rates vary from year to year (due to changes in hunting regulations, environmental factors, etc.), but that annual survival is constant from year to year.

Table 2.4. *Expected numbers of band recoveries under Model 2 for a banding study with $k=3$, $\ell=5$, and $s=2$.*

Year banded	Number banded	Year of recovery 1	2	3	4	$\ell=5$
1	N_1	N_1f_1	N_1Sf_2	N_1SSf_3	N_1SSSf_4	N_1SSSSf_5
2	N_2		N_2f_2	N_2Sf_3	N_2SSf_4	N_2SSSf_5
$k=3$	N_3			N_3f_3	N_3Sf_4	N_3SSf_5

Example 2.1a

MODEL 1
****

ANALYSIS UNDER THE ASSUMPTIONS OF TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(A SYNTHESIS OF MODELS DEVELOPED BY SEBER (1970. BIOMETRIKA) AND ROBSON AND YOUNGS (1971. CORNELL BIOMETRICS UNIT PAPER 369))

SPECIFICALLY, THE MODEL STRUCTURE IS:

F(1)	S(1)F(2)	S(1)S(2)F(3)	S(1)S(2)S(3)F(4)	S(1)S(2)S(3)S(4)F(5)
	F(2)	S(2)F(3)	S(2)S(3)F(4)	S(2)S(3)S(4)F(5)
		F(3)	S(3)F(4)	S(3)S(4)F(5)
			F(4)	S(4)F(5)

MALE WOOD DUCK BANDED PRESEASON IN A MID-WESTERN STATE, 1964-66

BANDING AND RECOVERY INPUT DATA

YEAR	NUMBER BANDED	RECOVERY MATRIX				
1964	1603	127	44	37	40	17
1965	1595	0	62	76	44	28
1966	1157	0	0	82	61	24

INTERMEDIATE STATISTICS

I	C(I)	R(I)	T(I)	N(I)	GAMMA(I)	RHO(I)
1964	127.0	265.0	265.0	1603.0	0.0	0.16532
1965	106.0	210.0	348.0	1595.0	0.0	0.13166
1966	195.0	167.0	409.0	1157.0	0.0	0.14434
1967	145.0	0.0	0.0	0.0	0.05117	0.0
1968	69.0	0.0	0.0	0.0	0.02435	0.0
TOTALS		642.0		4355.0		

I	RECOVERY RATE F(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL	DIRECT RECOVERY RATE R(I,I)/N(I) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
1964	7.923	0.675	6.600 - 9.245	7.923	0.675	6.600 - 9.245
1965	4.010	0.415	3.197 - 4.823	3.887	0.484	2.939 - 4.836
1966	6.882	0.608	5.690 - 8.073	7.087	0.754	5.609 - 8.566

ARITHMETIC MEAN RECOVERY RATE (EXCEPT YEAR K) = 5.967
STANDARD ERROR OF THE MEAN RECOVERY RATE = 0.396
95% CONFIDENCE INTERVAL FOR MEAN RECOVERY RATE = 5.19 - 6.74

Example 2.1b

MODEL 1 -- ANALYSIS UNDER THE ASSUMPTIONS OF TIME SPECIFIC SURVIVAL AND RECOVERY RATES

MALE WOOD DUCK BANDED PRESEASON IN A MID-WESTERN STATE, 1964-66

YEAR	SURVIVAL S(I) (%)			
	SURVIVAL	STANDARD ERROR	COEFFICIENTS OF VARIATION	95% CONFIDENCE INTERVAL
1964	65.12	6.75	10.37	51.88 - 78.35
1965	63.11	6.47	10.26	50.42 - 75.79

ARITHMETIC MEAN SURVIVAL (%) = 64.11
STANDARD ERROR OF ARITHMETIC MEAN = 3.66
95% CONFIDENCE INTERVAL FOR ARITHMETIC MEAN 56.94- 71.28

MEAN LIFE SPAN AS AN ADULT = 2.25
STANDARD ERROR OF THE MEAN LIFE SPAN = 0.29
95% CONFIDENCE INTERVAL OF LIFE SPAN 1.78 - 2.95

RECOVERY MATRIX

YEAR	NUMBER BANDED					
1964	1603	127.	44.	37.	40.	17.
1965	1595	0.	62.	76.	44.	28.
1966	1157	0.	0.	82.	61.	24.

MATRIX OF EXPECTED VALUES -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES (MODEL 1)

127.0	42.0	45.8	34.0	16.2
0.0	64.0	69.6	51.8	24.6
0.0	0.0	79.6	59.2	28.2

MATRIX OF CHI-SQUARE VALUES -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES (MODEL 1)

0.00	0.09	1.67	1.05	0.04
0.0	0.06	0.58	1.17	0.46
0.0	0.0	0.07	0.05	0.62

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT MODEL 1 -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES

CHI-SQUARED VALUE (SAMPLE) = 5.87
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 11.10
DEGREES OF FREEDOM = 5

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 5.87 = 0.31892308

ESTIMATED COVARIANCE AND CORRELATION COEFFICIENTS: S(I) = SURVIVAL RATE IN YEAR I
F(I) = RECOVERY RATE IN YEAR I

I	COVAR(S(I),F(I))	COVAR(F(I),F(I+1))	COVAR(S(I+1),F(I))
1964	-0.000032183	0.0	0.0
1965	0.000031925	0.0	0.0

I	COVAR(S(I),F(I+1))	COVAR(S(I),S(I+1))	COVAR(F(I),F(I+2))
1964	-0.000107982	-0.001672101	0.0
1965	-0.000222522	************	0.0

I	CORR(S(I),F(I))	CORR(S(I),S(I+1))	CORR(S(I),F(I+1))
1964	-0.070637676	-0.382528884	-0.385493345
1965	0.118931397	************	************

CORR(AVE SURVIVAL,AVE RECOVERY RATE) = -0.5710
COVAR(AVE SURVIVAL,AVE RECOVERY RATE) = -0.000082691

(THE ABOVE COVARIANCE AND CORRELATION COEFFICIENTS ARE ESTIMATES OF THE DEGREE TO WHICH THE SAMPLING VARIANCES OF SOME PARAMETER ESTIMATORS ARE RELATED)

In general, the same model structure applies for arbitrary values of k and ℓ. We will illustrate Model 2 with a set of adult male mallard *(Anas platyrhynchos)* data from birds banded before the hunting season (August and September) in the San Luis Valley of Colorado. Banding occurred each year from 1963 through 1971 ($k=9$) and recoveries are available through 1971 ($\ell=9, s=0$). The data and estimates are displayed in Example 2.2 in the form of output from the program ESTIMATE (*cf.* Example 2.2d,e).

Estimation of Parameters

Estimable parameters are of the same type as those obtained under Model 1: point estimates of the parameters f_i for $i=1,\ldots,\ell$ and S, and estimated sampling covariances, correlations, standard errors, and confidence intervals. The estimators are obtained by the principle of ML, and are essentially unbiased when Model 2 is the correct model. However, estimates of f_i for $i=k+1, k+2,\ldots,\ell$ tend to be poor because they are based on so little data.

No formulae will be given for estimators under Model 2, because they cannot be expressed in a simple, useful form. The likelihood equations must be solved numerically, for example by the Newton-Raphson technique or Fisher's method of scoring (cf. Seber 1973, Chapter 1). A discussion of iterative solutions to the ML equations is given in Appendix B. The likelihood equations are discussed in Appendix A. The FORTRAN program prints estimates of f_i for $i=1,\ldots,k$ only and the average recovery rate is computed from these first k estimates of f_i,

$$\hat{\bar{f}}=\frac{1}{k}\sum_{i=1}^{k}\hat{f}_i$$

The assumptions made under Model 2 are closely related to the hypothesis of compensatory natural mortality (cf. Anderson and Burnham 1976). For example, up to some level, the exploitation rate may not affect the annual survival rate of the population (also see Section 3.3 for the age-specific case). This implies that the population is regulated by factors other than exploitation, at least up to some level of exploitation. We note, however, that the power of the test of Model 2 vs. Model 1 is often fairly low if banded samples are small and/or only a few data sets are available for analysis. For these reasons and others, the study of compensatory natural mortality must be carefully approached (cf. Anderson and Burnham 1976, Nichols et al. 1984).

Goodness of Fit Test

A chi-square goodness of fit test is made to assess the fit of the model to the data. The observed data and the estimated expected values (using $\hat{f}_i$ and $\hat{S}$ in place of the unknown parameters in the expectations given in Table 2.4) are compared via the following statistic:

$$\chi^2=\sum_{ij}\sum\frac{(\mathrm{O}_{ij}-\mathbf{E}_{ij})^2}{\mathbf{E}_{ij}}$$

where O_{ij} represents the observed data and $\mathbf{E}_{ij}$ are ML estimates based on Model 2.

This is approximately distributed as chi-square under the null hypothesis that the model is correct. The test has $(k(k+1)/2)+k(\ell-k)-(\ell+1)$ degrees of freedom, when no cells are combined due to small expected values. Program ESTIMATE combines cells as needed and reduces the degrees of freedom accordingly. The pooled observed data and estimated expected values are printed along with the test statistic.

An Example

The results of a computer analysis of the recovery data for mallards banded in the San Luis Valley of Colorado are used to illustrate Model 2 (Example 2.2, specifically 2.2d and e). The structure of the model and the banding and recovery data are displayed. The ML estimates of f_i are given with their standard errors and 95% confidence intervals are printed. The estimate of the constant survival rate and its standard error and 95% confidence interval on S are printed (in Example 2.2d, $\hat{S}$ = 63.75% and se$(\hat{S})$ = 1.53%). Material related to the goodness of fit test follows. Expected values are printed to allow easy comparison with the observed data. A matrix of chi-square values $\frac{(O-\mathbf{E})^2}{\mathbf{E}}$ is printed. Large values (of magnitude 4-5 or more) tend to indicate lack of fit for the particular cell. An overall assessment of the fit is made by summing all the chi-square values, giving 40.7 in Example 2.2e with 32 df (some pooling was necessary). We see no reason to reject Model 2 based on this test. (Before deciding that Model 2 is adequate for these data, we must examine the results of tests *between* models. This subject will be treated in Section 2.6). Estimates of sampling covariances and correlations are printed last. We see that $\hat{f}_i$ and $\hat{S}$ are substantially negatively correlated as i increases. Note that these data are analyzed under Model 1 for comparison in Example 2.2a-c.

2.4 Model 3

Model 3 is the simplest possible age-independent model of band recoveries. It is based on the assumption that recovery rates (and therefore harvest rates H and band reporting rates λ) and survival rates are constant from year to year and independent of the age of the bird or its capture history. Model 3 is based on only two parameters, f the constant recovery rate, and S the constant survival rate. Note that Model 3 is a special case of Model 2 where $f_1 = f_2 = \cdots = f_\ell \equiv f$ is assumed.

The structure of Model 3 is given in Table 2.5 in terms of the numbers N_i banded each year and the parameters f and S. The expected values in Table 2.5 reflect the assumptions of the model.

Table 2.5. *Expected numbers of band recoveries under Model 3 for a banding study with $k=3$, $\ell=5$, and $s=2$.*

Year banded	Number banded	Year of recovery 1	2	3	4	$\ell=5$
1	N_1	N_1f	N_1Sf	N_1SSf	N_1SSSf	N_1SSSSf
2	N_2		N_2f	N_2Sf	N_2SSf	N_2SSSf
$k=3$	N_3			N_3f	N_3Sf	N_3SSf

As before, the structure of the model is unchanged for arbitrary values of k and ℓ. Example 2.3 illustrates this model with the mallard data examined under Model 2, Example 2.2 in the previous section.

Many methods have been proposed for estimating parameters from a model making these assumptions (e.g., Hickey 1952; Haldane 1955; Chapman and Robson 1960; and Seber 1973). The ML estimates under Model 3 are asymptotically fully efficient, essentially unbiased, and represent an improvement over other published methods.

Estimation of Parameters

Program ESTIMATE gives the Maximum Likelihood estimates of the parameters f and S, their standard errors, 95% confidence intervals, and the sampling correlation of $\hat{f}$ and $\hat{S}$. Although the structure of the model is quite simple, the ML equations are complex. Simple, closed-form expressions for the ML estimators do not exist. As in Model 2, the likelihood equations must be solved iteratively (see Appendices A and B).

Three summary statistics are defined for Model 3:

$$N = \sum_{i=1}^{k} N_i, \qquad T = \sum_{i=1}^{k} R_i, \qquad Q = \sum_{i=1}^{k} \sum_{j=i}^{\ell} (j-i) R_{ij}.$$

These are convenient working statistics and are computed and printed by program ESTIMATE.

Goodness of Fit Test

A chi-square goodness of fit test is made to allow the fit of the model to be assessed, as was done for Models 1 and 2. This test has $(k(k+1)/2)+k(\ell-k)-2$ degrees of freedom if no pooling is necessary for cells with small expectations. As with Models 1 and 2, the arrays of expected values and chi-square values printed show any pooling that was necessary. The array of chi-square values shows an additional column. It gives the values of $((N_i-R_i)-(N_i-\hat{\mathbf{E}}(R_i)))^2/(N_i-\hat{\mathbf{E}}(R_i))$. These quantities are necessary for a valid goodness of fit test in Models 2 and 3 (they are not printed in Model 2).

Example 2.2a

MODEL 1
****

ANALYSIS UNDER THE ASSUMPTIONS OF TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(A SYNTHESIS OF MODELS DEVELOPED BY SEBER (1970. BIOMETRIKA) AND ROBSON AND YOUNGS (1971. CORNELL BIOMETRICS UNIT PAPER 369))

SPECIFICALLY, THE MODEL STRUCTURE IS:

F(1)	S(1)F(2)	S(1)S(2)F(3)	S(1)S(2)S(3)F(4)	S(1)S(2)S(3)S(4)F(5)
	F(2)	S(2)F(3)	S(2)S(3)F(4)	S(2)S(3)S(4)F(5)
		F(3)	S(3)F(4)	S(3)S(4)F(5)
			F(4)	S(4)F(5)

MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

BANDING AND RECOVERY INPUT DATA

YEAR	NUMBER BANDED	RECOVERY MATRIX								
1963	231	10	13	6	1	1	3	1	2	0
1964	649	0	58	21	16	15	13	6	1	1
1965	885	0	0	54	39	23	18	11	10	6
1966	590	0	0	0	44	21	22	9	9	3
1967	943	0	0	0	0	55	39	23	11	12
1968	1077	0	0	0	0	0	66	46	29	18
1969	1250	0	0	0	0	0	0	101	59	30
1970	938	0	0	0	0	0	0	0	97	22
1971	312	0	0	0	0	0	0	0	0	21

INTERMEDIATE STATISTICS

I	C(I)	R(I)	T(I)	N(I)	GAMMA(I)	RHO(I)
1963	10.0	37.0	37.0	231.0	0.0	0.16017
1964	71.0	131.0	158.0	649.0	0.0	0.20185
1965	81.0	161.0	248.0	885.0	0.0	0.18192
1966	100.0	108.0	275.0	590.0	0.0	0.18305
1967	115.0	140.0	315.0	943.0	0.0	0.14846
1968	161.0	159.0	359.0	1077.0	0.0	0.14763
1969	197.0	190.0	388.0	1250.0	0.0	0.15200
1970	218.0	119.0	310.0	938.0	0.0	0.12687
1971	113.0	21.0	113.0	312.0	0.0	0.06731
TOTALS		1066.0		6875.0		

I	RECOVERY RATE F(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL	DIRECT RECOVERY RATE B(I,I)/N(I) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
1963	4.329	1.339	1.705 - 6.953	4.329	1.339	1.705 - 6.953
1964	9.070	1.067	6.978 - 11.163	8.937	1.120	6.742 - 11.132
1965	5.942	0.688	4.594 - 7.290	6.102	0.805	4.525 - 7.679
1966	6.656	0.786	5.117 - 8.196	7.458	1.082	5.338 - 9.577
1967	5.420	0.584	4.276 - 6.564	5.832	0.763	4.337 - 7.328
1968	6.621	0.621	5.404 - 7.837	6.128	0.731	4.696 - 7.561
1969	7.718	0.644	6.455 - 8.980	8.080	0.771	6.569 - 9.591
1970	8.922	0.832	7.291 - 10.552	10.341	0.994	8.392 - 12.290
1971	6.731	1.418	3.951 - 9.511	6.731	1.418	3.951 - 9.511

ARITHMETIC MEAN RECOVERY RATE (EXCEPT YEAR K) = 6.835
STANDARD ERROR OF THE MEAN RECOVERY RATE = 0.302
95% CONFIDENCE INTERVAL FOR MEAN RECOVERY RATE = 6.24 - 7.43

TEST OF THE NULL HYPOTHESIS THAT THE FIRST-YEAR (DIRECT) RECOVERY RATES ARE CONSTANT EACH YEAR:

CHI-SQUARED (SAMPLE) = 26.52
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 15.50
DEGREES OF FREEDOM = 8

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 26.52 = 0.00085376

Example 2.2b

MODEL 1 -- ANALYSIS UNDER THE ASSUMPTIONS OF TIME SPECIFIC SURVIVAL AND RECOVERY RATES

MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

	SURVIVAL S(I) (%)			
YEAR	SURVIVAL	STANDARD ERROR	COEFFICIENTS OF VARIATION	95% CONFIDENCE INTERVAL
1963	57.56	11.34	19.70	35.33 - 79.78
1964	60.79	7.77	12.78	45.56 - 76.02
1965	66.42	8.03	12.08	50.69 - 82.15
1966	77.99	9.78	12.54	58.82 - 97.16
1967	63.51	7.32	11.52	49.17 - 77.85
1968	53.33	5.86	10.99	41.84 - 64.82
1969	58.55	7.04	12.02	44.75 - 72.35
1970	53.57	13.05	24.37	27.98 - 79.15

ARITHMETIC MEAN SURVIVAL (%) = 61.46
STANDARD ERROR OF ARITHMETIC MEAN = 2.25
95% CONFIDENCE INTERVAL FOR ARITHMETIC MEAN 57.05- 65.88

MEAN LIFE SPAN AS AN ADULT = 2.05
STANDARD ERROR OF THE MEAN LIFE SPAN = 0.15
95% CONFIDENCE INTERVAL OF LIFE SPAN 1.78 - 2.40

YEAR	NUMBER BANDED	RECOVERY MATRIX								
1963	231	10.	13.	6.	1.	1.	3.	0.	0.	3.
1964	649	0.	58.	21.	16.	15.	13.	6.	0.	2.
1965	885	0.	0.	54.	39.	23.	18.	11.	10.	6.
1966	590	0.	0.	0.	44.	21.	22.	9.	9.	3.
1967	943	0.	0.	0.	0.	55.	39.	23.	11.	12.
1968	1077	0.	0.	0.	0.	0.	66.	46.	29.	18.
1969	1250	0.	0.	0.	0.	0.	0.	101.	59.	30.
1970	938	0.	0.	0.	0.	0.	0.	0.	97.	22.
1971	312	0.	0.	0.	0.	0.	0.	0.	0.	21.

MATRIX OF EXPECTED VALUES -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES (MODEL 1)

10.0	12.1	4.9	3.6	2.3	1.8	0.0	0.0	2.2
0.0	58.9	23.6	17.7	11.3	8.8	5.5	0.0	5.3
0.0	0.0	52.6	39.4	25.2	19.6	12.3	8.4	3.5
0.0	0.0	0.0	39.3	25.1	19.6	12.2	8.3	3.5
0.0	0.0	0.0	0.0	51.1	39.9	24.9	17.0	7.2
0.0	0.0	0.0	0.0	0.0	71.3	44.5	30.4	12.8
0.0	0.0	0.0	0.0	0.0	0.0	96.5	65.8	27.8
0.0	0.0	0.0	0.0	0.0	0.0	0.0	83.7	35.3
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21.0

MATRIX OF CHI-SQUARE VALUES -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES (MODEL 1)

0.00	0.06	0.27	1.92	0.76	0.78	0.0	0.0	0.27
0.0	0.01	0.28	0.16	1.22	2.00	0.05	0.0	2.08
0.0	0.0	0.04	0.00	0.19	0.14	0.13	0.32	1.73
0.0	0.0	0.0	0.57	0.67	0.30	0.85	0.05	0.08
0.0	0.0	0.0	0.0	0.30	0.02	0.14	2.10	3.27
0.0	0.0	0.0	0.0	0.0	0.39	0.05	0.06	2.10
0.0	0.0	0.0	0.0	0.0	0.0	0.21	0.70	0.18
0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.12	5.02
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00

(FREQUENCIES WERE COMBINED WHERE EXPECTED VALUES WERE SMALL)

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT MODEL 1 -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES

CHI-SQUARED VALUE (SAMPLE) = 31.57
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 37.70
DEGREES OF FREEDOM = 25

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 31.57 = 0.17076620

Example 2.2c

ESTIMATED COVARIANCE AND CORRELATION COEFFICIENTS: S(I) = SURVIVAL RATE IN YEAR I
F(I) = RECOVERY RATE IN YEAR I

I	COVAR(S(I),F(I))	COVAR(F(I),F(I+1))	COVAR(S(I+1),F(I))
1963	-0.000107862	0.0	0.0
1964	-0.000013032	0.0	0.0
1965	0.000041399	0.0	0.0
1966	0.000203910	0.0	0.0
1967	0.000100093	0.0	0.0
1968	0.000090932	0.0	0.0
1969	0.000085214	0.0	0.0
1970	0.000196484	0.0	0.0

I	COVAR(S(I),F(I+1))	COVAR(S(I),S(I+1))	COVAR(F(I),F(I+2))
1963	-0.000318077	-0.002239599	0.0
1964	-0.000183524	-0.002062818	0.0
1965	-0.000334440	-0.004204500	0.0
1966	-0.000257104	-0.003033377	0.0
1967	-0.000225412	-0.001806838	0.0
1968	-0.000183695	-0.001488372	0.0
1969	-0.000383268	-0.002955420	0.0
1970	-0.001601311	************	0.0

I	CORR(S(I),F(I))	CORR(S(I),S(I+1))	CORR(S(I),F(I+1))
1963	-0.071052211	-0.254235898	-0.262843735
1964	-0.015713043	-0.330776038	-0.343470253
1965	0.075007441	-0.535659298	-0.530433255
1966	0.265420435	-0.423992558	-0.450288826
1967	0.234346949	-0.421212710	-0.496489364
1968	0.249875857	-0.360534445	-0.486498548
1969	0.187963436	-0.321557236	-0.654251044
1970	0.180882272	************	************

CORR(AVE SURVIVAL,AVE RECOVERY RATE) = -0.6631
COVAR(AVE SURVIVAL,AVE RECOVERY RATE) = -0.000045151

(THE ABOVE COVARIANCE AND CORRELATION COEFFICIENTS ARE ESTIMATES OF THE DEGREE TO WHICH THE SAMPLING VARIANCES OF SOME PARAMETER ESTIMATORS ARE RELATED)

Example 2.2d

MODEL 2
****

ANALYSIS ASSUMING CONSTANT SURVIVAL BUT TIME-SPECIFIC RECOVERY RATES

THIS MODEL WAS DEVELOPED BY DRS. BROWNIE AND ROBSON AT THE CORNELL BIOMETRICS UNIT.

SPECIFICALLY, THE MODEL STRUCTURE IS:

F(1)	SF(2)	SSF(3)	SSSF(4)	SSSSF(5)
	F(2)	SF(3)	SSF(4)	SSSF(5)
		F(3)	SF(4)	SSF(5)
			F(4)	SF(5)

MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

YEAR	NUMBER BANDED	RECOVERY MATRIX								
1963	231	10	13	6	1	1	3	1	2	0
1964	649	0	58	21	16	15	13	6	1	1
1965	885	0	0	54	39	23	18	11	10	6
1966	590	0	0	0	44	21	22	9	9	3
1967	943	0	0	0	0	55	39	23	11	12
1968	1077	0	0	0	0	0	66	46	29	18
1969	1250	0	0	0	0	0	0	101	59	30
1970	938	0	0	0	0	0	0	0	97	22
1971	312	0	0	0	0	0	0	0	0	21

RECOVERY RATE F(I) (%)

I	ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
1	4.285	1.329	1.680 - 6.889
2	8.939	1.013	6.953 - 10.925
3	5.877	0.637	4.628 - 7.126
4	6.840	0.681	5.505 - 8.175
5	6.072	0.573	4.949 - 7.194
6	6.983	0.566	5.874 - 8.092
7	7.259	0.536	6.209 - 8.310
8	8.200	0.595	7.034 - 9.366
9	5.641	0.582	4.499 - 6.782

ARITHMETIC MEAN RECOVERY RATE = 6.68
STANDARD ERROR OF MEAN RECOVERY RATE = 0.29
95% CONFIDENCE INTERVAL FOR MEAN RECOVERY RATE = 6.10 - 7.25

CONSTANT SURVIVAL RATE (%) = 63.75
STANDARD ERROR OF THE CONSTANT SURVIVAL RATE = 1.53
95% CONFIDENCE INTERVAL FOR THE CONSTANT SURVIVAL RATE 60.75 - 66.75

MEAN LIFE SPAN AS AN ADULT = 2.22
STANDARD ERROR OF THE MEAN LIFE SPAN = 0.12
95% CONFIDENCE INTERVAL OF LIFE SPAN 2.01 - 2.47

Example 2.2e

MATRIX OF EXPECTED VALUES -- ASSUMING A CONSTANT SURVIVAL RATE AND TIME-SPECIFIC RECOVERY RATES (MODEL 2)

9.9	13.2	5.5	4.1	2.3	1.7	0.0	0.0	2.3
0.0	58.0	24.3	18.0	10.2	7.5	5.0	0.0	5.1
0.0	0.0	52.0	38.6	21.8	16.0	10.6	7.6	3.4
0.0	0.0	0.0	40.4	22.8	16.7	11.1	8.0	3.5
0.0	0.0	0.0	0.0	57.3	42.0	27.8	20.0	8.8
0.0	0.0	0.0	0.0	0.0	75.2	49.8	35.9	15.7
0.0	0.0	0.0	0.0	0.0	C.0	90.7	65.3	28.7
0.0	0.0	0.0	0.0	0.0	C.0	0.0	76.9	33.7
0.0	0.0	0.0	0.0	0.0	C.0	0.0	0.0	17.6

MATRIX OF CHI-SQUARE VALUES--ASSUMING A CONSTANT SURVIVAL RATE AND TIME-SPECIFIC RECOVERY RATE (MODEL 2)

0.0	0.0	0.0	2.3	0.7	1.0	0.0	0.0	0.2
0.0	0.0	0.5	0.2	2.2	4.1	0.2	0.0	1.9
0.0	0.0	0.1	0.0	0.1	C.2	0.0	0.7	2.1
0.0	0.0	0.0	0.3	0.1	1.6	0.4	0.1	0.1
0.0	0.0	0.0	0.0	0.1	0.2	0.8	4.1	1.2
0.0	0.0	0.0	0.0	0.0	1.1	0.3	1.3	0.3
0.0	0.0	0.0	0.0	0.0	C.0	1.2	0.6	0.1
0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.2	4.1
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7

(FREQUENCIES WERE COMBINED WHERE EXPECTED VALUES WERE LESS THAN 2.0)

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT THE MODEL ASSUMING A CONSTANT SURVIVAL AND TIME-SPECIFIC RECOVERY RATES

CHI-SQUARE VALUE (SAMPLE) = 40.70
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 46.19
DEGREES OF FREEDOM = 32

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 40.70 = 0.13922474

ESTIMATED COVARIANCE AND CORRELATION COEFFICIENTS:

I	COVAR(F(I),S)	CORR(F(I),S)	COVAR(F(I),F(I+1))	CORR(F(I),F(I+1))
1963	-0.000003933	-0.019319387	-0.000002906	-0.021585861
1964	-0.000012430	-C.080068213	-0.000001620	-0.025084596
1965	-0.000012686	-0.129934977	-0.000000193	-0.004446413
1966	-0.000024639	-0.236107416	0.000001310	0.033586147
1967	-0.000022398	-0.255260811	0.000001659	0.051184037
1968	-0.000026649	-0.307364417	0.000002312	0.076204540
1969	-0.000027985	-0.340668818	0.000003560	0.111569167
1970	-0.000038370	-0.420859319	0.000005487	0.158340386
1971	-0.000039472	-0.442426983	-0.000039472	-0.442426983

5 ITERATIONS

An Example

A computer example for mallard data is presented below to illustrate Model 3 (Example 2.3). The model structure and the input data are printed. ML estimates are printed as are their estimated standard errors, coefficients of variation, and the 95% confidence intervals on S and f. Also shown is the sampling correlation of $\hat{S}$ and $\hat{f}$. Annual survival rate was estimated to be 63.38±1.46% (i.e., se$(\hat{S})=1.46\%$) and the recovery rate was 6.94±0.26%. The estimated sampling correlation between the estimators of S and f is substantial, −0.66. Visual comparison of the observed data and the expected values (computed using the structure of Model 3, Table 2.5, and the estimates of f and S) suggests a poor fit of the model to the data. The poor fit is verified by the chi-square test with 40 df and a test statistic value of 65.58. The null hypothesis that the data fit Model 3 is rejected at the 0.01 level of significance. The rejection indicates that Model 3 and the assumptions upon which it is based are not satisfactory for the analysis of this data set. Actually, this model is rarely acceptable because it is too restrictive. Particularly, the assumption that recovery rates are constant from year to year is unrealistic. The model is useful in that it provides a simple starting base and is a convenient point to begin examining alternatives.

Several hundred sets of real data have been examined with Model 3 and a very high percentage of the data fail to fit the model. This is noteworthy because the assumptions are essentially the same as those of the commonly used composite dynamic method for analyzing adult data (cf. Anderson and Burnham 1976).

2.5 Model 0

Models 2 and 3 were introduced as restrictions on Model 1. We now present a model that is a generalization of Model 1. Model 0 allows year-specific recovery and survival rates, as does Model 1. The generalization involves the assumption that the recovery rates the first year after banding are different from recovery rates of previously banded birds. This model is useful in situations where the band reporting rate is different near banding sites, and therefore affects primarily newly banded birds (cf. Henny and Burnham 1976). Low reporting rates near banding sites are probably the result of hunter familiarity with banded birds. In other banding sites reporting rates are higher because conservation agency personnel actively solicit bands from hunters. Thus the reporting rate for newly banded birds, which tend to still be clustered around the banding site at the beginning of the hunting season, may be different from that of birds banded in previous years. Birds banded in previous years are more widely dispersed and are not especially affected by reporting rates near the banding sites. The recovery rates for newly banded birds are denoted as f_i^*. Other parameters, f_i and S_i, are those defined under Model 1; however their ML estimators are different under Model 0.

We also note that Model 0 can sometimes be used as an approximation for the analysis of in-season banding (banding conducted during the hunting season). (This subject is discussed in Section 3.5 for the age-specific case.)

The structure of Model 0 is given in Table 2.6 in terms of expected band recoveries expressed as functions of N_i, S_i, f_i, and f_i^* (the recovery rates the first year after banding). The information in Table 2.6 reflects the assumptions upon which the model is explicitly based.

Table 2.6. *Expected numbers of band recoveries under Model 0 for a banding study with $k=3$, $\ell=5$, and $s=2$.*

		Year of recovery				
Year banded	Number banded	1	2	3	4	$\ell=5$
1	N_1	$N_1f_1^*$	$N_1S_1f_2$	$N_1S_1S_2f_3$	$N_1S_1S_2S_3f_4$	$N_1S_1S_2S_3S_4f_5$
2	N_2		$N_2f_2^*$	$N_2S_2f_3$	$N_2S_2S_3f_4$	$N_2S_2S_3S_4f_5$
$k=3$	N_3			$N_3f_3^*$	$N_3S_3f_4$	$N_3S_3S_4f_5$

Estimation of Parameters

Model 0 has a large number of parameters, $3k+2s-2$, but they are not all separately estimable. For a triangular data array, i.e., $s=0$, f_k and S_{k-1} are not separately estimable. For the case $s>0$, the additional survival rates $S_k, \ldots, S_{\ell-1}$ and recovery rates $f_{k+1}, \ldots, f_\ell$ are not separately estimable. Also note that f_1 does not exist under Model 0.

Example 2.3

MODEL 3
****

ANALYSIS ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES
(A GENERALIZATION AND EXTENSION OF THE MODELS DEVELOPED BY CHAPMAN AND ROBSON (1960. BIOMETRICS) AND HALDANE (1955. PROC. XI INT. ORN. CONGR.) — SEE BOTTOM OF PAGE 245 OF BOOK BY SEBER))

SPECIFICALLY, THE MODEL STRUCTURE IS:

```
F     SF     SSF     SSSF     SSSSF
       F      SF      SSF      SSSF
              F       SF       SSF
                      F        SF
```

MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

YEAR	NUMBER BANDED				RECOVERY MATRIX					
1963	231	10	13	6	1	1	3	1	2	0
1964	649	0	58	21	16	15	13	6	1	1
1965	885	0	0	54	39	23	18	11	10	6
1966	590	0	0	0	44	21	22	9	9	3
1967	943	0	0	0	0	55	39	23	11	12
1968	1077	0	0	0	0	0	66	46	29	18
1969	1250	0	0	0	0	0	0	101	59	30
1970	938	0	0	0	0	0	0	0	97	22
1971	312	0	0	0	0	0	0	0	0	21

INTERMEDIATE STATISTICS

N = 6875. T = 1066. Q = 1137.

PARAMETER		ESTIMATE (%)	STD. ERR. (%)	COEF. VARIAT. (%)	95% CONFIDENCE INTERVAL
SURVIVAL RATE	(S)	63.38	1.46	2.31	60.51 -- 66.24
RECOVERY RATE	(F)	6.94	0.26	3.75	6.43 -- 7.46

CORRELATION(S,F) = -.66317656

MEAN LIFE SPAN AS AN ADULT = 2.19
STANDARD ERROR OF MEAN LIFE SPAN = 0.11
95% CONFIDENCE INTERVAL OF LIFE SPAN 1.99 - 2.43

MATRIX OF EXPECTED VALUES -- ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES (MODEL 3)

16.0	10.2	6.4	4.1	2.6	1.6	0.0	0.0	2.1
0.0	45.1	28.6	18.1	11.5	7.3	4.6	0.0	4.8
0.0	0.0	61.5	39.0	24.7	15.6	9.9	6.3	4.0
0.0	0.0	0.0	41.0	26.0	16.5	10.4	6.6	4.2
0.0	0.0	0.0	0.0	65.5	41.5	26.3	16.7	10.6
0.0	0.0	0.0	0.0	0.0	74.8	47.4	30.0	19.0
0.0	0.0	0.0	0.0	0.0	0.0	86.8	55.0	34.9
0.0	0.0	0.0	0.0	0.0	0.0	0.0	65.1	41.3
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21.7

MATRIX OF CHI-SQUARE VALUES -- ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES (MODEL 3)

2.28	0.79	0.03	2.33	0.97	1.13	0.0	0.0	0.37	0.20
0.0	3.71	2.00	0.24	1.08	4.51	0.42	0.0	1.61	0.23
0.0	0.0	0.91	0.00	0.12	0.35	0.12	2.20	1.02	0.00
0.0	0.0	0.0	0.22	0.95	1.87	0.20	0.86	0.34	0.02
0.0	0.0	0.0	0.0	1.68	0.15	0.41	1.93	0.19	0.54
0.0	0.0	0.0	0.0	0.0	1.03	0.04	0.04	0.06	0.17
0.0	0.0	0.0	0.0	0.0	0.0	2.32	0.29	0.68	0.16
0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.58	9.01	0.19
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.02	0.00

(FREQUENCIES WERE COMBINED WHERE EXPECTED VALUES WERE SMALL)

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT THE MODEL ASSUMMING CONSTANT SURVIVAL AND RECOVERY RATES

CHI-SQUARE VALUE (SAMPLE) = 65.58
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 55.76
DEGREES OF FREEDOM = 40

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 65.58 = 0.00655472

9 ITERATIONS.

The ML estimators of the parameters of interest are:

$$\hat{f}_i^* = \frac{R_{ii}}{N_i} \qquad , i = 1, \ldots, k,$$

$$\hat{f}_i = \frac{R_i - R_{ii}}{N_i} \frac{C_i - R_{ii}}{T_i - R_i - C_i + R_{ii}} \qquad , i = \begin{cases} 2, \ldots, k-1, \text{ if } s = 0 \\ 2, \ldots, k \quad , \text{ if } s > 0, \end{cases}$$

$$\hat{S}_i = \frac{R_i - R_{ii}}{N_i} \frac{N_{i+1}}{R_{i+1} - R_{i+1,i+1}} \left(1 - \frac{C_{i+1} - R_{i+1,i+1}}{T_{i+1} - R_{i+1}}\right) \qquad , i = \begin{cases} 1, \ldots, k-2, \text{ if } s = 0 \\ 1, \ldots, k-1, \text{ if } s > 0, \end{cases}$$

The ML estimators of S_i and f_i under Model 0 are slightly biased for values of N_i usually encountered. The following adjusted estimators, which are essentially unbiased, are printed by the FORTRAN program ESTIMATE:

$$\tilde{f}_i = \frac{R_i - R_{ii}}{N_i} \frac{C_i - R_{ii}}{T_i - R_i - C_i + R_{ii} + 1},$$

$$\tilde{S}_i = \frac{R_i - R_{ii}}{N_i} \frac{N_{i+1} + 1}{R_{i+1} - R_{i+1,i+1} + 1} \left(1 - \frac{C_{i+1} - R_{i+1,i+1}}{T_{i+1} - R_{i+1}}\right).$$

Other parameters are estimable only as products and, while of little biological interest, are required for the goodness of fit tests. The corresponding ML estimators are:

$$\widehat{S_{k-1} f_k} = \frac{R_{k-1} - R_{k-1,k-1}}{N_{k-1}} \qquad , \text{ if } s = 0,$$

$$\widehat{S_k \cdots S_{k+i-1} f_{k+i}} = \frac{R_k - R_{kk}}{N_k} \frac{C_{k+i}}{\sum_{j=1}^{s} C_{k+j}} \qquad , i = 1, \ldots, s \text{ if } s > 0.$$

Sampling Variances, Standard Errors, and Confidence Intervals

Estimators of the sampling variances of estimators of the parameters of interest are:

$$\text{var}(\hat{f}_i^*) = \hat{f}_i^*(1 - \hat{f}_i^*)/N_i \qquad , i = 1, \ldots, k,$$

$$\text{var}(\tilde{f}_i) = (\tilde{f}_i)^2 \left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i} + \frac{1}{T_i - R_i - C_i + R_{ii}} + \frac{1}{C_i - R_{ii}}\right] \qquad , i = \begin{cases} 2, \ldots, k-1, \text{ if } s = 0 \\ 2, \ldots, k \quad , \text{ if } s > 0, \end{cases}$$

$$\text{var}(\tilde{S}_i) = (\tilde{S}_i)^2 \left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i} + \frac{1}{R_{i+1} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} + \frac{1}{T_{i+1} - R_{i+1} - C_{i+1} + R_{i+1,i+1}} - \frac{1}{T_{i+1} - R_{i+1}}\right] \qquad , i = \begin{cases} 1, \ldots, k-2, \text{ if } s = 0. \\ 1, \ldots, k-1, \text{ if } s > 0. \end{cases}$$

Program ESTIMATE prints the estimated standard errors of the estimates, e.g., $\text{se}(\hat{f}_i^*) = \sqrt{\text{var}(\hat{f}_i^*)}$. Also, the approximate 95% confidence intervals on the various parameters are computed and printed, e.g., $\tilde{f}_i \pm 1.96\ \text{se}(\tilde{f}_i)$ or $\tilde{S}_i \pm 1.96\ \text{se}(\tilde{S}_i)$.

Sampling Covariances and Correlations

Because the estimates of f_i, f_i^*, and S_i are all derived from the same data set, sampling correlations exist between these estimators. These quantities are estimates of the linear relationship between the sampling variations of the estimators. These sampling covariances and correlations are important because they determine the relationships between parameter estimates much more than any structural relationships between the true parameters.

The estimators of the non-negligible covariances are:

$$\operatorname{cov}(\hat{f}_i^*, \tilde{f}_i) = -\hat{f}_i^* \tilde{f}_i / N_i \qquad , i = \begin{cases} 2, \ldots, k-1 & \text{if } s = 0 \\ 2, \ldots, k & \text{if } s > 0, \end{cases}$$

$$\operatorname{cov}(\hat{f}_i^*, \tilde{S}_i) = -\hat{f}_i^* \tilde{S}_i / N_i \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } s = 0 \\ 1, \ldots, k-1 & \text{if } s > 0, \end{cases}$$

$$\operatorname{cov}(\hat{f}_{i+1}^*, \tilde{S}_i) = \hat{f}_{i+1}^* \tilde{S}_i / N_i \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } s = 0 \\ 1, \ldots, k-1 & \text{if } s > 0, \end{cases}$$

$$\operatorname{cov}(\tilde{f}_i, \tilde{S}_i) = \tilde{f}_i \tilde{S}_i \left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i} \right] \qquad , i = \begin{cases} 2, \ldots, k-2 & \text{if } s = 0 \\ 2, \ldots, k-1 & \text{if } s > 0, \end{cases}$$

$$\operatorname{cov}(\tilde{f}_{i+1}, \tilde{S}_i) = -\tilde{S}_i \tilde{f}_{i+1} \left[\frac{1}{R_{i+1} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} \right] \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } s = 0 \\ 1, \ldots, k-1 & \text{if } s > 0, \end{cases}$$

$$\operatorname{cov}(\tilde{S}_i, \tilde{S}_{i+1}) = -\tilde{S}_i \tilde{S}_{i+1} \left[\frac{1}{R_{i+1} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} \right] \qquad , i = \begin{cases} 1, \ldots, k-3 & \text{if } s = 0 \\ 1, \ldots, k-2 & \text{if } s > 0. \end{cases}$$

Goodness of Fit Test

The estimators of f_i, f_i^* and S_i and particularly their sampling variances and covariances, are dependent upon the assumptions being made. These assumptions are specified in the model. It is crucial that realistic assumptions are made and that a correct model is used in the analysis of a particular data set.

The fit of the model to the data is assessed by a test statistic approximately distributed as chi-square under the null hypothesis that Model 0 is correct. Estimated expected numbers of band recoveries are computed with the formulae in Table 2.6 and the ML estimates $\hat{f}_i$, $\hat{f}_i^*$ and $\hat{S}_i$. The test statistic is

$$\chi^2 = \sum_{ij}\sum \frac{(\mathrm{O}_{ij} - \mathbf{E}_{ij})^2}{\mathbf{E}_{ij}}.$$

The degrees of freedom for this test are

$$(k(k+1)/2) - (3k-3) \qquad \text{if } s = 0,$$

$$(k(k+1)/2) + (ks) - (3k-2+s) \qquad \text{if } s > 0.$$

Some combining of cells may be necessary if corresponding estimated expected values $\mathbf{E}_{ij}$ are small (less than 2). When cells (estimated expected values $\mathbf{E}_{ij}$'s) are combined, the corresponding data values R_{ij}'s must also be combined; this is done by program ESTIMATE and the pooled data and expected values are printed along with the goodness of fit statistic. One degree of freedom is lost for each cell that is pooled.

Proper and Improper Use of Model 0

Proper use of Model 0 is limited to situations where parameters can reasonably be assumed age-independent (e.g., adult bandings) and the first year recovery rates f_i^* are different because of band reporting rate problems. Then Model 0 can be used to estimate the parameters S_i, f_i, and f_i^*.

Two other models, making quite different assumptions, are "confounded" with Model 0. Table 2.7 presents the structure of these models. The ML estimators of expected recoveries in each cell are the same under Model 0 and the two models depicted in Table 2.7. Therefore, a goodness of fit test or likelihood ratio test cannot distinguish among the three models. Model 0 and the two cases represented in Table 2.7 are all parameterized in such a way that they all have the same minimal sufficient statistic (cf. Brownie 1974b).

Table 2.7. *Expected numbers of band recoveries under two models similar to Model 0, with $k=3$, $\ell=5$, and $s=2$.*

Model Version	Year of recovery 1	2	3	4	5
1	N_1f_1	$N_1S_1^*f_2$ N_2f_2	$N_1S_1^*S_2f_3$ $N_2S_2^*f_3$ N_3f_3	$N_1S_1^*S_2S_3f_4$ $N_2S_2^*S_3f_4$ $N_3S_3^*f_4$	$N_1S_1^*S_2S_3S_4f_5$ $N_2S_2^*S_3S_4f_5$ $N_3S_3^*S_4f_5$
2	$N_1f_1^*$	$N_1S_1^*f_2$ $N_2f_2^*$	$N_1S_1^*S_2f_3$ $N_2S_2^*f_3$ $N_3f_3^*$	$N_1S_1^*S_2S_3f_4$ $N_2S_2^*S_3f_4$ $N_3S_3^*f_4$	$N_1S_1^*S_2S_3S_4f_5$ $N_2S_2^*S_3S_4f_5$ $N_3S_3^*S_4f_5$

It is difficult to find a meaningful biological interpretation for the first case represented in Table 2.7. The second case could arise as the result of age-specific survival and recovery rates (age-specific for two age classes: young and adults) or as a result of survival (and hence recovery rates) being affected by stress during banding or tagging. The latter may arise more often in fish tagging experiments or studies of small birds. Estimates of f_i^*, S_i^*, f_i, and S_i are not possible for the former case unless adults are also banded. This subject is discussed in detail in Section 3.9.

Unless we have knowledge of the biological conditions and are willing to believe that band reporting rates (but not harvest rates) are different the first year, we do not know which of the three models we are testing for fit. The improper use of Model 0 when one of the two alternative (and indistinguishable) models is in fact correct is illustrated in the following example.

An Example

Data for young male mallards banded before the hunting season in southern Ontario are analyzed in Example 2.4. The proper analysis of these data (Example 3.6, using the age-specific models in Chapter 3) confirms that survival and recovery rates of young and adults are different. The appropriate model for the Ontario mallard data is the second model in Table 2.7. The example was chosen to illustrate the improper use of Model 0 even though it appears to fit the data. The goodness of fit test will not detect the improper use of Model 0, because this test does not distinguish between Model 0 and the two alternatives of Table 2.7.

The structure of Model 0 and the banding and recovery data are displayed in Example 2.4a. The ML estimates of f_i and f_i^* are printed along with corresponding standard errors and confidence intervals. The estimates $\hat{\bar{f}}$ and $\hat{\bar{f}}^*$ are 4.48% and 10.12%, respectively. This large difference would tend to suggest that Model 0 is appropriate, and we also observe the chi-square goodness of fit test fails to reject Model 0 ($\chi^2=27.6$, 19 df). At this point we might be inclined to think that because Model 0 fits the data, the ML estimates of f_i, f_i^* and S_i are useful. Except for the estimates of f_i^* this is not true; the estimates and their sampling variances and covariances, etc., are useless because the correct model for the data is not Model 0. Instead, the correct model is version 2 of Table 2.7, a model that cannot be distinguished from Model 0 unless adult birds are also banded. (See Section 3.9 for a further illustration of this subject, and Example 3.6 for the correct analysis of these data.)

If survial rates are not age-specific, but band reporting rates significantly distort the first-year recovery rates, then Model 0 should be used, even though the sampling variances will be large. Use of the simpler Model 1 would produce biased estimates of f_i and S_i. However, one must weigh the loss of precision vs. this bias. In other words, it may be appropriate in some situations to gain precision and risk the consequences of somewhat biased estimates. Unless the null hypothesis that Model 1 fits the data is strongly rejected (e.g., at the 0.01 level of significance) Model 0 probably should not be used. The analysis of a large number of data sets has shown that Model 0 is extremely general, and it would be unusual that a given adult data set could not be well described by Model 0. Often one of the simpler models, such as Model 1 or Model 2, also fit the data satisfactorily, and therefore should be used.

2.6 Testing Between Models

We have emphasized that the determination of the adequacy of a model involves goodness of fit tests and specific tests between models. The goodness of fit tests were discussed in the above sections. This section presents tests between models.

Example 2.4a

MODEL 0
****

ANALYSIS UNDER THE ASSUMPTIONS OF TIME-SPECIFIC SURVIVAL AND RECOVERY RATES WHERE THE FIRST-YEAR RECOVERY RATES ARE DIFFERENT FROM RECOVERY RATES OF PREVIOUSLY-BANDED COHORTS.

THIS MODEL IS AN EXTENSION OF THE SEBER-ROBSON-YOUNGS MODEL (MODEL 1) IN THAT THE FIRST-YEAR RECOVERY RATES ARE ALLOWED TO DIFFER. THIS MODEL IS USEFUL IF THE BAND REPORTING RATE IS QUITE DIFFERENT THE FIRST YEAR AFTER BANDING (E.G., BAND COLLECTING ACTIVITIES CONDUCTED BY CONSERVATION AGENCY PERSONNEL OFTEN AFFECT PRIMARILY THE FIRST-YEAR RECOVERY RATES).

(MODEL DEVELOPED BY BROWNIE (1974. CORNELL BIOMETRICS UNIT PAPER 535-M)). REFER TO BROWNIE (1974) FOR A DISCUSSION OF THE USE OF THIS MODEL VS. MODEL 1.

SPECIFICALLY, THE MODEL STRUCTURE IS:

F*(1)	S(1)F(2)	S(1)S(2)F(3)	S(1)S(2)S(3)F(4)	S(1)S(2)S(3)S(4)F(5)
	F*(2)	S(2)F(3)	S(2)S(3)F(4)	S(2)S(3)S(4)F(5)
		F*(3)	S(3)F(4)	S(3)S(4)F(5)
			F*(4)	S(4)F(5)

YOUNG MALE MALLARDS BANDED PRESEASON IN SOUTHERN ONTARIO, 1965-72

BANDING AND RECOVERY INPUT DATA

YEAR	NUMBER BANDED	RECOVERY MATRIX								
1965	1570	132	48	33	13	8	9	8	5	1
1966	1462	0	175	33	8	11	10	4	7	1
1967	1611	0	0	165	39	23	12	13	5	7
1968	1733	0	0	0	193	51	24	13	12	5
1969	1848	0	0	0	0	193	43	39	15	9
1970	3456	0	0	0	0	0	367	113	56	32
1971	4488	0	0	0	0	0	0	392	176	70
1972	3584	0	0	0	0	0	0	0	342	101

I	RECOVERY RATE F(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL	FIRST-YEAR RECOVERY RATE F*(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
1965	*******	*******	******* - *******	8.408	0.700	7.035 - 9.780
1966	3.115	0.673	1.796 - 4.433	11.970	0.849	10.306 - 13.634
1967	4.716	0.900	2.953 - 6.480	10.242	0.755	8.761 - 11.723
1968	2.908	0.534	1.862 - 3.954	11.137	0.756	9.656 - 12.618
1969	3.894	0.640	2.640 - 5.148	10.444	0.711	9.049 - 11.838
1970	3.886	0.574	2.760 - 5.011	10.498	0.518	9.482 - 11.514
1971	6.676	0.833	5.044 - 8.308	8.734	0.421	7.908 - 9.560
1972	6.173	0.900	4.409 - 7.936	9.542	0.491	8.581 - 10.504
MEAN	4.481	0.324	3.845 - 5.117	10.122	0.269	9.595 - 10.649

(QUANTITIES SHOWN AS ******* ARE NOT ESTIMABLE UNDER MODEL 0)

TEST OF THE NULL HYPOTHESIS THAT THE FIRST-YEAR RECOVERY RATES AND/OR SURVIVAL RATES ARE THE SAME AS THOSE FROM COHORTS BANDED IN PREVIOUS YEARS

THIS IS A TEST OF MODEL 1 (THE NULL HYPOTHESIS) VS. MODEL 0 (AN ALTERNATIVE HYPOTHESIS)

I	CHI-SQUARE	NORMAL(0,1)
1966	35.14	5.93
1967	13.74	3.71
1968	47.17	6.87
1969	29.93	5.47
1970	40.37	6.35
1971	3.76	1.94
1972	7.53	2.74
TOTAL	177.63	33.01

PROBABILITY OF OBSERVING A VALUE LARGER THAN 33.01 IS 0.0

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 177.63 = 0.0

Example 2.4b

YOUNG MALE MALLARDS BANDED PRESEASON IN SOUTHERN ONTARIO, 1965-72

SURVIVAL S(I) (%)

YEAR	SURVIVAL	STANDARD ERROR	COEFFICIENTS OF VARIATION	95% CONFIDENCE INTERVAL
1965	95.67	15.18	15.87	65.92 - 125.42
1966	45.93	7.61	16.57	31.01 - 60.84
1967	67.75	9.83	14.51	48.48 - 87.01
1968	62.18	8.97	14.43	44.59 - 79.77
1969	59.09	7.57	12.80	44.26 - 73.92
1970	46.95	5.16	10.99	36.83 - 57.06
1971	60.05	8.27	13.77	43.84 - 76.27
MEAN	62.52	2.68	4.28	57.27 - 67.76

MEAN LIFE SPAN AS AN ADULT = 2.13
STANDARD ERROR OF THE MEAN LIFE SPAN = 0.19
95% CONFIDENCE INTERVAL OF LIFE SPAN 1.79 - 2.57

RECOVERY MATRIX

YEAR	NUMBER BANDED									
1965	1570	132	48	33	13	8	9	8	0	6
1966	1462	0	175	33	8	11	10	4	0	8
1967	1611	0	0	165	39	23	12	13	5	7
1968	1733	0	0	0	193	51	24	13	12	5
1969	1848	0	0	0	0	193	43	39	15	9
1970	3496	0	0	0	0	0	367	113	56	32
1971	4488	0	0	0	0	0	0	392	176	70
1972	3584	0	0	0	0	0	0	0	342	101

MATRIX OF EXPECTED VALUES UNDER MODEL 0

132.0	46.8	32.5	13.6	11.3	6.7	5.4	0.0	4.4
0.0	175.0	31.7	13.2	11.0	6.5	5.2	0.0	4.2
0.0	0.0	165.0	31.7	26.4	15.6	12.6	7.0	3.2
0.0	0.0	0.0	193.0	42.0	24.7	20.0	11.1	5.1
0.0	0.0	0.0	0.0	193.0	42.4	34.2	19.0	8.7
0.0	0.0	0.0	0.0	0.0	367.0	109.6	60.8	27.8
0.0	0.0	0.0	0.0	0.0	0.0	392.0	166.4	76.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	342.0	101.0

MATRIX OF CHI-SQUARE VALUES UNDER MODEL 0

0.00	0.03	0.01	0.03	0.97	0.81	1.27	0.0	0.62
0.0	0.00	0.06	2.07	0.00	1.89	0.29	0.0	3.34
0.0	0.0	0.00	1.66	0.44	0.82	0.01	0.56	4.57
0.0	0.0	0.0	0.00	1.95	0.02	2.42	0.08	0.00
0.0	0.0	0.0	0.0	0.00	0.01	0.67	0.84	0.01
0.0	0.0	0.0	0.0	0.0	0.00	0.11	0.39	0.64
0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.56	0.47
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00

(FREQUENCIES WERE COMBINED WHERE EXPECTED VALUES WERE SMALL)

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT MODEL 0

CHI-SQUARED VALUE (SAMPLE) = 27.63
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 30.10
DEGREES OF FREEDOM = 19

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 27.63 = 0.09086195

Example 2.4c

ESTIMATED COVARIANCE AND CORRELATION COEFFICIENTS: S(I) = SURVIVAL RATE IN YEAR I
F*(I) = FIRST-YEAR RECOVERY RATE IN YEAR I
F(I) = RECOVERY RATE IN YEAR I

I	COVAR(F*(I),F(I))	COVAR(F*(I),S(I))	COVAR(F*(I+1),S(I))
1965	************	-0.000051233	0.000078328
1966	-0.000002550	-0.000037604	0.000029200
1967	-0.000002998	-0.000043070	0.000043536
1968	-0.000001869	-0.000039958	0.000035140
1969	-0.000002200	-0.000033392	0.000017743
1970	-0.000001167	-0.000014097	0.000009136
1971	-0.000001299	-0.000011687	0.000015989

I	CORR(F*(I),F(I))	CORR(F*(I),S(I))	CORR(F*(I+1),S(I))
1965	************	-0.048187387	0.0
1966	-0.044650967	-0.058206191	0.050795258
1967	-0.044114387	-0.057998132	0.058603420
1968	-0.046338944	-0.058927604	0.055046339
1969	-0.048345842	-0.062041808	0.045237679
1970	-0.039196022	-0.052690527	0.042007287
1971	-0.037023185	-0.033527258	0.039389816

I	COVAR(F(I),S(I))	COVAR(F(I+1),S(I))	COVAR(S(I),S(I+1))
1965	************	-0.000382310	-0.005637361
1966	0.000183541	-0.000205351	-0.002949845
1967	0.000302896	-0.000176273	-0.003768778
1968	0.000161788	-0.000215305	-0.003267254
1969	0.000204598	-0.000107663	-0.001300688
1970	0.000085538	-0.000120416	-0.001083202
1971	0.000154039	-0.000356692	************

I	CORR(F(I),S(I))	CORR(F(I+1),S(I))	CORR(S(I),S(I+1))
1965	************	-0.374335362	-0.487976793
1966	0.358507539	-0.299919269	-0.394310302
1967	0.342448450	-0.335967767	-0.427238378
1968	0.337826278	-0.375033912	-0.481277833
1969	0.422694451	-0.247822746	-0.333143781
1970	0.288651216	-0.280225629	-0.253765227
1971	0.223656361	-0.479286826	************

(QUANTITIES SHOWN AS ******* ARE NOT ESTIMABLE UNDER MODEL 0)

Summary of Models 0, 1, 2, and 3

We begin by summarizing the basic structure and assumptions relating to each of the four models for adults.

MODEL 0—f_i, f_i^*, and S_i—recovery and survival rates are year-specific and the first-year recovery rates are allowed to differ from others.

$N_1f_1^*$	$N_1S_1f_2$	$N_1S_1S_2f_3$	$N_1S_1S_2S_3f_4$
	$N_2f_2^*$	$N_2S_2f_3$	$N_2S_2S_3f_4$
		$N_3f_3^*$	$N_3S_3f_4$

MODEL 1—f_i and S_i—recovery and survival rates are year-specific (a restriction from Model 0).

N_1f_1	$N_1S_1f_2$	$N_1S_1S_2f_3$	$N_1S_1S_2S_3f_4$
	N_2f_2	$N_2S_2f_3$	$N_2S_2S_3f_4$
		N_3f_3	$N_3S_3f_4$

MODEL 2—f_i and S—recovery rates are year-specific, but survival is assumed to be constant (a restriction from Model 1).

N_1f_1	N_1Sf_2	N_1SSf_3	N_1SSSf_4
	N_2f_2	N_2Sf_3	N_2SSf_4
		N_3f_3	N_3Sf_4

MODEL 3—f and S—recovery and survival rates are constant (a restriction from Model 2).

N_1f	N_1Sf	N_1SSf	N_1SSSf
	N_2f	N_2Sf	N_2SSf
		N_3f	N_3Sf

This sequence of models proceeds from a very general model to a very simple and restrictive model. Selection of the proper model is important to avoid biased estimators and to realistically estimate the precision of the estimates. Therefore, models that do not fit the data should not be used (e.g., Model 3 fits data only infrequently). On the other hand, models that are too general should not be used if simpler models (fewer parameters) also fit the data.

Model 1 vs. Model 0

A test of Model 1 (as the null hypothesis) vs. Model 0 (the alternative hypothesis) provides information on the fit of Model 1 and on the significance of band reporting rates the first year after banding. Specifically, this is a test of the hypothesis that $f_i^* = f_i$, $i = 2, \ldots, k$. A chi-square test statistic to test the individual hypothesis $f_i^* = f_i$ is given by

$$\chi_i^2 = \left(\frac{R_{ii}}{C_i} - \frac{R_i}{T_i}\right)^2 \Big/ \frac{R_i(T_i - R_i)(T_i - C_i)}{C_i(T_i)^3} \qquad \text{, for } i = \begin{cases} 2, \ldots, k-1 & \text{if } s = 0 \\ 2, \ldots, k & \text{if } s > 0 . \end{cases}$$

Each χ_i^2 is approximately distributed as chi-square with 1 degree of freedom. An overall test is based on the statistic $\chi^2 = \sum_i \chi_i^2$, which is chi-square with $k-2$ degrees of freedom under the null hypothesis if $s = 0$ and $k-1$ degrees of freedom if $s > 0$.

If we have prior reason to believe, for example, that $f_i^* > f_i$, a more powerful test is appropriate. It is computed as $\pm\sqrt{\chi_i^2}$: minus if $f_i^* < f_i$ and plus if $f_i^* > f_i$ ($i = 2, \ldots, k-1$ if $s = 0$ and $i = 2, \ldots, k$, if $s > 0$). To be specific, $z_i = \pm\sqrt{\chi_i^2}$ and is approximately a standard normal variate under the null hypothesis $f_i^* = f_i$. An overall test statistic is $z = \sum_i z_i$ which is normally distributed under the null hypothesis with variance $k-2$ if $s = 0$ and $k-1$ if $s > 0$. Both tests and their achieved significance levels are computed by the FORTRAN program. In Example 2.5 the San Luis Valley mallard data (see Section 2.3) are used to illustrate these tests and their interpretation.

Individual χ_i^2 and z_i values appear along with the total χ^2 and z of 16.33 and 5.35, respectively. Both the less specific chi-square test (7 df) and the one-sided normal test indicate a rejection of Model 1. In other words, the first-year recovery rates are significantly higher and, therefore, Model 0 appears to be appropriate. Although there is evidence that we should reject Model 1 (0.021 and 0.037 significance levels), we see that the large χ^2 and z values are largely attributed to just 1 year, 1970. In this situation we would suggest the use of Model 1 (assuming the goodness of fit test of Model 1 is satisfactory) and we would be cautious of the estimates of survival in 1969 and 1970 and the recovery rate in 1970. This is an example of an instance where it may be judicious to risk a little bias to gain increased precision. Compromises such as this must be carefully considered: if carried too far one might have a very precise but biased estimate.

Tests Between Models 1, 2, and 3

Tests between Models 1, 2, and 3 are likelihood ratio tests and will not be explained in detail because of the complexity involved (see Appendix A for mathematical background). These test statistics are approximately distributed as chi-square under the null hypothesis. The null hypothesis is that the more simple model fits the data. For example, if we test Model 3 vs. Model 2, we are testing the null hypothesis that Model 3 fits the data, against the specific alternative hypothesis that Model 2 is the true model. The degrees of freedom for these tests are merely the difference in the number of estimable parameters between the two models.

We use the San Luis Valley mallard data to illustrate these tests; the computer results appear in Example 2.5. Mallards were banded for 9 years and recovery data were summarized $(k=\ell=9)$. Model 1 has 17 parameters (9 f_i and 8 S_i), Model 2 has 10 parameters (9 f_i and 1 S), and Model 3 has 2 parameters (f and S). The degrees of freedom for each likelihood ratio test are summarized below.

Comparison	Degrees of freedom
Model 3 vs. Model 1	17 − 2 = 15
Model 3 vs. Model 2	10 − 2 = 8
Model 2 vs. Model 1	17 − 10 = 7

The first comparison (Example 2.5) relates to the test of Model 3 vs. Model 1. A chi-square value of 33.36 with 15 df is very unlikely $(P=0.004)$ if the null hypothesis is true. Therefore, we reject Model 3. The second comparison is similar in that Model 3 is again rejected $(P=0.003)$, this time in favor of Model 2.

The third comparison (Example 2.5) is the most useful and important and is a test of Model 2 vs. Model 1. Specifically, this is a test that annual survival rate varies. In the example, we see that a chi-square value of 10.03 with 7 df is not unusual $(P=0.187)$ and therefore we cannot reject Model 2 in favor of Model 1. At this point, Model 1 and Model 2 fit the data (in addition to the very general Model 0). We would recommend using Model 2 because it is the simplest model that adequately describes the data. It has only 10 parameters (as opposed to 17 for Model 1 and 24 for Model 0) and therefore substantially better precision of the estimators is achieved. Note that the choice of a model should not be based on these comparative tests alone. The goodness of fit tests must also be examined. For example, it is possible for the test between Model 2 vs. Model 1 to fail to be significant, yet Model 2 does not fit well, while Model 1 fits. In this case Model 1 is appropriate.

In the example discussed above Model 2 was indicated as being appropriate. Biologically, we are not saying the parameter S is a constant. We are saying that as far as estimation is concerned and considering the precision of the estimates, $\hat{S}$ from Model 2 is a better estimate of the parameter S_i than the estimate $\hat{S}_i$ from Model 1.

As caution we add that if only small samples of recovery data are available for analysis, the power and distributional properties of these tests may be unsatisfactory.

2.7 Mean Life Span

The arithmetic mean length of life of birds banded as adults is defined as the mean life span (MLS) (Seber 1973). An estimate of the MLS and measures of the precision of the estimate are computed by program ESTIMATE for all four models for adults. Strictly speaking, it applies only to Models 2 and 3. It is a useful approximation for Models 0 and 1 if the parameters S_i do not vary appreciably. MLS is expressed in years.

Under Models 2 and 3 the ML estimate of the mean life span is

$$\widehat{\text{MLS}}=\frac{1}{-\ell n(\hat{S})}.$$

The average survival is used in the approximation for Models 0 and 1

$$\widehat{\text{MLS}}=\frac{1}{-\ell n(\bar{\hat{S}})}.$$

Estimates of the sampling variance of this statistic are given by Cormack (1964). We have employed a better and simpler procedure to construct approximate confidence intervals on MLS. From the principle of Maximum Likelihood we know that the estimates of S_i are approximately normally distributed for large sample sizes. In particular, the average, $\bar{\hat{S}}$, is nearly normally distributed. Since $\widehat{\text{MLS}}$ is a simple one-to-one transformation of $\hat{S}$ or $\bar{\hat{S}}$, we merely

Example 2.5

MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

TEST OF THE NULL HYPOTHESIS THAT THE FIRST-YEAR RECOVERY RATES AND/OR SURVIVAL RATES ARE THE SAME AS THOSE FROM COHORTS BANDED IN PREVIOUS YEARS

THIS IS A TEST OF MODEL 1 (THE NULL HYPOTHESIS) VS. MODEL 0 (AN ALTERNATIVE HYPOTHESIS)

I	CHI-SQUARE	NORMAL(0,1)
1964	0.14	-0.37
1965	0.16	0.40
1966	1.47	1.21
1967	0.84	0.92
1968	1.29	-1.13
1969	0.85	0.92
1970	11.59	3.40
TOTAL	16.33	5.35

PROBABILITY OF OBSERVING A VALUE LARGER THAN 5.35 IS 0.02151668

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 16.33 = 0.03790844

TESTS OF VARIOUS MODELS AND ASSUMPTIONS

(IN EACH CASE THE NULL HYPOTHESIS BEING TESTED IS THAT THE SIMPLEST MODEL, THE ONE WITH THE FEWEST PARAMETERS, FITS THE DATA)

TEST OF THE MODEL ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES AGAINST THE MODEL
ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(F,S VS F(1),...,F(K) AND S(1),...,S(K-1) MODEL) -- MODEL 3 VS. MODEL 1

CHI-SQUARED VALUE = 33.36
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 25.00
DEGREES OF FREEDOM = 15

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 33.36 = 0.00418205

TEST OF THE MODEL ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES AGAINST THE MODEL
ASSUMING CONSTANT SURVIVAL BUT TIME-SPECIFIC RECOVERY RATES
(F,S VS F(1),...,F(K),S MODEL) -- MODEL 3 VS. MODEL 2

CHI-SQUARED VALUE = 23.33
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 15.50
DEGREES OF FREEDOM = 8

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 23.33 = 0.00296054

TEST OF THE MODEL ASSUMING CONSTANT SURVIVAL BUT TIME-SPECIFIC RECOVERY RATES AGAINST THE MODEL
ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(F(1),...,F(K),S VS F(1),...,F(K) AND S(1),...,S(K-1) MODEL) -- MODEL 2 VS. MODEL 1

CHI-SQUARED VALUE = 10.03
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 14.10
DEGREES OF FREEDOM = 7

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 10.03 = 0.18703687

make the same transformation of the lower and upper confidence interval bounds $\hat{S}_L$ and $\hat{S}_U$, computed on S. The lower and upper confidence intervals (nonsymmetric) on the MLS are

$$\widehat{\text{MLS}}_L = \frac{1}{-\ell n(\hat{S}_L)}, \quad \widehat{\text{MLS}}_U = \frac{1}{-\ell n(\hat{S}_U)} \quad \text{for Models 2 and 3}$$

and

$$\widehat{\text{MLS}}_L = \frac{1}{-\ell n(\hat{\bar{\bar{S}}}_L)}, \quad \widehat{\text{MLS}}_U = \frac{1}{-\ell n(\hat{\bar{\bar{S}}}_U)} \quad \text{for Models 0 and 1.}$$

This procedure is used in the FORTRAN program ESTIMATE.

2.8 A Computer Example Using All Four Models

In this section we will discuss the analysis and interpretation of an excellent set of data from male mallards banded during January and February in Illinois. Banding was done annually from 1963 through 1970 and recovery information was recorded through 1973 ($k=8$, $\ell=11$, $s=3$). These data represent a good example because large numbers of birds were banded and a large number of recoveries are available for critical analysis. Good data sets such as these allow confidence regarding the various point and interval estimates, power of tests, and approximate normality of the estimators.

The results of the computer analysis of the Illinois mallard data are presented in Example 2.6. The rest of this section discusses this example for each of the four models. The computer output begins with a page of definitions. Subscripts are parenthesized and all symbols are in upper case on the computer output, f_i=F(I), S_i=S(I), etc. The definitions are followed by an analysis of data under the four models. The sequence progresses from the general Model 0 to the very simple, but restrictive, Model 3. In each case, the specific structure of the model is given in terms of cell probabilities as functions of survival and recovery rates. Heading information used to identify the data set is printed next, followed by a display of the banding and recovery data. For all four models estimates, standard errors, and confidence intervals are given as a percentage. Covariances are, of course, not in percentages and correlations are unitless quantities.

Model 0

Estimates of the two types of recovery rates $\tilde{f}_i$ and $\hat{f}_i^*$ are printed under Model 0 with estimates of their standard errors and 95% confidence intervals (Example 2.6b). The recovery rate in 1967 was estimated to be 3.899±0.420% (i.e., the standard error of the estimate is 0.420), while the first-year recovery rate in 1967 (f^*) was estimated to be 3.645±0.337%. Estimates of the average recovery rates are nearly the same for each type of recovery rate, 3.870± 0.244% and 3.714±0.146%, for $\bar{\bar{f}}$ and $\hat{\bar{f}}^*$, respectively. A final remark regarding recovery rates is that f_1 is not defined in Model 0 and therefore cannot be estimated.

A test of Model 1 (the null hypothesis) vs. Model 0 (the alternative hypothesis) is printed following estimates of recovery rates. Individual chi-square and z values are printed and the final tests are based on their sums as described in Section 2.6. Both tests, one-sided and two-sided, fail to reject the null hypothesis. Neither test statistic was near a reasonable significance level (e.g., the 0.05 or 0.01 levels); significance levels were 0.47 and 0.18 for one- and two-sided tests, respectively. From this we conclude that the two types of recovery rates are similar and therefore Model 1, rather than the more general model, may be satisfactory. The results of this test could have been anticipated because the two average rates $\bar{\bar{f}}$ and $\bar{f}^*$ were nearly the same. At this point one would normally examine the simpler models and discard Model 0 as being too general. We will finish discussing Model 0 for illustrative purposes.

Estimates of annual survival are printed next with their standard errors, coefficients of variation, and 95% confidence intervals (Example 2.6c). For example, the estimated survival rate from the time of banding in 1966 to the time of banding in 1967 was 69.69 ± 5.97%. One further example, assuming the midpoint of banding was 1 February, the survival rate between 1 February 1969 and 31 January 1970 was estimated to be about 63.7±5.9%. The standard errors of the annual survival estimates tend to be fairly substantial under this model, although very large banded samples are involved. Some precision is sacrificed for the general assumptions allowed under Model 0. A total of 22 parameters are estimated under Model 0 for the Illinois data. In spite of this, the average annual survival rate is fairly well estimated, 68.51±1.31 %.

The mean life span as an adult was estimated at 2.64 ± 0.13 years after banding. Of course, many died within a few months while a few probably would have lived 13-20 years. An alternative way of perceiving "longevity" is the concept of a half-life, the estimated time at which one-half the banded birds are dead. This time is estimated to be 1.82 years for the Illinois mallard data (it is not computed by the program; hence it is not shown in Example 2.6). Therefore, we estimate that about one-half the adult male mallards banded in Illinois during 1963-70 died within 1.8 years after banding. The formula for half-life is MLS $\times$ $\ell n(2)$ = MLS $\times$ 0.69 (in the example, $1.82 = 2.64 \times 0.69$); the derivation of this formula is given in Appendix A.

Information relating to the goodness of fit test is given in four parts. First, the banding and recovery data are printed for easy comparison with the matrix of estimated expected values, printed second (Example 2.6c). After any combination necessary for the chi-square test, both the data matrix $\mathbf{R}_{ij}$ and the matrix of estimated expected values $\mathbf{E}_{ij}$ are printed. Only one cell was combined in the example data, $R_{1,10}$ with $R_{1,11}$. Thirdly, the matrix of individual chi-square values is printed. When the model is correct, values above 4 in this matrix will occur only in about 5% of the cells. Thus, we suggest that values greater than 4, if excessive, suggest lack of fit in those cells. Only one cell in the Illinois data exceeds 4 (row 3, column 7 has a χ^2 values of 4.23); this is not unusual. In general the observed and expected values are in good agreement.

The fourth part of the output is the goodness of fit test itself. By summing all the individual values in the chi-square matrix, a value of 29.98 is obtained. A value as large as 29.98 with 34 df is not at all unusual if Model 0 is the correct model. It is far from 48.60, the 0.05 level of significance. We conclude that we cannot reject Model 0 because it seems to fit the data quite well; however, we already suspect a simpler model may also fit very well.

The computer analysis of the Illinois data under Model 0 concludes with estimates of the various non-negligible sampling covariances and correlation coefficients (Example 2.6d). Several parameter estimators are highly correlated, e.g., corr($\tilde{f}_i$,$\tilde{S}_i$), corr($\tilde{f}_{i+1}$,$\tilde{S}_i$), and corr($\tilde{S}_i$,$\tilde{S}_{i+1}$) (indicated on the computer output as CORR(F(I),S(I)), CORR (F(I+1),S(I)), and CORR (S(I),S(I+1)), respectively). Several correlations are undefined or nonestimable under Model 0 and are indicated on the example computer output by asterisks. If $k = \ell$, then additional correlations are not estimable.

Model 1

The model structure, identification of the data set, and the banding and recovery data are displayed (Example 2.6e). Because Model 1 is so useful in the series, several intermediate statistics are displayed: column totals C(I), row totals R(I), block totals T(I), numbers banded N(I), estimates of the products $S_k \cdots S_{k+j-1} f_{k+j}$ denoted as GAMMA(I), and total recovery rates R_i/N_i from each banded cohort denoted as RHO(I). Generally these statistics are most useful in checking figures or quickly assessing the quantity of data being analyzed. For example, we see that 22,805 male mallards were banded in Illinois and 2,703 recoveries are available for analysis.

Estimates of recovery rates $\hat{f}_i$ are printed with their standard errors and 95 % confidence intervals. The average recovery rate $\hat{\bar{f}}$ is estimated at 3.511±0.109%. So-called "direct recovery rates" or "first-year recovery rates" are printed to the right of the ML estimates (these are the same as $\hat{f}_i^*$ under Model 0). These rates, R_{ii}/N_i, have been used historically and are printed merely for comparison. We do not recommend their use as they are merely inefficient estimates of f_i if Model 1 is the correct model, and therefore they have larger sampling variances (except $\hat{f}_1$). The direct recovery rates are useful in that a simple, though inefficient, test can be made to test the null hypothesis that they are constant. A more efficient test of constant recovery rates could be made based on the ML estimators of f_i, but it is much more complicated. The simple test resulted in a chi-square value of 43.09 on 7 df. This is a very unusual value if, in fact, the true recovery rates were constant each year ($P = 0.00000032$). We conclude from this that there was significant variation in recovery rates. This appears to be a very common result and it is unusual that we cannot reject this null hypothesis, even with the inefficient test.

Estimates of annual survival rates and associated statistics are printed next (Example 2.6f). The average survival rate under Model 1 is 68.87±0.92 and is similar to the average under Model 0, 68.51±1.31. The estimates of S_i are more precise than those under Model 0. The 95% confidence interval on the mean is small, 67.06 – 70.68. The mean life span is also estimated with good precision, 2.68±0.10 years.

Matrices of observed data, estimated expected values, and chi-square values follow. Here we see three cells with chi-square values of 4 or larger. Summing the individual values in the chi-square matrix we have a chi-square value of 49.18 on 42 df. This is a likely value ($P = 0.208$) and we conclude that Model 1 fits the data satisfactorily.

Often a very large chi-square value for a particular cell is from an error in the input data. A common mistake is to keypunch a number, or row of numbers, in the wrong column.

Finally, various non-negligible sampling covariance and correlation coefficients under Model 1 are printed (Example 2.6g). Generally, the substantial correlations are similar to those found under Model 0: corr($\tilde{S}_i$, $\hat{f}_i$), corr($\tilde{S}_i$, $\tilde{S}_{i+1}$), and corr($\tilde{S}_i$, $\hat{f}_{i+1}$) (these are denoted on the computer output as CORR(S(I),F(I)), CORR(S(I),S(I+1)), and CORR(S(I),F(I+1)), respectively). In particular, the estimators of average recovery rate and average annual survival rate are highly correlated, – 0.526 in the Illinois example.

Model 2

The interpretation of Model 2 is quite simple. The computer output (Example 2.6h) begins, as before, with the structure of the model, identification of the data set, and the input data. Estimates of the recovery rates and associated statistics are printed next. In the example, recovery rates ranged from 2.064% in 1965 to 5.003% in 1970. The average recovery rate for the 1963-70 period was estimated at 3.66±0.10%.

The survival rate, assumed constant under Model 2, was estimated at 70.44±0.85%, a value slightly higher than the average annual survival estimated under Model 0 and Model 1. This difference is reflected in a larger mean life span, 2.85±0.10 years.

Matrices of estimated expected values and individual chi-square values are printed next, followed by the final goodness of fit test (Example 2.6i). Again we find that the model fits the data ($\chi^2 = 61.18$, 48 df, $P = 0.096$. At this point we have three models, each making somewhat different assumptions which appear to adequately describe the mallard data. However, final judgment must await examination of the tests between models.

Sampling covariances and correlations are printed last. Note the corr$(\hat{f}_i, \hat{S})$ (printed on the output as CORR (F(I),S)) increases from year to year. Also, $\hat{f}$ and $\hat{S}$ estimators are highly negatively correlated.

Model 3

This is the simplest model and begins with a display of the structure of the model and the input data (Example 2.6j). The constant survival rate was estimated at 70.91±0.74% and the constant recovery rate at 3.89±0.10%. The estimators of these two parameters are highly correlated, −0.70.

Statistics related to the mean life span are printed, followed by matrices of estimated expected values and individual chi-square values. Here we note signs of a very poor fit. Seventeen cells have a chi-square value larger than 4. The final goodness of fit test confirms the suspicion ($\chi^2 = 217.43$, 58 df, $P = 0.0$). The probability of a value this large, if Model 3 is the correct model, is essentially zero (0 to 8 decimal places). Model 3 fits data sets only infrequently, so the above results were not surprising.

Tests Between Models

The final page of the computer analysis (Example 2.6k) presents tests between Models 1, 2, and 3 (a test of Models 0 vs. 1 was presented earlier because it is part of the output under Model 0). In each case the null hypothesis is that the simplest model fits the data. After all, our objective is to find and use the simplest model that adequately describes the observed data.

The test of Model 3 vs. Model 1 is rejected conclusively ($\chi^2 = 180.10$, 16 df). The test of Model 3 vs. Model 2 is also rejected beyond doubt ($\chi^2 = 164.19$, 10 df). The results of these tests and the goodness of fit test of Model 3 allow us to discard Model 3 and its assumptions. In addition, we rejected Model 0 earlier as being too general. The final question remaining is whether to use Model 1 or Model 2 for the Illinois mallard data.

The test of Model 2 vs. Model 1 is significant at about the 1% level ($\chi^2 = 15.91$, 6 df). Therefore, we reject Model 2 and conclude that there is evidence that annual survival rate has varied significantly. From this information we would want to use the parameter estimates under Model 1.

This concludes our detailed coverage of the four models using the Illinois mallard data. In practice, the whole procedure is quite simple. Often, one might want to examine the tests between models (the final page) first. If these tests tend to eliminate one or two models, then we could quickly examine the goodness of fit tests for the remaining models under consideration. Selection of an adequate model can often be done in a minute or two.

2.9 General Discussion

The final section in this chapter will discuss a few new points and review others. First of all, the computer analysis of data from adult bandings (Models 0, 1, 2, and 3) results in a large amount of information. We believe, however, that this information can be quickly and easily interpreted.

Fit of the data is assessed in two ways: tests between specific models and goodness of fit tests of a particular model. The use of a model that does not fit the data often results in seriously biased estimates of parameters, and the estimated sampling variances and covariances will be inappropriate. We have found that Model 1 or Model 2 is adequate for the analysis of most banding studies of adult game birds. Model 3 is too restrictive and somewhat unrealistic biologically, while Model 0 is usually too general for most data sets.

Once a computer analysis of a set of data has been obtained, we recommend the following steps be taken: (1) Carefully check the input data for accuracy. Any errors in summarizing the banding or recovery data or in keypunching

Example 2.6a

THIS COMPUTER OUTPUT PRESENTS A DETAILED ANALYSIS OF FOUR GENERAL STOCHASTIC MODELS FOR THE ANALYSIS OF BANDING EXPERIMENTS. EACH MODEL IS BASED ON SEVERAL ASSUMPTIONS. IN EACH CASE PARAMETERS ARE ESTIMATED USING THE THEORY OF MAXIMUM LIKELIHOOD AND THE ASSUMPTIONS ARE TESTED STATISTICALLY. THE SEQUENCE PROGRESSES FROM A GENERAL MODEL, BASED ON RELATIVELY FEW ASSUMPTIONS, TO A MODEL WHICH MAKES VERY SIMPLE, BUT QUITE RESTRICTIVE ASSUMPTIONS.

DEFINITIONS AND NOTATION

K THE NUMBER OF YEARS OF BANDING

L THE NUMBER OF YEARS OF RECOVERY

F(I) THE RECOVERY RATE IN YEAR I. SPECIFICALLY, THE PROBABILITY THAT A BANDED BIRD IS RECOVERED AND REPORTED TO THE BIRD BANDING LABORATORY IN YEAR I, GIVEN THAT IT WAS ALIVE AT THE BEGINNING OF YEAR I.
I=1,2,...,K EXCEPT MODEL 0, WHERE I=2,3,...,K OR 2,3,...,K-1

F THE CONSTANT RECOVERY RATE (NOT AN AVERAGE VALUE)

S(I) THE SURVIVAL RATE IN YEAR I. SPECIFICALLY, THE PROBABILITY OF SURVIVAL OF A BIRD IN YEAR I, GIVEN THAT IT WAS ALIVE AT THE BEGINNING OF YEAR I.
I=1,2,...,K-1

ESTIMATES OF SURVIVAL PERTAIN TO THE PERIOD FROM THE TIME OF BANDING IN YEAR I TO THE TIME OF BANDING IN YEAR I+1

S THE CONSTANT SURVIVAL RATE (NOT AN AVERAGE VALUE)

C(I) COLUMN TOTAL OF THE RECOVERY MATRIX IN YEAR I -- THE TOTAL RECOVERIES IN CALENDAR YEAR I
I=1,2,...,L

R(I) ROW TOTAL OF THE RECOVERY MATRIX IN YEAR I -- THE TOTAL RECOVERIES FROM BANDING IN YEAR I
I=1,2,...,K

T(I) A BLOCK TOTAL: THE TOTAL RECOVERIES IN AND AFTER YEAR I FROM ALL BIRDS BANDED PRIOR TO, AND IN, YEAR I
I=1,2,...,K

N(I) NUMBER BANDED IN YEAR I
I=1,2,...,K

RHO(I) TOTAL RECOVERY RATE FROM BIRDS BANDED IN YEAR I
I=1,2,...,K

Example 2.6b

MODEL 0
****

ANALYSIS UNDER THE ASSUMPTIONS OF TIME-SPECIFIC SURVIVAL AND RECOVERY RATES WHERE THE FIRST-YEAR RECOVERY RATES ARE DIFFERENT FROM RECOVERY RATES OF PREVIOUSLY-BANDED COHORTS.

THIS MODEL IS AN EXTENSION OF THE SEBER-ROBSON-YOUNGS MODEL (MODEL 1) IN THAT THE FIRST-YEAR RECOVERY RATES ARE ALLOWED TO DIFFER. THIS MODEL IS USEFUL IF THE BAND REPORTING RATE IS QUITE DIFFERENT THE FIRST YEAR AFTER BANDING (E.G., BAND COLLECTING ACTIVITIES CONDUCTED BY CONSERVATION AGENCY PERSONNEL OFTEN AFFECT PRIMARILY THE FIRST-YEAR RECOVERY RATES).

(MODEL DEVELOPED BY BROWNIE (1974. CORNELL BIOMETRICS UNIT PAPER 535-M)). REFER TO BROWNIE (1974) FOR A DISCUSSION OF THE USE OF THIS MODEL VS. MODEL 1.

SPECIFICALLY, THE MODEL STRUCTURE IS:

F*(1)	S(1)F(2)	S(1)S(2)F(3)	S(1)S(2)S(3)F(4)	S(1)S(2)S(3)S(4)F(5)
	F*(2)	S(2)F(3)	S(2)S(3)F(4)	S(2)S(3)S(4)F(5)
		F*(3)	S(3)F(4)	S(3)S(4)F(5)
			F*(4)	S(4)F(5)

MALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70

BANDING AND RECOVERY INPUT DATA

YEAR	NUMBER BANDED	RECOVERY MATRIX										
1963	2583	91	89	24	18	16	11	8	7	7	2	6
1964	3075	0	141	45	52	50	17	30	21	16	7	3
1965	1195	0	0	27	31	21	8	19	7	9	4	3
1966	3418	0	0	0	156	92	44	50	49	34	23	5
1967	3100	0	0	0	0	113	68	57	65	41	23	10
1968	2400	0	0	0	0	0	63	52	59	44	30	12
1969	2601	0	0	0	0	0	0	91	80	58	37	25
1970	4433	0	0	0	0	0	0	0	222	169	95	46

I	RECOVERY RATE F(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL	FIRST-YEAR RECOVERY RATE F*(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
1963	*******	*******	******* - *******	3.523	0.363	2.812 - 4.234
1964	6.975	1.106	4.807 - 9.144	4.585	0.377	3.846 - 5.325
1965	2.165	0.357	1.466 - 2.865	2.259	0.430	1.417 - 3.102
1966	3.215	0.415	2.402 - 4.028	4.564	0.357	3.864 - 5.264
1967	3.899	0.420	3.075 - 4.722	3.645	0.337	2.985 - 4.305
1968	2.396	0.277	1.853 - 2.940	2.625	0.326	1.985 - 3.265
1969	3.403	0.362	2.694 - 4.113	3.499	0.360	2.792 - 4.205
1970	5.035	0.477	4.100 - 5.970	5.008	0.328	4.366 - 5.650
MEAN	3.870	0.244	3.392 - 4.347	3.714	0.146	3.428 - 4.000

(QUANTITIES SHOWN AS ******* ARE NOT ESTIMABLE UNDER MODEL 0)

TEST OF THE NULL HYPOTHESIS THAT THE FIRST-YEAR RECOVERY RATES AND/OR SURVIVAL RATES ARE THE SAME AS THOSE FROM COHORTS BANDED IN PREVIOUS YEARS

THIS IS A TEST OF MODEL 1 (THE NULL HYPOTHESIS) VS. MODEL 0 (AN ALTERNATIVE HYPOTHESIS)

I	CHI-SQUARE	NORMAL(0,1)
1964	5.69	-2.39
1965	0.02	0.15
1966	5.17	2.27
1967	0.23	-0.48
1968	0.27	0.52
1969	0.03	0.17
1970	0.00	-0.07
TOTAL	11.42	0.18

PROBABILITY OF OBSERVING A VALUE LARGER THAN 0.18 IS 0.47136378

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 11.42 = 0.17888993

Example 2.6c

MALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70

SURVIVAL S(I) (%)

YEAR	SURVIVAL	STANDARD ERROR	COEFFICIENTS OF VARIATION	95% CONFIDENCE INTERVAL
1963	48.72	5.67	11.64	37.61 - 59.83
1964	72.54	8.44	11.64	55.99 - 89.08
1965	71.41	8.15	11.42	55.43 - 87.39
1966	69.69	5.97	8.57	57.99 - 81.40
1967	79.90	7.40	9.26	65.40 - 94.40
1968	73.61	7.33	9.95	59.25 - 87.97
1969	63.67	5.93	9.31	52.05 - 75.29
MEAN	68.51	1.31	1.91	65.94 - 71.08

MEAN LIFE SPAN AS AN ADULT = 2.64
STANDARD ERROR OF THE MEAN LIFE SPAN = 0.13
95% CONFIDENCE INTERVAL OF LIFE SPAN 2.40 - 2.93

RECOVERY MATRIX

YEAR	NUMBER BANDED											
1963	2583	91	89	24	18	16	11	8	7	7	0	8
1964	3075	0	141	45	52	50	17	30	21	16	7	3
1965	1195	0	0	27	31	21	8	19	7	9	4	3
1966	3418	0	0	0	156	92	44	50	49	34	23	5
1967	3100	0	0	0	0	113	68	57	65	41	23	10
1968	2400	0	0	0	0	0	63	52	59	44	30	12
1969	2601	0	0	0	0	0	0	91	80	58	37	25
1970	4433	0	0	0	0	0	0	0	222	169	95	46

MATRIX OF EXPECTED VALUES UNDER MODEL 0

91.0	87.8	19.8	21.0	17.7	8.7	9.1	8.6	6.3	0.0	5.6
0.0	141.0	48.3	51.2	43.3	21.3	22.2	20.9	15.5	9.1	4.5
0.0	0.0	27.0	27.4	23.2	11.4	11.9	11.2	8.3	4.9	2.4
0.0	0.0	0.0	156.0	92.9	45.6	47.7	44.9	33.3	19.4	9.7
0.0	0.0	0.0	0.0	113.0	59.3	62.1	58.4	43.3	25.3	12.6
0.0	0.0	0.0	0.0	0.0	63.0	60.1	56.6	41.9	24.5	12.2
0.0	0.0	0.0	0.0	0.0	0.0	91.0	83.4	61.7	36.1	18.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	222.0	165.3	96.6	48.1

MATRIX OF CHI-SQUARE VALUES UNDER MODEL 0

0.00	0.02	0.91	0.42	0.17	0.61	0.13	0.29	0.07	0.0	1.08
0.0	0.00	0.22	0.01	1.04	0.85	2.72	0.00	0.02	0.47	0.51
0.0	0.0	0.00	0.46	0.21	1.01	4.23	1.58	0.06	0.15	0.14
0.0	0.0	0.0	0.00	0.01	0.06	0.11	0.37	0.02	0.65	2.26
0.0	0.0	0.0	0.0	0.00	1.26	0.41	0.73	0.12	0.21	0.53
0.0	0.0	0.0	0.0	0.0	0.00	1.10	0.10	0.10	1.23	0.00
0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.14	0.23	0.02	2.75
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.08	0.03	0.09

(FREQUENCIES WERE COMBINED WHERE EXPECTED VALUES WERE SMALL)

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT MODEL 0

CHI-SQUARED VALUE (SAMPLE) = 29.98
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 48.60
DEGREES OF FREEDOM = 34

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 29.98 = 0.66501882

Example 2.6d

ESTIMATED COVARIANCE AND CORRELATION COEFFICIENTS: S(I) = SURVIVAL RATE IN YEAR I
F*(I) = FIRST-YEAR RECOVERY RATE IN YEAR I
F(I) = RECOVERY RATE IN YEAR I

I	COVAR(F*(I),F(I))	COVAR(F*(I),S(I))	COVAR(F*(I+1),S(I))
1963	************	-0.000006645	0.000007265
1964	-0.000001040	-0.000010816	0.000013715
1965	-0.000000409	-0.000013502	0.000009536
1966	-0.000000429	-0.000009306	0.000008195
1967	-0.000000458	-0.000009395	0.000008739
1968	-0.000000262	-0.000008051	0.000009901
1969	-0.000000458	-0.000008564	0.000007193

I	CORR(F*(I),F(I))	CORR(F*(I),S(I))	CORR(F*(I+1),S(I))
1963	************	-0.032315183	0.0
1964	-0.024922511	-0.033973902	0.037797667
1965	-0.026692697	-0.038520063	0.032760066
1966	-0.028986946	-0.043641333	0.040757213
1967	-0.032412505	-0.037731105	0.036198790
1968	-0.028960376	-0.033673470	0.037511655
1969	-0.035123865	-0.040104564	0.037043767

I	COVAR(F(I),S(I))	COVAR(F(I+1),S(I))	COVAR(S(I),S(I+1))
1963	************	-0.000129952	-0.001351374
1964	0.000193489	-0.000140838	-0.004644976
1965	0.000138656	-0.000070580	-0.001530147
1966	0.000068881	-0.000094157	-0.001929635
1967	0.000107945	-0.000089205	-0.002740431
1968	0.000082184	-0.000115634	-0.002163230
1969	0.000100020	-0.000096181	************

I	CORR(F(I),S(I))	CORR(F(I+1),S(I))	CORR(S(I),S(I+1))
1963	************	-0.207202347	-0.282454375
1964	0.207186290	-0.467688310	-0.674918050
1965	0.476626853	-0.208662257	-0.314151728
1966	0.277969498	-0.375125126	-0.436679782
1967	0.347273144	-0.434875249	-0.505648080
1968	0.404537234	-0.436277709	-0.498143567
1969	0.466434424	-0.340115948	************

(QUANTITIES SHOWN AS ******* ARE NOT ESTIMABLE UNDER MODEL 0)

Example 2.6e

```
MODEL 1
********

ANALYSIS UNDER THE ASSUMPTIONS OF TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(A SYNTHESIS OF MODELS DEVELOPED BY SEBER (1970. BIOMETRIKA) AND ROBSON AND YOUNGS (1971. CORNELL BIOMETRICS
UNIT PAPER 369))


SPECIFICALLY, THE MODEL STRUCTURE IS:

   F(1)      S(1)F(2)     S(1)S(2)F(3)     S(1)S(2)S(3)F(4)     S(1)S(2)S(3)S(4)F(5)
                 F(2)         S(2)F(3)         S(2)S(3)F(4)         S(2)S(3)S(4)F(5)
                                  F(3)             S(3)F(4)             S(3)S(4)F(5)
                                                       F(4)                 S(4)F(5)
```

MALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70

BANDING AND RECOVERY INPUT DATA

YEAR	NUMBER BANDED	RECOVERY MATRIX										
1963	2583	91	89	24	18	16	11	8	7	7	2	6
1964	3075	0	141	45	52	50	17	30	21	16	7	3
1965	1195	0	0	27	31	21	8	19	7	9	4	3
1966	3418	0	0	0	156	92	44	50	49	34	23	5
1967	3100	0	0	0	0	113	68	57	65	41	23	10
1968	2400	0	0	0	0	0	63	52	59	44	30	12
1969	2601	0	0	0	0	0	0	91	80	58	37	25
1970	4433	0	0	0	0	0	0	0	222	169	95	46

INTERMEDIATE STATISTICS

I	C(I)	R(I)	T(I)	N(I)	GAMMA(I)	RHO(I)
1963	91.0	279.0	279.0	2583.0	0.0	0.10801
1964	230.0	382.0	570.0	3075.0	0.0	0.12423
1965	96.0	129.0	469.0	1195.0	0.0	0.10795
1966	257.0	453.0	826.0	3418.0	0.0	0.13253
1967	292.0	377.0	946.0	3100.0	0.0	0.12161
1968	211.0	260.0	914.0	2400.0	0.0	0.10833
1969	307.0	291.0	994.0	2601.0	0.0	0.11188
1970	510.0	532.0	1219.0	4433.0	0.0	0.12001
1971	378.0	0.0	0.0	0.0	0.03721	0.0
1972	221.0	0.0	0.0	0.0	0.02176	0.0
1973	110.0	0.0	0.0	0.0	0.01083	0.0
TOTALS		2703.0		22805.0		

I	RECOVERY RATE F(I) (%)				DIRECT RECOVERY RATE B(I,I)/N(I)			
	ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL		ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL	
1963	3.523	0.363	2.812 -	4.234	3.523	0.363	2.812 -	4.234
1964	5.013	0.350	4.326 -	5.699	4.585	0.377	3.846 -	5.325
1965	2.210	0.272	1.676 -	2.744	2.259	0.430	1.417 -	3.102
1966	4.124	0.280	3.576 -	4.672	4.564	0.357	3.864 -	5.264
1967	3.754	0.257	3.250 -	4.258	3.645	0.337	2.985 -	4.305
1968	2.501	0.210	2.089 -	2.913	2.625	0.326	1.985 -	3.265
1969	3.455	0.252	2.962 -	3.949	3.499	0.360	2.792 -	4.205
1970	5.021	0.265	4.501 -	5.541	5.008	0.328	4.366 -	5.650

```
ARITHMETIC MEAN RECOVERY RATE (EXCEPT YEAR K) =      3.511
STANDARD ERROR OF THE MEAN RECOVERY RATE =      0.109
95% CONFIDENCE INTERVAL FOR MEAN RECOVERY RATE =      3.30 -      3.72


TEST OF THE NULL HYPOTHESIS THAT THE FIRST-YEAR (DIRECT) RECOVERY RATES ARE CONSTANT EACH YEAR:

   CHI-SQUARED (SAMPLE) =      43.09
   THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL =      14.10
   DEGREES OF FREEDOM =      7

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN    43.09  =    0.00000032
```

Example 2.6f

MODEL 1 -- ANALYSIS UNDER THE ASSUMPTIONS OF TIME SPECIFIC SURVIVAL AND RECOVERY RATES

MALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70

YEAR	SURVIVAL S(I) (%)			
	SURVIVAL	STANDARD ERROR	COEFFICIENTS OF VARIATION	95% CONFIDENCE INTERVAL
1963	58.45	4.97	8.50	48.72 - 68.19
1964	68.17	6.95	10.20	54.55 - 81.80
1965	64.65	6.26	9.68	52.38 - 76.93
1966	74.90	5.18	6.92	64.74 - 85.06
1967	77.34	6.11	7.89	65.38 - 89.31
1968	74.25	6.13	8.25	62.24 - 86.26
1969	64.33	4.62	7.18	55.27 - 73.38

ARITHMETIC MEAN SURVIVAL (%) = 68.87
STANDARD ERROR OF ARITHMETIC MEAN = 0.92
95% CONFIDENCE INTERVAL FOR ARITHMETIC MEAN 67.06- 70.68

MEAN LIFE SPAN AS AN ADULT = 2.68
STANDARD ERROR OF THE MEAN LIFE SPAN = 0.10
95% CONFIDENCE INTERVAL OF LIFE SPAN 2.50 - 2.88

YEAR	NUMBER BANDED	RECOVERY MATRIX										
1963	2583	91.	89.	24.	18.	16.	11.	8.	7.	7.	2.	6.
1964	3075	0.	141.	45.	52.	50.	17.	30.	21.	16.	7.	3.
1965	1195	0.	0.	27.	31.	21.	8.	19.	7.	9.	4.	3.
1966	3418	0.	0.	0.	156.	92.	44.	50.	49.	34.	23.	5.
1967	3100	0.	0.	0.	0.	113.	68.	57.	65.	41.	23.	10.
1968	2400	0.	0.	0.	0.	0.	63.	52.	59.	44.	30.	12.
1969	2601	0.	0.	0.	0.	0.	0.	91.	80.	58.	37.	25.
1970	4433	0.	0.	0.	0.	0.	0.	0.	222.	169.	95.	46.

MATRIX OF EXPECTED VALUES -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES (MODEL 1)

91.0	75.9	23.0	27.7	19.0	9.8	10.1	9.4	7.0	4.1	2.0
0.0	154.1	46.6	56.4	38.5	19.9	20.5	19.2	14.2	8.3	4.1
0.0	0.0	26.4	31.9	21.8	11.3	11.6	10.9	8.1	4.7	2.3
0.0	0.0	0.0	140.9	96.3	49.8	51.2	48.0	35.6	20.8	10.3
0.0	0.0	0.0	0.0	116.4	60.2	61.9	58.0	43.0	25.1	12.5
0.0	0.0	0.0	0.0	0.0	60.0	61.8	57.8	42.9	25.1	12.5
0.0	0.0	0.0	0.0	0.0	0.0	89.9	84.1	62.4	36.5	18.1
0.0	0.0	0.0	0.0	0.0	0.0	0.0	222.6	165.0	96.4	48.0

MATRIX OF CHI-SQUARE VALUES -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES (MODEL 1)

0.00	2.28	0.05	3.43	0.46	0.15	0.43	0.63	0.00	1.07	7.71
0.0	1.12	0.06	0.34	3.41	0.43	4.40	0.17	0.22	0.21	0.31
0.0	0.0	0.01	0.03	0.03	0.95	4.71	1.38	0.11	0.11	0.18
0.0	0.0	0.0	1.61	0.19	0.68	0.03	0.02	0.07	0.23	2.76
0.0	0.0	0.0	0.0	0.10	1.02	0.39	0.85	0.09	0.18	0.50
0.0	0.0	0.0	0.0	0.0	0.15	1.54	0.02	0.03	0.97	0.02
0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.20	0.31	0.01	2.59
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.10	0.02	0.08

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT MODEL 1 -- ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES

CHI-SQUARED VALUE (SAMPLE) = 49.18
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 58.12
DEGREES OF FREEDOM = 42

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 49.18 = 0.20775297

Example 2.6g

ESTIMATED COVARIANCE AND CORRELATICN COEFFICIENTS: S(I) = SURVIVAL RATE IN YEAR I
F(I) = RECOVERY RATE IN YEAR I

I	COVAR(S(I),F(I))	COVAR(F(I),F(I+1))	COVAR(S(I+1),F(I))
1963	-0.000007973	0.0	0.0
1964	0.000018392	0.0	0.0
1965	0.000068330	0.0	0.0
1966	0.000021752	0.0	0.0
1967	0.000036955	0.0	0.0
1968	0.000043367	0.0	0.0
1969	0.000045476	0.0	0.0

I	COVAR(S(I),F(I+1))	COVAR(S(I),S(I+1))	COVAR(F(I),F(I+2))
1963	-0.000067177	-0.000948583	0.0
1964	-0.000104167	-0.002947986	0.0
1965	-0.000051055	-0.000946454	0.0
1966	-0.000065506	-0.001383923	0.0
1967	-0.000066335	-0.001965232	0.0
1968	-0.000078303	-0.001440766	0.0
1969	-0.000053424	************	0.0

I	CORR(S(I),F(I))	CORR(S(I),S(I+1))	CORR(S(I),F(I+1))
1963	-0.044236175	-0.274685410	-0.385871364
1964	0.075521897	-0.677388957	-0.550150794
1965	0.400592834	-0.291548938	-0.291690127
1966	0.150087411	-0.437208493	-0.491095682
1967	0.235262558	-0.525338898	-0.516513313
1968	0.336462994	-0.509092588	-0.507854645
1969	0.391270001	************	************

CORR(AVE SURVIVAL,AVE RECOVERY RATE) = -0.5264
COVAR(AVE SURVIVAL,AVE RECOVERY RATE) = -0.000005299

(THE ABOVE COVARIANCE AND CORRELATICN COEFFICIENTS ARE ESTIMATES OF THE DEGREE TO WHICH THE SAMPLING VARIANCES OF SOME PARAMETER ESTIMATORS ARE RELATED)

Example 2.6h

```
MODEL 2
********

ANALYSIS ASSUMING CONSTANT SURVIVAL BUT TIME-SPECIFIC RECOVERY RATES

THIS MODEL WAS DEVELOPED BY DRS. BROWNIE AND ROBSON AT THE CORNELL BIOMETRICS UNIT.

SPECIFICALLY, THE MODEL STRUCTURE IS:

   F(1)     SF(2)     SSF(3)     SSSF(4)     SSSSF(5)
             F(2)      SF(3)      SSF(4)      SSSF(5)
                        F(3)       SF(4)       SSF(5)
                                    F(4)        SF(5)
```

MALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70

YEAR	NUMBER BANDED	RECOVERY MATRIX										
1963	2583	91	89	24	18	16	11	8	7	7	2	6
1964	3075	0	141	45	52	50	17	30	21	16	7	3
1965	1195	0	0	27	31	21	8	19	7	9	4	3
1966	3418	0	0	0	156	92	44	50	49	34	23	5
1967	3100	0	0	0	0	113	68	57	65	41	23	10
1968	2400	0	0	0	0	0	63	52	59	44	30	12
1969	2601	0	0	0	0	0	0	91	80	58	37	25
1970	4433	0	0	0	0	0	0	0	222	169	95	46

I	RECOVERY RATE F(I) (%) ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
1	3.475	0.360	2.769 - 4.180
2	4.685	0.304	4.089 - 5.280
3	2.064	0.211	1.650 - 2.478
4	3.866	0.242	3.391 - 4.340
5	3.760	0.224	3.321 - 4.198
6	2.677	0.189	2.306 - 3.048
7	3.752	0.225	3.311 - 4.194
8	5.003	0.235	4.543 - 5.463

```
ARITHMETIC MEAN RECOVERY RATE =      3.66
STANDARD ERROR OF MEAN RECOVERY RATE =      0.10
95% CONFIDENCE INTERVAL FOR MEAN RECOVERY RATE =      3.45 -      3.87


CONSTANT SURVIVAL RATE (%) =    70.44
STANDARD ERROR OF THE CONSTANT SURVIVAL RATE =      0.85
95% CONFIDENCE INTERVAL FOR THE CONSTANT SURVIVAL RATE    68.77 -     72.10


MEAN LIFE SPAN AS AN ADULT =    2.85
STANDARD ERROR OF THE MEAN LIFE SPAN =     0.10
95% CONFIDENCE INTERVAL OF LIFE SPAN     2.67 -     3.06
```

Example 2.6i

MATRIX OF EXPECTED VALUES -- ASSUMING A CONSTANT SURVIVAL RATE AND TIME-SPECIFIC RECOVERY RATES (MODEL 2)

89.8	85.2	26.5	34.9	23.9	12.0	11.8	11.1	8.2	4.8	2.4
0.0	144.1	44.7	59.0	40.4	20.3	20.0	18.8	13.9	8.1	4.1
0.0	0.0	24.7	32.5	22.3	11.2	11.0	10.4	7.7	4.5	2.2
0.0	0.0	0.0	132.1	90.5	45.4	44.8	42.1	31.2	18.2	9.1
0.0	0.0	0.0	0.0	116.6	58.5	57.7	54.2	40.2	23.5	11.7
0.0	0.0	0.0	0.0	0.0	64.3	63.4	59.6	44.2	25.8	12.8
0.0	0.0	0.0	0.0	0.0	0.0	97.6	91.7	67.9	39.7	19.8
0.0	0.0	0.0	0.0	0.0	0.0	0.0	221.8	164.4	96.1	47.8

MATRIX OF CHI-SQUARE VALUES--ASSUMING A CONSTANT SURVIVAL RATE AND TIME-SPECIFIC RECOVERY RATE (MODEL 2)

0.0	0.2	0.2	8.2	2.6	0.1	1.2	1.5	0.2	1.6	5.4
0.0	0.1	0.0	0.8	2.3	0.5	5.0	0.3	0.3	0.2	0.3
0.0	0.0	0.2	0.1	0.1	0.9	5.7	1.1	0.2	0.1	0.3
0.0	0.0	0.0	4.3	0.0	0.0	0.6	1.1	0.3	1.2	1.8
0.0	0.0	0.0	0.0	0.1	1.6	0.0	2.2	0.0	0.0	0.2
0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.7	0.1
0.0	0.0	0.0	0.0	0.0	0.0	0.4	1.5	1.5	0.2	1.4
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT THE MODEL ASSUMING A CONSTANT SURVIVAL AND TIME-SPECIFIC RECOVERY RATES

CHI-SQUARE VALUE (SAMPLE) = 61.18
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 65.16
DEGREES OF FREEDOM = 48

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 61.18 = 0.09593084

ESTIMATED COVARIANCE AND CORRELATION COEFFICIENTS:

I	COVAR(F(I),S)	CORR(F(I),S)	COVAR(F(I),F(I+1))	CORR(F(I),F(I+1))
1963	0.000000107	0.009590297	-0.000000217	-0.019811420
1964	0.000000437	0.046523153	-0.000000050	-0.007825487
1965	0.000000437	0.066882480	0.000000084	0.016499264
1966	0.000000751	0.100388263	0.000000187	0.034633406
1967	0.000000844	0.121949214	0.000000223	0.052715433
1968	0.000000742	0.126839018	0.000000334	0.078254829
1969	0.000001118	0.160418519	0.000000595	0.112509907
1970	0.000001174	0.161906535	0.000001174	0.161906535

4 ITERATIONS

Example 2.6j

MODEL 3
****

ANALYSIS ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES
(A GENERALIZATION AND EXTENSION OF THE MODELS DEVELOPED BY CHAPMAN AND ROBSON (1960. BIOMETRICS) AND HALDANE (1955. PROC. XI INT. ORN. CONGR.) -- SEE BOTTOM OF PAGE 245 OF BOOK BY SEBER))

SPECIFICALLY, THE MODEL STRUCTURE IS:

```
F      SF      SSF      SSSF      SSSSF
       F       SF       SSF       SSSF
               F        SF        SSF
                        F         SF
```

MALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70

YEAR	NUMBER BANDED	RECOVERY MATRIX										
1963	2583	91	89	24	18	16	11	8	7	7	2	6
1964	3075	0	141	45	52	50	17	30	21	16	7	3
1965	1195	0	0	27	31	21	8	19	7	9	4	3
1966	3418	0	0	0	156	92	44	50	49	34	23	5
1967	3100	0	0	0	0	113	68	57	65	41	23	10
1968	2400	0	0	0	0	0	63	52	59	44	30	12
1969	2601	0	0	0	0	0	0	91	80	58	37	25
1970	4433	0	0	0	0	0	0	0	222	169	95	46

INTERMEDIATE STATISTICS

N = 22805.   T = 2703.   Q = 4664.

PARAMETER		ESTIMATE (%)	STD. ERR. (%)	COEF. VARIAT. (%)	95% CONFIDENCE INTERVAL
SURVIVAL RATE	(S)	70.91	0.74	1.04	69.47 -- 72.35
RECOVERY RATE	(F)	3.89	0.10	2.54	3.69 -- 4.08

CORRELATION(S,F) = -.70374912

MEAN LIFE SPAN AS AN ADULT = 2.91
STANDARD ERROR OF MEAN LIFE SPAN = 0.09
95% CONFIDENCE INTERVAL OF LIFE SPAN 2.75 - 3.09

MATRIX OF EXPECTED VALUES -- ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES (MODEL 3)

100.4	71.2	50.5	35.8	25.4	18.0	12.8	9.1	6.4	4.6	3.2
0.0	119.5	84.8	60.1	42.6	30.2	21.4	15.2	10.8	7.6	5.4
0.0	0.0	46.5	32.9	23.4	16.6	11.7	8.3	5.9	4.2	3.0
0.0	0.0	0.0	132.9	94.2	66.8	47.4	33.6	23.8	16.9	12.0
0.0	0.0	0.0	0.0	120.5	85.5	60.6	43.0	30.5	21.6	15.3
0.0	0.0	0.0	0.0	0.0	93.3	66.2	46.9	33.3	23.6	16.7
0.0	0.0	0.0	0.0	0.0	0.0	101.1	71.7	50.8	36.1	25.6
0.0	0.0	0.0	0.0	0.0	0.0	0.0	172.3	122.2	86.7	61.5

MATRIX OF CHI-SQUARE VALUES -- ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES (MODEL 3)

0.88	4.44	13.90	8.85	3.47	2.72	1.78	0.47	0.05	1.43	2.38	1.52
0.0	3.85	18.66	1.09	1.28	5.79	3.42	2.21	2.53	0.05	1.08	0.09
0.0	0.0	8.15	0.11	0.24	4.43	4.48	0.21	1.62	0.01	0.00	0.53
0.0	0.0	0.0	4.02	0.05	7.79	0.14	7.06	4.35	2.21	4.07	0.22
0.0	0.0	0.0	0.0	0.47	3.57	0.21	11.29	3.64	0.09	1.85	0.00
0.0	0.0	0.0	0.0	0.0	9.84	3.03	3.11	3.46	1.74	1.34	0.19
0.0	0.0	0.0	0.0	0.0	0.0	1.01	0.96	1.01	0.02	0.01	0.01
0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.31	17.91	0.80	3.88	2.00

TEST OF THE NULL HYPOTHESIS THAT THE DATA FIT THE MODEL ASSUMMING CONSTANT SURVIVAL AND RECOVERY RATES

CHI-SQUARE VALUE (SAMPLE) = 217.43
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL = 76.78
DEGREES OF FREEDOM = 58

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN 217.43 = 0.0

8 ITERATIONS

Example 2.6k

```
                         TESTS OF VARIOUS MODELS AND ASSUMPTIONS

(IN EACH CASE THE NULL HYPOTHESIS BEING TESTED IS THAT THE SIMPLEST MODEL, THE ONE WITH THE FEWEST PARAMETERS, FITS THE DATA)


TEST OF THE MODEL ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES AGAINST THE MODEL
ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(F,S  VS  F(1),...,F(K) AND S(1),...,S(K-1) MODEL)     --     MODEL 3 VS. MODEL 1

CHI-SQUARED VALUE =     180.10
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL =      26.30
DEGREES OF FREEDOM =    16

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  180.10  =   0.0



TEST OF THE MODEL ASSUMING CONSTANT SURVIVAL AND RECOVERY RATES AGAINST THE MODEL
ASSUMING CONSTANT SURVIVAL BUT TIME-SPECIFIC RECOVERY RATES
(F,S  VS  F(1),...,F(K),S MODEL)     --     MODEL 3 VS. MODEL 2

CHI-SQUARED VALUE =     164.19
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL =      18.30
DEGREES OF FREEDOM =    10

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  164.19  =   0.0



TEST OF THE MODEL ASSUMING CONSTANT SURVIVAL BUT TIME-SPECIFIC RECOVERY RATES AGAINST THE MODEL
ASSUMING TIME-SPECIFIC SURVIVAL AND RECOVERY RATES
(F(1),...,F(K),S  VS  F(1),...,F(K) AND S(1),...,S(K-1) MODEL)     --     MODEL 2 VS. MODEL 1

CHI-SQUARED VALUE =      15.91
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL =      12.60
DEGREES OF FREEDOM =     6

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   15.91  =   0.01425950
```

this information must be eliminated. (2) Examine the tests between specific models (the test of Model 1 vs. Model 0 on page 3 or 4 of the computer output and the three tests on the final page of the output). Interpretation of these test statistics will usually eliminate Model 3 and sometimes Model 2. (3) Examine the goodness of fit tests for the remaining models. Select the simplest model that fits the data (based on tests between models and the goodness of fit tests for a particular model). (4) Use the estimates of parameters under the model suggested from steps 2 and 3, above.

Adult males and adult females usually have different recovery and survival rates and should, therefore, be analyzed separately. A statistical test of this difference is available as an option in program BROWNIE and is discussed in Section 5.1. If this test is made and we find that males and females have similar parameters, the banding and recovery data should be combined for the final analysis. We have seen several sets of data on geese where adult males and adult females appear to have similar survival and recovery rates.

Generally, the survival rate relates to the period between bandings (*not* to the period between hunting seasons). An example for migratory waterfowl follows:

Banding period	Approximate mid-point of banding	Survival period
Winter	1 February	1 February-31 January
Molting	20 July	20 July-19 July
Prehunting season	1 September	1 September-31 August

The theory underlying Models 0 and 1 allows these periods to be of unequal length; however, the computer program assumes the periods to be of equal length (days, months, years, groups of years, whatever, as long as the intervals between banding are equal in length). Many entomological studies are conducted on a daily or weekly basis, whereas most ornithological studies are conducted on an annual basis. Banding of Canada geese *(Branta canadensis)* in several western States has been done every third year. A series of such bandings can be analyzed to estimate 3-year survival rates. Average annual survival rates can be estimated by taking the cube root of these 3-year rates.

Occasionally we observe an estimated survival rate greater than 100 %. This condition is usually associated with poor data, or in more adequate data sets where only a few birds were banded in a particular year. It can also happen in long-lived species where the annual survival rate is normally high. In either case these estimates have very large sampling variances and indicate that the estimate is a very poor one.

Many early methods for estimating parameters from banding data dealt with mortality rather than survival rates. Estimates can, of course, be easily compared because mortality rate is merely the complement of survival rate, $M = 1 - S$. We might also mention that the variance of the mortality rate estimate equals the variance of the survival rate estimate, $\mathrm{var}(\hat{M}) = \mathrm{var}(\hat{S})$.

The examples we have used in this chapter to illustrate various models and points have tended to be based on good sets of data. Studies where only a few birds are banded each year, for instance less than 300, generally represent wasted effort. The estimates from such data are poor, the variances are very large, the power of the various tests are low, and computer problems often arise. Not only the numbers banded have to be considered, but also the recovery rate, f. Species having a recovery rate of 1% or less (e.g., blue-winged teal *(Anas discors)*, several blackbirds, and American woodcock *(Philohela minor)* require very large banded samples (2,000 to 3,000 per year) before the information is useful in estimating population parameters. This subject is discussed in more detail in Chapter 9.

We discussed the subject of statistical bias in the ML estimators of some parameters with reluctance. This is a minor point and should be viewed with this perspective. If the unadjusted ML estimators are used, the statistical bias is usually in the 1-4% range. However, the standard error of the estimate is often an order of magnitude larger. The bias becomes important only with very poor data sets and here the estimates themselves are poor and sampling variance is large. Bias-adjusted estimators are used by the FORTRAN program ESTIMATE.