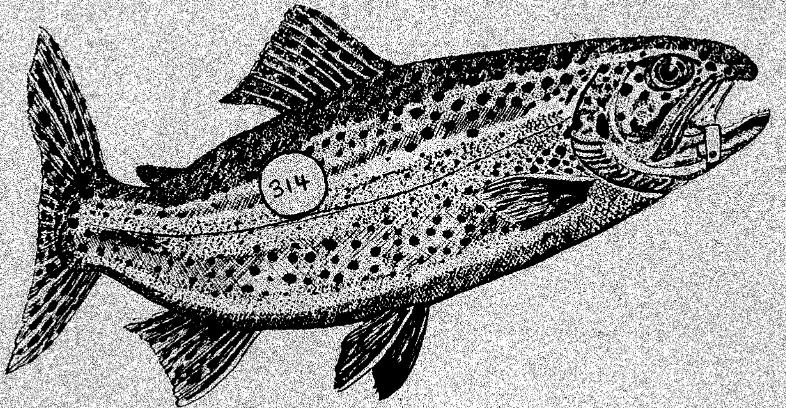
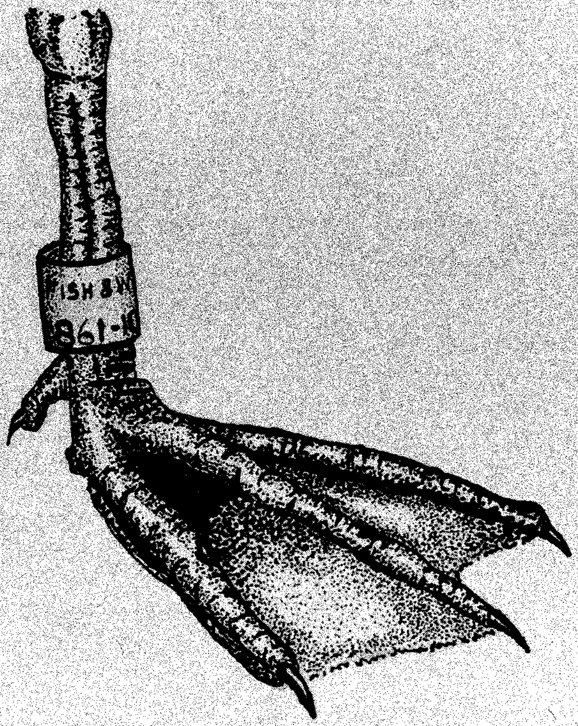
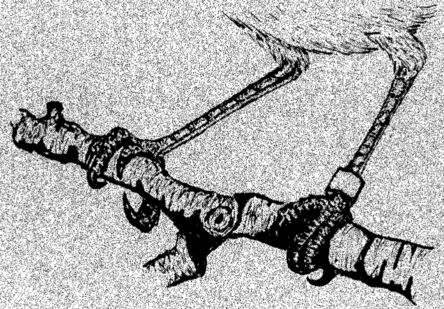
STATISTICAL INFERENCE FROM BAND RECOVERY DATA — A HANDBOOK

Second Edition



$$\begin{array}{rrrr}
N_1f_1 & N_1S_1f_2 & N_1S_1S_2f_3 & N_1S_1S_2S_3f_4 \\
 & N_2f_2 & N_2S_2f_3 & N_2S_2S_3f_4 \\
 & & N_3f_3 & N_3S_3f_4
\end{array}$$





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

$$\hat{S}_i=\left(\frac{R_i}{N_i}-\hat{f}_i\right)\Big/\frac{N_{i+1}+1}{R_{i+1}+1} \qquad ,i=1,\ldots,k-1$$

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

UNITED STATES DEPARTMENT OF THE INTERIOR
FISH AND WILDLIFE SERVICE / RESOURCE PUBLICATION 156

RESOURCE PUBLICATIONS

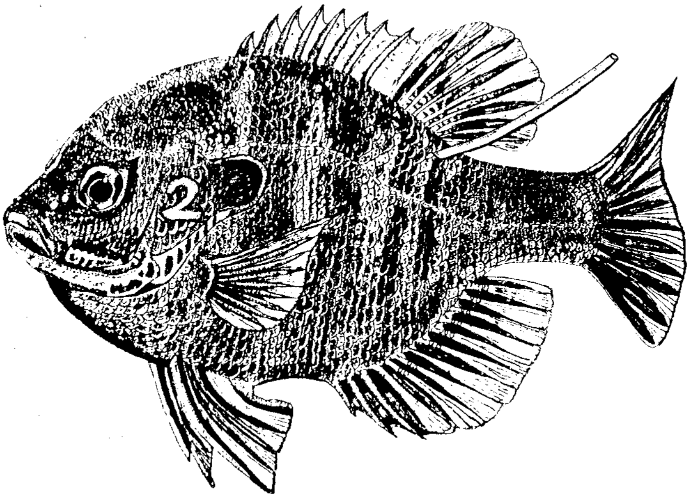
This publication of the Fish and Wildlife Service is one of a series of semitechnical or instructional materials dealing with investigations related to wildlife and fish. Each is published as a separate paper. The Service distributes a limited number of these reports for the use of Federal and State agencies and cooperators. A list of recent issues appears on inside back cover.

Copies of this publication may be obtained from the Publications Unit, U.S. Fish and Wildlife Service, Washington, DC 20240 or may be purchased from the Government Printing Office, Superintendent of Documents, Washington, DC 20402 and National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161

Library of Congress Cataloging-in-Publication Data
Main entry under title:

Statistical inference from band recovery data.

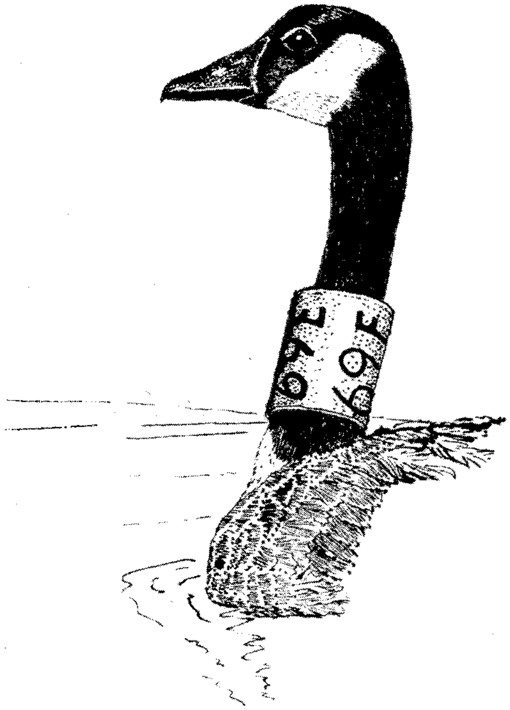
(Resource publication/United States Department of the Interior, Fish and Wildlife Service; no. 156)
Bibliography: p.
Supt. of Docs. no.: I 49.66:156
1. Bird-banding-Statistical methods. 2. Animal marking-Statistical methods. I. Brownie, Cavell. II. U.S. Fish and Wildlife Service. III. Series: Resource publication (U.S. Fish and Wildlife Service); no.156.
S914.A3 no. 156a 333.95'4'0973 s 85-600160
(QL677.5) (596'.0524'0723)



STATISTICAL INFERENCE FROM BAND RECOVERY DATA — A HANDBOOK

Second Edition

By Cavell Brownie
David R. Anderson
Kenneth P. Burnham
Douglas S. Robson



UNITED STATES
DEPARTMENT OF THE INTERIOR
Fish and Wildlife Service
Resource Publication No. 156
Washington, D.C. • 1985

Preface to the Second Edition

A second edition to *Statistical Inference from Band Recovery Data — A Handbook* was prompted by the exhaustion of the existing copies. Funds were sought for a second printing and plans were made to update certain sections. We believe this second edition will serve users for many years to come.

There has been relatively little new theory on the class of estimation and testing problems addressed in the *Handbook* since its publication in 1978. However, there have been significant developments in four areas. First, the assumptions have been investigated further and studies made on estimator robustness to partial failure of certain assumptions. These studies provide a firm setting for assessing the validity of estimates, and other inferences, from banding data. Second, various extensions have been developed to model survival rates as functions of auxillary information and thereby use banding data to explore population dynamic processes. Third, a large amount of new theory has been developed since 1978 on the analysis of multiple recapture data. This important subject was just entering a period of rapid development at the time Section 8.2 was written. Many of the new references added relate to these new developments. Fourth, many changes have been made to computer software. Programs ESTIMATE and BROWNIE have been written to be interactive and made to function on a host of microcomputers. In addition, two new, very general, programs allow the sophisticated user to fully explore arbitrarily complex models for the analysis of banding and recovery data.

The material in Chapters 1 - 5, 7, and 9 are little changed, but the developments listed above have led to rewriting material in Chapters 6 and 8. Chapter 6 has been rewritten to reflect the many changes in available computer software. Section 8.2 has been modified extensively to provide the reader insight into the recently developed methods for analyzing multiple recapture data (open population models for capture-recapture data). Sample size calculations described in Section 9.3 are now available in program ESTIMATE as an option. The *Bibliography* and *Other Literature Cited* sections have been merged and updated with the inclusion of 69 new references. New references are starred to allow the reader quick access to the new citations. Two minor errors have been corrected in Appendix A dealing with log-likelihood tests. Finally, we have added Appendix C to include 10 recent papers that are very important to the subject of banding data analysis. These are reproduced exactly as they appeared in various journals, including the original pagination. While those papers selected for reprinting reflect our personal judgement, we feel their inclusion provides a state-of-the-art Handbook. We believe the present theory has matured and stabilized and sophisticated computer algorithms now exist for easy application and analysis of almost all band recovery data.

We appreciate the efforts of Dr. Rollin D. Sparrowe and Richard S. Pospahala of the Office of Migratory Bird Management, U.S. Fish and Wildlife Service for authorizing the funds for the printing of this edition. We thank the editors of *Biometrics, Ecology, Journal of Animal Ecology, Journal of Field Ornithology,* and the *Journal of Wildlife Management* for permission to reprint the papers in Appendix C.

As we complete the second edition, Cavell Brownie has returned to the United States to a postion in the Department of Statistics at North Carolina State University, Raleigh, NC; David Anderson is the Unit Leader of the Colorado Cooperative Fish and Wildlife Research Unit at Colorado State University, after spending nine years in Utah; Ken Burnham left Colorado after eight years to take a position with the Agricultural Research Service and the Department of Statistics, North Carolina State University, in Raleigh NC; and Doug Robson has remained at Cornell University.

C. Brownie, D.R. Anderson,
K.P. Burnham, and D.S. Robson

June 1, 1985

Preface to the First Edition

This handbook presents a discussion of modern methods for the detailed analysis of certain types of marking studies of animal populations. The discussion and examples focus on bird banding studies, which are a common and important application and permit a consistent terminology. The estimation methods and statistical testing procedures presented here are potentially applicable to fish tagging experiments, entomological investigations, and studies of certain reptiles and amphibians. Bird banding studies, as these of game and nongame birds both migratory and resident, are perhaps the most extensive field of application. Indeed, probably over 40 million birds have been banded in North America alone.

The material presented here represents research done by the authors over a 4-year period. Interest in the specific subject was engendered by G. A. F. Seber's paper in *Biometrika* (1970). (Robson and Youngs had developed the same model independently while Seber's paper was in press.) At that time D. R. Anderson was studying population ecology of the mallard at Patuxent Wildlife Research Center and wanted to extend Seber's stochastic model to admit age-specificity. Subsequent contact with D. S. Robson and C. Brownie at the Cornell Biometrics Unit soon produced a contract funded by the U.S. Fish and Wildlife Service to explore the age-specific estimation problem.

This work led to the development of a series of models that we believe will be useful to many persons in years ahead. Under the initial contract, three new models (herein called H_1, H_2, and H_3) were developed for the analysis of age-dependent banding data. The contract was then extended and additional models were developed to allow further generalizations and new experimental situations. Eight of these models (H_1 through H_8) formed the basis for Brownie's Ph.D. dissertation in biometry written under Robson at Cornell University. K. P. Burnham's arrival at Patuxent during the contract work stimulated further thought and development particularly with regard to additional age-independent models and further testing procedures. Comprehensive computer programs were developed at Patuxent for the age-independent models and at Cornell for the age-dependent models.

With the analysis of 4 million mallard bandings in progress at Patuxent, we could still see the need for additional models. Five more models (M_0, M_2, M_3, H_{01} and H_{02}) were developed and incorporated into the existing computer programs. We began to consider other issues as well: additional statistical tests required, estimation of instantaneous mortality rates under various assumptions, the power of certain tests, statistical bias of estimators, geometric means of survival rates, estimators of band reporting rates, and sampling correlations between certain estimators.

Our efforts in this area have reached a convenient stopping point. The analysis procedures for the most common field experiments are now well developed. Still, there are several additional areas deserving attention. Most notable are the age-specific marking experiments of animal populations captured and released alive. Also, estimation and testing procedures for experiments involving both recovery information and data on live recaptures need to be developed and computer programs written to facilitate their use.

It is important for users of this handbook to know and understand the assumptions of a particular method, test, or model. The results are dependent, sometimes critically, upon the assumptions being made. Consequently, we have given considerable emphasis to these underlying assumptions. Furthermore, we present tests to assess the goodness of fit of each model to the data and tests between models. The assumptions and the procedures to statistically test these assumptions are as important as the estimation methods themselves.

The most useful results of our efforts are presented in this handbook, which we hope is a simple, easy-to-read primer. The subject of estimating parameters from marking experiments of animal populations is now very advanced. We have tried to simplify the presentation by employing a number of examples from real data. Of course, the availability of the computer programs alleviates many of the technical difficulties faced by biologists who use these methods. A guide to the mathematical theory underlying these methods is presented in two appendices.

The handbook is written on a level that should be understood by biologists who have taken two or three courses in applied statistics and a course in differential calculus. We have had to assume the reader is familiar with concepts such as random variables, estimators, sampling variances, confidence intervals, and chi-square test statistics. We make no apologies for these fundamental requirements.

As we complete work on this manuscript, C. Brownie has returned to Jamaica to live, D. R. Anderson has taken a position in Utah, K. P. Burnham has taken a position in Colorado, and D. S. Robson is heading for Australia on Sabbatical leave.

Cavell Brownie, David R. Anderson,
Kenneth P. Burnham, and Douglas S. Robson
June 1, 1976

Contents

Contents – Continued

Contents – Continued

Contents – Continued

Contents – Continued

List of Tables

List of Examples

List of Figures

Statistical Inference from Band Recovery Data
– A Handbook –

by

Cavell Brownie

Department of Statistics
North Carolina State University, Raleigh, NC 27695

David R. Anderson

Colorado Cooperative Fish & Wildlife Research Unit
Colorado State University, Fort Collins, CO 80523

Kenneth P. Burnham

Department of Statistics
North Carolina State University, Raleigh, NC 27695

Douglas S. Robson

Biometrics Unit
Cornell University, Ithaca, NY 14850

Chapter 1. Introduction

1.1 Purpose and Scope of Handbook

This handbook was prepared as an aid to those engaged in the analysis of several kinds of bird banding and other animal tagging studies. A common objective in most of these studies is the estimation of parameters which will reflect population survival. Here we focus considerable attention on the estimation of survival rates and specifically concentrate on inference procedures (estimation and hypothesis tests) regarding time- and age-specific survival rates.

We focus on migratory bird banding studies because such studies motivated our research and the development of most of the techniques discussed in the following chapters. In addition, bird banding studies are relatively widely read and employ a common, simple terminology. We hope the consistent use of bird banding terminology and examples will aid rather than confuse people involved in other types of animal tagging experiments. We feel that the methods presented in this handbook are potentially applicable to a wide range of field studies in addition to bird banding studies: fish tagging; bat banding; marking studies of certain reptiles and amphibians, several marine and terrestrial mammals; and a wide variety of entomological investigations.

We examine animal banding or tagging studies where: (1) A number of animals are caught and banded (tag, mark) each time period (year, month, week) for k such equal time periods (often the time period is a year). Generally, each band or tag carries a unique number or code. (2) Records are kept of bands or tags reported from dead animals for each time period from each batch banded. Therefore, the recovery data conveniently form an array representing the number of bands or tags recovered in time period j from those animals originally banded in time period i.

We consider only bands from dead animals. However, the methods we discuss in the following chapters can often be used as a good approximation to the analysis of data from studies of banded animals recaptured alive and subsequently released again, if the recapture rate is low. This subject, which is discussed in Section 8.2, extends the usefulness of the methods described. The development of comprehensive, efficient inference methods for many common types of recapture studies is incomplete and, therefore, it may be appropriate to use the techniques presented here as useful approximations until more general and efficient methods are developed and made available.

Banding studies of exploited species of birds are typical of the types of field research covered in this handbook. For example, a large sample of birds such as adult male pintail *(Anas acuta)*, are banded in August each year for k consecutive years in one general area. Records are kept by the Bird Banding Laboratory on the number of bands reported from dead birds in the j^{th} year after banding. (For game bird data, it is common to analyze only recoveries from normal, wild birds that were shot during the legal hunting season.)

Several methods for the analysis of such data have been proposed in the literature over the last 4 decades. The early methods were unsophisticated and poorly developed and usually based on the deterministic life table concept taken from human demography. Only in recent years have the proper methods for analysis of these data been developed. Seber (1962) probably made the original contribution to the correct conceptualization of analysis procedures (although he drew from the results of Darroch [1959] and others). More recent papers of importance include **Jolly (1965), Robson and Youngs (1971), and Seber (1965, 1972). Cormack (1979) and Seber (1982) present detailed** reviews of the various methods for such data as well as analysis techniques for live recapture experiments.

This handbook covers the analysis of banding studies for one, two, or three identifiable age classes; it also presents methods for use when banding is done twice a year on the same population. In all, we discuss 14 models, each allowing different and testable assumptions. For each model we present optimal estimators of certain parameters, the most important of which are annual survival and recovery rates (other parameters include mean life span and average annual survival and recovery rates). Estimates of sampling variation (precision) are given for all parameter estimators. Confidence intervals on parameters are presented and, for models currently of practical value, goodness of fit tests are presented. Also, tests between models are presented which are useful for selection of the appropriate model and for pooling data sets. The last chapter is devoted to the subject of planning a banding study.

1.2 Basic Structure of Band Recovery Data

We will illustrate the basic nature of experimental banding data with a study from a midwestern State where adult male birds were banded in August each year in 1964-66. Banding was initiated in 1964 when 1,603 adults were banded. In August 1965, 1,595 birds were banded. Some banded birds from these two cohorts were killed during the hunting seasons of 1964 and 1965 and their bands were reported to the Banding Laboratory. After 2 years, the recovery data can be summarized in the following triangular array:

Year banded	Number banded	Recoveries by hunting season 1964	1965
1964	1,603	127	44
1965	1,595		62

The interpretation of this table is straightforward. Of the 1,603 birds banded and released before the 1964 hunting season, 127 bands were recovered from birds killed during the 1964 season. From survivors, of this original banded cohort of 1,603, alive at the start of the 1965 hunting season, 44 bands were recovered. Sixty-two bands were recovered during the 1965 hunting season from the 1,595 birds banded just before the 1965 season.

In 1966, 1,157 new birds were banded, and from this cohort 82 bands were recovered during the 1966 hunting season. The recovery data, through 1966, can now be summarized in the following triangular array:

Year banded	Number banded	Recoveries by hunting season 1964	1965	1966
1964	1,603	127	44	37
1965	1,595		62	76
1966	1,157			82

From the 1966 hunting season a total of $195(= 37 + 76 + 82)$ bands were recovered from the three banded cohorts.

The analysis of this data could be accomplished by considering only the recovery data from the 1964-66 hunting seasons. The methods described in this handbook are free of "truncation" problems that are associated with many life table methods (i.e., analyzing the recoveries before all banded birds are dead). Although no bias results from the use of only the 1964-66 data, additional precision in the parameter estimates will be realized by using recovery data that are obtained after the 1966 hunting season. For example, if the recovery data from the 1967 and 1968 hunting seasons were also available, the analysis could be improved (greater precision) by considering the extended, nontriangular data array below:

Year banded	Number banded	Recoveries by hunting season				
		1964	1965	1966	1967	1968
1964	1,603	127	44	37	40	17
1965	1,595		62	76	44	28
1966	1,157			82	61	24

Experimental data, such as those shown above, could result from banding programs conducted during the winter (when birds are on the wintering grounds) or in late summer (when birds are still on the breeding grounds). Birds are generally not banded during the breeding seasons so as not to disturb them during this critical period. Banding during migration or during the hunting season should be avoided because of difficulties in interpretation and analysis of such data.

In general, banding studies will involve catching, banding, and releasing a sample from some population at regular intervals. (In bird banding studies, intervals are usually 1 calendar year.) We introduce the following terminology to facilitate developments of this chapter. For k banding occasions let $N_1, \ldots, N_k$ be the numbers banded and released back into the population. Equal time intervals between bandings are assumed. Let the band recovery data be represented by R_{ij} defined as

R_{ij} = the number of band recoveries in hunting season j from birds originally banded in year i.

In the example above $R_{11}=127$, $R_{12}=44$, $R_{23}=76$, etc. The general method we recommend for displaying data is shown below in this symbolic notation

Year banded	Number banded	Recoveries by hunting season				
		1	2	3,	...	,k
1	N_1	R_{11}	R_{12}	R_{13},	···	,R_{1k}
2	N_2		R_{22}	R_{23},	···	,R_{2k}
3	N_3			R_{33},	···	,R_{3k}
.	.				.	.
.	.				.	.
.	.				.	.
k	N_k					R_{kk}

We make the following definitions:

k = the number of years of banding
ℓ = the total number of years of recovery ($\ell \geqslant k$).

Often one has $k=\ell$ as in the triangular data array above. In general, however, band recoveries continue to be reported after the banding program has ceased. In the example above, the extended, nontriangular data array has $k=3$ and $\ell=5$.

For adults from one species and sex, there will be just one such recovery data array per study. Often, however, both young and adults are banded from the same species and population. Then there will be two such data arrays to consider for statistical analysis. Less often, three age classes are banded: young, subadults, and adults. Also, there may be birds of both sexes banded leading to yet more data arrays to consider. Models (14 in all) and data analysis procedures for all these situations are dealt with in chapters to follow.

1.3 Modeling Band Recovery Data

General Concepts

Explicitly-stated models are essential to any statistical inference problem; it is the model which hypothesizes the relationship between what we know (the data) and what we do not know but want to estimate, e.g., the relationship of band recovery data to the animal population under study. Because the detailed statistical analyses and their recommended interpretation are critically dependent upon the various models used, it is imperative that the reader understand the assumptions of these models. For this reason, in the following chapters, we have stressed the nature and assumptions of each model. This section gives an introduction to how these models are conceived and structured, and to the basic underlying assumptions.

When an animal is banded or tagged we cannot predict when, if ever, the band or tag will be recovered; thus the event "the band is recovered" must be conceptualized as a random event. The correct model to describe these events requires the use of probability statements, and consequently must be stochastic (probabilistic) as opposed to the older deterministic descriptions of such data.

For example, consider just the recoveries R_{11} from the first hunting season after banding N_1 birds. The appropriate model is a probabilistic one, which treats R_{11} as a random variable. If we let the parameter f_1 be the band recovery rate (i.e., the probability of a band recovery from any bird) for this first year, an appropriate model takes R_{11} as a binomial random variable with sample size N_1, and rate parameter f_1. In abbreviated form we say R_{11} is binomial (N_1,f_1). We note this model has both stochastic and structural components, with the structural component being represented by the average or expected value of R_{11}, symbolized as $\mathbf{E}(R_{11})=N_1f_1$. Let R_{12} be the number of bands recovered in the second hunting season from the N_1 banded birds. Let S_1 be the survival rate during this first year of the study (the calendar year between bandings). Let f_2 be the band recovery rate from birds alive at the time of banding in the second year of the study. Then the model structure for the expected number of recoveries in year 2 $(\mathbf{E}(R_{12}))$ is $N_1S_1f_2$. A model for R_{12}, as a random variable, is the binomial (N_1,S_1f_2) model.

In practice we do not deal with just 1 year of recoveries from one banded cohort. As explained in the previous section, bandings will be made at approximately yearly intervals, for some period of years, and band recoveries obtained over several years from each banded cohort. The model used must relate to the entire study whether there is one or more array of recovery data (one or more age classes, and one or both sexes). There are three components to these models: (1) the model structure, which expresses the expected recoveries in terms of numbers banded and survival and recovery rate parameters; (2) the stochastic component which describes the sampling probability distribution of the recovery data (i.e., recognizing the recoveries are random variables and are from a sample); and (3) a component that is never explicity used. The third component is the assumption that the banded sample is representative of the larger population, hence inferences (such as survival rate estimates) apply to this population, not just the sample at hand. This last component can be broken into numerous assumptions, most of them not testable from the recovery data, and their validity generally depends upon the sampling design and field procedures used in banding.

In this section and throughout most of the handbook, when we refer to a model, we mean only the structural component. Only in Section 1.5 do we discuss the general sampling probability model for these types of data. The next two subsections should give the reader a better understanding of model structures and assumptions.

Model Structures

The structure of band recovery models is generally based on two types of parameters: an annual survival rate S and an annual band recovery rate f. These parameters have probability interpretations as follows: S = the probability that a bird alive when a given cohort is banded will survive 1 calendar year to the time of next banding, and f = the probability that a banded bird alive when a given cohort is banded will be shot and its band reported during the next hunting season.

We emphasize that the critical or key assumptions in constructing a model for the analysis of banding data relate to these survival and recovery rate parameters. For example, do the values of these parameters vary with the age or sex of the bird, its capture and banding history, or the calendar year? It is of biological importance to know if the survival rate for a particular species or population segment is age- or sex-specific. In an attempt both to answer these questions and to obtain estimates of the parameters, a series of models (hypotheses) have been developed, each with a different model structure. These models differ depending on the hypothesis concerning how S and f vary.

Consider the hypothesis (tentative assumption) that recovery and survival rates are constant (i.e., they do not vary by age of the bird or calendar year). This hypothesis leads to the following model structure representing the expected number of band recoveries:

Year banded	Number banded	Expected recoveries by hunting season, $\mathbf{E}(R_{ij})$			
		1	2	3	$\ell=4$
1	N_1	N_1f	N_1Sf	N_1SSf	N_1SSSf
2	N_2		N_2f	N_2Sf	$[N_2SSf]$
$k=3$	N_3			N_3f	N_3Sf

As above, our representation of a model will entail its structure only and be expressed in a form analogous to the recovery data themselves.

Consider the expression in brackets in the table above. This term represents the expected number of band recoveries the fourth year of the banding study, from the sample of N_2 birds banded in the second year of the study, $\mathbf{E}(R_{24}) = N_2SSf$. Its meaning is simple and logical. The expected number of survivors during year 2 of the study is N_2S. From the start of year 3, we expect a proportion S of these survivors to live to year 4 of the study. Hence, there are N_2SS survivors expected at the start of year 4. A proportion f of these survivors will, on the average, be shot and their bands reported. Thus the expected number of band recoveries in the fourth year of the study from the N_2 birds banded in year 2 is N_2SSf. In this particular model we are hypothesizing that S and f are constant from year to year and independent of age.

Now, let us specify, for illustrative purposes, that the underlying population parameters are:

$$f = 0.10 \text{ or a } 10\% \text{ annual recovery rate,}$$
$$S = 0.50 \text{ or a } 50\% \text{ annual survival rate,}$$

and that, in fact, f and S are indeed constant. We also specify (for the purpose of this example):

$$\left.\begin{array}{l} N_1 = 2{,}000 \\ N_2 = \phantom{1{,}}400 \\ N_3 = 1{,}200 \end{array}\right\} \text{Number banded in the } i^{th} \text{ year.}$$

Then, on the average we would expect the following recovery data:

Year banded	Number banded	Expected recoveries by hunting season, $\mathbf{E}(R_{ij})$			
		1	2	3	$\ell = 4$
1	2,000	200	100	50	25
2	400		40	20	[10]
$k = 3$	1,200			120	60

Return again to the value that is bracketed: $N_2SSf = 400 \times 0.50 \times 0.50 \times 0.10 = 10$ recoveries.

The expectations of the recovery data can be easily expressed for this simple model in terms of the two parameters S and f (for arbitrary values of k and ℓ, the number of years of banding and recovery, respectively):

$$\mathbf{E}(R_{ii}) = N_if \qquad ,i = 1, \ldots, k,$$
$$\mathbf{E}(R_{ij}) = N_iS^{j-i}f \qquad ,i = 1, \ldots, k, \qquad j = i+1, \ldots, \ell \ .$$

These concepts could be made more general and realistic if, for instance, recovery and survival rates were hypothesized to vary each year. The model structure representing these tentative assumptions would be:

Year banded	Number banded	Expected recoveries by hunting seasion, $\mathbf{E}(R_{ij})$			
		1	2	3	$\ell = 4$
1	N_1	N_1f_1	$N_1S_1f_2$	$N_1S_1S_2f_3$	$N_1S_1S_2S_3f_4$
2	N_2		N_2f_2	$N_2S_2f_3$	$[N_2S_2S_3f_4]$
$k = 3$	N_3			N_3f_3	$N_3S_3f_4$

Here, the parameters of the model are subscripted to indicate the dependence on calendar year (i.e., the parameters S and f are year-specific).

Again, for purposes of example, let us specify that the underlying population parameters are:

Year	Recovery rates	Survival rates
1	$f_1 = 0.05$ or 5%	$S_1 = 0.50$ or 50%
2	$f_2 = 0.10$ or 10%	$S_2 = 0.50$ or 50%
3	$f_3 = 0.06$ or 6%	$S_3 = 0.70$ or 70%
4	$f_4 = 0.05$ or 5%	

Given the above, on the average we would expect the following recovery data (the number banded will remain the same as the previous example):

Year banded	Number banded	Expected recoveries by hunting season, $\mathbf{E}(R_{ij})$			
		1	2	3	$\ell=4$
1	2,000	100	100	30	18
2	400		40	12	[7]
$k=3$	1,200			72	42

Here, for example, the term in brackets is $N_2S_2S_3f_4 = 400 \times 0.50 \times 0.70 \times 0.05 = 7$ recoveries. Notice that the expected number of recoveries is quite different under the two models. In general, the expectations under the assumptions that recovery and survival rates are time-specific can be expressed in terms of these parameters as:

$$\mathbf{E}(R_{ii}) = N_i f_i \qquad ,i = 1, \ldots, k,$$
$$\mathbf{E}(R_{ij}) = N_i S_i \cdots S_{j-1} f_j \qquad ,i = 1, \ldots, k, \qquad j = i+1, \ldots, \ell.$$

In this case, the parameters $S_1, \ldots, S_{k-1}$ and $f_1, \ldots, f_k$ are estimable whether $\ell = k$ or not. If, however, $\ell > k$ the parameters $S_k, \ldots, S_\ell$ and $f_{k+1}, \ldots, f_\ell$ are not separately estimable; only the products such as $S_k f_{k+1}$, $S_k S_{k+1} f_{k+2}, \ldots,$ $S_k \cdots S_{\ell-1} f$ are estimable. This subject is discussed in Section 2.2 of Chapter 2.

The above examples show the relationships among the specific hypotheses about survival and recovery rates, model structures representing these hypotheses, and the values that would be expected on the average under such a model. Of course, the actual recovery data will vary somewhat from the expected or average values because survival and band reporting are random events. In practice, the biologist has only the observed data (plus some knowledge of the biology of the species), but usually does not have *a priori* knowledge of the underlying process and, therefore, does not know what model is most appropriate.

Assumptions

As mentioned above, there are numerous assumptions involved in making inferences from banding data; important ones are listed below.

Assumptions relating to study planning, field procedures, and type of species:

(1) The sample is representative of the target population;
(2) age and sex of individuals are correctly determined;
(3) there is no band loss;
(4) survival rates are not affected by the banding or tagging itself; and
(5) the year (hunting season) of band recoveries is correctly tabulated.

Assumptions relating to the stochastic model component:

(6) The fate of each banded animal is independent of (not correlated with) the fate of other banded individuals;
(7) the fate of a given banded animal (i.e., band recovery in year 1, 2,... after banding) is a multinomial random variable.

Assumptions relating to model structure:

(8) All banded individuals of an identifiable class (e.g., by species, age, sex) in the sample have the same annual survival and recovery rates (except model H_3 of Chapter 3); and
(9) annual survival and recovery rates may vary by calendar year, and/or by age and sex of individuals (variation of survival and recovery rates by area [population] is also possible).

In practice, assumption 9 is a series of very specific assumptions which can be investigated and tested in detail from the recovery data themselves. These assumptions specify the exact model structure, they constitute the focus of this handbook, and in the usual type of banding study they are the only testable assumptions.

It is beyond the scope of this handbook to discuss field procedures in detail because they will vary by species and study. Sections 9.1 and 9.2 offer some guidelines on this subject. Obtaining a representative sample is important; in studying populations. This is always of potential concern (see Weatherhead and Ankney 1984, 1985 and Burnham and Nichols 1985). Recently, Pollock and Raveling (1982) and Nichols et al. (1982) presented valuable information on these assumptions (see Appendix C). Nelson et al. (1980) examines the effect of band loss (assumption 3)

on survival estimators and Anderson and Burnham (1980) investigate the effect of delayed reporting of bands (assumption 5) on these methods (both of these papers appear in Appendix C).

Based on assumptions 6, 7, and 8, we model the recovery data from any banded cohort as multinomially distributed random variables. This distribution is the basis for deriving sampling variances, and covariances of estimators, tests for goodness of fit, and tests between models. We believe these tests do not critically depend upon this presumed sampling distribution; hence, some degree of failure of this assumption (which occurs if assumption 6 is not true) does not destroy the usefulness of the methods presented here. Assumption 7 is shown to be mathematically true in Section 1.5. Assumption 8 is important but it is generally not testable if there is no method even in principle, to further partition the identified class of animals. However, the test between models H_1 and H_3 (Chapter 3) is a case where assumption 8 is testable.

Given assumptions 1 through 8, we have a model in which the only unknown is the structure of the expected band recoveries as a function of annual survival and recovery rates, i.e., $\mathbf{E}(R_{ij}) = N \times$ (a function of annual survival and recovery rates). In subsequent chapters when we discuss model assumptions, we mean those distinct, detailed assumptions which specify a model structure for expected band recoveries in terms of variations in S and f (survival and recovery rates) with respect to time (calendar year), area, age, or sex.

The reader should be aware of one more assumption. We have limited ourselves to instances of equal time periods (usually 1 year) between bandings, the most common situation. All the analysis methods presented herein assume equal time periods, and thus certain of these models and tests are not appropriate for use on data with unequal time periods between banding or tagging. This subject is discussed further in Section 8.5.

1.4 Data Analysis

Estimation with a Given Model

Given an explicit, biologically reasonable stochastic model representing a specific hypothesis about survival and band recovery rates, we are faced with estimating the parameters of the model. Fortunately, a well-developed theory for efficient estimation of parameters from such models exists. This theory of parameter estimation has had a long history; most of the initial developments and their early elaboration are credited to the famous statistician Sir Ronald A. Fisher. In the 1920's, Fisher published extensively on the method he called "Maximum Likelihood." This method forms one fundamental approach to statistical estimation and inference theory. For many practical models, Maximum Likelihood (ML) estimators are optimal in many respects (e.g., for a given model, no other method will produce consistent estimators with a smaller sampling variance). Most modern methods for estimating parameters from animal marking experiments are based on the ML method, and it is the basis for the estimation procedures in this handbook.

It is not necessary for the reader to understand the statistical theory underlying any of the data analysis methods presented in this handbook (e.g., parameter estimation, estimation of sampling variances and covariances, confidence interval construction, goodness of fit tests, or tests between models). All the general underlying statistical theory is fairly standard, and its properties are known. Thus we can assert that most of the data analyses methods presented here can not be improved upon. These methods make optimal use of the data under any given model. We give some minimal introduction to the relevant mathematical methods behind these models in Section 1.5 and in the Appendices. However, for the most part, the interested reader will have to consult the original references for the details of estimation and inference developments for each model.

Our strategy in the following chapters that discuss given models is to first present the model structure and the key assumptions on survival and recovery rates which yield this structure. Then, where possible, the formulae for parameter estimators are given; finally, the formulae are given for sampling variances and covariances of these estimators. The following type of notation is used: If X represents a parameter, then its ML estimator is denoted by $\hat{X}$. Theoretical sampling variances, covariances, etc., and their estimators are denoted as shown below:

Quantity	Theoretical	Estimated
variance	$\text{VAR}(\hat{X})$	$\text{var}(\hat{X})$
standard error	$\text{SE}(\hat{X})$	$\text{se}(\hat{X})$
covariance	$\text{COV}(\hat{X},\hat{Y})$	$\text{cov}(\hat{X},\hat{Y})$
correlation	$\text{CORR}(\hat{X},\hat{Y})$	$\text{corr}(\hat{X},\hat{Y})$

Often ML estimators are slightly biased. This is true with some estimators (mainly estimators of survival rates) under several models. We have developed adjusted estimators that are essentially free of this small sample statistical bias. The ML estimators that are adjusted for statistical bias are represented by a tilde (˜) over the parameter, e.g., $\tilde{X}$. Correspondingly, $\tilde{X}$ replaces $\hat{X}$ in formulae for sampling variances, covariances, and correlations, e.g., $\text{VAR}(\tilde{X})$ and $\text{var}(\tilde{X})$. (We do assume that the reader has had some basic statistics, and is familiar with concepts of estimation, bias, accuracy, sampling variation, models, etc. A useful reference on this subject is Overton 1969.)

As part of the analysis of recovery data under any given model we also usually give a goodness of fit test to judge whether the model is an acceptable fit to the data. These goodness of fit tests are meant to be used in the processes of selecting the appropriate model for any given data set.

Selecting the Best Model

The proper model must be used if sound inferences are to be made from the analysis of banding or tagging data. We want to adopt the simplest model which adequately fits the given data set. This philosophy follows the general principle in science of using the simplest acceptable model to describe a phenomenon (Occam's razor).

First we will examine the alternatives to this philosphy: models that are too simple and models that are too general. In using models of banding studies, we have found that overly simple and restrictive models are worse than overly general models. Here the simplified model, and the hypothesis (tentative set of assumptions) it represents, does not fit the observed data. Here we risk substantial bias (regardless of how large the sample size) in the estimates of parameters and/or their sampling variances and covariances. Worse yet, the estimates of sampling variances are almost always too small. Consequently, the investigator has a false sense of security because he/she believes the estimates are very precise. This bias can lead to apparently significant differences if, for example, survival rates are compared between two geographic areas, where actually no differences can properly be shown. Use of a model that is too simple is almost always nonconservative. The composite dynamic method has been used in the analysis of bird banding data for 35 years but few investigators had ever bothered to see if this "model" fit the data being analyzed. These assumptions were recently tested for waterfowl (Burnham and Anderson 1979) and nongame birds (Anderson et al. 1981) and were soundly rejected. These facts further emphasize the need for proper testing procedures to aid the selection of an adequate set of models.

The second alternative is the use of a model that is too general, a model that allows more parameters than are necessary. Here, the point estimators of parameters are inefficient, although they are unbiased. The estimated sampling variances and covariances are large, indicating that precision has been sacrificed. Testing procedures tend to be conservative, failing to detect significant differences that, in fact, exist.

We are recommending two types of procedures to statistically test various models and to select the proper one. The first test is a goodness of fit test. Here the null hypothesis is that the model fits the data, while the alternative hypothesis is a general one – the model does not fit the data. If the null hypothesis is rejected for a particular data set, that model should no longer be considered. The second test is more specific and tests one model against another. Here the alternative is specific. The null hypothesis is that the simplest model, the one with the fewest parameters, fits. Again, if the null hypothesis is rejected, the simple model should not be considered further in the analysis of a particular data set.

Choice of a "proper" and "adequate" model is somewhat relative since other models may also be "good." For example, Model H_1 (given in Chapter 3 and used for the analysis of banding of both young and adults) is often a good model of a bird population with age-specific parameters. Results of the goodness of fit test usually fail to reject H_1 whereas simpler models are often rejected, and more general models are often shown to be unnecessary for a particular data set. However, although Model H_1 might be adequate for the analysis of the data set, there may be other models that are equally appropriate. For example, Johnson (1974) developed a model similar to H_1, but containing fewer parameters because of two restrictions. In many situations, both models would be satisfactory and we have little preference as to which of the two is better except that Johnson's has fewer parameters. (Note: we do not discuss Johnson's model in this handbook because it is not fully developed; Model H_1 is fully developed, complete with intensive testing features and a comprehensive computer code).

Computer Algorithms

We have developed two large computer programs to alleviate the tedious task of computing the estimators, test statistics, etc., presented in this handbook. Program ESTIMATE is for age-independent populations and is discussed in Chapter 2. Program BROWNIE was written for several age-dependent models and to test for sex-specific parameters in adult recovery data. Output from BROWNIE is discussed in Chapters 3, 4, and 5. The use of programs ESTIMATE and BROWNIE is discussed in Chapter 6.

1.5 Mathematical Overview

This section can be omitted by the reader with little knowledge or interest in mathematical statistics.

Model Conceptualization

Consider a study where banding is done for k consecutive years, and recoveries are accumulated for $\ell \geqslant k$ years. For a given bird banded in year i, exactly one of $\ell - i + 2$ discrete events regarding the band will occur. Either the band is recovered in one of the $\ell - i + 1$ years after banding, or it is never recovered. For recovery to occur in year j the bird must survive until the start of year j, then be shot during that year, and the band must be reported to the Bird Banding Laboratory. To represent these outcomes, define an indicator random variable for a bird banded in year i $(i = 1, \ldots, k)$ as

$$X_{ij} = \begin{cases} 1 \text{ if the band is recovered in year } j,\ j = i, \ldots, \ell, \\ 0 \text{ otherwise.} \end{cases}$$

For notational convenience, define $X_{i,\ell+1} = 1 - \sum_{j=i}^{\ell} X_{ij}$. Hence, $X_{i,\ell+1}$ is 1 if and only if the band is never recovered. Finally define the vector valued random variable $\underline{X}_i = (X_{ii}, \ldots, X_{i,\ell+1})$.

The random variable $\underline{X}_i$ has the classic multinominal probability distribution with $\ell - i + 2$ cells (Johnson and and Kotz 1969). Let $P\{X_{ij} = 1\} = \pi_{ij}$ be the probability the band is recovered in year $j, j = i, \ldots, \ell$. Then $\pi_{i,\ell+1} = 1 - \sum_{j=i}^{\ell} \pi_{ij}$ is the probability the band is never recovered. The probability of any of the $\ell - i + 2$ possible outcomes is represented as $P\{\underline{X}_i\} = \prod_{j=i}^{\ell+1} (\pi_{ij})^{X_{ij}}$.

A general structure on the π_{ij} cell probabilities can be used without loss of generality. A band can be recovered in year j only if the bird survived and retained its band until the start of year j. Then it must be shot and the band reported to the Bird Banding Laboratory. Consequently, define the following conditional annual survival, band retention, and recovery rates:

S_{ji} = probability of surviving year j, given the bird was alive at the start of year j and was banded at the start of year $i \leqslant j$;

θ_{ji} = probability of band retention in year j given the bird still has the band at the start of year j and was banded in year i; and

f_{ji} = probability that a bird alive at the start of year j, which was banded in year i, is shot in year j and its band reported to the Banding Laboratory.

Now a general structural representation for any π_{ij} is

$$\pi_{ij} = \begin{cases} f_{ii} & ,j = i \\ \left[\prod_{h=i}^{j-1} (S_{hi}\theta_{hi})\right] f_{ji} & ,j > i. \end{cases}$$

It should be clear that survival rates are confounded with band retention rates. This is indeed true; no useful models can be developed for band recovery data unless we assume there is no band loss, i.e., $\theta_{hi} \equiv 1$. We do make this assumption; it is implicit in all the models presented in this handbook. Given that no band loss occurs, the general structural representation for π_{ij} is

$$\pi_{ij} = \begin{cases} f_{ii} & ,j = i \\ \left[\prod_{h=i}^{j-1} S_{hi}\right] f_{ji} & ,j > i. \end{cases} \tag{1.5.1}$$

Now assume N_i birds from the same population are banded and released in year i. We make two more assumptions to obtain a general model structure for banding studies. First, assume all these N_i birds suffer statistically independent fates. Second, assume that all the N_i banded birds have the same annual survival and recovery probabilities. From a practical standpoint this second assumption is unavoidable if the N_i birds cannot be partitioned, even in principle, into subgroups whose survival and recovery parameters might differ. Thus, with a banded sample from one species, age, sex, and area obtained at one brief time interval this second assumption is acceptable.

Define new summary random variables (for k years of banding)

$$R_{ij} = \text{all recoveries in year } j \text{ of birds banded in year } i, j=i,\dots,\ell, i=1,\dots,k.$$

If we use m to index the variables $\underline{X}_i$ over the N_i birds, then the raw data are $\underline{X}_{i,m}$, $m=1,\dots,N_i$, $i=1,\dots,k$. But for each i, $\underline{X}_{i,1},\dots,\underline{X}_{i,Ni}$ are identically and independently distributed as multinomial random variables; hence, their sum

$$(R_{ii},\dots,R_{i,\ell+1}) = \sum_{m=1}^{N_i} \underline{X}_{i,m}$$

is a sufficient statistic with the multinomial distribution $\text{Mult}(N_i; \pi_{ii},\dots,\pi_{i,\ell+1})$. Consequently,

$$P\{R_{ii},\dots,R_{i\ell}\} = \binom{N_i}{R_{ii},\dots,R_{i\ell},R_{i,\ell+1}} \prod_{j=i}^{\ell+1} [\pi_{ij}]^{R_{ij}},$$

where $R_{i,\ell+1} = N_i - \sum_{j=i}^{\ell} R_{ij}$ and the π_{ij} have the general structure displayed in formula (1.5.1).

The above development shows we can reduce our data to the recovery frequencies $\{R_{ij}\}$ which may be conveniently displayed in an array as:

Year banded	Number banded	Recoveries by hunting season						
		1	2	3	,...,	k	,...,	ℓ
1	N_1	R_{11}	R_{12}	R_{13}	...,	R_{1k},	...,	$R_{1\ell}$
2	N_2		R_{22}	R_{23}	...,	R_{2k},	...,	$R_{2\ell}$
3	N_3			R_{33}	...,	R_{3k},	...,	$R_{3\ell}$
⋮	⋮				⋮	⋮		⋮
k	N_k					R_{kk},	...,	$R_{k\ell}$

where if $\ell = k$ the above array of recoveries is triangular. The number of bands never recovered for a given cohort is $N_i - R_{i\cdot}$ where $R_{i\cdot} = \sum_{j=1}^{\ell} R_{ij}$, hence, their number need not be displayed.

Bearing in mind that each set of recovery data $(R_{ii},\dots,R_{i\ell})$ are independent multinomial random variables for $i=1,\dots,k$; the statistical model for these data is completely specified by displaying the cell probabilities π_{ij} as:

Year banded	Number banded	Model structure, by year of recovery						
		1	2	3	,...,	k	,...,	ℓ
1	N_1	π_{11}	π_{12}	π_{13},	...,	π_{1k},	...,	$\pi_{1\ell}$
2	N_2		π_{22}	π_{23},	...,	π_{2k},	...,	$\pi_{2\ell}$
3	N_3			π_{33},	...,	π_{3k},	...,	$\pi_{3\ell}$
⋮	⋮				⋮	⋮		⋮
k	N_k					π_{kk},	...,	$\pi_{k\ell}$.

The representation $\pi_{ij}=\left(\prod_{h=1}^{j-1} S_{hi}\right)f_{hi}$ is too general to be useful; none of the individual survival or recovery rates are estimable. In order to obtain meaningful, useful models we must make additional assumptions about the nature of the annual survival probabilities and about the annual recovery probabilities. For example, if we band only adults, then perhaps the survival rate is age-independent. This implies S_{hi} need not depend upon the year of banding; rather, adult survival may depend only upon environmental conditions in each year. This implies assuming $S_{hi}\equiv S_h$ independent of the year the bird was banded. Similarly, adult recovery rates may be independent of the year of banding, which implies assuming $f_{hi}\equiv f_h$. Given these assumptions of age-independent survival and recovery rates, the model structure is given by

$$\pi_{ij}=\begin{cases} f_i & ,j=i \\ S_i\cdots S_{j-1}f_j & ,j>i. \end{cases}$$

This model is displayed in array form below:

Year banded	Number banded	Model structure, by year of recovery						
		1	2	3	,...,	k	,...,	ℓ
1	N_1	f_1	S_1f_2	$S_1S_2f_3$,	...,	$S_1S_2\cdots S_{k-1}f_k$,	...,	$S_1S_2\cdots S_{\ell-1}f_\ell$
2	N_2		f_2	S_2f_3,	...,	$S_2\cdots S_{k-1}f_k$,	...,	$S_2\cdots S_{\ell-1}f_\ell$
3	N_3			f_3,	...,	$S_3\cdots S_{k-1}f_k$,	...,	$S_3\cdots S_{\ell-1}f_\ell$
.	.				.	.		.
.	.				.	.		.
.	.				.	.		.
k	N_k					f_k,	...,	$S_k\cdots S_{\ell-1}f_\ell$

Analysis of adult banding data based on the above model (which we call Model 1 here) has been thoroughly developed by Seber (1970) and Robson and Youngs (1971).

The above discusses only one data set. Often birds will be identified as to sex or age, giving rise to multiple data sets that may have some survival or recovery rates in common. The modeling problem in these cases is to specify what parameters may be the same or different across ages and/or sexes. The theory of the problem is not fundamentally altered, just made more complex because of additional parameters and data.

The Likelihood Function

The recoveries from any given banded cohort in the data matrix $\{R_{ij}\}$ are modeled as multinomial random variables. From this fact we can write the joint probability function of the data as

$$P(\{R_{ij}\}\,|\,\underline{S},\underline{f},\underline{N})=\prod_{i=1}^{k}\left\{\binom{N_i}{R_{ii},\ldots,R_{i\ell},R_{i,\ell+1}}\prod_{j=i}^{\ell+1}[\pi_{ij}]^{R_{ij}}\right\}, \tag{1.5.2}$$

where

$$R_{i,\ell+1}=N_i-\sum_{j=i}^{\ell}R_{ij}\equiv N_i-R_{i.}$$

$$\pi_{i,\ell+1}=1-\sum_{j=i}^{\ell}\pi_{ij},$$

$$\pi_{ij}=\left(\prod_{h=i}^{j-1}S_{hi}\right)f_{hi}\qquad ,j=i,\ldots,\ell,\ i=1,\ldots,k,$$

and $\underline{S}$, f are the vectors of parameters. For any given model used in this handbook the actual likelihood function can easily be derived from the above symbolic representation. For example, assuming $S_{hi} \equiv S$, and $f_{hi} \equiv f$ we have

$$\pi_{ij} = S^{(j-i)}f \quad ,j = i,\ldots,\ell \quad ,i = 1,\ldots,k$$

and

$$\pi_{i,\ell+1} = 1 - \left(\frac{1 - S^{\ell - i + 1}}{1 - S}\right)f.$$

Therefore, the likelihood is

$$\mathcal{L}(S,f) = \left[\prod_{i=1}^{k} \left(1 - \left(\frac{1 - S^{\ell - i + 1}}{1 - S}\right)f\right)^{N_i - R_{i.}}\right] f^T S^Q,$$

where

$$T = \sum_{i=1}^{k} \sum_{j=i}^{\ell} R_{ij}, \text{ and } Q = \sum_{i=1}^{k} \sum_{j=i}^{\ell} (j-i)\, R_{ij}.$$

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.

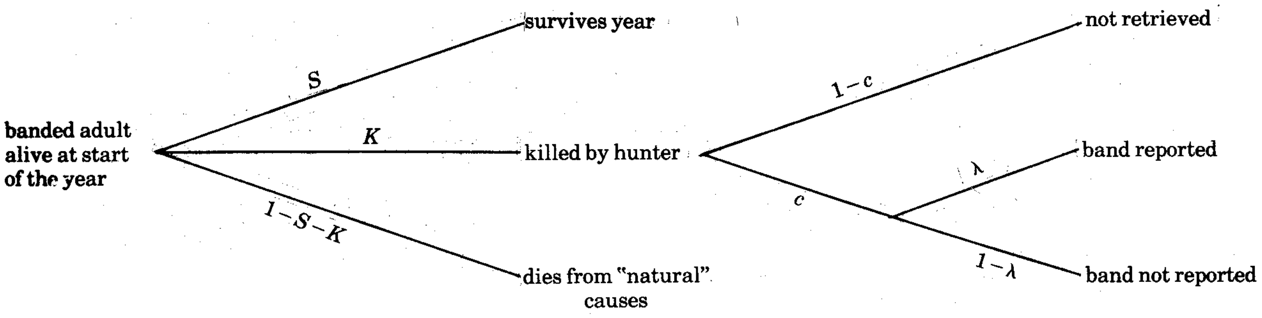


Fig. 2.1 Potential fates of a banded bird alive at the start of the year.

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:

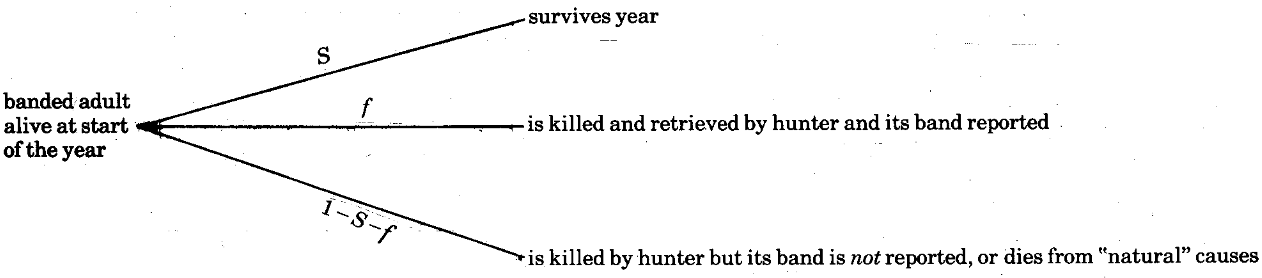


Fig. 2.2 Three critical fates of a banded bird alive at the start of the year.

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

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

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

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

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, 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

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

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.

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

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 $\overline{S}=(\sum_{i=1}^{k-1} S_i)/(k-1)$. 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., $\mathrm{corr}(\tilde{S}_i,\hat{f}_{i+1})$ and $\mathrm{corr}(\bar{\tilde{S}},\bar{\hat{f}})$. 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:

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

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

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

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

Estimated sampling correlation coefficients are derived by definition,

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

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

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

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

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

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

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

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:

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

where

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

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

and

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

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 $1603\times0.6539\times0.0401=42.0$. 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,

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

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 ± 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                                    RECOVERY MATRIX
____  BANDED                                    _______________

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

               _____________RECOVERY RATE    F(I)  (%)________________      _______DIRECT RECOVERY RATE  R(I,I)/N(I)__
    I          ESTIMATE               STANDARD            95% CONFIDENCE      ESTIMATE         STANDARD     95% CONFIDENCE
               ________               __ERROR__           ___INTERVAL___      ________         __ERROR__    ___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


               ___________________SURVIVAL_S(I)_(%)______________________
   YEAR        SURVIVAL       STANDARD      COEFFICIENTS    95% CONFIDENCE
               ________       __ERROR__     OF_VARIATION    ___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




YEAR NUMBER                                  RECOVERY MATRIX
---- 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                                   RECOVERY MATRIX
---- BANDED                                   ---------------

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

            _____________RECOVERY RATE    F(I)  (%)___________________        ________DIRECT RECOVERY RATE  R(I,I)/N(I)__
    I       ESTIMATE               STANDARD                 95% CONFIDENCE       ESTIMATE         STANDARD     95% CONFIDENCE
            --------               __ERROR__                ___INTERVAL___       --------         __ERROR__    ___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       COEFFICIENTS      95% CONFIDENCE
             ________         ERROR         OF VARIATION         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                                         RECOVERY MATRIX
---- BANDED                                         ---------------

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 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            -0.000107862                      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 CORRELATICN 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                                              RECOVERY MATRIX
---- BANDED                                              ---------------

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                 95% CONFIDENCE
                  --------               __ERROR__                ___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   0.0   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 CORRELATICN 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 $63.38 \pm 1.46\%$ (i.e., $se(\hat{S}) = 1.46\%$) and the recovery rate was $6.94 \pm 0.26\%$. 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 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			$N_3f_3^*$	$N_3S_3f_4$	$N_3S_3S_4f_5$

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                                            RECOVERY MATRIX
---- BANDED                                            ---------------

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:

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

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

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

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

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

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

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

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

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

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 survival 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                                    RECOVERY MATRIX
---- BANDED                                    ---------------

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  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


                ---------------RECOVERY RATE    F(I)  (%)--------------------          ----FIRST-YEAR RECOVERY RATE    F*(I)  (%)----
      I         ESTIMATE                 STANDARD                95% CONFIDENCE          ESTIMATE          STANDARD          95% CONFIDENCE
                --------                 --ERROR--               ---INTERVAL---          --------          --ERROR--         ---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        COEFFICIENTS        95% CONFIDENCE
                --------        __ERROR__       OF_VARIATION        ___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




YEAR NUMBER                                          RECOVERY MATRIX
---- 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  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

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 CORRELATICN 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.

$$\begin{array}{cccc} N_1f_1^* & N_1S_1f_2 & N_1S_1S_2f_3 & N_1S_1S_2S_3f_4 \\ & N_2f_2^* & N_2S_2f_3 & N_2S_2S_3f_4 \\ & & N_3f_3^* & N_3S_3f_4 \end{array}$$

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

$$\begin{array}{cccc} N_1f_1 & N_1S_1f_2 & N_1S_1S_2f_3 & N_1S_1S_2S_3f_4 \\ & N_2f_2 & N_2S_2f_3 & N_2S_2S_3f_4 \\ & & N_3f_3 & N_3S_3f_4 \end{array}$$

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

$$\begin{array}{cccc} N_1f_1 & N_1Sf_2 & N_1SSf_3 & N_1SSSf_4 \\ & N_2f_2 & N_2Sf_3 & N_2SSf_4 \\ & & N_3f_3 & N_3Sf_4 \end{array}$$

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

$$\begin{array}{cccc} N_1f & N_1Sf & N_1SSf & N_1SSSf \\ & N_2f & N_2Sf & N_2SSf \\ & & N_3f & N_3Sf \end{array}$$

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

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

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

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

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, $\bar{\bar{S}}$, is nearly normally distributed. Since $\widehat{\text{MLS}}$ is a simple one-to-one transformation of $\hat{S}$ or $\bar{\bar{S}}$, 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 $\hat{\bar{f}}^*$ 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 × $\ell n(2) =$ MLS × 0.69 (in the example, 1.82 = 2.64 × 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                                 RECOVERY MATRIX
---- BANDED                                 ---------------

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


              ____________RECOVERY_RATE___F(I)__(%)_________________     ___FIRST-YEAR_RECOVERY_RATE___F*(I)__(%)____
      I      ESTIMATE            STANDARD              95% CONFIDENCE     ESTIMATE         STANDARD       95% CONFIDENCE
             --------            __ERROR__             ___INTERVAL___     --------         __ERROR__      ___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        COEFFICIENTS      95% CONFIDENCE
                 --------        __ERROR__       OF VARIATION      ___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




YEAR NUMBER                                               RECOVERY MATRIX
---- 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                                       RECOVERY MATRIX
---- BANDED                                       ---------------

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

              ______________RECOVERY RATE____F(I)__(%)_________________         ________DIRECT RECOVERY RATE__B(I,I)/N(I)__
      I       ESTIMATE               STANDARD              95% CONFIDENCE        ESTIMATE         STANDARD      95% CONFIDENCE
              --------               __ERROR__             ___INTERVAL___        --------         __ERROR__     ___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


               ___________________SURVIVAL_S(I)_(%)_______________________
 YEAR          SURVIVAL       STANDARD       COEFFICIENTS      95% CONFIDENCE
               ________       __ERROR__      OF_VARIATION      ___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                                               RECOVERY MATRIX
---- BANDED                                               ---------------

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 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            -0.000007973                          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 CORRELATION 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                                         RECOVERY MATRIX
---- BANDED                                         ---------------

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

               ----------------RECOVERY RATE  F(I)  (%)-------------------
       I       ESTIMATE             STANDARD               95% CONFIDENCE
               --------              ERROR                   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                                            RECOVERY MATRIX
---- BANDED                                            ---------------

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 VARICUS MODELS AND ASSUMPTICNS


(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-SQUAREC VALUE =     180.10
THEORETICAL CHI-SQUARE VALUE AT THE 5% LEVEL =      26.30
DEGREES CF FREEDOM =    16

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




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

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

PROBABILITY OF A CHI-SCUARE 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 CF FREEDOM =     6

PROBABILITY OF A CHI-SCUARE 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, $\text{var}(\hat{M}) = \text{var}(\hat{S})$.

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.

Chapter 3. Models for Birds Banded as Young and Adults

3.1 Introduction

It has often been observed in banding studies that young birds have different survival rates from adults and are more vulnerable to hunters. Thus, the assumption in Chapter 2 that survival and hunting mortality rates are age-independent limits application of the models of Chapter 2 to recoveries from birds banded as adults only. Frequently when banding programs are carried out, both young and adult birds are banded at the same time, and band recoveries are recorded for both age groups. This chapter presents methods for utilizing the records from birds banded as adults and as young to obtain estimates of both adult and first-year survival and recovery rates, and to make tests concerning the age-dependence of these rates.

The Experimental Situation

This situation is basically the same as that of Chapter 2, except that each batch of banded individuals released includes a known number of young birds in addition to the known number of adults. It is important that both young and adults are banded each year or parameters of interest will not be estimable. In particular, banding programs where only young are banded are virtually useless because survival rates of young are not estimable (see Section 3.9). Records of the numbers of birds banded and of the annual band recoveries are kept separately for the two groups of birds. As with the situation discussed in Chapter 2, banding is carried out each year at about the same time for a number of years, and annual recoveries are recorded for each of these years, and possibly for several additional years. Again a "year" of the experiment or banding study is the period between successive bandings, or between anniversaries of the last banding date.

Notation and Definitions

k = number of years at the start of which a release of banded birds is made
ℓ = number of years during which annual band recoveries are recorded, $\ell \geqslant k$, and we sometimes write $\ell = k + s$, where s = the number of years when no release is made but recoveries are recorded, $s \geqslant 0$
N_i = number of adults banded in year i, $i = 1, \ldots, k$
M_i = number of young birds banded in year i, $i = 1, \ldots, k$
R_{ij} = number of bands recovered in year j from the *adults* released in year i, $i = 1, \ldots, k, j = i, \ldots, \ell$
Q_{ij} = number of bands recovered in year j from the *young* birds released in year i, $i = 1, \ldots, k, j = i, \ldots, \ell$

The data are presented in the form of two tables, the first for recoveries from birds banded as adults and the second for recoveries from birds banded as young. This is illustrated in terms of N_i, M_i, R_{ij}, Q_{ij} for a banding study with $k = 3$, $\ell = 5$, $s = 2$, in Table 3.1 below. Summary totals which are used in the calculation of estimators and tests are indicated in the table and defined below.

Table 3.1. *Symbolic representation of the data for a banding study with $k=3$, $\ell=5$, $(s=2)$, when both young and adult birds are banded and released each year.*

Year banded	Number banded	Year of recovery					Row totals
		1	2	3	4	$\ell=5$	
				Birds banded as adults			
1	N_1	T_1 R_{11}	R_{12}	R_{13}	R_{14}	R_{15}	$R_{1\cdot}=T_1$
2	N_2		T_2 R_{22}	R_{23}	R_{24}	R_{25}	$R_{2\cdot}$
$k=3$	N_3			T_3 R_{33}	T_4 R_{34}	R_{35}	$R_{3\cdot}$
	Column totals	$R_{\cdot 1}$	$R_{\cdot 2}$	$R_{\cdot 3}$	$R_{\cdot 4}$	$R_{\cdot 5}=T_5$	
				Birds banded as young			
1	M_1	U_1 Q_{11}	Q_{12}	Q_{13}	Q_{14}	Q_{15}	$Q_{1\cdot}=U_1$
2	M_2		U_2 Q_{22}	Q_{23}	Q_{24}	Q_{25}	$Q_{2\cdot}$
3	M_3			U_3 Q_{33}	U_4 Q_{34}	Q_{35}	$Q_{3\cdot}$
	Column totals	$Q_{\cdot 1}$	$Q_{\cdot 2}$	$Q_{\cdot 3}$	$Q_{\cdot 4}$	$Q_{\cdot 5}=U_5$	

The array R_{ij} is called the recovery matrix for adults and the array Q_{ij} is called the recovery matrix for young. Row totals, column totals, and certain block totals of these arrays are used in summarizing the data, as indicated in Table 3.1. Thus

$R_{i\cdot}=i^{th}$ row total of the recovery matrix for adults, or equivalently the total number of recoveries from the adults banded in year i, $i=1,\ldots,k$,

$Q_{i\cdot}=i^{th}$ row total of the recovery matrix for adults, or equivalently the total number of recoveries from the young banded in year i, $i=1,\ldots,k$,

$R_{\cdot i}=i^{th}$ column total of the recovery matrix for adults, $i=1,\ldots,\ell$,

$Q_{\cdot i}=i^{th}$ column total of the recovery matrix for young, $i=1,\ldots,\ell$,

$T_1=R_{1\cdot}$

$$T_i=\begin{cases} R_{i\cdot}+T_{i-1}-R_{\cdot i-1} & ,i=2,\ldots,k \\ T_{i-1}-R_{\cdot i-1} & ,i=k+1,\ldots,\ell,\text{ if } \ell>k \end{cases}$$

$U_1=Q_{1\cdot}$

$$U_i=\begin{cases} Q_{i\cdot}+U_{i-1}-Q_{\cdot i-1} & ,i=2,\ldots,k \\ U_{i-1}-Q_{\cdot i-1} & ,i=k+1,\ldots,\ell \text{ if } \ell>k. \end{cases}$$

Other useful subtotals which involve elements from both of the recovery matrices are

$W_i=R_{\cdot i}+Q_{\cdot i}-Q_{ii}=$ the total number of recoveries from adults in year i (including recoveries from adult survivors of young banded before year i), $i=1,\ldots,k$,

$T_i+U_i-Q_{i\cdot}=$ the total number of recoveries in years i to ℓ inclusive from all banded adults alive at the start of year i (whether banded as adults or as young), $i=1,\ldots,k$.

Real data for this situation are presented in Table 3.2.

Table 3.2. *Data from a banding study of young and adult male mallards banded preseason in the San Luis Valley, Colorado, 1963-1971* ($k=\ell=9, s=0$).

Year banded	Number banded	Year of recovery								
		1963	1964	1965	1966	1967	1968	1969	1970	1971
					Banded as Adults					
1963	231	10	13	6	1	1	3	1	2	0
1964	649		58	21	16	15	13	6	1	1
1965	885			54	39	23	18	11	10	6
1966	590				44	21	22	9	9	3
1967	943					55	39	23	11	12
1968	1,077						66	46	29	18
1969	1,250							101	59	30
1970	938								97	22
1971	312									21
					Banded as Young					
1963	962	83	35	18	16	6	8	5	3	1
1964	702		103	21	13	11	8	6	6	0
1965	1,132			82	36	26	24	15	18	4
1966	1,201				153	39	22	21	16	8
1967	1,199					109	38	31	15	1
1968	1,155						113	64	29	22
1969	1,131							124	45	22
1970	906								95	25
1971	353									38

				Subtotals				
i	$R_{i\cdot}$	$R_{\cdot i}$	$Q_{i\cdot}$	$Q_{\cdot i}$	T_i	U_i	W_i	$T_i+U_i-Q_{i\cdot}$
1	37	10	175	83	37	175	10	37
2	131	71	168	138	158	260	106	250
3	161	81	205	121	248	327	120	370
4	108	100	259	218	275	465	165	481
5	140	115	194	191	315	441	197	562
6	159	161	228	213	359	478	261	609
7	190	197	191	266	388	456	339	653
8	119	218	120	227	310	310	350	500
9	21	113	38	121	113	121	196	196

Parameters and Assumptions

The parameters and models for the experimental situation of this chapter and that of Chapter 2 are analogous in many ways. Thus, for the reasons given in Chapter 2, the actual harvest rate H and the reporting rate λ are not separately estimable, but their product $H\lambda$ is estimable, and the models are therefore defined in terms of S, the annual survival rate, and $f=H\lambda$, the annual recovery rate. Subscripts on f and S are used to denote year-specificity (i.e., dependence on a particular year), and superscripts (usually primes) on f and S denote age-dependence, or for f, superscripts may indicate a possible dependence of the reporting rate on the year of banding.

As in Chapter 2, assumptions about the specificity of the annual rate parameters can be made progressively more general giving rise to a series of models of increasing complexity. In this chapter, assumptions based on biological and ecological information are made about the parameters H, λ, and S. However, as H and λ are not separately identifiable (estimable), these assumptions must be translated in terms of the parameters f and S into an appropriate model structure. Usually there is not a one-to-one relationship between the model based on the parameters f and S and the original assumptions about H, λ, and S. Because of this a set of assumptions about H, λ, and S is referred to as a hypothesis and denoted H_i, e.g., H_1, and the derived model is referred to as the model under the hypothesis; the model under H_1, or sometimes, the H_1 model. This terminology, which is also used in Chapter 4, has not been made to conform exactly to that of Chapter 2 as it has been used consistently in material of this chapter and the next (cf. Brownie and Robson 1976) and appears in the documentation of related computer output. Specific assumptions, and precise definition of parameters, are described separately for each model in Sections 3.2 to 3.6.

The assumption that banded birds in the population suffer independent fates, on which the development of sampling variance and testing procedures is based, is questionable for species such as geese where young and adults tend to maintain distinct family units for at least part of the first year.

3.2 The Model Under H_1

As in Chapter 2, we begin with one of the most useful models for this situation, which we call the model under H_1. The assumptions of H_1 are:

(1) Annual survival, reporting and harvest rates (hence recovery rates) are year-specific;
(2) annual survival and harvest rates are age-dependent for the first year of life only (i.e., young and adult birds have different survival and harvest rates); and
(3) reporting rates are not dependent on time of release.

In terms of f and S these assumptions lead to defining the following parameters:

f_i' = recovery rate in year i for birds banded and released as young in year i, $i=1,\ldots,k$
S_i' = survival rate for year i for birds banded and released as young in year i, $i=1,\ldots,k-1$ if $\ell=k$, and $1,\ldots,k$ if $\ell>k$
f_i = recovery rate for adults in year i, $i=1,\ldots,\ell$
S_i = survival rate for adults in year i, $i=1,\ldots,\ell-1$.

Tables of expected band recoveries in terms of N_i, M_i, f_i, f_i', S_i and S_i' are used to express the structure of this model. For banded adults, the assumptions of H_1 are the same as those of Model 1 in Chapter 2, so expected recoveries from birds banded as adults in terms of N_i, f_i, S_i are the same under H_1 and Model 1 (see Table 2.3). For birds banded as young, if M_1 are released in the first year, on the average we would expect M_1f_1' bands to be recovered that year, and M_1S_1' of the cohort to survive to adulthood. At the start of the second year M_2 young are banded and released, and there are M_1S_1' survivors (now adults) from the first batch released. Thus, in the second year the expected number of band recoveries from the M_2 new releases is M_2f_2', and the expected number of survivors M_2S_2'. Also in the second year the M_1S_1' survivors of the first batch will reflect the adult rates f_2 and S_2 giving on the average $M_1S_1'f_2$ recoveries and $M_1S_1'S_2$ survivors. Continuing to the third year when M_3 new young are released, we would expect $M_1S_1'S_2f_3$ recoveries from the first batch released, $M_2S_2'f_3$ from the second batch, and M_3f_3' from the new releases. In this way we arrive at the entries in Table 3.3.

Table 3.3 *Expected numbers of band recoveries under H_1 for a banding study with $k=3$, $\ell=5$, $s=2$.*

Year banded	Number banded	Year of recovery 1	2	3	4	5
				Birds banded and released as adults		
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	$N_2S_2f_3$	$N_2S_2S_3f_4$	$N_2S_2S_3S_4f_5$
3	N_3			N_3f_3	$N_3S_3f_4$	$N_3S_3S_4f_5$
				Birds banded and released as young		
1	M_1	M_1f_1'	$M_1S_1'f_2$	$M_1S_1'S_2f_3$	$M_1S_1'S_2S_3f_4$	$M_1S_1'S_2S_3S_4f_5$
2	M_2		M_2f_2'	$M_2S_2'f_3$	$M_2S_2'S_3f_4$	$M_2S_2'S_3S_4f_5$
3	M_3			M_3f_3'	$M_3S_3'f_4$	$M_3S_3'S_4f_5$

From Table 3.3 we see that observations corresponding to the off-diagonal elements of the recovery matrix for young (i.e., values of Q_{ij} for $j>i$) will provide information about the adult rate parameters, and any efficient estimation procedure should exploit this fact. We note also that if $\ell>k$, the parameters $S_k,\ldots,S_{\ell-1}$, S_k', $f_{k+1},\ldots,f_\ell$ are not separately estimable, though products such as S_kf_{k+1}, $S_k'f_{k+1}$, etc., are estimable.

Estimation of Parameters

ML estimators of adult and young recovery rates are

$$\hat{f}_i=\frac{R_{i\cdot}}{N_i}\frac{W_i}{T_i+U_i-Q_{i\cdot}} \qquad ,i=1,\ldots,k,$$

$$\hat{f}_i'=\frac{Q_{ii}}{M_i} \qquad ,i=1,\ldots,k,$$

The data in Table 3.2 give, for example,

$$\hat{f}_1=\frac{R_{1\cdot}}{N_1}\frac{W_1}{T_1+U_1-Q_{1\cdot}}=\frac{37\times 10}{231\times 37}=0.0433$$

$$\hat{f}_1'=\frac{Q_{11}}{M_1}=\frac{83}{962}=0.0863$$

$$\hat{f}_2=\frac{R_{2\cdot}}{N_2}\frac{W_2}{T_2+U_2-Q_{2\cdot}}=\frac{131\times 106}{649\times 250}=0.0856$$

$$\hat{f}_2'=\frac{Q_{22}}{M_2}=\frac{103}{702}=0.1467.$$

Bias-adjusted ML estimators of the survival rates are

$$\tilde{S}_i=\frac{R_{i\cdot}}{N_i}\frac{T_i+U_i-Q_{i\cdot}-W_i}{T_i+U_i-Q_{i\cdot}}\frac{N_{i+1}+1}{R_{i+1\cdot}+1} \qquad ,i=1,\ldots,k-1,$$

$$\tilde{S}_i'=\frac{Q_{i\cdot}-Q_{ii}}{M_i}\frac{N_{i+1}+1}{R_{i+1\cdot}+1} \qquad ,i=1,\ldots,k-1.$$

Data in Table 3.2 give

$$\tilde{S}_1=\frac{R_{1\cdot}}{N_1}\frac{T_1+U_1-Q_{1\cdot}-W_1}{T_1+U_1-Q_{1\cdot}}\frac{N_2+1}{R_{2\cdot}+1}=\frac{37\times(37-10)\times 650}{231\times 37\times 132}=0.5756$$

$$\tilde{S}_1'=\frac{Q_{1\cdot}-Q_{11}}{M_1}\frac{N_2+1}{R_{2\cdot}+1}=\frac{(175-83)\times 650}{962\times 132}=0.4709$$

$$\tilde{S}_2=\frac{131\times(250-106)\times 886}{649\times 250\times 162}=0.6359$$

$$\tilde{S}_2'=\frac{(168-103)\times 886}{702\times 162}=0.5064.$$

The FORTRAN program BROWNIE prints the estimates $\hat{f}_i$, $\hat{f}_i'$, $i=1,\ldots,k$, and $\tilde{S}_i$, $\tilde{S}_i'$, $i=1,\ldots,k-1$, as illustrated in Example 3.1a. The unadjusted ML estimators $\hat{S}_i$ and $\hat{S}_i'$, of the adult and young survival rates are computed by the program but are not printed. They are used in the program to obtain the "matrices of expected values" (i.e., the matrices of ML estimates of the expected values) and are defined below.

$$\hat{S}_i=\frac{R_{i\cdot}}{N_i}\frac{T_i+U_i-Q_{i\cdot}-W_i}{T_i+U_i-Q_{i\cdot}}\frac{N_{i+1}}{R_{i+1\cdot}} \qquad ,i=1,\ldots,k-1,$$

$$\hat{S}_i'=\frac{Q_{i\cdot}-Q_{ii}}{M_i}\frac{N_{i+1}}{R_{i+1\cdot}} \qquad ,i=1,\ldots,k-1.$$

The data in Table 3.2 give, for example

$$\hat{S}_1=\frac{37\times(37-10)\times 649}{231\times 37\times 131}=0.5791,$$

$$\hat{S}_1'=\frac{(175-83)\times 649}{962\times 131}=0.4738$$

$$\hat{S}_2=\frac{131\times(250-106)\times 885}{649\times 250\times 161}=0.6391$$

$$\hat{S}_2'=\frac{(168-103)\times 885}{702\times 161}=0.5090\,.$$

Note that these estimates differ only slightly from the bias-adjusted estimates. In general, for sample sizes for which the precision of estimators is good, the difference between $\hat{S}_i$ and $\tilde{S}_i$ is negligible.

Additional Estimates if $\ell > k$

As in Chapter 2, if $\ell > k$, annual rate parameters in years k to ℓ are not separately identifiable, though certain products of these parameters are estimable, e.g., $S_k f_{k+1}$, $S_k S_{k+1} f_{k+2}, \ldots, S_k S_{k+1} \cdots S_{k+s-1} f_{k+s}$. The corresponding ML estimators are

$$\widehat{S_k \cdots S_{k+j-1} f_{k+j}} = \frac{R_{\cdot k+j} + Q_{\cdot k+j}}{T_{k+1} + U_{k+1}} \frac{R_{k\cdot}}{N_k} \left(\frac{T_k + U_k - Q_{k\cdot} - W_k}{T_k + U_k - Q_{k\cdot}} \right) \qquad ,j = 1, \ldots, s,$$

$$\widehat{S_k' f_{k+1}} = \frac{Q_{k\cdot} - Q_{kk}}{M_k} \frac{R_{\cdot k+1} + Q_{\cdot k+1}}{T_{k+1} + U_{k+1}} \qquad ,s = 2$$

and, if $s > 2$,

$$\widehat{S_k' S_{k+1} \cdots S_{k+j-1} f_{k+j}} = \frac{Q_{k\cdot} - Q_{kk}}{M_k} \frac{R_{\cdot k+j} + Q_{\cdot k+j}}{T_{k+1} + U_{k+1}} \qquad ,j = 2, \ldots, s.$$

For the data of Table 3.2, $k = \ell$ and $s = 0$, so the above estimators are not needed.

Sampling Variances, Standard Errors, and Confidence Intervals

Estimates of standard errors, and approximate 95% confidence intervals for f_i, S_i, f_i' and S_i' are contained in the output of the program BROWNIE (see Example 3.1a). The formulae used and some numerical illustrations, again from the data in Table 3.2, are given below.

The notation is similar to that of Chapter 2, thus var$(\hat{f}_i)$ denotes an estimator of the sampling variance of $\hat{f}_i$, and se$(\hat{f}_i) = \sqrt{\text{var}(\hat{f}_i)}$ is the corresponding estimator of the standard error of $\hat{f}_i$.

We have

$$var\,(\hat{f}_i) = (\hat{f}_i)^2 \left[\frac{1}{R_{i\cdot}} - \frac{1}{\mathbf{N_i}} + \frac{1}{W_i} - \frac{1}{T_i + U_i - Q_{i\cdot}} \right] \qquad ,i = 1, \ldots, k,$$

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

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

$$\text{var}(\tilde{S}_i') = (\tilde{S}_i')^2 \left[\frac{1}{Q_{i\cdot} - Q_{ii}} - \frac{1}{M_i} + \frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}} \right] \qquad ,i = 1, \ldots, k-1.$$

Confidence intervals are constructed as in Chapter 2. For example, for the parameter f_i, the approximate 95% confidence interval is given by $(\hat{f}_i - 1.96\ \text{se}(\hat{f}_i), \hat{f}_i + 1.96\ \text{se}(\hat{f}_i))$. This is illustrated for the data of Table 3.2 as follows:

$$\text{var}(\hat{f}_2) = (\hat{f}_2)^2 \left[\frac{1}{R_{2\cdot}} - \frac{1}{N_2} + \frac{1}{W_2} - \frac{1}{T_2 + U_2 - Q_{2\cdot}} \right] = (0.0856)^2 \left[\frac{1}{131} - \frac{1}{649} + \frac{1}{106} - \frac{1}{250} \right] = 0.00008446,$$

$$\text{se}(\hat{f}_2) = \sqrt{0.00008446} = 0.0092,$$

$$1.96 \times \text{se}(\hat{f}_2) = (1.96) \times (0.0092) = 0.0180,$$

and the 95% confidence interval for f_2 is (0.0856 − 0.0180, 0.0856 + 0.0180) or (0.0676, 0.1036). Similarly,

$$\text{var}(\tilde{S}_2) = (0.6359)^2 \left[\frac{1}{131} - \frac{1}{649} + \frac{1}{161} - \frac{1}{885} + \frac{1}{144} - \frac{1}{250} \right] = 0.00570906,$$

$$\text{se}(\tilde{S}_2) = \sqrt{0.00570906} = 0.0756,$$

and $\qquad 1.96 \times \text{se}(\tilde{S}_2) = 0.1482,$

thus the 95% confidence interval for S_2 is (0.6359 − 0.1482, 0.6359 + 0.1482) or (0.4877, 0.7841). For $\hat{f}_2'$,

$$\text{var}(\hat{f}_2') = \frac{(0.1467) \times (1 - 0.1467)}{702} = 0.00017832,$$

$$\text{se}(\hat{f}_2') = \sqrt{0.00017832} = 0.0134,$$

and, $1.96 \times \text{se}(\hat{f}_2') = 0.0263$, giving the 95% confidence interval for f_2' as (0.1467 − 0.0263, 0.1467 + 0.0263) or (0.1204, 0.1730).

For $\tilde{S}_2'$

$$\text{var}(\tilde{S}_2') = (0.5064)^2\left[\frac{1}{65}-\frac{1}{702}+\frac{1}{161}-\frac{1}{885}\right] = 0.00488298,$$

$$\text{se}(\tilde{S}_2') = \sqrt{0.00488298} = 0.0699,$$

and $1.96 \times \text{se}(\tilde{S}_2') = 0.1370$, giving the 95% confidence interval for S_2' as $(0.5064-0.1370,\ 0.5064+0.1370)$ or $(0.3694,\ 0.6434)$.

These results can be compared with the computer output in Example 3.1a. Small differences between the calculations above and the computer output are due to the greater degree of accuracy of the latter.

Sampling Covariances and Correlations

Estimates of the sampling covariances between the estimators $\hat{f}_i$, $\tilde{S}_i$, $\hat{f}_i'$ and $\tilde{S}_i'$ are also contained in the output of the program BROWNIE. Estimates are calculated only for those covariances which are not exactly or approximately (for large N_i and M_i) equal to zero. Estimates of corresponding correlations are obtained with these covariance estimates and the variance estimates described above, and are printed out as shown in Example 3.1b. Appropriate formulae and some numerical examples with the data of Table 3.2 are given below.

Let $\text{cov}(\hat{f}_i,\tilde{S}_i)$ be the estimator of the covariance between $\hat{f}_i$ and $\tilde{S}_i$ defined by

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

An estimate of the correlation between the estimators $\hat{f}_i$ and $\tilde{S}_i$ is

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

where $\text{se}(\hat{f}_i)$ and $\text{se}(\tilde{S}_i)$ are as defined above. Thus,

$$\text{cov}(\hat{f}_2,\tilde{S}_2) = \hat{f}_2\tilde{S}_2\left[\frac{1}{R_{2\cdot}}-\frac{1}{N_2}-\frac{1}{T_2+U_2-Q_{2\cdot}}\right] = (0.0856)\,(0.6359)\left[\frac{1}{131}-\frac{1}{649}-\frac{1}{250}\right] = 0.0001141,$$

$$\text{corr}(\hat{f}_2,\tilde{S}_2) = \frac{0.0001141}{(0.0092)\,(0.0756)} = 0.1640.$$

Estimates of other covariances and correlations are obtained in a similar way using the following:

$$\text{cov}(\tilde{S}_i,\tilde{S}_{i+1}) = -\tilde{S}_i\tilde{S}_{i+1}\left[\frac{1}{R_{i+1\cdot}}-\frac{1}{N_{i+1}}\right] \qquad ,i=1,\ldots,k-2,$$

$$\text{cov}(\hat{f}_{i+1},\tilde{S}_i) = -\hat{f}_{i+1}\tilde{S}_i\left[\frac{1}{R_{i+1\cdot}}-\frac{1}{N_{i+1}}\right] \qquad ,i=1,\ldots,k-1,$$

$$\text{cov}(\tilde{S}_i,\tilde{S}_i') = \tilde{S}_i\tilde{S}_i'\left[\frac{1}{R_{i+1\cdot}}-\frac{1}{N_{i+1}}\right] \qquad ,i=1,\ldots,k-1,$$

$$\text{cov}(\hat{f}_i',\tilde{S}_i') = -\frac{\hat{f}_i'\tilde{S}_i'}{M_i} \qquad ,i=1,\ldots,k-1,$$

$$\text{cov}(\tilde{S}_i',\tilde{S}_{i+1}) = -\tilde{S}_i'\tilde{S}_{i+1}\left[\frac{1}{R_{i+1\cdot}}-\frac{1}{N_{i+1}}\right] \qquad ,i=1,\ldots,k-2$$

$$\text{cov}(\hat{f}_{i+1},\tilde{S}_i') = -\tilde{S}_i'\hat{f}_{i+1}\left[\frac{1}{R_{i+1\cdot}}-\frac{1}{N_{i+1}}\right] \qquad ,i=1,\ldots,k-1.$$

It is emphasized that the above computations provide estimates of the sampling correlations between *estimators*, that is, they reflect a relationship which is a property of the estimators themselves, and which cannot be interpreted as evidence of a relationship between the corresponding parameters. This subject is discussed in more detail in Section 8.4.

Goodness of Fit Test

For the experimental situation of this chapter, we have so far defined a set of assumptions, a corresponding model, and the estimators of parameters of the model. For a given data set, the validity of the estimates obtained will thus depend on whether the assumptions of the model are satisfied. This question cannot usually be answered with complete certainty for a given data set but indications of the adequacy of the model are obtained by means of a goodness of fit test and by tests which compare this model with other models to be defined in Sections 3.3 through 3.7.

A goodness of fit test to H_1 is therefore computed by the program BROWNIE, though not in the conventional way described, for example, in Section 2.2. Under the assumption that the H_1 model holds, the test statistic is approximately distributed as a chi-square variable, and as usual, "large" chi-square values (i.e., values associated with a small probability of occurrence) suggest the model may be incorrect.

The FORTRAN program BROWNIE also provides a means for examining the agreement between the model and the data on the basis of individual observations (i.e., individual R_{ij} or Q_{ij} values), in the following manner. The ML estimate of the expected value under H_1 of each entry in the recovery matrices for adults and young is computed (in terms of the N_i, $\hat{f}_i$, $\hat{S}_i$ for adults, and the M_i, $\hat{f}_i'$ $\hat{S}_i'$, $\hat{f}_i$, $\hat{S}_i$ for young), giving two arrays (cf. Example 3.1c) which are labeled in the output "MATRIX OF EXPECTED VALUES – ADULTS" and "MATRIX OF EXPECTED VALUES – YOUNG." Each observation in the recovery matrices is compared with the estimate of its expected value and the difference, suitably normalized, is approximately a standard normal variable under the assumption that the model holds. The resulting values are printed out in two matrices (one for adults, one for young) each labeled "MATRIX OF STANDARD NORMAL DEVIATES." Inspection of these deviates on an individual basis will help to determine where departures from the model lie, as indicated by unusually small (for instance less than -2), or unusually large (for instance, greater than $+2$), values.

This procedure is illustrated below from the data of Table 3.2. The following notation is needed: Let $\mathbf{E}_{ij}$ represent the ML estimate of the expected value of R_{ij}, and $\mathbf{E}_{ij}'$ represent the ML estimate of the expected value of Q_{ij}; then the corresponding standard normal deviates are

$$Z_{ij} = \frac{R_{ij} - \mathbf{E}_{ij}}{\sqrt{\dfrac{\mathbf{E}_{ij}}{N_i}(N_i - \mathbf{E}_{ij})}},$$

and

$$Z_{ij}' = \frac{Q_{ij} - \mathbf{E}_{ij}'}{\sqrt{\dfrac{\mathbf{E}_{ij}'}{M_i}(M_i - \mathbf{E}_{ij}')}}.$$

From Table 3.2, the ij^{th} entry in the recovery matrix for adults, with $i=1$, and $j=2$ is $R_{12}=13$. From Table 3.3 the expected value of R_{12} is $N_1\hat{S}_1\hat{f}_2$ and therefore $\mathbf{E}_{12} = N_1\hat{S}_1\hat{f}_2 = (231)\times(0.5791)\times(0.0856) = 11.45$, giving

$$Z_{12} = \frac{R_{12} - \mathbf{E}_{12}}{\sqrt{\dfrac{\mathbf{E}_{12}}{N_1}(N_1 - \mathbf{E}_{12})}} = \frac{13 - 11.45}{\sqrt{\dfrac{11.45}{231}(231 - 11.45)}} = \frac{1.55}{\sqrt{10.8825}} = 0.470$$

which is not an extreme value.

The data value 13, the estimate of its expected value 11.45, and the corresponding normal deviate 0.47, are easily found in the respective matrices for adults, in the computer output of Example 3.1c.

Similarly, from Table 3.2, $Q_{35} = 26$, and based on Table 3.3, $\mathbf{E}_{35}' = M_3\hat{S}_3'\hat{S}_4\hat{f}_5 = (1132)\times(0.5936)\times(0.8100)\times(0.0520) = 28.30$, and

$$Z_{35}' = \frac{26 - 28.30}{\sqrt{\dfrac{28.30}{1132}(1132 - 28.30)}} = -0.44.$$

The difference in the above value of $\mathbf{E}_{35}'$ and the corresponding value in the computer output of Example 3.1c is due to the greater accuracy of the computer calculations. Note that $\mathbf{E}_{ij}$ and $\mathbf{E}_{ij}'$ are obtained with the unadjusted ML estimators $\hat{S}_i$ and $\hat{S}_i'$, rather than the bias-adjusted $\tilde{S}_i$ and $\tilde{S}_i'$.

Examination of the deviates in the matrices for adults and young in the output of Example 3.1c shows that there are a few extreme values but on the whole, the agreement between the model and the data seems good.

An Example

Output from the computer analysis of the mallard data of Table 3.2 appears in Example 3.1. The output is on the whole self-explanatory. For obvious reasons N_i appears as N(I), M_i as M(I), f_i and $\hat{f}_i$ as F(I), S_i and $\tilde{S}_i$ as S(I), cov$(\hat{f}_i, \tilde{S}_i)$ as COVAR(F(I),S(I)), corr$(\hat{f}_i, \tilde{S}_i)$ as CORR(F(I),S(I)), and so on. In this example we note that in any year, the estimate of the recovery rate for young is *higher* than the estimate of the adult recovery rate, and the estimate of survival for young is *lower* than the estimate of adult survival.

The quantity $\bar{F}=0.0668$, which appears (Example 3.6a) below the column of estimates F(I) $\equiv \hat{f}_i$, is simply the average of the $\hat{f}_i$. It is an estimate of the average adult recovery rate during 1963-71. Similarly $\bar{F}'=0.1049$, $\bar{S}=0.6248$ and $\bar{S}'=0.5179$ are averages of the appropriate year-specific estimates, and are estimates of the corresponding average recovery and survival rates. Confidence intervals based on these average estimates are smaller than those based on individual (annual) estimates and $\bar{S}$ and $\bar{S}'$, in particular, may provide useful information in data sets where N_i and M_i are not large.

Examination of the matrices of standard normal deviates suggests that the agreement between model and data is reasonable. There are no really extreme entries in the matrices, and no obvious trends to suggest departure from the model.

Tests related to the model under H_1 appear in another part of the output as can be seen in the complete output for the same data in Example 3.5 at the end of this chapter.

3.3 The Model Under H_{02}

We now define another useful model for the two-age-class situation. The assumptions on which this model is based are collectively referred to as the hypothesis H_{02}. This hypothesis is more restrictive than H_1 in that survival rates are assumed to be constant from year to year, but is otherwise the same as H_1.

The assumptions of H_{02} are:

(1) Annual reporting and harvest rates (and hence recovery rates) are year-specific;
(2) annual survival and harvest rates are age-dependent for the first year of life only; and
(3) annual survival rates are otherwise constant from year to year.

The parameters of the model H_{02} are:

f_i'= recovery rate in year i for birds banded and released as young in year $i, i=1,\ldots,k$,
S'= constant annual survival rate for young,
f_i= recovery rate for adults in year $i, i=1,\ldots,\ell$,
S= constant annual survival rate for adults.

As before, tables of expected numbers of band returns, in terms of N_i, M_i, f_i, f_i', S, and S' are used to express the structure of this model.

Table 3.4. *Expected numbers of band recoveries under H_{02} for a banding study with $k=3$, $\ell=5$, $s=0$.*

Year banded	Number banded	Year of recovery				
		1	2	3	4	5
		Birds banded and released as adults				
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
3	N_3			N_3f_3	N_3Sf_4	N_3SSf_5
		Birds banded and released as young				
1	M_1	M_1f_1'	$M_1S'f_2$	$M_1S'Sf_3$	$M_1S'SSf_4$	$M_1S'SSSf_5$
2	M_2		M_2f_2'	$M_2S'f_3$	$M_2S'Sf_4$	$M_2S'SSf_5$
3	M_3			M_3f_3'	$M_3S'f_4$	$M_3S'Sf_5$

Example 3.1a

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO


THE HYPOTHESIS H1.   (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
------------
              (2)  YOUNG BIRDS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES FROM THOSE OF ADULTS.


PARAMETERS:
-----------
F(I)   = BAND RECOVERY RATE FOR ADULTS IN YEAR I.
S(I)   = SURVIVAL RATE FOR ADULTS IN YEAR I.
F'(I)  = BAND RECOVERY RATE FOR YOUNG IN YEAR I.
S'(I)  = SURVIVAL RATE FOR YOUNG IN YEAR I.


STRUCTURE OF THE MODEL UNDER H1  (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
-------------------------------

                             BANDED AS ADULTS

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

                             BANDED AS YOUNG

           M(1)F'(1)     M(1)S'(1)F(2)   M(1)S'(1)S(2)F(3)  M(1)S'(1)S(2)S(3)F(4)
                         M(2)F'(2)       M(2)S'(2)F(3)      M(2)S'(2)S(3)F(4
                                         M(3)F'(3)          M(3)S'(3)F(4)


ESTIMATES UNDER H1
------------------


                         F(I)                                                      S(I)
         ---------------------------------------------------       ----------------------------------------------------
 I   YR  ESTIMATE         STANDARD          95% CONFIDENCE          ESTIMATE         STANDARD          95% CONFIDENCE
         --------         __ERROR__         ___INTERVAL___          --------         __ERROR__         ___INTERVAL___

 1  1963  0.0433           0.0134          0.0170 -   0.0695         0.5756           0.1134          0.3533 -   0.7978

 2  1964  0.0856           0.0092          0.0676 -   0.1036         0.6359           0.0756          0.4878 -   0.7840

 3  1965  0.0590           0.0061          0.0470 -   0.0710         0.6665           0.0787          0.5122 -   0.8207

 4  1966  0.0628           0.0067          0.0496 -   0.0760         0.8051           0.0977          0.6136 -   0.9967

 5  1967  0.0520           0.0050          0.0422 -   0.0619         0.6496           0.0724          0.5078 -   0.7914

 6  1968  0.0633           0.0055          0.0525 -   0.0740         0.5525           0.0581          0.4387 -   0.6664

 7  1969  0.0789           0.0061          0.0670 -   0.0908         0.5719           0.0663          0.4419 -   0.7020

 8  1970  0.0888           0.0080          0.0730 -   0.1046         0.5415           0.1286          0.2894 -   0.7936

 9  1971  0.0673           0.0142          0.0395 -   0.0951
         AVERAGE          STANDARD          95% CONFIDENCE          AVERAGE          STANDARD          95% CONFIDENCE
         ESTIMATE         ERROR___          ___INTERVAL___          ESTIMATE         ERROR___          ___INTERVAL___

  F̄      = 0.0668          0.0029           0.0610 -   0.0726  S̄    = 0.6248            0.0214          0.5828 -   0.6668

                         F'(I)                                                     S'(I)
         ---------------------------------------------------       ----------------------------------------------------
 I   YR  ESTIMATE         STANDARD          95% CONFIDENCE          ESTIMATE         STANDARD          95% CONFIDENCE
         --------         __ERROR__         ___INTERVAL___          --------         __ERROR__         ___INTERVAL___

 1  1963  0.0863           0.0091          0.0685 -   0.1040         0.4709           0.0594          0.3545 -   0.5874

 2  1964  0.1467           0.0134          0.1205 -   0.1729         0.5064           0.0699          0.3694 -   0.6434

 3  1965  0.0724           0.0077          0.0573 -   0.0875         0.5891           0.0717          0.4486 -   0.7297

 4  1966  0.1274           0.0096          0.1085 -   0.1463         0.5909           0.0716          0.4506 -   0.7312

 5  1967  0.0909           0.0083          0.0746 -   0.1072         0.4776           0.0610          0.3581 -   0.5971

 6  1968  0.0978           0.0087          0.0807 -   0.1150         0.6521           0.0723          0.5104 -   0.7939

 7  1969  0.1096           0.0093          0.0914 -   0.1278         0.4635           0.0678          0.3307 -   0.5964

 8  1970  0.1049           0.0102          0.0849 -   0.1248         0.3926           0.1133          0.1705 -   0.6147

 9  1971  0.1076           0.0165          0.0753 -   0.1400
         AVERAGE          STANDARD          95% CONFIDENCE          AVERAGE          STANDARD          95% CONFIDENCE
         ESTIMATE         ERROR___          ___INTERVAL___          ESTIMATE         ERROR___          ___INTERVAL___

  F̄'     = 0.1049          0.0035           0.0979 -   0.1118  S̄'   = 0.5179            0.0265          0.4659 -   0.5699
```

Example 3.1b

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H1


I      YR  COVAR(F(I),S(I))          CORR(F(I),S(I))           COVAR(S(I),S(I+1))          CORR(S(I),S(I+1))
1    1963      -0.000107861              -0.071052198              -0.002229831               -0.260315312
2    1964       0.000113888               0.164048642              -0.002153369               -0.362159372
3    1965       0.000093530               0.194625663              -0.004058998               -0.527787699
4    1966       0.000277319               0.420584456              -0.003181363               -0.449952175
5    1967       0.000145477               0.398951049              -0.001924280               -0.457790708
6    1968       0.000130008               0.407034557              -0.001410437               -0.365958111
7    1969       0.000132313               0.329578633              -0.002272297               -0.266319993
8    1970       0.000256653               0.248228325


I      YR  COVAR(F(I+1),S(I))        CORR(F(I+1),S(I))         COVAR(F'(I),S'(I))          CORR(F'(I),S'(I))
1    1963      -0.000300121              -0.288098030              -0.000042236               -0.078513900
2    1964      -0.000190633              -0.413183861              -0.000105842               -0.113419513
3    1965      -0.000316565              -0.596183595              -0.000037700               -0.068248954
4    1966      -0.000254852              -0.517437047              -0.000062679               -0.090986152
5    1967      -0.000220346              -0.553958592              -0.000036215               -0.071550000
6    1968      -0.000194598              -0.553483970              -0.000055240               -0.087394643
7    1969      -0.000372666              -0.698801335              -0.000044936               -0.071363550
8    1970      -0.001618715              -0.887260327              -0.000045436               -0.039394739


I      YR  COVAR(S'(I),S(I))         CORR(S'(I),S(I))          COVAR(S'(I),S(I+1))         CORR(S'(I),S(I+1))
1    1963       0.001651412               0.245122223              -0.001824454               -0.406359058
2    1964       0.001636182               0.309903856              -0.001714925               -0.311848868
3    1965       0.002970106               0.526383420              -0.003588038               -0.512085954
4    1966       0.002893728               0.413544294              -0.002334868               -0.450695864
5    1967       0.001663408               0.377116138              -0.001414800               -0.399447721
6    1968       0.001608233               0.382845799              -0.001664660               -0.347049369
7    1969       0.001945246               0.432637488              -0.001841692               -0.211269655
8    1970       0.009441458               0.647830523


I      YR  COVAR(S'(I),F(I+1))       CORR(S'(I),F(I+1))
1    1963      -0.000245560              -0.449728519
2    1964      -0.000151819              -0.355785182
3    1965      -0.000279834              -0.578446973
4    1966      -0.000187041              -0.518292501
5    1967      -0.000162006              -0.483359803
6    1968      -0.000229674              -0.524885804
7    1969      -0.000302045              -0.554353714
8    1970      -0.001173591              -0.730147293


     THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.


         COVAR(S̄,F̄)   = -0.000033711                           CORR(S̄,F̄)   = -0.533301536

         COVAR(S̄',F̄')= -0.000005976                            CORR(S̄',F̄')= -0.063481008
```

Example 3.1c

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

MATRIX OF DATA VALUES -- ADULTS              H1
----------------------------------------------
        1    10.00   13.00    6.00    1.00    1.00    3.00    1.00    2.00    0.0
        2     0.0    58.00   21.00   16.00   15.00   13.00    6.00    1.00    1.00
        3     0.0     0.0    54.00   39.00   23.00   18.00   11.00   10.00    6.00
        4     0.0     0.0     0.0    44.00   21.00   22.00    9.00    9.00    3.00
        5     0.0     0.0     0.0     0.0    55.00   39.00   23.00   11.00   12.00
        6     0.0     0.0     0.0     0.0     0.0    66.00   46.00   29.00   18.00
        7     0.0     0.0     0.0     0.0     0.0     0.0   101.00   59.00   30.00
        8     0.0     0.0     0.0     0.0     0.0     0.0     0.0    97.00   22.00
        9     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    21.00


MATRIX OF EXPECTED VALUES -- ADULTS
----------------------------------------------
        1    10.00   11.45    5.04    3.60    2.42    1.92    1.33    0.86    0.37
        2     0.0    55.54   24.47   17.49   11.74    9.32    6.45    4.18    1.79
        3     0.0     0.0    52.22   37.32   25.05   19.89   13.77    8.93    3.83
        4     0.0     0.0     0.0    37.05   24.87   19.75   13.67    8.86    3.80
        5     0.0     0.0     0.0     0.0    49.07   38.97   26.97   17.49    7.50
        6     0.0     0.0     0.0     0.0     0.0    68.14   47.17   30.58   13.11
        7     0.0     0.0     0.0     0.0     0.0     0.0    98.64   63.95   27.41
        8     0.0     0.0     0.0     0.0     0.0     0.0     0.0    83.30   35.70
        9     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    21.00


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
----------------------------------------------
        1     0.00    0.47    0.43   -1.38   -0.92    0.78   -0.29    1.23   -0.61
        2     0.0     0.34   -0.72   -0.36    0.96    1.21   -0.18   -1.56   -0.59
        3     0.0     0.0     0.25    0.28   -0.42   -0.43   -0.75    0.36    1.11
        4     0.0     0.0     0.0     1.18   -0.79    0.52   -1.28    0.05   -0.41
        5     0.0     0.0     0.0     0.0     0.87    0.01   -0.78   -1.57    1.65
        6     0.0     0.0     0.0     0.0     0.0    -0.27   -0.17   -0.29    1.36
        7     0.0     0.0     0.0     0.0     0.0     0.0     0.25   -0.64    0.50
        8     0.0     0.0     0.0     0.0     0.0     0.0     0.0     1.57   -2.34
        9     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00



MATRIX OF DATA VALUES -- YOUNG               H1
----------------------------------------------
        1    83.00   35.00   18.00   16.00    6.00    8.00    5.00    3.00    1.00
        2     0.0   103.00   21.00   13.00   11.00    8.00    6.00    6.00    0.0
        3     0.0     0.0    82.00   36.00   26.00   24.00   15.00   18.00    4.00
        4     0.0     0.0     0.0   153.00   39.00   22.00   21.00   16.00    8.00
        5     0.0     0.0     0.0     0.0   109.00   38.00   31.00   15.00    1.00
        6     0.0     0.0     0.0     0.0     0.0   113.00   64.00   29.00   22.00
        7     0.0     0.0     0.0     0.0     0.0     0.0   124.00   45.00   22.00
        8     0.0     0.0     0.0     0.0     0.0     0.0     0.0    95.00   25.00
        9     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    38.00


MATRIX OF EXPECTED VALLES -- YOUNG
----------------------------------------------
        1    83.00   39.01   17.19   12.28    8.25    6.55    4.53    2.94    1.26
        2     0.0   103.00   21.08   15.07   10.11    8.03    5.56    3.60    1.54
        3     0.0     0.0    82.00   42.19   28.33   22.49   15.57   10.09    4.33
        4     0.0     0.0     0.0   153.00   37.16   29.50   20.42   13.24    5.67
        5     0.0     0.0     0.0     0.0   109.00   36.43   25.22   16.35    7.01
        6     0.0     0.0     0.0     0.0     0.0   113.00   59.70   38.71   16.59
        7     0.0     0.0     0.0     0.0     0.0     0.0   124.00   46.90   20.10
        8     0.0     0.0     0.0     0.0     0.0     0.0     0.0    95.00   25.00
        9     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    38.00


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
----------------------------------------------
        1     0.0    -0.66    0.20    1.07   -0.79    0.57    0.22    0.04   -0.23
        2     0.0     0.0    -0.02   -0.54    0.28   -0.01    0.19    1.27   -1.24
        3     0.0     0.0     0.0    -0.97   -0.44    0.32   -0.15    2.50   -0.16
        4     0.0     0.0     0.0     0.0     0.31   -1.40    0.13    0.76    0.98
        5     0.0     0.0     0.0     0.0     0.0     0.26    1.16   -0.34   -2.28
        6     0.0     0.0     0.0     0.0     0.0     0.0     0.57   -1.59    1.34
        7     0.0     0.0     0.0     0.0     0.0     0.0     0.0    -0.28    0.43
        8     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00
        9     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0
```

It is readily seen, for example by comparing Table 3.4 with Table 3.3, that there are fewer parameters to be estimated under H_{02} than under H_1. All the parameters of interest under H_{02}, that is $f_1, f_2, \ldots, f_\ell, f_1', f_2', \ldots, f_k'$, S and S', are identifiable, and hence, estimable.

Before discussing estimation, we note that there is a relationship between the assumptions of H_{02} and the hypothesis of compensatory natural mortality, in that harvest or exploitation rates are assumed to vary from year to year, but total mortality (the complement of S) is assumed to be constant. However, the resulting model cannot be unambiguously associated with compensatory mortality because of the nonidentifiability of the actual harvest rate H. Thus variable f_i and constant S arise if H and S are both constant, but λ_i (the reporting rate) varies. The assumption of a constant exploitation rate is usually known to be unreasonable, however. Testing the validity of the H_{02} model, and in particular comparing H_{02} with H_1 are therefore of interest (see Section 3.7).

Estimation of Parameters

The ML estimators of the parameters cannot be expressed explicitly as simple functions of the data in contrast to the ML estimators under H_1. Program BROWNIE contains an algorithm for obtaining these estimates which are printed out together with estimates of their standard errors and 95% confidence intervals. More precisely, the likelihood equations cannot be solved explicitly for $\hat{f}_i$, $\hat{f}_i'$, $\hat{S}$, and $\hat{S}'$, and solutions must be obtained by numerical methods. The method of scoring is used and is described in more detail in Appendix B and in Seber (1973:16-18).

Estimates of certain sampling covariances and correlations between the estimators (those most likely to be of interest) are also printed. Estimation under this model is illustrated in the computer output in Example 3.2. Note that if $\ell > k$ $(s > 0)$, although the parameters $f_{k+1}, f_{k+2}, \ldots, f_{k+s}$ are each estimable, their estimates are not printed out. Also, only the first k estimates $\hat{f}_i$ are averaged to obtain $\bar{f}$, an estimate of the average adult recovery rate in years 1 to k. This is because for data sets with typical values of N_i and M_i (i.e., values less than 500) the variances of the estimators $\hat{f}_{k+1}, \ldots, \hat{f}_{k+s}$ and corresponding confidence intervals tend to be so large that the estimates are of little use.

Goodness of Fit Test

A goodness of fit test to the model under H_{02} is computed in the conventional way described in Section 2.2. The resulting chi-square value and associated degrees of freedom are printed, and as usual, large chi-square values indicate that some of the assumptions of the model may be violated.

If $\mathbf{E}_{ij}$ and $\mathbf{E}_{ij}'$ represent the ML estimates of the expected values of R_{ij} and Q_{ij}, respectively, then the goodness of fit chi-square statistic is

$$\sum_{i=1}^{k} \sum_{j=i}^{\ell} \left\{ \frac{(R_{ij} - \mathbf{E}_{ij})^2}{\mathbf{E}_{ij}} + \frac{(Q_{ij} - \mathbf{E}_{ij}')^2}{\mathbf{E}_{ij}'} \right\} + \sum_{i=1}^{k} \left\{ \frac{[(N_i - R_{i\cdot}) - (N_i - \mathbf{E}_{i\cdot})]^2}{N_i - \mathbf{E}_{i\cdot}} + \frac{[(M_i - Q_{i\cdot}) - (M_i - \mathbf{E}_{i\cdot}')]^2}{M_i - \mathbf{E}_{i\cdot}'} \right\},$$

where $N_i - R_{i\cdot}$ and $M_i - Q_{i\cdot}$ are the number of bands never recovered from the i^{th} releases of adults and young, respectively, and $N_i - \mathbf{E}_{i\cdot}$ and $M_i - \mathbf{E}_{i\cdot}'$ are the ML estimates of the corresponding expected numbers. The degrees of freedom are $k^2 + 2ks - k - s - 2$.

In practice, some of the $\mathbf{E}_{ij}$ and $\mathbf{E}_{ij}'$ values will be too small to justify the chi-square approximation, and some pooling is necessary. As necessary, program BROWNIE will pool expected recoveries within a row from right to left in the matrix until a combined value of at least 2 is obtained. At the same time, corresponding R_{ij}'s or Q_{ij}'s are pooled also. A degree of freedom is lost for each R_{ij} or Q_{ij} pooled.

Again, a rough idea of the agreement between model and data can be obtained by comparing individual data values with their estimated expected values.

Thus, as under H_1, corresponding to each of the recovery matrices, a matrix of "expected values," and a matrix of "standard normal deviates" are printed.

An Example

The analysis described above for the model under H_{02} is illustrated for data from a study on male blue-winged teal banded preseason in Saskatchewan during 1962-68, with recoveries recorded through the 1973 season. For these data, $k = 7$, $\ell = 12$, $s = 5$, and the data with basic subtotals appear in the form of sample output from program BROWNIE in Example 3.2a.

For reasons given above, although estimates of each of the annual recovery rates $f_1, f_2, \ldots, f_{12}$ are computed, only the first $k(=7)$ are printed (see Example 3.2a). Note that estimates of recovery rates for young are higher than those for adults, and the estimate of the constant survival rate for young is lower than that for adults, indicating there is age-dependence of these parameters.

The estimates of both young and adult recovery rates are low compared with estimates from the mallard data used for Examples 3.1 and 3.5. For example, for the blue-winged teal, averaging the H_{02} estimates of adult recovery gives $\hat{f}=0.0067$, while for the mallard data of Example 3.5c, $\hat{f}=0.0646$.

These low recovery rates are reflected in the low precision of $\hat{S}$, and $\hat{S}'$ (and hence in the large confidence intervals for S, and S'), even though the number of birds banded in most years is over 2,000. Thus comparison of Example 3.2a with Example 3.5c shows that for the blue-winged teal, $\hat{S}=0.6356$ with standard error 0.0240 and for the mallards, $\hat{S}=0.6515$ with standard error 0.0120, although the total number of releases for the blue-winged teal is considerably greater than that for the mallards. This should indicate that the number of birds which must be banded to obtain a given level of precision for the survival estimates will depend greatly on the effective recovery rate for the species and flyway of interest. This subject is discussed further in Section 9.3.

Low recovery rates are also reflected in the "MATRICES OF EXPECTED VALUES," the entries of which are frequently less than 1 (see Example 3.2c). Examination of the "MATRICES OF STANDARD NORMAL DEVIATES" shows that there are few extreme values and agreement between model and data seems reasonably good. This is borne out by the goodness of fit test which yields (after much pooling) a chi-square value of 64.79 with 57 df. If H_{02} is true, there is a better than 20% chance of a chi-square value at least this extreme (64.79) so there is no reason to suspect that the model is incorrect. The likelihood ratio test comparing H_{02} with H_1 (inadvertently not photographed as part of Example 3.2c) also suggests that the model under H_{02} is adequate for these data ($\chi^2=15.14$, 11 df, $P<0.18$). This test is discussed in Section 3.7.

3.4 The Model Under H_{01}

The model discussed in this section is simple but very restrictive. As under H_{02}, the young and adult annual survival rates are assumed to be constant, but in addition young and adult recovery rates are also assumed to be constant. In most situations this last assumption is not appropriate and consequently the model is inadequate (see Examples 3.3 and 3.5b).

The assumptions of H_{01} are:

(1) Annual recovery and survival rates are age-dependent for the first year of life only; and

(2) annual recovery and survival rates are otherwise constant from year to year.

The parameters of the model under H_{01} are:

f' = constant annual recovery rate for young,
S' = constant annual survival rate for young,
f = constant annual recovery rate for adults,
S = constant annual survival rate for adults.

The four parameters f, f', S, S' are each estimable.

Tables of expected numbers of band recoveries, in terms of N_i, M_i, f, f', S, S', are used to express the structure of this model.

Table 3.5. *Expected numbers of band recoveries under H_{01} for a banding study with $k=3, \ell=5, s=2$.*

Year banded	Number banded	Year of recovery 1	2	3	4	5
		Birds banded and released as adults				
1	N_1	N_1f	N_1Sf	N_1SSf	N_1SSSf	N_1SSSSf
2	N_2		N_2f	N_2Sf	N_2SSf	N_2SSSf
3	N_3			N_3f	N_3Sf	N_3SSf
		Birds banded and released as young				
1	M_1	M_1f'	$M_1S'f$	$M_1S'Sf$	$M_1S'SSf$	$M_1S'SSSf$
2	M_2		M_2f'	$M_2S'f$	$M_2S'Sf$	$M_2S'SSf$
3	M_3			M_3f'	$M_3S'f$	$M_3S'Sf$

Example 3.2a

```
                     MALE BLUE WING TEAL BANDED PRESEASON IN SASKATCHEWAN, 1962-68
ADULTS         INPUT MATRIX


1962  1033.     3.    3.    3.    3.    1.    3.    0.    0.    0.    0.    0.    0.
1963  2233.     0.   14.    5.   11.    6.    5.    2.    0.    1.    0.    0.    0.
1964  2658.     0.    0.   17.   18.    7.   10.    1.    7.    1.    1.    0.    0.
1965  1705.     0.    0.    0.    9.   14.    6.    2.    0.    4.    0.    0.    0.
1966  4699.     0.    0.    0.    0.   44.   23.    5.    9.    2.    5.    0.    4.
1967  4542.     0.    0.    0.    0.    0.   47.   14.   19.   19.    8.    5.    2.
1968  2892.     0.    0.    0.    0.    0.    0.   11.   25.   12.    4.    2.    2.

YOUNG          INPUT MATRIX


1962   910.     6.    2.    1.    1.    0.    2.    1.    0.    0.    0.    0.    0.
1963  1157.     0.   11.    5.    6.    1.    1.    1.    1.    0.    0.    0.    1.
1964  1394.     0.    0.   19.    4.    4.    4.    0.    0.    1.    1.    0.    0.
1965  3554.     0.    0.    0.   65.   25.    8.    4.    2.    4.    4.    1.    0.
1966  4849.     0.    0.    0.    0.   65.   17.    2.   10.    6.    2.    3.    1.
1967  2555.     0.    0.    0.    0.    0.   52.    9.    8.    3.    4.    2.    1.
1968   305.     0.    0.    0.    0.    0.    0.    3.    1.    0.    1.    0.    0.



BASIC SUBTOTALS

 I  RROW(I)  RCOL(I)  CROW(I)  QCOL(I)     T(I)     U(I)     W(I)     Z(I)
 1   16.00     3.00    13.00     6.00    16.00    13.00     3.00     0.0
 2   44.00    17.00    27.00    13.00    57.00    34.00    19.00    13.00
 3   62.00    25.00    33.00    25.00   102.00    54.00    31.00    45.00
 4   35.00    41.00   113.00    76.00   112.00   142.00    52.00    92.00
 5   92.00    72.00   106.00    95.00   163.00   172.00   102.00    89.00
 6  114.00    94.00    79.00    84.00   205.00   156.00   126.00   127.00
 7   56.00    35.00     5.00    20.00   167.00    77.00    52.00   156.00
                                         132.00    57.00             187.00



 THE HYPOTHESIS H02
 ------------------

 ASSUMPTIONS:  (1)  YOUNG ANC ADULTS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES
 ------------
               (2)  SURVIVAL RATES ARE OTHERWISE CONSTANT FROM YEAR TO YEAR

               (3)  RECOVERY RATES ARE YEAR-SPECIFIC


 PARAMETERS:
 -----------
 S     = CCNSTANT ANNUAL SURVIVAL RATE FOR ADULTS
 F(I)  = BAND RECOVERY RATE IN YEAR I FOR ACULTS
 S'    = CONSTANT ANNUAL SURVIVAL RATE FCR YCUNG
 F'(I) = BAND RECOVERY RATE IN YEAR I FOR YCUNG

 STRUCTURE OF THE MCDEL UNDER H02    (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
 ------------------------------------

               BANDEC AS ACULTS                                     BANDED AS YOUNG

 N(1)F(1)    N(1)SF(2)    N(1)SSF(3)   N(1)SSSF(4)        M(1)F'(1)   M(1)S'F(2)   M(1)S'SF(3)   M(1)S'SSF(4)
             N(2)F(2)     N(2)SF(3)    N(2)SSF(4)                     M(2)F'(2)    M(2)S'F(3)    M(2)S'SF(4)
                          N(3)F(3)     N(3)SF(4)                                   M(3)F'(3)     M(3)S'F(4)



 ESTIMATES UNDER H02
 -------------------


         F(I)   RECOVERY RATE FOR ADULTS                                F'(I)  RECOVERY RATE FOR YOUNG
         ---------------------------------------------------------      ----------------------------------------------------------
         ESTIMATE           STANDARD           95% CONFIDENCE           ESTIMATE           STANDARD           95% CCNFIDENCF
  I YEAR                      ERROR                 INTERVAL                                 ERROR                 INTERVAL
         --------           ---------          -------------            --------           ---------          -------------
  1 1962  0.0029             0.0017       -0.0004 -   0.0062             0.0066             0.0027        0.0013 -   0.0118
  2 1963  0.0057             0.0013        0.0031 -   0.0083             0.0095             0.0029        0.0039 -   0.0151
  3 1964  0.0058             0.0011        0.0038 -   0.0079             0.0136             0.0031        0.0075 -   0.0197
  4 1965  0.0090             0.0013        0.0065 -   0.0115             0.0183             0.0022        0.0139 -   0.0227
  5 1966  0.0101             0.0011        0.0080 -   0.0122             0.0134             0.0016        0.0101 -   0.0166
  6 1967  0.0095             0.0009        0.0076 -   0.0113             0.0204             0.0028        0.0149 -   0.0259
  7 1968  0.0041             0.0006        0.0029 -   0.0053             0.0098             0.0056       -0.0013 -   0.0208

         AVERAGE            STANDARD           95% CONFIDENCE           AVERAGE            STANDARD           95% CCNFIDENCE
         ESTIMATE           ERROR                 INTERVAL              ESTIMATE           ERROR                 INTERVAL
         --------           --------           -------------            --------           --------           -------------
   -                                                               -
   F     = 0.0067            0.0005          0.0058 -   0.0077     F'    = 0.0131             0.0012          0.0107 -   0.0155

         S      SURVIVAL RATE FOR ADULTS                                S'     SURVIVAL RATE FOR YOUNG
         ---------------------------------------------------------      ----------------------------------------------------------
         ESTIMATE           STANDARD           95% CONFIDENCE           ESTIMATE           STANDARD           95% CONFIDENCE
                              ERROR                INTERVAL                                  ERROR                INTERVAL
         --------           ---------          -------------            --------           ---------          -------------

          0.6356             0.0240        0.5885 -   0.6827             0.4859             0.0460        0.3957 -   0.5760
```

Example 3.2b

```
SELECTED ESTIMATED COVARIANCES AND CORRELATIONS UNDER H02

I     YEAR   COV(F(I),S)             CORR(F(I),S)        COV(F(I),S')          CORR(F(I),S')
1     1962    -0.00000007               -0.0017            0.00000001              0.0002
2     1963    -0.00000133               -0.0424           -0.00000350             -0.0581
3     1964    -0.00000284               -0.1125           -0.00000451             -0.0933
4     1965    -0.00000768               -0.2464           -0.00000973             -0.1630
5     1966    -0.00000690               -0.2705           -0.00001213             -0.2482
6     1967    -0.00000766               -0.3405           -0.00001364             -0.3167
7     1968    -0.00000484               -0.3268           -0.00000630             -0.2220

I     YEAR   COV(F'(I),S)            CORR(F'(I),S)       COV(F'(I),S')         CORR(F'(I),S')
1     1962    -0.00000008               -0.0013           -0.00000023             -0.0019
2     1963    -0.00000009               -0.0014           -0.00000035             -0.0027
3     1964    -0.00000008               -0.0010           -0.00000054             -0.0038
4     1965    -0.00000004               -0.0008           -0.00000067             -0.0065
5     1966     0.00000003                0.0008           -0.00000042             -0.0055
6     1967     0.00000017                0.0025           -0.00000053             -0.0041
7     1968     0.00000020                0.0015           -0.00000029             -0.0011

I     YEAR   COV(F(I),F'(I))         CORR(F(I),F'(I))    COV(F(I+1),F'(I))     CORR(F(I+1),F'(I))
1     1962     0.00000000                0.0000           -0.00000000             -0.0014
2     1963     0.00000000                0.0002           -0.00000000             -0.0013
3     1964     0.00000000                0.0004           -0.00000001             -0.0018
4     1965     0.00000000                0.0011           -0.00000000             -0.0021
5     1966     0.00000000                0.0010           -0.00000000             -0.0017
6     1967     0.00000000                0.0004           -0.00000000             -0.0019
7     1968    -0.00000000               -0.0002           -0.00000001             -0.0013

I     YEAR   COV(F(I),F'(I+1))       CORR(F(I),F'(I+1))  COV(F'(I),F'(I+1))    CORR(F'(I),F'(I+1))
1     1962     0.00000000                0.0000            0.00000000              0.0000
2     1963     0.00000000                0.0002            0.00000000              0.0000
3     1964     0.00000000                0.0006            0.00000000              0.0000
4     1965     0.00000000                0.0006            0.00000000              0.0000
5     1966     0.00000000                0.0003            0.00000000              0.0000
6     1967    -0.00000000               -0.0001            0.00000000              0.0000

I     YEAR   COV(F(I),F(I+1))        CORR(F(I),F(I+1))   COV(F(I),F(I+2))      CORR(F(I),F(I+2))
1     1962    -0.00000000               -0.0014           -0.00000000             -0.0006
2     1963     0.00000001                0.0066            0.00000003              0.0159
3     1964     0.00000005                0.0354            0.00000005              0.0465
4     1965     0.00000013                0.0937            0.00000015              0.1205
5     1966     0.00000015                0.1501            0.00000008              0.1276
6     1967     0.00000009                0.1611

             COV(S,S')               CORR(S,S')
              0.00012635                 0.1144


THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.

NUMBER OF ITERATIONS COMPLETED =  4
```

Example 3.2c

```
MALE BLUE WING TEAL BANDED PRESEASON IN SASKATCHEWAN, 1962-68

H02


MATRIX OF DATA VALUES -- ADULTS
-----------------------------------------------
       1      3.00    3.00    3.00    3.00    1.00    3.00    0.0     0.0     0.0     0.0     0.0     0.0
       2      0.0    14.00    5.00   11.00    6.00    5.00    2.00    0.0     1.00    0.0     0.0     0.0
       3      0.0     0.0    17.00   18.00    7.00   10.00    1.00    7.00    1.00    1.00    0.0     0.0
       4      0.0     0.0     0.0     9.00   14.00    6.00    2.00    0.0     4.00    0.0     0.0     0.0
       5      0.0     0.0     0.0     0.0    44.00   23.00    5.00    9.00    2.00    5.00    0.0     4.00
       6      0.0     0.0     0.0     0.0     0.0    47.00   14.00   19.00   19.00    8.00    5.00    2.00
       7      0.0     0.0     0.0     0.0     0.0     0.0    11.00   25.00   12.00    4.00    2.00    2.00


MATRIX OF EXPECTED VALUES -- ADULTS
-----------------------------------------------
       1      3.00    3.75    2.43    2.39    1.70    1.01    0.28    0.44    0.28    0.16    0.07    0.06
       2      0.0    12.74    8.26    8.12    5.78    3.45    0.96    1.48    0.96    0.54    0.23    0.20
       3      0.0     0.0    15.46   15.21   10.83    6.46    1.79    2.77    1.79    1.01    0.44    0.37
       4      0.0     0.0     0.0    15.35   10.93    6.52    1.81    2.80    1.81    1.02    0.44    0.38
       5      0.0     0.0     0.0     0.0    47.39   28.26    7.84   12.14    7.84    4.44    1.92    1.63
       6      0.0     0.0     0.0     0.0     0.0    42.98   11.92   18.46   11.93    6.75    2.93    2.48
       7      0.0     0.0     0.0     0.0     0.0     0.0    11.94   18.49   11.95    6.76    2.93    2.48


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
-----------------------------------------------
       1     -0.00   -0.39    0.37    0.40   -0.54    1.97   -0.53   -0.66   -0.53   -0.40   -0.26   -0.24
       2      0.0     0.35   -1.14    1.01    0.09    0.84    1.07   -1.22    0.04   -0.74   -0.48   -0.45
       3      0.0     0.0     0.39    0.72   -1.17    1.40   -0.59    2.54   -0.59   -0.01   -0.66   -0.61
       4      0.0     0.0     0.0    -1.63    0.93   -0.20    0.14   -1.67    1.63   -1.01   -0.67   -0.61
       5      0.0     0.0     0.0     0.0    -0.49   -0.99   -1.01   -0.50   -2.09    0.27   -1.39    1.86
       6      0.0     0.0     0.0     0.0     0.0     0.62    0.60    0.13    2.05    0.48    1.21   -0.30
       7      0.0     0.0     0.0     0.0     0.0     0.0    -0.27    1.52    0.01   -1.06   -0.54   -0.31



MATRIX OF DATA VALUES -- YOUNG
-----------------------------------------------
       1      6.00    2.00    1.00    1.00    0.0     2.00    1.00    0.0     0.0     0.0     0.0     0.0
       2      0.0    11.00    5.00    6.00    1.00    1.00    1.00    1.00    0.0     0.0     0.0     1.00
       3      0.0     0.0    19.00    4.00    4.00    4.00    0.0     0.0     1.00    1.00    0.0     0.0
       4      0.0     0.0     0.0    65.00   25.00    8.00    4.00    2.00    4.00    4.00    1.00    0.0
       5      0.0     0.0     0.0     0.0    65.00   17.00    2.00   10.00    6.00    2.00    3.00    1.00
       6      0.0     0.0     0.0     0.0     0.0    52.00    9.00    8.00    3.00    4.00    2.00    1.00
       7      0.0     0.0     0.0     0.0     0.0     0.0     3.00    1.00    0.0     1.00    0.0     0.0


MATRIX OF EXPECTED VALUES -- YOUNG
-----------------------------------------------
       1      5.99    2.52    1.63    1.61    1.14    0.68    0.19    0.29    0.19    0.11    0.05    0.04
       2      0.0    11.04    3.27    3.22    2.29    1.37    0.38    0.59    0.38    0.21    0.09    0.08
       3      0.0     0.0    18.97    6.10    4.34    2.59    0.72    1.11    0.72    0.41    0.18    0.15
       4      0.0     0.0     0.0    65.13   17.42   10.39    2.88    4.46    2.88    1.63    0.71    0.60
       5      0.0     0.0     0.0     0.0    64.87   22.30    6.18    9.57    6.19    3.50    1.52    1.28
       6      0.0     0.0     0.0     0.0     0.0    52.07    5.13    7.94    5.13    2.90    1.26    1.06
       7      0.0     0.0     0.0     0.0     0.0     0.0     2.99    1.49    0.96    0.55    0.24    0.20


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
-----------------------------------------------
       1      0.00   -0.33   -0.50   -0.48   -1.07    1.59    1.86   -0.54   -0.44   -0.33   -0.22   -0.20
       2      0.0    -0.01    0.96    1.55   -0.85   -0.31    1.01    0.54   -0.62   -0.46   -0.30    3.28
       3      0.0     0.0     0.01   -0.85   -0.16    0.88   -0.85   -1.05    0.33    0.93   -0.42   -0.39
       4      0.0     0.0     0.0    -0.02    1.82   -0.74    0.66   -1.17    0.66    1.85    0.35   -0.77
       5      0.0     0.0     0.0     0.0     0.02   -1.12   -1.68    0.14   -0.08   -0.80    1.20   -0.25
       6      0.0     0.0     0.0     0.0     0.0    -0.01    1.71    0.02   -0.94    0.64    0.66   -0.06
       7      0.0     0.0     0.0     0.0     0.0     0.0     0.01   -0.40   -0.98    0.62   -0.49   -0.45


THESE STANDARD NORMAL DEVIATES ARE USEFUL IN EXAMINING THE AGREEMENT BETWEEN THE MODEL AND THE OBSERVED DATA IN A PARTICULAR
CELL. FOR EXAMPLE, A VALUE OF SAY -5 FOR A PARTICULAR CELL MAY INDICATE AN UNUSUAL OBSERVATION OR, PERHAPS, A MISTAKE WAS MADE
IN SUMMARIZING THE DATA. IF THE MODEL IS CORRECT, ABOUT 95% OF THESE STATISTICS SHOULD LIE WITHIN THE INTERVAL -2 TO 2.



CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H02.
-------------------------------------------------------

CHI-SQUARE VALUE =   64.79
DEGREES OF FREEDOM =        57
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   64.79  = 0.22335
```

Estimation of Parameters

Simple expressions cannot be obtained for the ML estimators of the four parameters f, f', S, and S', and evaluation of the estimates for a given data set using a desk calculator would be extremely tedious. However, estimates are easily obtained by using program BROWNIE as illustrated in Example 3.3 (see Appendices A and B for a discussion of the method used in the program). Estimates of the standard errors and confidence intervals are also printed, as well as estimates of the covariances and correlations between the estimators.

No goodness of fit test is computed for this model in program BROWNIE because the performance of this model is on the whole very poor, and the H_{02} model will almost always be preferred. That is, comparison of H_{01} with H_{02} (by means of a likelihood ratio test) almost always results in a clear-cut rejection of H_{01}, in favor of H_{02}, and testing fit to H_{01} then seems unnecessary.

An Example

The blue-winged teal data of Example 3.2 are also used here to illustrate the analysis under H_{01}. The sample output is self-explanatory. Note again the low recovery rates for these data, and the age-dependence of both recovery and survival rates.

The likelihood ratio test (discussed further in Section 3.7) comparing H_{01} with H_{02} is very significant, and therefore the model under H_{01} is rejected in favor of the model under H_{02}; thus the assumption that adult and young recovery rates are constant from year to year is also rejected.

3.5 The Model Under H_2

The reasons that led to considering Model 0 as an extension of Model 1 in Chapter 2 also suggest a similar extension of the H_1 model of this chapter. As discussed in Section 2.5, reporting rates may be different for hunters near banding sites, being increased by band solicitation or decreased by hunters who are more accustomed to seeing bands. A difference in the reporting rate near banding sites affects primarily newly banded birds that are more concentrated near banding sites at the beginning of the hunting season than birds banded in previous years.

In defining the model under H_2, we note that there are many similarities between the H_2 model of this chapter and Model 0 of Chapter 2. Thus the assumptions of H_2 are appropriate in the bird banding context, but may not be valid for similar studies on other types of populations, such as fish-tagging experiments. Also, there are other sets of assumptions which give rise to models that cannot be distinguished from the model under H_2 on the basis of the type of data collected. These points are discussed further below.

The assumptions of H_2 are:

(1) Annual survival, harvest and reporting rates are year-specific;

(2) annual survival and harvest rates are age-dependent for the first year of life only; and

(3) in any year, the reporting rate for newly released birds is different from that for survivors of previous releases.

Assumptions 2 and 3 lead to defining three recovery rates for the model under H_2. The parameters of this model are:

f_i' = recovery rate in year i for birds banded and released as young in year i, $i = 1, \ldots, k$,

S_i' = survival rate for young in year i, $i = \begin{cases} 1, \ldots, k-1 \text{ if } \ell = k \\ 1, \ldots, k \text{ if } \ell > k, \end{cases}$

f_i''' = recovery rate in year i for adults released in year i, $i = 1, \ldots, k$,

f_i = recovery rate in year i for survivors of birds released before year i, $i = 2, \ldots, \ell$,

S_i = survival rate for adults in year i, $i = 1, \ldots, \ell - 1$.

Note that although assumption 3 above applies to the reporting rate for bands taken from young as well as from adults, the rate parameters defined for young under H_2 are the same as those under H_1. This is because the actual reporting and harvest rates are not separately identifiable, so that the harvest rate for young and the different reporting rate for new releases are both reflected in the recovery rate f_i'.

Tables of expected numbers of band recoveries are used to express the structure of this model in terms of N_i, M_i, f_i''', f_i, f_i', S_i, and S_i'. Note that, as suggested by the paragraph above, the expected recoveries for birds released as young are the same under H_2 and H_1 (cf. Tables 3.6 and 3.3).

Example 3.3

```
MALE BLUE WING TEAL BANDED PRESEASON IN SASKATCHEWAN, 1962-68


THE HYPOTHESIS H01
------------------

ASSUMPTIONS:  (1)  YOUNG AND ADULTS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES
------------
              (2)  OTHERWISE, SURVIVAL AND RECOVERY RATES ARE CONSTANT FROM YEAR TO YEAR


PARAMETERS:
----------
S      = CONSTANT ANNUAL SURVIVAL RATE FOR ADULTS
F      = CONSTANT BAND RECOVERY RATE FOR ADULTS
S'     = CONSTANT ANNUAL SURVIVAL RATE FOR YOUNG
F'     = CONSTANT BAND RECOVERY RATE FOR YOUNG

STRUCTURE OF THE MODEL UNDER H01    (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
--------------------------------

                 BANDED AS ADULTS                                          BANDED AS YOUNG

N(1)F        N(1)SF        N(1)SSF        N(1)SSSF            M(1)F'       M(1)S'F       M(1)S'SF       M(1)S'SSF
             N(2)F         N(2)SF         N(2)SSF                          M(2)F'        M(2)S'F        M(2)S'SF
                           N(3)F          N(3)SF                                         M(3)F'         M(3)S'F



ESTIMATES UNDER H01
-------------------

          F       RECOVERY RATE FOR ADULTS                        F'      RECOVERY RATE FOR YOUNG
          ------------------------------------------------       ------------------------------------------------------
          ESTIMATE        STANDARD          95% CONFIDENCE        ESTIMATE        STANDARD             95% CONFIDENCE
                            ERROR              INTERVAL                             ERROR                 INTERVAL
          --------        ---------         --------------        --------        ---------            --------------

           0.0077          0.0005         0.0068 -   0.0086        0.0150           0.0010           0.0130 -   0.0170

          S       SURVIVAL RATE FOR ADULTS                        S'      SURVIVAL RATE FOR YOUNG
          ------------------------------------------------       ------------------------------------------------------
          ESTIMATE        STANDARD          95% CONFIDENCE        ESTIMATE        STANDARD             95% CONFIDENCE
                            ERROR              INTERVAL                             ERROR                 INTERVAL
          --------        ---------         --------------        --------        ---------            --------------

           0.6492          0.0152         0.6193 -   0.6791        0.5007           0.0468           0.4090 -   0.5923



ESTIMATED COVARIANCES AND CORRELATIONS UNDER H01

              COV(F,F')               CORR(F,F')               COV(F,S')               CORR(F,S')
             -0.00000000               -0.0000               -0.00000917                -0.4172

              COV(F,S)                CORR(F,S)                COV(S,F')               CORR(S,F')
             -0.00000438               -0.6112               -0.00000000                -0.0000

              COV(S,S')               CORR(S,S')               COV(F',S')              CORR(F',S')
              0.00000875                0.0123               -0.00000051                -0.0109



THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.

NUMBER OF ITERATIONS COMPLETED =   4


LIKELIHOOD RATIO TEST OF H01 VS H02.
-----------------------------------


THIS TEST COMPARES THE MODEL UNDER H01 WITH THAT UNDER H02 AND THUS TESTS THE ASSUMPTION THAT ADULT
AND YOUNG RECOVERY RATES ARE CONSTANT FROM YEAR TO YEAR.  A 'LARGE' CHI-SQUARE VALUE INDICATES THAT
H02 BETTER DESCRIBES THE DATA AND THAT RECOVERY RATES ARE NOT CONSTANT FROM YEAR TO YEAR.

CHI-SQUARE VALUE =    72.72
DEGREES OF FREEDOM =       17
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN    72.72   = 0.00000
```

Table 3.6. *Expected numbers of band recoveries under H_2 for a banding study with $k=3, \ell=5, s=2$.*

Year banded	Number banded	Year of recovery 1	2	3	4	5
				Birds banded and released as adults		
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'''	$N_2S_2f_3$	$N_2S_2S_3f_4$	$N_2S_2S_3S_4f_5$
3	N_3			N_3f_3'''	$N_3S_3f_4$	$N_3S_3S_4f_5$
				Birds banded and released as young		
1	M_1	M_1f_1'	$M_1S_1'f_2$	$M_1S_1'S_2f_3$	$M_1S_1'S_2S_3f_4$	$M_1S_1'S_2S_3S_4f_5$
2	M_2		M_2f_2'	$M_2S_2'f_3$	$M_2S_2'S_3f_4$	$M_2S_2'S_3S_4f_5$
3	M_3			M_3f_3'	$M_3S_3'f_4$	$M_3S_3'S_4f_5$

Estimation of Parameters

ML estimators of the different recovery rates are:

$$\hat{f}_i=\frac{R_{i\cdot}-R_{ii}}{N_i}\,\frac{W_i-R_{ii}}{T_i+U_i-Q_{i\cdot}-R_{i\cdot}-W_i+R_{ii}} \qquad ,i=\begin{cases}2,\ldots,k-1 \text{ if } \ell=k\\ 2,\ldots,k \text{ if } \ell>k\end{cases}$$

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

$$\hat{f}_i'=\frac{Q_{ii}}{M_i} \qquad ,i=1,\ldots,k.$$

The data of Table 3.2 give

$$\hat{f}_1'''=\frac{R_{11}}{N_1}=\frac{10}{231}=0.0433,$$

$$\hat{f}_1' =\frac{Q_{11}}{M_1}=\frac{83}{962}=0.0863$$

$$\hat{f}_2=\frac{R_{2\cdot}-R_{22}}{N_2}\,\frac{W_2-R_{22}}{T_2+U_2-Q_{2\cdot}-R_{2\cdot}-W_2+R_{22}}=\frac{(131-58)\times(106-58)}{649\times(250-131-106+58)}=0.0760$$

$$\hat{f}_2''' =\frac{58}{649}=0.0894,$$

$$\hat{f}_2'=\frac{103}{702}=0.1467$$

$$\hat{f}_3=\frac{(161-54)\times(120-54)}{885\times(370-161-120+54)}=0.0558.$$

Bias-adjusted ML estimators of survival rates are:

$$\tilde{S}_i=\frac{R_{i\cdot}-R_{ii}}{N_i}\,\frac{N_{i+1}+1}{R_{i+1\cdot}-R_{i+1,i+1}+1}\left(\frac{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}-W_{i+1}+R_{i+1,i+1}}{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}}\right) \qquad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k\end{cases}$$

$$\tilde{S}_i'=\frac{Q_{i\cdot}-Q_{ii}}{M_i}\,\frac{N_{i+1}+1}{R_{i+1\cdot}-R_{i+1,i+1}+1}\left(\frac{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}-W_{i+1}+R_{i+1,i+1}}{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}}\right) \qquad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k.\end{cases}$$

The data of Table 3.2 give

$$\tilde{S}_1 = \frac{R_{1\cdot} - R_{11}}{N_1} \frac{N_2 + 1}{R_{2\cdot} - R_{22} + 1} \left(\frac{T_2 + U_2 - Q_{2\cdot} - R_{2\cdot} - W_2 + R_{22}}{T_2 + U_2 - Q_{2\cdot} - R_{2\cdot}} \right) = \frac{(37 - 10) \times (649 + 1) \times (250 - 131 - 106 + 58)}{231 \times (131 - 58 + 1) \times (250 - 131)} = 0.6126$$

$$\tilde{S}'_1 = \frac{Q_{1\cdot} - Q_{11}}{M_1} \frac{N_2 + 1}{R_{2\cdot} - R_{22} + 1} \left(\frac{T_2 + U_2 - Q_{2\cdot} - R_{2\cdot} - W_2 + R_{22}}{T_2 + U_2 - Q_{2\cdot} - R_{2\cdot}} \right) = \frac{(175 - 83) \times (649 + 1) \times (250 - 131 - 106 + 58)}{962 \times (131 - 58 + 1) \times (250 - 131)} = 0.5012.$$

These calculations can be compared with the complete evaluation of estimates in the printout from the FORTRAN program presented in Example 3.5i and 3.5j.

For completeness we define the unadjusted ML estimators of survival and additional ML estimators, most of which are appropriate when the data are nontriangular (i.e., when $\ell > k$). Thus

$$\hat{S}_i = \frac{R_{i\cdot} - R_{ii}}{N_i} \frac{N_{i+1}}{R_{i+1\cdot} - R_{i+1,i+1}} \left(\frac{T_{i+1} + U_{i+1} - Q_{i+1\cdot} - R_{i+1\cdot} - W_{i+1} + R_{i+1,i+1}}{T_{i+1} + U_{i+1} - Q_{i+1\cdot} - R_{i+1\cdot}} \right) \qquad , i = \begin{cases} 1, \ldots, k-2 \text{ if } \ell = k \\ 1, \ldots, k-1 \text{ if } \ell > k. \end{cases}$$

$$\hat{S}'_i = \frac{Q_{i\cdot} - Q_{ii}}{M_i} \frac{N_{i+1}}{R_{i+1\cdot} - R_{i+1,i+1}} \left(\frac{T_{i+1} + U_{i+1} - Q_{i+1\cdot} - R_{i+1\cdot} - W_{i+1} + R_{i+1,i+1}}{T_{i+1} + U_{i+1} - Q_{i+1\cdot} - R_{i+1\cdot}} \right) \qquad , i = \begin{cases} 1, \ldots, k-2 \text{ if } \ell = k \\ 1, \ldots, k-1 \text{ if } \ell > k. \end{cases}$$

In addition, if $\ell = k$,

$$\widehat{S_{k-1} f_k} = \frac{R_{k-1\cdot} - R_{k-1,k-1}}{N_{k-1}}$$

$$\widehat{S'_{k-1} f_k} = \frac{Q_{k-1\cdot} - Q_{k-1,k-1}}{M_{k-1}}$$

and if $\ell = k + 1$,

$$\widehat{S_k f}_{k+1} = \frac{R_{k\cdot} - R_{kk}}{N_k}$$

$$\widehat{S'_k f}_{k+1} = \frac{Q_{k\cdot} - Q_{kk}}{M_k}$$

and finally if $\ell > k + 1$,

$$\widehat{S_k \cdots S_{k+i-1} f}_{k+i} = \frac{R_{\cdot k+i} + Q_{\cdot k+i}}{T_{k+1} + U_{k+1}} \; \frac{R_{k\cdot} - R_{kk}}{N_k} \qquad , i = 1, \ldots, s,$$

$$\widehat{S'_k f}_{k+1} = \frac{R_{\cdot k+1} + Q_{\cdot k+1}}{T_{k+1} + U_{k+1}} \; \frac{Q_{k\cdot} - Q_{kk}}{M_k}$$

$$\widehat{S'_k S_{k+1} \cdots S_{k+i-1} f}_{k+1} = \frac{R_{\cdot k+i} + Q_{\cdot k+i}}{T_{k+1} + U_{k+1}} \; \frac{Q_{k\cdot} - Q_{kk}}{M_k} \qquad , i = 2, \ldots, s.$$

These additional estimates are used in program BROWNIE in obtaining the matrices of "expected values" and of "standard normal deviates," but they are not all contained in the printout because they are not of biological interest. Examining the output in Example 3.5j shows that, for the data of Table 3.2 where $k = \ell = 9$, and $s = 0$,

$$\widehat{S_{k-1} f_k} = \widehat{S_8 f_9} = \frac{R_{8\cdot} - R_{88}}{N_8} = \frac{R_{89}}{N_8} = \frac{22}{938} = 0.0235.$$

However, $\widehat{S'_{k-1} f_k} = \dfrac{25}{906} = 0.0276$ is not printed.

Sampling Variances, Standard Errors, and Confidence Intervals

Estimators of the large-sample variances of $\hat{f}_i$, $\hat{f}_i'''$, $\hat{f}_i'$, $\tilde{S}_i$ and $\tilde{S}_i'$ are given below. The notation is similar to that introduced in Section 3.2 for the H_1 model. These variance estimators are used as described in Section 3.2 to obtain estimates of the corresponding standand errors and 95% confidence intervals, all of which are contained in the printout of the FORTRAN program BROWNIE (see Example 3.5).

$$\text{var}(\hat{f}_i) = (\hat{f}_i)^2\left[\frac{1}{R_{i\cdot}-R_{ii}} - \frac{1}{N_i} + \frac{1}{T_i+U_i-Q_{i\cdot}-R_{i\cdot}-W_i+R_{ii}} + \frac{1}{W_i-R_{ii}}\right] \quad ,i=\begin{cases}2,\ldots,k-1 \text{ if } \ell=k\\ 2,\ldots,k \quad\text{ if } \ell>k,\end{cases}$$

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

$$\text{var}(\hat{f}_i') = \hat{f}_i'\,(1-\hat{f}_i')/M_i \quad ,i=1,\ldots,k,$$

$$\text{var}(\tilde{S}_i) = (\tilde{S}_i)^2\left[\frac{1}{R_{i\cdot}-R_{ii}} - \frac{1}{N_i} + \frac{1}{R_{i+1\cdot}-R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right.$$
$$\left. + \frac{1}{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}-W_{i+1}+R_{i+1,i+1}} - \frac{1}{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}}\right] \quad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k,\end{cases}$$

$$\text{var}(\tilde{S}_i') = (\tilde{S}_i')^2\left[\frac{1}{Q_{i\cdot}-Q_{ii}} - \frac{1}{M_i} + \frac{1}{R_{i+1\cdot}-R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right.$$
$$\left. + \frac{1}{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}-W_{i+1}+R_{i+1,i+1}} - \frac{1}{T_{i+1}+U_{i+1}-Q_{i+1\cdot}-R_{i+1\cdot}}\right] \quad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k.\end{cases}$$

The computation of these variances and corresponding confidence intervals without using the FORTRAN program is somewhat tedious, as we illustrate below, again using the data of Table 3.2

$$\text{var}(\hat{f}_3) = (\hat{f}_3)^2\left[\frac{1}{R_{3\cdot}-R_{33}} - \frac{1}{N_3} + \frac{1}{T_3+U_3-Q_{3\cdot}-R_{3\cdot}-W_3+R_{33}} + \frac{1}{W_3-R_{33}}\right]$$

$$= (0.0558)^2\left[\frac{1}{161-54} - \frac{1}{885} + \frac{1}{370-161-120+54} + \frac{1}{120-54}\right] = 0.00009453,$$

$$\text{se}(\hat{f}_3) = 0.0097,$$

$$1.96 \times \text{se}(\hat{f}_3) = 0.0190$$

and the 95 % confidence interval for f_3 is (0.0368, 0.0748). Comparison with the corresponding results in the printout of Example 3.5i shows a slight difference in the confidence interval obtained there, which is again due to the greater accuracy of the calculations performed by the computer.

Sampling Covariances and Correlations

Formulae for obtaining estimates of the covariances between the estimators $\hat{f}_i$, $\hat{f}_i'''$, $\hat{f}_i'$, $\tilde{S}_i$, and $\tilde{S}_i'$ are given below, again using the notation of Section 3.2. Estimates of the correlations between corresponding pairs of estimators are obtained using the covariances and variances above in the manner described in Section 3.2 for the H_1 estimators. Use of the FORTRAN program is recommended to avoid the time-consuming computations involved in the evaluation of all the formulae below which are presented here mainly for reference purposes or for use by the reader with a single data set with small k, (e.g., $k=4$ or 5).

Thus,

$$\mathrm{cov}(\hat{f}_i''',\hat{f}_i) = -\hat{f}_i'''\hat{f}_i/N_i \qquad , i = \begin{cases} 2,\ldots,k-1 \text{ if } \ell = k \\ 2,\ldots,k \quad \text{if } \ell > k, \end{cases}$$

$$\mathrm{cov}(\hat{f}_i''',\tilde{S}_i) = -\hat{f}_i'''\tilde{S}_i/N_i \qquad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

$$\mathrm{cov}(\hat{f}_i'''{}_1,\tilde{S}_i) = \hat{f}_i'''{}_1\tilde{S}_i/N_{i+1} \qquad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

$$\mathrm{cov}(\hat{f}_{i+1}''',\tilde{S}_i') = \hat{f}_i'''{}_1\tilde{S}_i'/N_{i+1} \qquad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

$$\mathrm{cov}(\hat{f}_i,\tilde{S}_i) = \hat{f}_i\tilde{S}_i\left[\frac{1}{R_{i\cdot} - R_{ii}} - \frac{1}{N_i}\right] \qquad , i = \begin{cases} 2,\ldots,k-2 \text{ if } \ell = k \\ 2,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

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

$$\mathrm{cov}(\hat{f}_{i+1},\tilde{S}_i') = -\hat{f}_{i+1}\tilde{S}_i'\left[\frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

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

$$\mathrm{cov}(\tilde{S}_i,\tilde{S}_i') = \tilde{S}_i\tilde{S}_i'\left[\frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} + \frac{1}{T_{i+1} + U_{i+1} - Q_{i+1\cdot} - R_{i+1\cdot} - W_{i+1} + R_{i+1,i+1}}\right.$$
$$\left. - \frac{1}{T_{i+1} + U_{i+1} - Q_{i+1\cdot} - R_{i+1\cdot}}\right] \qquad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

$$\mathrm{cov}(\tilde{S}_{i+1},\tilde{S}_i') = -\tilde{S}_{i+1}\tilde{S}_i'\left[\frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = \begin{cases} 1,\ldots,k-3 \text{ if } \ell = k \\ 1,\ldots,k-2 \text{ if } \ell > k, \end{cases}$$

$$\mathrm{cov}(\hat{f}_i',\tilde{S}_i') = -\hat{f}_i'\tilde{S}_i'/M_i \qquad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

For the mallard data of Table 3.2 we have obtained above $\hat{f}_3 = 0.0558$, $\mathrm{se}(\hat{f}_3) = 0.0097$, and from Example 3.5i we obtain $\tilde{S}_3 = 0.7427$, and $\mathrm{se}(\tilde{S}_3) = 0.1137$. Thus, as shown in example 3.5k.

$$\mathrm{cov}(\hat{f}_3,\tilde{S}_3) = \hat{f}_3\tilde{S}_3\left[\frac{1}{R_{3\cdot} - R_{33}} - \frac{1}{N_3}\right] = (0.0558)\times(0.7427)\times\left[\frac{1}{161-54} - \frac{1}{885}\right] = 0.00034049.$$

The estimate of the correlation between $\hat{f}_3$ and $\tilde{S}_3$ using the method described in Section 3.2, is

$$\mathrm{corr}(\hat{f}_3,\tilde{S}_3) = \frac{\mathrm{cov}(\hat{f}_3,\tilde{S}_3)}{\mathrm{se}(\hat{f}_3)\mathrm{se}(\tilde{S}_3)} = \frac{0.00034049}{(0.0097)\times(0.1137)} = 0.3087.$$

Again this is slightly different from the corresponding result in the output in Example 3.5k, because of the greater accuracy of the latter.

As stated in Section 3.2, it must be recognized that these correlations reflect a property of the estimators themselves, and not of the unknown parameters. The reader is referred to the discussion in Section 8.4.

Goodness of Fit Test

A goodness of fit test to the model under H_2 can be obtained in the conventional way (described, for example, in Section 3.3) or by the alternative method referred to in Section 3.2 in relation to testing fit to the model under H_1. The method used in program BROWNIE is the latter, and the resulting chi-square value, degrees of freedom, and associated significance level are printed as shown in Example 3.4e. As usual, large chi-square values (values associated with a small probability under H_2) suggest that the model is not appropriate.

In Example 3.5p, the goodness of fit test to the model under H_2 yields a chi-square value of 37.66 with 41 df. The probability under H_2 of observing a value larger than this is 0.62; hence, there is no reason to suspect that the model is inadequate.

Again a rough idea of fit to the model can be obtained by examining the differences (suitably normalized) between each individual observation and the ML estimate of its expectation under H_2. This is done, as described in Sections 3.2 and 3.3, by obtaining matrices of $\mathbf{E}_{ij}$'s, $\mathbf{E}'_{ij}$'s, Z_{ij}'s, and Z'_{ij}'s, using the ML estimates under H_2 of the expected values presented in Table 3.6 (see Example 3.4d).

Proper and Improper Use of the Model under H_2

We now examine models which are based on different sets of assumptions and parameterizations from those of the H_2 model but which, on the basis of the type of data considered in this chapter, are indistinguishable from the H_2 model. Consider a fish-tagging experiment involving young and adults, where tagging is known to affect the recovery and survival of both age groups in the year immediately following tagging. This would give rise to a model where the expected returns for fish tagged as young would be represented as in Table 3.6, but the expected returns for fish tagged as adults would be

$$\begin{array}{ccccc}
N_1 f_1''' & N_1 S_1''' f_2 & N_1 S_1''' S_2 f_3 & N_1 S_1''' S_2 S_3 f_4 & N_1 S_1''' S_2 S_3 S_4 f_5 \\
 & N_2 f_2''' & N_2 S_2''' f_3 & N_2 S_2''' S_3 f_4 & N_2 S_2''' S_3 S_4 f_5 \\
 & & N_3 f_3''' & N_3 S_3''' f_4 & N_3 S_3''' S_4 f_5\,.
\end{array}$$

Under this model, the parameters f_i, S_i, S_i', and S_i''' are not separately estimable for reasons similar to those given in Section 3.9 in relation to the nonidentifiability of f_i and S_i' if only young are banded. Also the ML estimators of the expected values of R_{ij} and Q_{ij} are exactly the same functions of the data under this model as under the H_2 model. Thus it is not possible to devise a test to distinguish between these models, and the goodness of fit test to H_2 above tests fit to the H_2 model or this alternative model which we will call H_{2a}. Thus if data are analyzed using the FORTRAN program, the tests computed may indicate that the model under H_2 is appropriate when the true model is in fact the H_{2a} model. The H_2 estimates of f_i, S_i, and S_i' will not be valid in this case. Suppose the unknown parameters S_i''' are substantially smaller than the S_i (for instance due to the detrimental effect of tagging) and H_2 estimates of S_i are computed, then these estimates will, on the average, be too small (i.e., negatively biased).

Another model which is also indistinguishable from the H_2 model is characterized by the same expected recoveries for individuals banded as young, but expected recoveries for individuals banded as adults would be

$$\begin{array}{ccccc}
N_1 f_1 & N_1 S_1''' f_2 & N_1 S_1''' S_2 f_3 & N_1 S_1''' S_2 S_3 f_4 & N_1 S_1''' S_2 S_3 S_4 f_5 \\
 & N_2 f_2 & N_2 S_2''' f_3 & N_2 S_2''' S_3 f_4 & N_2 S_2''' S_3 S_4 f_5 \\
 & & N_3 f_3 & N_3 S_3''' f_4 & N_3 S_3''' S_4 f_5\,.
\end{array}$$

Under this model the parameters are separately estimable, but there does not seem to be a meaningful biological interpretation for this parameterization, and so this model is not given further consideration.

In the bird banding context, banding during the hunting season may give rise to the H_{2a} model. This is because recovery rates for new releases reflect recoveries from part of the hunting season, and survival rates for new releases relate to a period which is less than a year. Thus we would expect f_i''' to be less than f_i, and S_i''' to be higher than S_i in this situation, and the H_{2a} model rather than the H_2 model is appropriate. If the relative difference between S_i''' and S_i is negligible compared with the relative difference between f_i''' and f_i, then the H_2 model may be a reasonable approximation for in-season banding and the H_2 estimates will be only slightly biased. This is not unreasonable, because the relative difference in the period of survival for new releases may be close to 1 month out of 12, whereas the difference in the effective hunting season will be more like 1 month out of 3.

To use the H_2 model as an approximation in this situation, all recoveries must be recorded for the hunting season during which they occur, regardless of whether they occur before or after the time of banding. With the exception of survival rates for newly banded birds, an annual survival rate then applies to the period between the start of one hunting season and the next. If the H_2 model is not a good approximation, i.e., if S_i''' is appreciably greater than S_i, the H_2 estimators $\tilde{S}_i$ will be positively biased (i.e., too large on the average).

This use of the model under H_2 is illustrated with a data set for Canada Geese in Example 3.4 below. Other examples containing discussion of the model under H_2 are Examples 3.5 and 3.6. The former contains the mallard data of Table 3.2 and has already been referred to several times in this section. The latter contains a data set for which the model under H_2 seems to be the most appropriate model.

An Example

The model under H_2 is used as an approximation for analyzing data for Canada geese banded "in-season" (i.e., during the hunting season) at Swan Lake Refuge, Missouri, 1949-57 (Vaught and Kirsch 1966). Appropriate portions of the output from program BROWNIE are presented in Example 3.4.

Banding usually occurred during the final weeks of the hunting season, so recovery rates in any season should be lower for the newly banded adults. In Example 3.4b we see that except for years 1954 and 1956, the estimates reflect this; i.e., except for $i=6$ and 8, $\hat{f}_i$ is greater than $\hat{f}_i'''$.

Confidence intervals for annual survival rates are large, but the average survival rates are estimated with reasonable precision (see confidence intervals based on $\bar{\bar{S}}$, $\bar{\bar{S}}'$).

The matrices of standard normal deviates (Example 3.4d) contain few extreme values, and the goodness of fit test yields a chi-square value of 95.17 with 94 df (Example 3.4e), suggesting that agreement between model and data is good. However, both methods of assessing fit are misleading in this situation as neither provides information about how closely the H_2 model approximates the H_{2a} model. By using the H_2 model we are assuming that the survival rate for newly banded adults is not appreciably different from that for previously banded birds. As discussed above, we have no way to test this assumption, and the validity of the H_2 estimators is therefore questionable.

In-season banding of certain species is a common practice because of the convenience of banding at a time when birds are clustered together. This use of the H_2 model is suggested as an approximate method for analyzing some of the data of this sort which already exist. However, we do not thereby intend to encourage the practice of in-season banding in future programs. A decision to band during the hunting season must take into account the tradeoff between the gain in precision, due to the ease of banding large numbers of birds, and the increase in bias of the estimators. We have no idea of the magnitude of the bias likely to be incurred by using the H_2 model to describe in-season banding data.

3.6 The Model Under H_3

The assumption concerning the age-dependence of survival and recovery rates, common to H_{01}, H_{02}, H_1, and H_2, can be made more general. This leads to the hypothesis H_3, the assumptions of which are the same as those of H_2 except that survival and recovery rates are assumed to be different for three age classes, i.e., for young, subadults (birds between 1 and 2 years old), and "older birds" (birds more than 2 years old).

The experimental situation remains the same as that of previous sections in that only two age classes are recognized during banding, and we will continue to refer to these two groups as "young" and "adults." Thus every cohort of banded "adults" will contain an unknown number of subadults.

The assumptions of H_3 are:

(1) Annual survival and recovery rates are age-dependent for the first 2 years of life (i.e., are different for young, subadults, and older birds). Note that this embraces assumption 3 of model H_2 for the experimental situation we are concerned with, because of the nonidentifiability of the reporting rate λ; and

(2) annual survival and recovery rates are year-specific.

Each cohort of banded adults released contains an unknown number of subadults and under assumption 1 above, the survival and recovery rates for this mixed batch will be different from the corresponding rates which are characteristic of the groups consisting of subadults alone and all older birds. We define f_i''' and S_i''' as the rates which apply to this mixed group in the year immediately after release. Thus the parameters of H_3 are as follows:

f_i''' and S_i''' are the recovery and survival rates, respectively, in year i for the adults (i.e., subadults and older birds) banded and released in year i,

f_i' and S_i' are the recovery and survival rates, respectively, in year i for the young released in year i;

Example 3.4a

```
                                 CANADA GEESE BANDED IN-SEASON AT SWAN LAKE REFUGE, 1949-57

ADULTS        INPUT MATRIX


1949   828.   46.   56.   37.   32.   32.    8.   19.   12.    7.    7.    2.    3.
1950   881.    0.   35.   36.   46.   26.   17.   39.   11.    9.    7.    4.    5.
1951   379.    0.    0.   13.   17.   18.   16.   17.   10.    4.    2.    1.    1.
1952   317.    0.    0.    0.   12.   12.    9.   21.    8.    3.    3.    4.    1.
1953   358.    0.    0.    0.    0.   22.   17.   20.   13.    6.    4.    4.    4.
1954   425.    0.    0.    0.    0.    0.   36.   33.    8.    7.   10.    8.    2.
1955   833.    0.    0.    0.    0.    0.    0.   57.   40.   25.   21.   18.   15.
1956   455.    0.    0.    0.    0.    0.    0.    0.   33.   26.   25.    9.    4.
1957   433.    0.    0.    0.    0.    0.    0.    0.    0.   15.   27.   17.    8.

YOUNG         INPUT MATRIX


1949   662.   59.   31.   29.   17.   19.   11.   20.    5.    5.    6.    3.    2.
1950   596.    0.   29.   36.   19.   17.   16.    9.    4.    1.    5.    0.    1.
1951   573.    0.    0.   37.   33.   19.   14.   21.    7.    4.   10.    6.    2.
1952   676.    0.    0.    0.   53.   39.   31.   37.   14.    5.   10.    3.    2.
1953   601.    0.    0.    0.    0.   51.   36.   39.   12.    7.   13.   12.    3.
1954  1160.    0.    0.    0.    0.    0.  140.   81.   22.   16.   16.   10.    6.
1955   994.    0.    0.    0.    0.    0.    0.  156.   38.   22.   25.   13.   11.
1956   850.    0.    0.    0.    0.    0.    0.    0.  105.   33.   31.   21.   16.
1957   400.    0.    0.    0.    0.    0.    0.    0.    0.   39.   27.   11.   10.


BASIC SUBTOTALS

 I  RROW(I)   RCOL(I)   CROW(I)   QCOL(I)      T(I)      U(I)      W(I)      Z(I)
 1   261.00     46.00    207.00     59.00    261.00    207.00     46.00      0.0
 2   235.00     91.00    137.00     60.00    450.00    285.00    122.00    215.00
 3    99.00     86.00    153.00    102.00    458.00    378.00    151.00    476.00
 4    73.00    107.00    194.00    122.00    445.00    470.00    176.00    532.00
 5    90.00    110.00    173.00    145.00    428.00    521.00    204.00    545.00
 6   104.00    103.00    291.00    248.00    422.00    667.00    211.00    572.00
 7   176.00    206.00    265.00    363.00    495.00    684.00    413.00    587.00
 8    97.00    135.00    206.00    207.00    386.00    527.00    237.00    501.00
 9    67.00    102.00     87.00    132.00    318.00    407.00    195.00    470.00
                                             216.00    275.00              443.00


THE HYPOTHESIS H2.   (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
------------
              (2)  YOUNG BIRDS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES FROM THOSE OF ADULTS.

              (3)  IN ANY YEAR, THE REPORTING RATE FOR NEW RELEASES IS DIFFERENT FROM THAT FOR SURVIVORS OF PREVIOUSLY
                   BANDED COHORTS, AND HENCE THE CORRESPONDING RECOVERY RATES ARE DIFFERENT.

  H2 IS AN EXTENSION OF H1 IN THAT THE FIRST YEAR ADULT RECOVERY RATE IN YEAR I IS DIFFERENT FROM THE RECOVERY RATE IN

YEAR I OF PREVIOUSLY BANDED ADULTS. (THE SOLICITING OF BANDS FROM HUNTERS BY CONSERVATION OFFICERS NEAR BANDING SITES

MAY GIVE RISE TO THIS SITUATION).


PARAMETERS:
-----------
F'''(I)= BAND RECOVERY RATE IN YEAR I FOR ADULTS BANDED IN YEAR I.
F(I)   = BAND RECOVERY RATE IN YEAR I FOR SURVIVORS OF COHORTS BANDED BEFORE YEAR I.
S(I)   = SURVIVAL RATE FOR ADULTS IN YEAR I.
F'(I)  = BAND RECOVERY RATE FOR YOUNG IN YEAR I.
S'(I)  = SURVIVAL RATE FOR YOUNG IN YEAR I.


STRUCTURE OF THE MODEL UNDER H2  (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
-------------------------------

                                   BANDED AS ADULTS

               N(1)F'''(1)  N(1)S(1)F(2)     N(1)S(1)S(2)F(3)    N(1)S(1)S(2)S(3)F(4)
                            N(2)F'''(2)      N(2)S(2)F(3)        N(2)S(2)S(3)F(4)
                                             N(3)F'''(3)         N(3)S(3)F(4)

                                   BANDED AS YOUNG

               M(1)F'(1)    M(1)S'(1)F(2)    M(1)S'(1)S(2)F(3)   M(1)S'(1)S(2)S(3)F(4)
                            M(2)F'(2)        M(2)S'(2)F(3)       M(2)S'(2)S(3)F(4
                                             M(3)F'(3)           M(3)S'(3)F(4)
```

Example 3.4b

```
ESTIMATES UNDER H2
------------------


                             F(I)                                                      S(I)
             ----------------------------------------------------        ----------------------------------------------------
I    YR   ESTIMATE         STANDARD            95% CONFIDENCE            ESTIMATE         STANDARD            95% CONFIDENCE
          --------         __ERROR__           ___INTERVAL___            --------         __ERROR__           ___INTERVAL___

1  1949                                                                   0.8663           0.0783             0.7128 -  1.0199

2  1950   0.0716            0.0099            0.0522 -  0.0909            0.7573           0.0876             0.5855 -  0.9290

3  1951   0.0702            0.0095            0.0515 -  0.0889            0.8693           0.1311             0.6123 -  1.1263

4  1952   0.0652            0.0095            0.0465 -  0.0839            0.7356           0.1179             0.5045 -  0.9666

5  1953   0.0686            0.0096            0.0499 -  0.0873            0.8770           0.1380             0.6066 -  1.1474

6  1954   0.0539            0.0076            0.0390 -  0.0689            0.5756           0.0830             0.4128 -  0.7384

7  1955   0.1331            0.0150            0.1038 -  0.1625            0.6670           0.0977             0.4755 -  0.8585

8  1956   0.0707            0.0102            0.0507 -  0.0907            0.7887           0.1392             0.5159 -  1.0616

9  1957   0.0553            0.0087            0.0381 -  0.0724

          AVERAGE          STANDARD            95% CONFIDENCE            AVERAGE          STANDARD            95% CONFIDENCE
          ESTIMATE         ERROR___            ___INTERVAL___            ESTIMATE         ERROR___            ___INTERVAL___

  F̄     = 0.0736            0.0036             0.0665 -  0.0806     S̄    = 0.7671            0.0170             0.7338 -  0.8003

                             F'(I)                                                     S'(I)
             ----------------------------------------------------        ----------------------------------------------------
I    YR   ESTIMATE         STANDARD            95% CONFIDENCE            ESTIMATE         STANDARD            95% CONFIDENCE
          --------         __ERROR__           ___INTERVAL___            --------         __ERROR__           ___INTERVAL___

1  1949   0.0891            0.0111            0.0674 -  0.1108            0.7459           0.0745             0.5999 -  0.8919

2  1950   0.0487            0.0088            0.0314 -  0.0659            0.6045           0.0790             0.4495 -  0.7594

3  1951   0.0646            0.0103            0.0444 -  0.0847            0.7755           0.1114             0.5572 -  0.9939

4  1952   0.0784            0.0103            0.0581 -  0.0987            0.7973           0.1071             0.5873 -  1.0073

5  1953   0.0849            0.0114            0.0626 -  0.1071            0.9372           0.1304             0.6816 -  1.1929

6  1954   0.1207            0.0096            0.1019 -  0.1394            0.4683           0.0559             0.3588 -  0.5778

7  1955   0.1569            0.0115            0.1343 -  0.1796            0.5120           0.0767             0.3618 -  0.6623

8  1956   0.1235            0.0113            0.1014 -  0.1457            0.6663           0.1084             0.4539 -  0.8787

9  1957   0.0975            0.0148            0.0684 -  0.1266

          AVERAGE          STANDARD            95% CONFIDENCE            AVERAGE          STANDARD            95% CONFIDENCE
          ESTIMATE         ERROR___            ___INTERVAL___            ESTIMATE         ERROR___            ___INTERVAL___

  F̄'    = 0.0960            0.0037             0.0888 -  0.1033     S̄'   = 0.6884            0.0339             0.6220 -  0.7548

                             F'''(I)                                                   SK...SK+I-1FK+I
             ----------------------------------------------------        ----------------------------------------------------
I    YR   ESTIMATE         STANDARD            95% CONFIDENCE            ESTIMATE         STANDARD            95% CONFIDENCE
          --------         __ERROR__           ___INTERVAL___            --------         __ERROR__           ___INTERVAL___

1  1949   0.0556            0.0080            0.0400 -  0.0712            0.0609           0.0084             0.0445 -  0.0773

2  1950   0.0397            0.0066            0.0268 -  0.0526            0.0357           0.0053             0.0254 -  0.0460

3  1951   0.0343            0.0093            0.0160 -  0.0526            0.0235           0.0037             0.0162 -  0.0308

4  1952   0.0379            0.0107            0.0168 -  0.0589

5  1953   0.0615            0.0127            0.0366 -  0.0863

6  1954   0.0847            0.0135            0.0582 -  0.1112

7  1955   0.0684            0.0087            0.0513 -  0.0856

8  1956   0.0725            0.0122            0.0487 -  0.0964

9  1957   0.0346            0.0088            0.0174 -  0.0519

          AVERAGE          STANDARD            95% CONFIDENCE
          ESTIMATE         ERROR___            ___INTERVAL___

  F̄''' = 0.0544            0.0034             0.0476 -  0.0611
```

Example 3.4c

```
     CANADA GEESE BANDED IN-SEASON AT SWAN LAKE REFUGE, 1949-57

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H2


I         YR    COVAR(F'''(I),F(I))       CORR(F'''(I),F(I))        COVAR(F'''(I),S(I))       CORR(F'''(I),S(I))
1       1949                                                        -0.000058127              -0.093211773
2       1950    -0.000003227              -0.049737690              -0.000034147              -0.059233279
3       1951    -0.000006354              -0.071217249              -0.000078674              -0.064181501
4       1952    -0.000007786              -0.076151510              -0.000087838              -0.069523514
5       1953    -0.000011774              -0.097117528              -0.000150540              -0.085960804
6       1954    -0.000010753              -0.104359439              -0.000114719              -0.102271894
7       1955    -0.000010936              -0.083557977              -0.000054794              -0.064110916
8       1956    -0.000011266              -0.090916909              -0.000125723              -0.074273399
9       1957    -0.000004423              -0.057536717


I         YR    COVAR(F'''(I+1),S(I))     CORR(F'''(I+1),S(I))      COVAR(F'''(I+1),S'(I))    CORR(F'''(I+1),S'(I))
1       1949     0.000039066               0.075782919               0.000033635               0.068597296
2       1950     0.000068534               0.083678970               0.000054705               0.074029646
3       1951     0.000103807               0.073860008               0.000092613               0.077550584
4       1952     0.000126264               0.084395831               0.000136861               0.100655835
5       1953     0.000174791               0.093795504               0.000186800               0.106022676
6       1954     0.000047282               0.065081605               0.000038468               0.078730115
7       1955     0.000106326               0.089504179               0.000081617               0.087555889
8       1956     0.000063101               0.051576477               0.000053305               0.055973691


I         YR    COVAR(F(I),S(I))          CORR(F(I),S(I))           COVAR(S(I),F(I+1))        CORR(S(I),F(I+1))
1       1949                                                        -0.000239601              -0.310224790
2       1950     0.000209433               0.242477212              -0.000477940              -0.571621450
3       1951     0.000548653               0.438431588              -0.000750385              -0.599952081
4       1952     0.000634953               0.564727629              -0.000601028              -0.533856266
5       1953     0.000716583               0.543755784              -0.000584460              -0.555291211
6       1954     0.000383594               0.605472811              -0.000551958              -0.444205843
7       1955     0.000639652               0.437580896              -0.000633005              -0.635745482
8       1956     0.000748481               0.527561525              -0.000737848              -0.605903168


I         YR    COVAR(F(I+1),S'(I))       CORR(F(I+1),S'(I))        COVAR(S(I),S(I+1))        CORR(S(I),S(I+1))
1       1949    -0.000206292              -0.280809760              -0.002535504              -0.369450782
2       1950    -0.000381502              -0.505705398              -0.005917452              -0.515149064
3       1951    -0.000669466              -0.629930166              -0.008465208              -0.547734082
4       1952    -0.000651472              -0.636710910              -0.007684615              -0.472527400
5       1953    -0.000624618              -0.627678792              -0.006235566              -0.544182876
6       1954    -0.000449060              -0.537362166              -0.002765458              -0.340822761
7       1955    -0.000485898              -0.621906553              -0.007064112              -0.519363711
8       1956    -0.000623306              -0.657560223


I         YR    COVAR(S(I),S'(I))         CORR(S(I),S'(I))          COVAR(S(I+1),S'(I))       CORR(S(I+1),S'(I))
1       1949     0.003058611               0.523986295              -0.002183030              -0.334420054
2       1950     0.004357167               0.629215732              -0.004723437              -0.455745538
3       1951     0.009277858               0.635106565              -0.007552341              -0.575102865
4       1952     0.007295091               0.577755662              -0.008329593              -0.563566130
5       1953     0.010552902               0.586329501              -0.006664004              -0.615122417
6       1954     0.002281828               0.491917489              -0.002249913              -0.412297956
7       1955     0.004867200               0.649801508              -0.005422451              -0.508058438
8       1956     0.009315956               0.617527848


I         YR    COVAR(F'(I),S'(I))        CORR(F'(I),S'(I))
1       1949    -0.000100419              -0.121698325
2       1950    -0.000049348              -0.070839992
3       1951    -0.000087398              -0.076403828
4       1952    -0.000092471              -0.083494728
5       1953    -0.000132335              -0.089244904
6       1954    -0.000048722              -0.091199354
7       1955    -0.000080842              -0.091398080
8       1956    -0.000096829              -0.079173023


      THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.
```

Example 3.4d

```
         CANADA GEESE BANDED IN-SEASON AT SWAN LAKE REFUGE, 1949-57

MATRIX OF DATA VALUES -- ADULTS
------------------------------------------
        1     46.00   56.00   37.00   32.00   32.00    8.00   19.00   12.00    7.00    7.00    2.00    3.00
        2      0.0    35.00   36.00   46.00   26.00   17.00   39.00   11.00    9.00    7.00    4.00    5.00
        3      0.0     0.0    13.00   17.00   18.00   16.00   17.00   10.00    4.00    2.00    1.00    1.00
        4      0.0     0.0     0.0    12.00   12.00    9.00   21.00    8.00    3.00    3.00    4.00    1.00
        5      0.0     0.0     0.0     0.0    22.00   17.00   20.00   13.00    6.00    4.00    4.00    4.00
        6      0.0     0.0     0.0     0.0     0.0    36.00   33.00    8.00    7.00   10.00    8.00    2.00
        7      0.0     0.0     0.0     0.0     0.0     0.0    57.00   40.00   25.00   21.00   18.00   15.00
        8      0.0     0.0     0.0     0.0     0.0     0.0     0.0    33.00   26.00   25.00    9.00    4.00
        9      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    15.00   27.00   17.00    8.00


MATRIX OF EXPECTED VALUES -- ADULTS
------------------------------------------
        1     46.00   51.53   38.63   31.60   24.74   17.27   24.71    8.87    5.56    6.13    3.59    2.36
        2      0.0    35.00   47.26   38.66   30.27   21.14   30.24   10.85    6.81    7.50    4.40    2.89
        3      0.0     0.0    13.00   21.77   17.04   11.90   17.02    6.11    3.83    4.22    2.48    1.63
        4      0.0     0.0     0.0    12.00   16.18   11.30   16.17    5.80    3.64    4.01    2.35    1.55
        5      0.0     0.0     0.0     0.0    22.00   17.15   24.53    8.80    5.52    6.08    3.57    2.35
        6      0.0     0.0     0.0     0.0     0.0    36.00   32.80   11.77    7.38    8.14    4.77    3.14
        7      0.0     0.0     0.0     0.0     0.0     0.0    57.00   39.80   24.97   27.50   16.13   10.60
        8      0.0     0.0     0.0     0.0     0.0     0.0     0.0    33.00   20.18   22.22   13.03    8.57
        9      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    15.00   26.37   15.46   10.17


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
------------------------------------------
        1      0.00    0.64   -0.27    0.07    1.48   -2.26   -1.17    1.06    0.61    0.35   -0.84    0.41
        2      0.0     0.00   -1.68    1.21   -0.79   -0.91    1.62    0.05    0.84   -0.18   -0.19    1.24
        3      0.0     0.0     0.00   -1.05    0.24    1.21   -0.01    1.59    0.09   -1.09   -0.94   -0.49
        4      0.0     0.0     0.0     0.00   -1.07   -0.70    1.23    0.92   -0.34   -0.51    1.08   -0.44
        5      0.0     0.0     0.0     0.0     0.00   -0.04   -0.95    1.43    0.20   -0.85    0.23    1.08
        6      0.0     0.0     0.0     0.0     0.0     0.00    0.04   -1.11   -0.14    0.66    1.49   -0.64
        7      0.0     0.0     0.0     0.0     0.0     0.0     0.00    0.03    0.01   -1.26    0.47    1.36
        8      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00    1.33    0.60   -1.13   -1.58
        9      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00    0.13    0.40   -0.69


MATRIX OF DATA VALUES -- YOUNG
------------------------------------------
        1     59.00   31.00   29.00   17.00   19.00   11.00   20.00    5.00    5.00    6.00    3.00    2.00
        2      0.0    29.00   36.00   19.00   17.00   16.00    9.00    4.00    1.00    5.00    0.0     1.00
        3      0.0     0.0    37.00   33.00   19.00   14.00   21.00    7.00    4.00   10.00    6.00    2.00
        4      0.0     0.0     0.0    53.00   39.00   31.00   37.00   14.00    5.00   10.00    3.00    2.00
        5      0.0     0.0     0.0     0.0    51.00   36.00   39.00   12.00    7.00   13.00   12.00    3.00
        6      0.0     0.0     0.0     0.0     0.0   140.00   81.00   22.00   16.00   16.00   10.00    6.00
        7      0.0     0.0     0.0     0.0     0.0     0.0   156.00   38.00   22.00   25.00   13.00   11.00
        8      0.0     0.0     0.0     0.0     0.0     0.0     0.0   105.00   33.00   31.00   21.00   16.00
        9      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    39.00   27.00   11.00   10.00


MATRIX OF EXPECTED VALUES -- YOUNG
------------------------------------------
        1     59.00   35.47   26.59   21.75   17.03   11.89   17.01    6.10    3.83    4.22    2.47    1.63
        2      0.0    29.00   25.52   20.87   16.34   11.41   16.33    5.86    3.68    4.05    2.37    1.56
        3      0.0     0.0    37.00   29.36   22.99   16.05   22.96    8.24    5.17    5.70    3.34    2.20
        4      0.0     0.0     0.0    53.00   37.41   26.12   37.37   13.41    8.41    9.27    5.43    3.57
        5      0.0     0.0     0.0     0.0    51.00   30.76   44.01   15.79    9.91   10.92    6.40    4.21
        6      0.0     0.0     0.0     0.0     0.0   140.00   72.84   26.14   16.40   18.06   10.59    6.96
        7      0.0     0.0     0.0     0.0     0.0     0.0   156.00   36.45   22.87   25.19   14.77    9.71
        8      0.0     0.0     0.0     0.0     0.0     0.0     0.0   105.00   31.84   35.07   20.57   13.52
        9      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    39.00   24.34   14.27    9.38


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
------------------------------------------
        1      0.0    -0.77    0.48   -1.04    0.48   -0.26    0.73   -0.45    0.60    0.87    0.34    0.29
        2      0.0     0.0     2.12   -0.42    0.16    1.37   -1.84   -0.77   -1.40    0.47   -1.54   -0.45
        3      0.0     0.0     0.0     0.69   -0.85   -0.52   -0.42   -0.44   -0.52    1.81    1.46   -0.13
        4      0.0     0.0     0.0     0.0     0.27    0.97   -0.06    0.16   -1.18    0.24   -1.05   -0.83
        5      0.0     0.0     0.0     0.0     0.0     0.97   -0.78   -0.97   -0.93    0.64    2.23   -0.59
        6      0.0     0.0     0.0     0.0     0.0     0.0     0.99   -0.82   -0.10   -0.49   -0.18   -0.37
        7      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.26   -0.18   -0.04   -0.46    0.42
        8      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.21   -0.70    0.10    0.68
        9      0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.56   -0.88    0.20
```

Example 3.4e

```
        CANADA GEESE BANDED IN-SEASON AT SWAN LAKE REFUGE, 1949-57

CHI-SQUARE TEST OF H1 VS H2
---------------------------


       2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                                 WITH 1 DEGREE OF FREEDOM
       -----------------------   ----------------------------------

I= 2          35     200
              87     276                       7.232

I= 3          13      86
             138     446                       5.418

I= 4          12      61
             164     484                       2.798

I= 5          22      68
             182     504                       0.179

I= 6          36      68
             175     519                       4.108

I= 7          57     119
             356     382                      14.418

I= 8          33      64
             204     406                       0.013

I= 9          15      52
             180     351                       2.358
|------------------------------------------------------|
|TOTAL CHI-SQUARE WITH  8 DEGREES OF FREEDOM =  36.523|
|------------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   36.52  = 0.00001


THIS TEST OF THE HYPOTHESIS H1 AGAINST THE HYPOTHESIS H2 TESTS THE ASSUMPTION THAT RECOVERY RATES FOR NEWLY
RELEASED ADULTS ARE THE SAME AS FOR SURVIVORS OF PREVIOUSLY BANDED COHORTS.


CHI-SQUARE TEST OF H2 VS H3
---------------------------

|------------------------------------------------------|
|TOTAL CHI-SQUARE WITH  9 DEGREES OF FREEDOM =  13.532|
|------------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   13.53  = 0.13999



CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H0
-----------------------------------------------------

                                   |----------------------------------------|
                                   |TOTAL CHI-SQUARE     252.54 WITH120 D.F.|
                                   |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  252.54  = 0.0


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H1
-----------------------------------------------------

                                   |----------------------------------------|
                                   |TOTAL CHI-SQUARE     131.69 WITH102 D.F.|
                                   |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  131.69  = 0.02551


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H2
-----------------------------------------------------

                                   |----------------------------------------|
                                   |TOTAL CHI-SQUARE      95.17 WITH 94 D.F.|
                                   |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   95.17  = 0.44686

  FROM THE MODELS ABOVE, ONE SHOULD CHOOSE THE SIMPLEST MODEL (FEWEST PARAMETERS) THAT ADEQUATELY DESCRIBES THE DATA.

ADEQUACY MAY BE JUDGED BY EXAMINING THE RESULTS OF (1) THE GOODNESS OF FIT TESTS, AND (2) THE TESTS BETWEEN SPECIFIC

MODELS. FREQUENTLY, H02 OR H1 IS ADEQUATE.
```

f_i'' and S_i'' are the recovery and survival rates, respectively, in year i for the subadults (i.e., survivors of young released in year $i-1$); and

f_i and S_i are the recovery and survival rates, respectively, in year i for birds over 2 years old.

The structure of the model under H_3 is expressed in Table 3.7 by the expected numbers of band recoveries in terms of $N_i, M_i, f_i, f_i', f_i'', f_i''', S_i, S_i', S_i''$ and S_i'''.

Under this model the only parameters which are separately estimable are $f_i', f_i''', i=1,\ldots,k;$ thus estimation under this model is of little interest and is omitted here. If assumption 1 of H_3 is true, then the H_1 and H_2 estimates will be biased, and it is therefore of interest to attempt to determine whether this assumption is necessary. To this end, the data can be used to compute a goodness of fit test to the H_3 model and a test to compare H_2 with the more general H_3 (see Section 3.7).

Table 3.7. *Expected numbers of band recoveries under H_3 for a banding study with $k=3$, $\ell=5$, $s=0$.*

Year banded	Number banded	Year of recovery				
		1	2	3	4	5
		Birds banded and released as adults (i.e., as subadults or older)				
1	N_1	N_1f_1'''	$N_1S_1'''f_2$	$N_1S_1'''S_2f_3$	$N_1S_1'''S_2S_3f_4$	$N_1S_1'''S_2S_3S_4f_5$
2	N_2		N_2f_2'''	$N_2S_2'''f_3$	$N_2S_2'''S_3f_4$	$N_2S_2'''S_3S_4f_5$
3	N_3			N_3f_3'''	$N_3S_3'''f_4$	$N_3S_3'''S_4f_5$
		Birds banded and released as young				
1	M_1	M_1f_1'	$M_1S_1'f_2''$	$M_1S_1'S_2''f_3$	$M_1S_1'S_2''S_3f_4$	$M_1S_1'S_2''S_3S_4f_5$
2	M_2		M_2f_2'	$M_2S_2'f_3''$	$M_2S_2'S_3''f_4$	$M_2S_2'S_3''S_4f_5$
3	M_3			M_3f_3'	$M_3S_3'f_4''$	$M_3S_3'S_4''f_5$

Goodness of Fit Test

A goodness of fit test to H_3 is computed by the FORTRAN program BROWNIE. The test statistic is obtained by summing the chi-square values from a series of contingency tables, the rows of which are obtained from rows of the recovery matrices in the manner described in Brownie (1973). Again, because observations in the upper right portion of the recovery matrices are frequently 0 or 1, pooling is usually necessary to justify the chi-square approximation, and the FORTRAN program does this using an algorithm adapted from a method described in Robson (1971c). The program prints the contingency tables, as they appear after pooling, together with the chi-square value and associated degrees of freedom (see Example 3.5p). Pooling, if necessary, is accomplished by combining entries in the same row. Occasionally, if the data are very scant, pooling may result in a single column and no chi-square value can be computed. In this situation, the single column is printed together with the message "NO CHI-SQUARE COMPUTED."

Summing the chi-square values and (separately) the degrees of freedom from the contingency tables gives the test statistic (printed as the "TOTAL CHI-SQUARE") and its degrees of freedom. In Example 3.5p, the goodness of fit test to H_3 for the mallard data of Table 3.2 yields a total chi-square of 34.07 with 34 df, and under H_3, the probability of a value at least this large is 0.46. Thus there is no reason to reject the H_3 model.

3.7 Testing Between Models

The models corresponding to H_{01}, H_{02}, H_1, H_2, and H_3 become increasingly more complex as the assumptions of the hypotheses become more general. For a given data set, if we use estimators from too simple a model (a model with too few parameters) they will be biased, but if we use too general a model (a model with more parameters than are necessary) the estimators will not be efficient (i.e., their precision will be unnecessarily low). In that we want to choose the simplest model that adequately describes the data, we must examine the results of the goodness of fit tests and of tests which compare one model with another as described below.

Each of the tests in this section compares a given model with a more general alternative, and "large" values of the test statistic indicate that the simpler model should be rejected in favor of the more general alternative.

Likelihood Ratio Test of H_{01} vs. H_{02}

The simplest model introduced so far in this chapter is the H_{01} model which is compared with the H_{02} model by means of a likelihood ratio test. For large N_i and M_i the test statistic is chi-square distributed with $k + \ell - 2$ degrees of freedom if H_{01} is true. In the FORTRAN program, this test is printed out after the goodness of fit test to H_{02} and before the H_1 estimates (see Example 3.5e).

As mentioned in Section 3.4 the H_{01} model is apparently too simple for most data sets, and the test usually results in a "large" chi-square value and a clearcut rejection of the H_{01} model in favor of the H_{02} model. The assumption that adult and young recovery rates are constant from year to year is usually too restrictive.

For the blue-winged teal data of Example 3.3 the chi-square value for this test is 72.72 with 17 df, and the probability of a value this large, if H_{01} is true, is virtually zero.

For the mallard data of Example 3.5e, the chi-square value is 77.35 with 16 df and again H_{01} is rejected in favor of H_{02}.

Likelihood Ratio Test of H_{02} vs. H_1

The model under H_{02} is compared with the model under H_1 by means of a likelihood ratio test. If H_{02} is true, for large N_i and M_i, the test statistic is chi-square with degrees of freedom given by

$$2k - 4 \text{ if } s = 0 \ (\ell = k)$$
$$2k - 3 \text{ if } s > 0 \ (\ell > k).$$

This test is of interest because, as mentioned earlier, results obtained by the analysis of a large number of data sets indicate that these two models are frequently the most useful, with respect to the criteria above. The test is also of interest because of the relationship of H_{02} to the "hypothesis of compensatory natural mortality." Thus, the test of H_{02} against H_1 essentially tests the assumption that, in the presence of variable recovery rates, adult and young survival rates are constant from year to year. Rejection of the H_{02} assumption of a constant survival rate is necessary before it is possible to say that natural mortality does not compensate fully for variations in hunting losses. The proper formulation of, and testing for, additivity vs. compensatory mortality is dealt with in Anderson and Burnham (1976).

Unfortunately for values of N_i and M_i, usually encountered in practice, the power of this test is low, i.e., the test is not very likely to result in rejection of H_{02} when differences between the annual survival rates are small. The test is computed in the FORTRAN program and the result is printed following a series of tests between H_0, H_1, H_2, and H_3 (where H_0 is defined below) as indicated in Examples 3.5n and 3.5o.

For the blue-winged teal data of Example 3.2c, the test results in a chi-square value of 15.14 with 11 df (inadvertently not photographed), and the H_{02} model is not rejected. In view of this result, the low recovery rates and hence low precision of the estimators, the H_{02} model is probably the most appropriate model for this particular set of data.

For the mallard data in Example 3.5, this test is significant at approximately the 2% level (chi-square value = 26.94 with 14 df, cf. Example 3.5o), and the H_1 model is preferred to the H_{02} model for these data. Having rejected the model under H_{02}, we must compare the model under H_1 with the still more general models under H_2 and H_3, as described below, before we can determine which is the most appropriate model for these data.

Testing Between H_0, H_1, H_2, and H_3

Common to all the models introduced so far in this chapter is the assumption that survival and recovery rates are different for young and adults, so that testing between these models will not provide a test of this assumption. We therefore define the H_0 model, the assumptions of which are that survival, hunting, reporting, and hence recovery, rates are year-specific but independent of age. The model under H_0 is equivalent to assuming Model 1 of Chapter 2 applies to the pooled adult and young recovery data. Like H_{01} and H_{02}, this leads to a model which is simpler than the H_1 model, but which is not comparable with the H_{01} and H_{02} models. The test of H_0 against H_1, tests the assumption that annual survival and recovery rates are independent of age. We also test H_1 against the more general H_2, and H_2 against the still more general H_3.

These three tests are not likelihood ratio tests; rather they are analogous to the goodness of fit tests to the models under H_1, H_2, and H_3 referred to above. The derivation of both these tests between models and the goodness of fit tests to these three models is outlined in Brownie and Robson (1976).

H_0 *vs.* H_1

The test of the null hypothesis that the H_0 model fits the data against the alternative H_1 model (or H_0 vs. H_1) is based on a statistic obtained by summing the single degree of freedom chi-square statistics from each of the 2x2 contingency tables given by

$$\begin{array}{|c:c|l} \hline R_{i\cdot} & N_i - R_{i\cdot} & N_i \\ \hdashline Q_{i\cdot} & M_i - Q_{i\cdot} & M_i \\ \hline \multicolumn{2}{c|}{} & N_i + M_i \end{array} \quad , i = 1, \ldots, k, \text{ and}$$

$$\begin{array}{|c:c|l} \hline W_i & Z_{i+1} & T_i + U_i - Q_{i\cdot} \\ \hdashline Q_{ii} & Q_{i\cdot} - Q_{ii} & Q_{i\cdot} \\ \hline \multicolumn{2}{c|}{} & T_i + U_i \end{array} \quad , i = \begin{cases} 1, \ldots, k-1 & \text{if } s = 0 \\ 1, \ldots, k & \text{if } s > 0, \end{cases}$$

where $Z_{i+1} = T_i + U_i - Q_{i\cdot} - W_i$, $i = 1, \ldots, k$.

The contingency tables and chi-square values are printed together with the total chi-square and degrees of freedom as in Example 3.5n. For the mallard data of Table 3.2 as shown in Example 3.5n, the contingency tables for $i = 2$ are

$$\begin{array}{|c:c|} \hline R_{2\cdot} & N_2 - R_{2\cdot} \\ \hdashline Q_{2\cdot} & M_2 - Q_{2\cdot} \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c:c|} \hline W_2 & T_2 + U_2 - Q_{2\cdot} - W_2 \\ \hdashline Q_{22} & Q_{2\cdot} - Q_{22} \\ \hline \end{array}$$

or

131	518
168	534

and

106	144
103	65

with chi-square values 2.747 and 14.371, respectively. The total chi-square for these data is 129.853 with 17 df, and if the model under H_0 is true, a value this extreme is highly unlikely. In Example 3.5, the model under H_0 is rejected, and we infer that survival and recovery rates are different for young and adults.

Note that although the test is based on the total chi-square value, information concerning particular years of the experiment may be obtained by examining the chi-square values from the contingency tables individually. This may be of more value in testing H_1 against H_2 (see below), but we indicate here how each of the contingency tables above is associated with a particular year of the experiment.

The first series of contingency tables tests the equality of the ratios $R_{i\cdot}/N_i$ and $Q_{i\cdot}/M_i$, where $R_{i\cdot}/N_i$ is the proportion of the adults released in year i which are subsequently recovered, and $Q_{i\cdot}/M_i$ is the proportion of the young released in year i which are subsequently recovered.

Clearly, if survival and recovery rates are the same for adults and young in subsequent years (i.e., $f_j' = f_j$, $S_j = S_j'$, $j = 1, \ldots, \ell$), these two fractions will have the same average or expected value.

The second series of tables tests the equality of $W_i/(T_i + U_i - Q_i)$ and $Q_{ii}/Q_{i\cdot}$, where the former ratio is the proportion of the total recoveries from banded *adults* alive at the start of year i (including survivors of previously released young), which were recovered in year i, and $Q_{ii}/Q_{i\cdot}$ is the proportion of the total recoveries from the *young* released in year i, which were recovered in year i. If $f_i = f_i'$ and $S_i = S_i'$, these fractions should not differ significantly.

For example, the contingency table

W_i	Z_{i+1}
Q_{ii}	$Q_{i\cdot} - Q_{ii}$

for $i=4$ (year 1966) of Example 3.5n gives a large chi-square value of 42.149. Examination of the H_1 estimates (Example 3.5f) suggests this may be due to an unusually high estimated adult survival rate ($\tilde{S}_4=0.8051$) in 1966-67, reflected in the fraction $Z_5/(T_4+U_4-Q_{4\cdot})=316/481$; and also to a high estimated recovery rate for young $\hat{f}_4'=0.1274$, reflected in the fraction $Q_{44}/Q_{4\cdot}=153/259$.

The analysis of a large number of data sets with program BROWNIE indicates that this test is always extremely significant, and hence that there are real differences in the survival and recovery rates for young and adults of a broad range of species of waterfowl. If, however, this test did not result in rejection of H_0, the data for young and adults should be pooled and analyzed by the methods of Chapter 2.

H_1 *vs.* H_2

The test of the H_1 model against H_2 is based on the sum of the single degrees of freedom chi-square statistics obtained from the contingency tables.

R_{ii}	$R_{i\cdot} - R_{ii}$	$R_{i\cdot}$
$W_i - R_{ii}$	$Z_{i+1} - R_{i\cdot} + R_{ii}$	$T_i + U_i - Q_{i\cdot} - R_{i\cdot}$
		$T_i + U_i - Q_{i\cdot} = W_i + Z_{i+1}$

$$, i = \begin{cases} 2,\ldots,k-1 & \text{if } \ell = k \\ 2,\ldots,k & \text{if } \ell > k. \end{cases}$$

This sum is printed by the FORTRAN program as the "TOTAL CHI-SQUARE" with the degrees of freedom which are $k-2$ if $\ell=k$ and $k-1$ if $\ell>k$. Each of the contingency tables and corresponding chi-square values are also printed, and as described above, the individual chi-square values may contain useful information.

The individual chi-squares test the equality of the ratios $R_{ii}/R_{i\cdot}$ and $(W_i-R_{ii})/(T_i+U_i-Q_{i\cdot}-R_{i\cdot})$, where $R_{ii}/R_{i\cdot}$ is the proportion of the total recoveries from the adults released in year i which were recovered in the year immediately after release (i.e., in year i), and $(W_i-R_{ii})/(T_i+U_i-Q_{i\cdot}-R_{i\cdot})$ is the proportion of the total recoveries from the adults alive at the start of year i, but released before year i, which were also recovered in year i. Under H_1, if $f_i'''=f_i$, these two fractions should not differ significantly.

In Example 3.5o, the test of H_1 vs. H_2 gives a total chi-square value of 14.766 with 7 df. The probability under H_1 of observing a value larger than this is 0.04 and we would usually reject H_1 in favor of H_2. However, examining the individual chi-square values we see that only one of these is significantly large, namely the value from the $i=8$ table (corresponding to the year 1970-71) which is 9.856, and which constitutes a large proportion of the total chi-square value of 14.766. Note that this corresponds to the year for which the difference between the H_2 estimates $\hat{f}_i'''$ and $\hat{f}_i$ is greatest (viz $\hat{f}_8=0.0464$, $\hat{f}_8'''=0.1034$). This suggests that the bias of the H_1 estimators of f_i and S_i is probably small except for this year. In view of the poor precision of the H_2 estimators the H_1 model is probably better for these data, although the test of H_1 vs. H_2 would, at first glance, suggest that H_1 is not appropriate. Before deciding that H_1 is the model to use, however, the goodness of fit tests and the test of H_2 vs. H_3 (see below) should also be examined.

Example 3.6 contains a data set for which the H_2 model does seem to be appropriate, and this test of H_1 vs. H_2 is also discussed there.

H_2 *vs.* H_3

This test of H_2 against the alternative H_3 is based on the sum of the single degrees of freedom chi-square statistics obtained from the contingency tables

$Q_{i-1,i}$	$Q_{i-1\cdot} - Q_{i-1,i-1} - Q_{i-1,i}$	$Q_{i-1\cdot} - Q_{i-1,i-1}$
$W_i - R_{ii} - Q_{i-1,i}$	$Z_{i+1} - R_{i\cdot} + R_{ii} - Q_{i-1,i-1} + Q_{i-1,i}$	$T_i + U_i - Q_{i\cdot} - R_{i\cdot}$ $- Q_{i-1\cdot} + Q_{i-1,i-1}$
		$T_i + U_i - Q_{i\cdot} - R_{i\cdot}$

$$, i = \begin{cases} 2, \ldots, k-1 \text{ if } \ell = k \\ 2, \ldots, k \quad\ \ \text{if } \ell > k \end{cases},$$

and if $\ell > k+1$, there is one additional table,

$Q_{k,k+1}$	$Q_{k\cdot} - Q_{kk} - Q_{k,k+1}$	$Q_{k\cdot} - Q_{kk}$
$R_{\cdot k+1} + Q_{\cdot k+1} - Q_{k,k+1}$	$T_{k+1} + U_{k+1} - Q_{k\cdot} + Q_{kk} - R_{\cdot k+1} - Q_{\cdot k+1} + Q_{k,k+1}$	$T_{k+1} + U_{k+1} - Q_{k\cdot} + Q_{kk}$
		$T_{k+1} + U_{k+1}$

The above tables and corresponding chi-square values are printed by the FORTRAN program together with the total chi-square value and degrees of freedom which are $k-2$, $k-1$ or k if $\ell = k$, $\ell = k+1$ or $\ell > k+1$, respectively.

The individual chi-squares are sensitive to differences between the ratios f_i''/S_i'' and f_i/S_i, where f_i'' and S_i'', as defined in Section 3.6, are rate parameters for subadults. Thus, if survival and recovery rates are different for subadults as well as for young (i.e., if $f_i'' \neq f_i$, $S_i'' \neq S_i$) and if $f_i''/S_i'' \neq f_i/S_i$, then the test should usually result in rejection of H_2 in favor of H_3 (if banded samples N_i and M_i are sufficiently large).

The test of H_2 vs. H_3 for the data of Table 3.2, as given in Example 3.5o shows that none of the individual chi-square values nor the total chi-square value, is significantly large. Thus, there is no reason to reject H_2. This result is typical of the analysis of most data sets for a broad variety of waterfowl species, and there does not seem to be an indication that subadults have different survival and recovery rates than older birds. The H_3 model appears to be too general for these data.

An Example

Example 3.5 consists of a complete output from the FORTRAN program containing the analysis of the data of Table 3.2 and has been referred to in most of the preceding sections. Here we illustrate how the complete output is used to determine which model is the most appropriate for a given data set.

Note that the H_1 adult recovery rate estimates $\hat{f}_i$ are slightly larger than the corresponding H_2 estimates, except for $\hat{f}_8 = 0.0888$ which is considerably larger than the H_2 estimate of 0.0464. This is accounted for by noting that except for $i = 6$, the H_2 estimates of recovery rates for newly released adults, $\hat{f}_i'''$, are larger than the corresponding H_2 estimates $\hat{f}_i$, with $\hat{f}_8'''$ considerably larger than $\hat{f}_8$. There is therefore some indication that recovery rates are higher for newly banded birds.

Note also that the confidence intervals based on the H_2 estimates of survival rates are large, and in some instances contain inadmissable values (i.e., values greater than 1). Comparison of the standard errors and confidence intervals for the H_1 and H_2 estimators shows that the loss in precision incurred by using the H_2 estimators is considerable. For example, under H_1, $\text{se}(\hat{f}_2) = 0.0092$, $\text{se}(\tilde{S}_1) = 0.1134$, $\text{se}(\tilde{S}_1') = 0.0594$, while under H_2, $\text{se}(\hat{f}_2) = 0.0165$, $\text{se}(\tilde{S}_1) = 0.1377$, $\text{se}(\tilde{S}_1') = 0.0834$. The decision to choose model H_2 over model H_1 for this data set must take into account the tradeoff between the decrease in bias versus the loss in precision associated with use of the H_2 estimators.

Note the high estimated sampling correlations between $\hat{f}_i$ and $\tilde{S}_i$, denoted by CORR(F(I),S(I)) in Example 3.5g. These correlations make it virtually impossible to use the estimates $\hat{f}_i$ and $\tilde{S}_i$ to obtain information about the relationship between parameters f and S.

The goodness of fit tests to H_0, H_1, H_2, and H_3 on Example 3.5p show that the H_0 model is inappropriate (chi-square = 182.27 with 65 df) but do not indicate a lack of fit to any of the other models. The test of H_{01} vs. H_{02} in Example 3.5e rules out H_{01}. The goodness of fit test to H_{02} (Example 3.5e) and the test of H_{02} vs. H_1 (Example 3.5o) show that H_1 is preferable to H_{02}. Comparing H_1 with the more general H_2 shows that except for 1 extreme year (or possibly 2 years), the H_1 model seems adequate. Lastly, comparing H_2 and H_3 does not indicate any reason for using H_3 instead of H_2. Thus the best model for these data (with regard to the conflicting aims of minimizing bias and maximizing precision of the estimators) is that under either H_1 or H_2. In view of the low precision of the H_2 estimators, H_1 seems preferable.

Example 3.5a

```
THIS OUTPUT REPRESENTS A STATISTICAL ANALYSIS OF BANDING DATA WHEN ANNUAL RECOVERIES ARE RECORDED SEPARATELY FOR BIRDS
BANDED AS ADULTS AND FOR BIRDS BANDED AS YOUNG. DIFFERENT ASSUMPTIONS ABOUT ANNUAL SURVIVAL AND RECOVERY RATES GIVE RISE
TO A SERIES OF HYPOTHESES AND A CORRESPONDING SERIES OF STOCHASTIC MODELS. ESTIMATES OF SURVIVAL AND RECOVERY RATES ARE
COMPUTED FOR MOST MODELS AND THE ASSUMPTIONS OF THE HYPOTHESES ARE EXAMINED BY TESTS OF GOODNESS OF FIT TO THE MODELS AND
BY TESTS TO DISCRIMINATE BETWEEN ALTERNATIVE MODELS.


DEFINITIONS AND NOTATION
------------------------

THE MODELS ARE DEFINED IN TERMS OF TWO BASIC PARAMETERS- S, THE ANNUAL SURVIVAL RATE, AND F, THE ANNUAL BAND RECOVERY
RATE. THAT IS, FOR A BIRD ALIVE AT THE START OF A YEAR, S IS THE PROBABILITY IT SURVIVES THE YEAR, AND F IS THE PROBABILITY
IT IS SHOT AND ITS BAND RETURNED TO THE BIRD BANDING LABORATORY WITHIN THE YEAR. A 'YEAR' IS THE PERIOD BETWEEN SUCCESSIVE
BANDINGS, OR BETWEEN ANNIVERSARIES OF THE BANDING TIME. DIFFERENT ASSUMPTIONS ARE MADE ABOUT THE YEAR-SPECIFICITY AND
AGE-DEPENDENCE OF S AND F UNDER EACH HYPOTHESIS AND THE RESULTING PARAMETERS ARE DEFINED SEPARATELY FOR EACH HYPOTHESIS.
IN GENERAL, AN INDEX ON S OR F DENOTES YEAR-SPECIFICITY, AND A SUPERSCRIPT ' DENOTES AGE-DEPENDENCE. FOR EXAMPLE, S(I)
IS A SURVIVAL RATE FOR YEAR I, AND S'(I) IS A SURVIVAL RATE FOR YOUNG IN YEAR I.

OTHER NOTATION WHICH IS COMMON TO ALL MODELS FOLLOWS.

K          THE NUMBER OF YEARS OF BANDING

L          THE NUMBER OF YEARS OF RECOVERY

N(I)       THE NUMBER OF ADULTS BANDED AND RELEASED AT THE START OF YEAR I, I=1,...,K.

M(I)       THE NUMBER OF YOUNG BANDED AND RELEASED AT THE START OF YEAR I, I=1,...,K.

RROW(I)    ROW TOTAL OF THE RECOVERY MATRIX FOR ADULTS, OR THE TOTAL NUMBER OF RECOVERIES FROM THE ADULTS
           RELEASED IN YEAR I, I=1,...,K.

R(I.)      SAME AS RROW(I)

QROW(I)    ROW TOTAL OF THE RECOVERY MATRIX FOR YOUNG, OR THE TOTAL NUMBER OF RECOVERIES FROM THE YOUNG
           RELEASED IN YEAR I, I=1,...,K.

Q(I.)      SAME AS QROW(I)

RCOL(I)    COLUMN TOTAL OF THE RECOVERY MATRIX FOR ADULTS, I=1,...,L.

QCOL(I)    COLUMN TOTAL OF THE RECOVERY MATRIX FOR YOUNG , I=1,...,L.

Q(I,I)     NUMBER OF RECOVERIES IN YEAR I FROM THE YOUNG RELEASED IN YEAR I, I=1,...,K.

W(I)       =RCOL(I)+QCOL(I)-Q(I,I)= TOTAL NUMBER OF ADULTS RECOVERED IN YEAR I, (INCLUDING RECOVERIES
           FROM SURVIVORS OF YOUNG RELEASED BEFORE YEAR I), I=1,...,K

T(I)       BLOCK TOTAL OF THE RECOVERY MATRIX FOR ADULTS, OR THE TOTAL NUMBER OF RECOVERIES IN YEARS I TO L
           INCLUSIVE, FROM ALL THE ADULTS RELEASED IN YEARS 1 TO I INCLUSIVE.

U(I)       BLOCK TOTAL OF THE RECOVERY MATRIX FOR YOUNG , OR THE TOTAL NUMBER OF RECOVERIES IN YEARS I TO L
           INCLUSIVE, FROM ALL THE YOUNG  RELEASED IN YEARS 1 TO I INCLUSIVE.

Z(I)       =T(I)+U(I)-RROW(I)-QROW(I)-QROW(I-1)+Q(I-1,I-1), A TOTAL INVOLVING BLOCKS FROM
           BOTH DATA ARRAYS, I=2,...,K
```

Example 3.5b

```
                    MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

ADULTS          INPUT 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   550.    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.

YOUNG           INPUT MATRIX

1963   962.   83.   35.   18.   16.    6.    8.    5.    3.    1.
1964   702.    0.  103.   21.   13.   11.    8.    6.    6.    0.
1965  1132.    0.    0.   82.   36.   26.   24.   15.   18.    4.
1966  1201.    0.    0.    0.  153.   39.   22.   21.   16.    8.
1967  1159.    0.    0.    0.    0.  109.   38.   31.   15.    1.
1968  1155.    0.    0.    0.    0.    0.  113.   64.   29.   22.
1969  1131.    0.    0.    0.    0.    0.    0.  124.   45.   22.
1970   906.    0.    0.    0.    0.    0.    0.    0.   95.   25.
1971   353.    0.    0.    0.    0.    0.    0.    0.    0.   38.


BASIC SUBTOTALS

 I   RROW(I)   RCOL(I)   CROW(I)   QCOL(I)      T(I)      U(I)      W(I)      Z(I)
 1     37.00     10.00    175.00     83.00     37.00    175.00     10.00      0.0
 2    131.00     71.00    168.00    138.00    158.00    260.00    106.00     27.00
 3    161.00     81.00    205.00    121.00    248.00    327.00    120.00    144.00
 4    108.00    100.00    259.00    218.00    275.00    465.00    165.00    250.00
 5    140.00    115.00    194.00    191.00    315.00    441.00    197.00    316.00
 6    159.00    161.00    228.00    213.00    359.00    478.00    261.00    365.00
 7    190.00    197.00    191.00    266.00    388.00    456.00    339.00    348.00
 8    119.00    218.00    120.00    227.00    310.00    310.00    350.00    314.00
 9     21.00    113.00     38.00    121.00    113.00    121.00    196.00    150.00


 THE HYPOTHESIS H01
 ------------------

 ASSUMPTIONS:  (1)  YOUNG AND ADULTS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES
 ------------
               (2)  OTHERWISE, SURVIVAL AND RECOVERY RATES ARE CONSTANT FROM YEAR TO YEAR

 PARAMETERS:
 -----------
 S      = CONSTANT ANNUAL SURVIVAL RATE FOR ADULTS
 F      = CONSTANT BAND RECOVERY RATE FOR ADULTS
 S'     = CONSTANT ANNUAL SURVIVAL RATE FOR YOUNG
 F'     = CONSTANT BAND RECOVERY RATE FOR YOUNG

 STRUCTURE OF THE MODEL UNDER H01    (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
 --------------------------------

           BANDED AS ADULTS                                         BANDED AS YOUNG

 N(1)F      N(1)SF      N(1)SSF      N(1)SSSF           M(1)F'      M(1)S'F     M(1)S'SF      M(1)S'SSF
            N(2)F       N(2)SF       N(2)SSF                        M(2)F'      M(2)S'F       M(2)S'SF
                        N(3)F        N(3)SF                                     M(3)F'        M(3)S'F


 ESTIMATES UNDER H01
 -------------------

      F     RECOVERY RATE FOR ADULTS                       F'    RECOVERY RATE FOR YOUNG
      ---------------------------------------------        ---------------------------------------------
      ESTIMATE      STANDARD         95% CONFIDENCE        ESTIMATE      STANDARD         95% CONFIDENCE
                      ERROR             INTERVAL                           ERROR             INTERVAL
      --------      ---------       -------------          --------      ---------       -------------

       0.0677         0.0023        0.0631 -  0.0722        0.1030         0.0033        0.0966 -  0.1093

      S     SURVIVAL RATE FOR ADULTS                       S'    SURVIVAL RATE FOR YOUNG
      ---------------------------------------------        ---------------------------------------------
      ESTIMATE      STANDARD         95% CONFIDENCE        ESTIMATE      STANDARD         95% CONFIDENCE
                      ERROR             INTERVAL                           ERROR             INTERVAL
      --------      ---------       -------------          --------      ---------       -------------

       0.6489         0.0115        0.6264 -  0.6715        0.5403         0.0250        0.4913 -  0.5892


 ESTIMATED COVARIANCES AND CORRELATIONS UNDER H01

           COV(F,F')                CORR(F,F')               COV(F,S')                CORR(F,S')
          -0.00000000                -0.0000               -0.00002921                -0.5032

           COV(F,S)                 CORR(F,S)                COV(S,F')                CORR(S,F')
          -0.00001545                -0.5776                0.00000000                 0.0000

           COV(S,S')                CORR(S,S')               COV(F',S')               CORR(F',S')
           0.00000370                 0.0129               -0.00000636                -0.0784

 THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.

 NUMBER OF ITERATIONS COMPLETED =  4
```

Example 3.5c

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO


THE HYPOTHESIS H02
------------------

ASSUMPTIONS:  (1)  YOUNG AND ADULTS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES
------------
              (2)  SURVIVAL RATES ARE OTHERWISE CONSTANT FROM YEAR TO YEAR

              (3)  RECOVERY RATES ARE YEAR-SPECIFIC

PARAMETERS:
-----------
S     = CONSTANT ANNUAL SURVIVAL RATE FOR ADULTS
F(I)  = BAND RECOVERY RATE IN YEAR I FOR ADULTS
S'    = CONSTANT ANNUAL SURVIVAL RATE FOR YOUNG
F'(I) = BAND RECOVERY RATE IN YEAR I FOR YOUNG

STRUCTURE OF THE MODEL UNDER H02    (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
--------------------------------

               BANDED AS ADULTS                                          BANDED AS YOUNG

N(1)F(1)    N(1)SF(2)    N(1)SSF(3)    N(1)SSSF(4)          M(1)F'(1)    M(1)S'F(2)    M(1)S'SF(3)    M(1)S'SSF(4)
            N(2)F(2)     N(2)SF(3)     N(2)SSF(4)                        M(2)F'(2)     M(2)S'F(3)     M(2)S'SF(4)
                         N(3)F(3)      N(3)SF(4)                                       M(3)F'(3)      M(3)S'F(4)


ESTIMATES UNDER H02
-------------------

        F(I)   RECOVERY RATE FOR ADULTS                         F'(I)  RECOVERY RATE FOR YOUNG
        ------------------------------------------------        ------------------------------------------------
        ESTIMATE        STANDARD           95% CONFIDENCE       ESTIMATE        STANDARD           95% CONFIDENCE
I YEAR                    ERROR                  INTERVAL                         ERROR                  INTERVAL
        --------        --------           --------------       --------        --------           --------------
1 1963   0.0429          0.0133         0.0169 -  0.0690         0.0855          0.0090         0.0679 -  0.1031
2 1964   0.0804          0.0077         0.0654 -  0.0955         0.1468          0.0133         0.1208 -  0.1728
3 1965   0.0569          0.0052         0.0467 -  0.0670         0.0736          0.0077         0.0585 -  0.0888
4 1966   0.0644          0.0051         0.0543 -  0.0745         0.1270          0.0096         0.1083 -  0.1458
5 1967   0.0600          0.0045         0.0512 -  0.0687         0.0891          0.0082         0.0730 -  0.1051
6 1968   0.0669          0.0044         0.0581 -  0.0756         0.0998          0.0088         0.0825 -  0.1170
7 1969   0.0769          0.0046         0.0680 -  0.0859         0.1093          0.0093         0.0912 -  0.1275
8 1970   0.0794          0.0048         0.0700 -  0.0888         0.1047          0.0102         0.0848 -  0.1246
9 1971   0.0534          0.0043         0.0450 -  0.0618         0.1076          0.0165         0.0753 -  0.1400

        AVERAGE         STANDARD           95% CONFIDENCE       AVERAGE         STANDARD           95% CONFIDENCE
        ESTIMATE        ERROR                INTERVAL           ESTIMATE        ERROR                INTERVAL
        --------        --------           --------------       --------        --------           --------------
 -                                                          -
 F     = 0.0646          0.0026           0.0595 -  0.0696   F'   = 0.1048          0.0035           0.0979 -  0.1118



        S       SURVIVAL RATE FOR ADULTS                        S'      SURVIVAL RATE FOR YOUNG
        ------------------------------------------------        ------------------------------------------------
        ESTIMATE        STANDARD           95% CONFIDENCE       ESTIMATE        STANDARD           95% CONFIDENCE
                          ERROR               INTERVAL                          ERROR               INTERVAL
        --------        --------           --------------       --------        --------           --------------

         0.6515          0.0120         0.6280 -  0.6751         0.5415          0.0251         0.4922 -  0.5907
```

Example 3.5d

```
SELECTED ESTIMATED COVARIANCES AND CORRELATIONS UNDER H02

I     YEAR   COV(F(I),S)                CORR(F(I),S)         COV(F(I),S')              CORR(F(I),S')
1     1963    -0.00000251                  -0.0157             0.00000250                  0.0075
2     1964    -0.00000545                  -0.0591            -0.00003531                 -0.1833
3     1965    -0.00000770                  -0.1242            -0.00002093                 -0.1615
4     1966    -0.00001287                  -0.2085            -0.00003055                 -0.2367
5     1967    -0.00001325                  -0.2468            -0.00002809                 -0.2503
6     1968    -0.00001622                  -0.3039            -0.00002999                 -0.2686
7     1969    -0.00001984                  -0.3609            -0.00003200                 -0.2784
8     1970    -0.00002403                  -0.4176            -0.00003428                 -0.2849
9     1971    -0.00002176                  -0.4212            -0.00002673                 -0.2474

I     YEAR   COV(F'(I),S)               CORR(F'(I),S)        COV(F'(I),S')             CORR(F'(I),S')
1     1963    -0.00000199                  -0.0185            -0.00000736                 -0.0327
2     1964    -0.00000250                  -0.0157            -0.00001118                 -0.0336
3     1965    -0.00000071                  -0.0077            -0.00000549                 -0.0283
4     1966    -0.00000048                  -0.0042            -0.00000922                 -0.0384
5     1967     0.00000040                   0.0041            -0.00000643                 -0.0313
6     1968     0.00000126                   0.0120            -0.00000655                 -0.0297
7     1969     0.00000184                   0.0166            -0.00000515                 -0.0222
8     1970     0.00000123                   0.0101            -0.00000207                 -0.0081
9     1971     0.0                          0.0                0.0                         0.0

I     YEAR   COV(F(I),F'(I))            CORR(F(I),F'(I))     COV(F(I+1),F'(I))         CORR(F(I+1),F'(I))
1     1963     0.00000012                   0.0010            -0.00000239                 -0.0347
2     1964     0.00000087                   0.0085            -0.00000168                 -0.0246
3     1965     0.00000026                   0.0066            -0.00000070                 -0.0177
4     1966     0.00000054                   0.0110            -0.00000085                 -0.0199
5     1967     0.00000028                   0.0077            -0.00000060                 -0.0165
6     1968     0.00000021                   0.0053            -0.00000080                 -0.0199
7     1969     0.00000005                   0.0012            -0.00000110                 -0.0248
8     1970    -0.00000007                  -0.0014            -0.00000093                 -0.0212
9     1971     0.0                          0.0                0.0                         0.0

I     YEAR   COV(F(I),F'(I+1))          CORR(F(I),F'(I+1))   COV(F'(I),F'(I+1))        CORR(F'(I),F'(I+1))
1     1963     0.00000006                   0.0004             0.00000027                  0.0023
2     1964     0.00000037                   0.0063             0.00000019                  0.0019
3     1965     0.00000037                   0.0074             0.00000014                  0.0019
4     1966     0.00000030                   0.0070             0.00000016                  0.0020
5     1967     0.00000020                   0.0050             0.00000012                  0.0017
6     1968     0.00000006                   0.0015             0.00000012                  0.0015
7     1969    -0.00000005                  -0.0010             0.00000007                  0.0007
8     1970     0.0                          0.0                0.0                         0.0

I     YEAR   COV(F(I),F(I+1))           CORR(F(I),F(I+1))    COV(F(I),F(I+2))          CORR(F(I),F(I+2))
1     1963    -0.00000183                  -0.0179            -0.00000043                 -0.0063
2     1964     0.00000031                   0.0078             0.00000150                  0.0380
3     1965     0.00000098                   0.0368             0.00000129                  0.0559
4     1966     0.00000201                   0.0876             0.00000256                  0.1120
5     1967     0.00000235                   0.1184             0.00000292                  0.1428
6     1968     0.00000323                   0.1584             0.00000395                  0.1856
7     1969     0.00000442                   0.2013             0.00000399                  0.2026
8     1970     0.00000458                   0.2219

             COV(S,S')                  CORR(S,S')
               0.00000629                   0.0209


THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.

NUMBER OF ITERATIONS COMPLETED =  4
```

Example 3.5e

```
                                                  HO2
MATRIX OF DATA VALUES -- ADULTS
-------------------------------------------------------
      1        10.00  13.00   6.00   1.00   1.00   3.00   1.00   2.00   0.0
      2         0.0   58.00  21.00  16.00  15.00  13.00   6.00   1.00   1.00
      3         0.0    0.0   54.00  39.00  23.00  18.00  11.00  10.00   6.00
      4         0.0    0.0    0.0   44.00  21.00  22.00   9.00   9.00   3.00
      5         0.0    0.0    0.0    0.0   55.00  39.00  23.00  11.00  12.00
      6         0.0    0.0    0.0    0.0    0.0   66.00  46.00  29.00  18.00
      7         0.0    0.0    0.0    0.0    0.0    0.0  101.00  59.00  30.00
      8         0.0    0.0    0.0    0.0    0.0    0.0    0.0   97.00  22.00
      9         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   21.00


MATRIX OF EXPECTED VALUES -- ADULTS
-------------------------------------------------------
      1         9.91  12.11   5.57   4.12   2.50   1.81   1.36   0.91   0.40
      2         0.0   52.21  24.04  17.75  10.76   7.82   5.86   3.94   1.73
      3         0.0    0.0   50.31  37.14  22.52  16.37  12.27   8.25   3.61
      4         0.0    0.0    0.0   38.00  23.05  16.75  12.56   8.44   3.70
      5         0.0    0.0    0.0    0.0   56.54  41.08  30.80  20.70   9.07
      6         0.0    0.0    0.0    0.0    0.0   72.01  53.99  36.29  15.90
      7         0.0    0.0    0.0    0.0    0.0    0.0   96.19  64.65  28.33
      8         0.0    0.0    0.0    0.0    0.0    0.0    0.0   74.46  32.63
      9         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   16.66


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
-------------------------------------------------------
      1         0.03   0.26   0.18  -1.55  -0.95   0.88  -0.31   1.14  -0.63
      2         0.0    0.84  -0.63  -0.42   1.30   1.86   0.06  -1.49  -0.55
      3         0.0    0.0    0.54   0.31   0.10   0.41  -0.37   0.61   1.26
      4         0.0    0.0    0.0    1.01  -0.43   1.30  -1.01   0.19  -0.36
      5         0.0    0.0    0.0    0.0   -0.21  -0.33  -1.43  -2.16   0.98
      6         0.0    0.0    0.0    0.0    0.0   -0.73  -1.12  -1.23   0.53
      7         0.0    0.0    0.0    0.0    0.0    0.0    0.51  -0.72   0.32
      8         0.0    0.0    0.0    0.0    0.0    0.0    0.0    2.72  -1.89
      9         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0    1.09


MATRIX OF DATA VALUES -- YOUNG                    HO2
-------------------------------------------------------
      1        83.00  35.00  18.00  16.00   6.00   8.00   5.00   3.00   1.00
      2         0.0  103.00  21.00  13.00  11.00   8.00   6.00   6.00   0.0
      3         0.0    0.0   82.00  36.00  26.00  24.00  15.00  18.00   4.00
      4         0.0    0.0    0.0  153.00  39.00  22.00  21.00  16.00   8.00
      5         0.0    0.0    0.0    0.0  109.00  38.00  31.00  15.00   1.00
      6         0.0    0.0    0.0    0.0    0.0  113.00  64.00  29.00  22.00
      7         0.0    0.0    0.0    0.0    0.0    0.0  124.00  45.00  22.00
      8         0.0    0.0    0.0    0.0    0.0    0.0    0.0   95.00  25.00
      9         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   38.00


MATRIX OF EXPECTED VALUES -- YOUNG
-------------------------------------------------------
      1        82.27  41.90  19.29  14.24   8.64   6.28   4.71   3.16   1.39
      2         0.0  103.06  21.61  15.95   9.67   7.03   5.27   3.54   1.55
      3         0.0    0.0   83.34  39.48  23.94  17.40  13.04   8.77   3.84
      4         0.0    0.0    0.0  152.57  38.99  28.33  21.24  14.28   6.26
      5         0.0    0.0    0.0    0.0  106.81  43.41  32.55  21.88   9.59
      6         0.0    0.0    0.0    0.0    0.0  115.21  48.12  32.35  14.17
      7         0.0    0.0    0.0    0.0    0.0    0.0  123.66  48.61  21.30
      8         0.0    0.0    0.0    0.0    0.0    0.0    0.0   94.87  26.19
      9         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   38.00


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
-------------------------------------------------------
      1         0.08  -1.09  -0.30   0.47  -0.90   0.69   0.14  -0.09  -0.33
      2         0.0   -0.01  -0.13  -0.75   0.43   0.37   0.32   1.31  -1.25
      3         0.0    0.0   -0.15  -0.56   0.43   1.60   0.54   3.13   0.08
      4         0.0    0.0    0.0    0.04   0.00  -1.20  -0.05   0.46   0.70
      5         0.0    0.0    0.0    0.0    0.22  -0.84  -0.27  -1.48  -2.78
      6         0.0    0.0    0.0    0.0    0.0   -0.22   2.34  -0.60   2.09
      7         0.0    0.0    0.0    0.0    0.0    0.0    0.03  -0.53   0.15
      8         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.01  -0.24
      9         0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.00


THESE STANDARD NORMAL DEVIATES ARE USEFUL IN EXAMINING THE AGREEMENT BETWEEN THE MODEL AND THE OBSERVED DATA IN A PARTICULAR
CELL. FOR EXAMPLE, A VALUE OF SAY -5 FOR A PARTICULAR CELL MAY INDICATE AN UNUSUAL OBSERVATION OR, PERHAPS, A MISTAKE WAS MADE
IN SUMMARIZING THE DATA. IF THE MODEL IS CORRECT, ABOUT 95% OF THESE STATISTICS SHOULD LIE WITHIN THE INTERVAL -2 TO 2.


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER HO2.
-------------------------------------------------------

CHI-SQUARE VALUE =   83.57
DEGREES OF FREEDOM =        64
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN     83.57  = 0.05082


LIKELIHOOD RATIO TEST OF HO1 VS HO2.
------------------------------------

THIS TEST COMPARES THE MODEL UNDER HO1 WITH THAT UNDER HO2 AND THUS TESTS THE ASSUMPTION THAT ADULT
AND YOUNG RECOVERY RATES ARE CONSTANT FROM YEAR TO YEAR.  A 'LARGE' CHI-SQUARE VALUE INDICATES THAT
HO2 BETTER DESCRIBES THE DATA AND THAT RECOVERY RATES ARE NOT CONSTANT FROM YEAR TO YEAR.

CHI-SQUARE VALUE =   77.35
DEGREES OF FREEDOM =        16
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN     77.35  = 0.0
```

Example 3.5f

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO


THE HYPOTHESIS H1.    (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
------------
              (2)  YOUNG BIRDS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES FROM THOSE OF ADULTS.


PARAMETERS:
-----------
F(I)    = BAND RECOVERY RATE FOR ADULTS IN YEAR I.
S(I)    = SURVIVAL RATE FOR ADULTS IN YEAR I.
F'(I)   = BAND RECOVERY RATE FOR YOUNG IN YEAR I.
S'(I)   = SURVIVAL RATE FOR YOUNG IN YEAR I.


STRUCTURE OF THE MODEL UNDER H1  (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
--------------------------------

                              BANDED AS ADULTS

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

                              BANDED AS YOUNG

           M(1)F'(1)     M(1)S'(1)F(2)   M(1)S'(1)S(2)F(3)   M(1)S'(1)S(2)S(3)F(4)
                         M(2)F'(2)       M(2)S'(2)F(3)       M(2)S'(2)S(3)F(4
                                         M(3)F'(3)           M(3)S'(3)F(4)


ESTIMATES UNDER H1
------------------


                           F(I)                                                  S(I)
        ------------------------------------------------------   ------------------------------------------------------
 I   YR  ESTIMATE          STANDARD          95% CONFIDENCE        ESTIMATE          STANDARD          95% CONFIDENCE
         --------          __ERROR__         ___INTERVAL___        --------          __ERROR__         ___INTERVAL___

 1  1963  0.0433            0.0134          0.0170 -  0.0695        0.5756            0.1134          0.3533 -  0.7978

 2  1964  0.0856            0.0092          0.0676 -  0.1036        0.6359            0.0756          0.4878 -  0.7840

 3  1965  0.0590            0.0061          0.0470 -  0.0710        0.6665            0.0787          0.5122 -  0.8207

 4  1966  0.0628            0.0067          0.0496 -  0.0760        0.8051            0.0977          0.6136 -  0.9967

 5  1967  0.0520            0.0050          0.0422 -  0.0619        0.6496            0.0724          0.5078 -  0.7914

 6  1968  0.0633            0.0055          0.0525 -  0.0740        0.5525            0.0581          0.4387 -  0.6664

 7  1969  0.0789            0.0061          0.0670 -  0.0908        0.5719            0.0663          0.4419 -  0.7020

 8  1970  0.0888            0.0080          0.0730 -  0.1046        0.5415            0.1286          0.2894 -  0.7936

 9  1971  0.0673            0.0142          0.0395 -  0.0951

         AVERAGE           STANDARD          95% CONFIDENCE        AVERAGE           STANDARD          95% CONFIDENCE
         ESTIMATE          ERROR___          ___INTERVAL___        ESTIMATE          ERROR___          ___INTERVAL___
   _                                                           _
   F      = 0.0668          0.0029           0.0610 -  0.0726   S    = 0.6248           0.0214           0.5828 -  0.6668

                           F'(I)                                                 S'(I)
        ------------------------------------------------------   ------------------------------------------------------
 I   YR  ESTIMATE          STANDARD          95% CONFIDENCE        ESTIMATE          STANDARD          95% CONFIDENCE
         --------          __ERROR__         ___INTERVAL___        --------          __ERROR__         ___INTERVAL___

 1  1963  0.0863            0.0091          0.0685 -  0.1040        0.4709            0.0594          0.3545 -  0.5874

 2  1964  0.1467            0.0134          0.1205 -  0.1729        0.5064            0.0699          0.3694 -  0.6434

 3  1965  0.0724            0.0077          0.0573 -  0.0875        0.5891            0.0717          0.4486 -  0.7297

 4  1966  0.1274            0.0096          0.1085 -  0.1463        0.5909            0.0716          0.4506 -  0.7312

 5  1967  0.0909            0.0083          0.0746 -  0.1072        0.4776            0.0610          0.3581 -  0.5971

 6  1968  0.0978            0.0087          0.0807 -  0.1150        0.6521            0.0723          0.5104 -  0.7939

 7  1969  0.1096            0.0093          0.0914 -  0.1278        0.4635            0.0678          0.3307 -  0.5964

 8  1970  0.1049            0.0102          0.0849 -  0.1248        0.3926            0.1133          0.1705 -  0.6147

 9  1971  0.1076            0.0165          0.0753 -  0.1400

         AVERAGE           STANDARD          95% CONFIDENCE        AVERAGE           STANDARD          95% CONFIDENCE
         ESTIMATE          ERROR___          ___INTERVAL___        ESTIMATE          ERROR___          ___INTERVAL___
   _                                                           _
   F'     = 0.1049          0.0035           0.0979 -  0.1118   S'   = 0.5179           0.0265           0.4659 -  0.5699
```

Example 3.5g

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO            33570

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H1


I      YR  COVAR(F(I),S(I))            CORR(F(I),S(I))             COVAR(S(I),S(I+1))          CORR(S(I),S(I+1))
1    1963      -0.000107861                -0.071052198                -0.002229831                -0.260315312
2    1964       0.000113888                 0.164048642                -0.002153369                -0.362159372
3    1965       0.000093530                 0.194625663                -0.004058998                -0.527787699
4    1966       0.000277319                 0.420584456                -0.003181363                -0.449952175
5    1967       0.000145477                 0.398951049                -0.001924280                -0.457790708
6    1968       0.000130008                 0.407034557                -0.001410437                -0.365958111
7    1969       0.000132313                 0.329578633                -0.002272297                -0.266319993
8    1970       0.000256653                 0.248228325


I      YR  COVAR(F(I+1),S(I))          CORR(F(I+1),S(I))           COVAR(F'(I),S'(I))          CORR(F'(I),S'(I))
1    1963      -0.000300121                -0.288098030                -0.000042236                -0.078513900
2    1964      -0.000190633                -0.413183861                -0.000105842                -0.113419513
3    1965      -0.000316565                -0.596183595                -0.000037700                -0.068248954
4    1966      -0.000254852                -0.517437047                -0.000062679                -0.090986152
5    1967      -0.000220346                -0.553958592                -0.000036215                -0.071550000
6    1968      -0.000194598                -0.553483970                -0.000055240                -0.087394643
7    1969      -0.000372666                -0.698801335                -0.000044936                -0.071363550
8    1970      -0.001618715                -0.887260327                -0.000045436                -0.039394739


I      YR  COVAR(S'(I),S(I))           CORR(S'(I),S(I))            COVAR(S'(I),S(I+1))         CORR(S'(I),S(I+1))
1    1963       0.001651412                 0.245122223                -0.001824454                -0.406359058
2    1964       0.001636182                 0.309903856                -0.001714925                -0.311848868
3    1965       0.002970106                 0.526383420                -0.003588038                -0.512085954
4    1966       0.002893728                 0.413544294                -0.002334868                -0.450695864
5    1967       0.001663408                 0.377116138                -0.001414800                -0.399447721
6    1968       0.001608233                 0.382845799                -0.001664660                -0.347049369
7    1969       0.001945246                 0.432637488                -0.001841692                -0.211269655
8    1970       0.009441458                 0.647830523


I      YR  COVAR(S'(I),F(I+1))         CORR(S'(I),F(I+1))
1    1963      -0.000245560                -0.449728519
2    1964      -0.000151819                -0.355785182
3    1965      -0.000279834                -0.578446973
4    1966      -0.000187041                -0.518292501
5    1967      -0.000162006                -0.483359803
6    1968      -0.000229674                -0.524885804
7    1969      -0.000302045                -0.554353714
8    1970      -0.001173591                -0.730147293


      THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.


          COVAR(S̄,F̄)   = -0.000033711                           CORR(S̄,F̄)   = -0.533301536

          COVAR(S̄',F̄')= -0.000005976                            CORR(S̄',F̄')= -0.063481008
```

Example 3.5h

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

MATRIX OF DATA VALUES -- ADULTS              H1
----------------------------------------------
        1      10.00  13.00   6.00   1.00   1.00   3.00   1.00   2.00   0.0
        2       0.0   58.00  21.00  16.00  15.00  13.00   6.00   1.00   1.00
        3       0.0    0.0   54.00  39.00  23.00  18.00  11.00  10.00   6.00
        4       0.0    0.0    0.0   44.00  21.00  22.00   9.00   9.00   3.00
        5       0.0    0.0    0.0    0.0   55.00  39.00  23.00  11.00  12.00
        6       0.0    0.0    0.0    0.0    0.0   66.00  46.00  29.00  18.00
        7       0.0    0.0    0.0    0.0    0.0    0.0  101.00  59.00  30.00
        8       0.0    0.0    0.0    0.0    0.0    0.0    0.0   97.00  22.00
        9       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   21.00


MATRIX OF EXPECTED VALUES -- ADULTS
----------------------------------------------
        1      10.00  11.45   5.04   3.60   2.42   1.92   1.33   0.86   0.37
        2       0.0   55.54  24.47  17.49  11.74   9.32   6.45   4.18   1.79
        3       0.0    0.0   52.22  37.32  25.05  19.89  13.77   8.93   3.83
        4       0.0    0.0    0.0   37.05  24.87  19.75  13.67   8.86   3.80
        5       0.0    0.0    0.0    0.0   49.07  38.97  26.97  17.49   7.50
        6       0.0    0.0    0.0    0.0    0.0   68.14  47.17  30.58  13.11
        7       0.0    0.0    0.0    0.0    0.0    0.0   98.64  63.95  27.41
        8       0.0    0.0    0.0    0.0    0.0    0.0    0.0   83.30  35.70
        9       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   21.00


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
----------------------------------------------
        1       0.00   0.47   0.43  -1.38  -0.92   0.78  -0.29   1.23  -0.61
        2       0.0    0.34  -0.72  -0.36   0.96   1.21  -0.18  -1.56  -0.59
        3       0.0    0.0    0.25   0.28  -0.42  -0.43  -0.75   0.36   1.11
        4       0.0    0.0    0.0    1.18  -0.79   0.52  -1.28   0.05  -0.41
        5       0.0    0.0    0.0    0.0    0.87   0.01  -0.78  -1.57   1.65
        6       0.0    0.0    0.0    0.0    0.0   -0.27  -0.17  -0.29   1.36
        7       0.0    0.0    0.0    0.0    0.0    0.0    0.25  -0.64   0.50
        8       0.0    0.0    0.0    0.0    0.0    0.0    0.0    1.57  -2.34
        9       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.00


MATRIX OF DATA VALUES -- YOUNG               H1
----------------------------------------------
        1      83.00  35.00  18.00  16.00   6.00   8.00   5.00   3.00   1.00
        2       0.0  103.00  21.00  13.00  11.00   8.00   6.00   6.00   0.0
        3       0.0    0.0   82.00  36.00  26.00  24.00  15.00  18.00   4.00
        4       0.0    0.0    0.0  153.00  39.00  22.00  21.00  16.00   8.00
        5       0.0    0.0    0.0    0.0  109.00  38.00  31.00  15.00   1.00
        6       0.0    0.0    0.0    0.0    0.0  113.00  64.00  29.00  22.00
        7       0.0    0.0    0.0    0.0    0.0    0.0  124.00  45.00  22.00
        8       0.0    0.0    0.0    0.0    0.0    0.0    0.0   95.00  25.00
        9       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   38.00


MATRIX OF EXPECTED VALUES -- YOUNG
----------------------------------------------
        1      83.00  39.01  17.19  12.28   8.25   6.55   4.53   2.94   1.26
        2       0.0  103.00  21.08  15.07  10.11   8.03   5.56   3.60   1.54
        3       0.0    0.0   82.00  42.19  28.33  22.49  15.57  10.09   4.33
        4       0.0    0.0    0.0  153.00  37.16  29.50  20.42  13.24   5.67
        5       0.0    0.0    0.0    0.0  109.00  36.43  25.22  16.35   7.01
        6       0.0    0.0    0.0    0.0    0.0  113.00  59.70  38.71  16.59
        7       0.0    0.0    0.0    0.0    0.0    0.0  124.00  46.90  20.10
        8       0.0    0.0    0.0    0.0    0.0    0.0    0.0   95.00  25.00
        9       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0   38.00


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
----------------------------------------------
        1       0.0   -0.66   0.20   1.07  -0.79   0.57   0.22   0.04  -0.23
        2       0.0    0.0   -0.02  -0.54   0.28  -0.01   0.19   1.27  -1.24
        3       0.0    0.0    0.0   -0.97  -0.44   0.32  -0.15   2.50  -0.16
        4       0.0    0.0    0.0    0.0    0.31  -1.40   0.13   0.76   0.98
        5       0.0    0.0    0.0    0.0    0.0    0.26   1.16  -0.34  -2.28
        6       0.0    0.0    0.0    0.0    0.0    0.0    0.57  -1.59   1.34
        7       0.0    0.0    0.0    0.0    0.0    0.0    0.0   -0.28   0.43
        8       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.00
        9       0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0
```

Example 3.5i

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO


THE HYPOTHESIS H2.   (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
-----------
              (2)  YOUNG BIRDS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES FROM THOSE OF ADULTS.

              (3)  IN ANY YEAR, THE REPORTING RATE FOR NEW RELEASES IS DIFFERENT FROM THAT FOR SURVIVORS OF PREVIOUSLY
                   BANDED COHORTS, AND HENCE THE CORRESPONDING RECOVERY RATES ARE DIFFERENT.

   H2 IS AN EXTENSION OF H1 IN THAT THE FIRST YEAR ADULT RECOVERY RATE IN YEAR I IS DIFFERENT FROM THE RECOVERY RATE IN

YEAR I OF PREVIOUSLY BANDED ADULTS. (THE SOLICITING OF BANDS FROM HUNTERS BY CONSERVATION OFFICERS NEAR BANDING SITES

MAY GIVE RISE TO THIS SITUATION).


PARAMETERS:
-----------
F'''(I)= BAND RECOVERY RATE IN YEAR I FOR ADULTS BANDED IN YEAR I.
F(I)   = BAND RECOVERY RATE IN YEAR I FOR SURVIVORS OF COHORTS BANDED BEFORE YEAR I.
S(I)   = SURVIVAL RATE FOR ADULTS IN YEAR I.
F'(I)  = BAND RECOVERY RATE FOR YOUNG IN YEAR I.
S'(I)  = SURVIVAL RATE FOR YOUNG IN YEAR I.


STRUCTURE OF THE MODEL UNDER H2  (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
-------------------------------

                            BANDED AS ADULTS

          N(1)F'''(1)  N(1)S(1)F(2)   N(1)S(1)S(2)F(3)  N(1)S(1)S(2)S(3)F(4)
                       N(2)F'''(2)    N(2)S(2)F(3)      N(2)S(2)S(3)F(4)
                                      N(3)F'''(3)       N(3)S(3)F(4)

                            BANDED AS YOUNG

          M(1)F'(1)    M(1)S'(1)F(2)  M(1)S'(1)S(2)F(3) M(1)S'(1)S(2)S(3)F(4)
                       M(2)F'(2)      M(2)S'(2)F(3)     M(2)S'(2)S(3)F(4
                                      M(3)F'(3)         M(3)S'(3)F(4)


ESTIMATES UNDER H2
------------------


                       F(I)                                                   S(I)
       -------------------------------------------------       ----------------------------------------------------
 I  YR  ESTIMATE        STANDARD           95% CONFIDENCE        ESTIMATE          STANDARD           95% CONFIDENCE
        --------        __ERROR__          ___INTERVAL___        --------          __ERROR__          ___INTERVAL___

 1  1963                                                         0.6126            0.1377         0.3426 -   0.8825

 2  1964  0.0760          0.0165         0.0437 -   0.1084        0.6314            0.0949         0.4454 -   0.8173

 3  1965  0.0558          0.0097         0.0367 -   0.0749        0.7427            0.1137         0.5199 -   0.9655

 4  1966  0.0521          0.0084         0.0356 -   0.0686        0.7900            0.1270         0.5411 -   1.0389

 5  1967  0.0457          0.0067         0.0326 -   0.0588        0.5858            0.0873         0.4146 -   0.7569

 6  1968  0.0660          0.0091         0.0483 -   0.0838        0.5833            0.0876         0.4116 -   0.7549

 7  1969  0.0753          0.0104         0.0549 -   0.0957        0.9766            0.2392         0.5077 -   1.4455

 8  1970  0.0464          0.0110         0.0248 -   0.0679
        AVERAGE         STANDARD           95% CONFIDENCE        AVERAGE           STANDARD           95% CONFIDENCE
        ESTIMATE        ERROR___           ___INTERVAL___        ESTIMATE          ERROR___           ___INTERVAL___

   F̄    = 0.0596          0.0040         0.0517 -   0.0675    S̄   = 0.7032             0.0371        0.6304 -   0.7760
```

Example 3.5j

```
                              F'(I)                                                                    S'(I)
         ------------------------------------------------------------------          ------------------------------------------------------------------
I    YR  ESTIMATE           STANDARD              95% CONFIDENCE                     ESTIMATE           STANDARD              95% CONFIDENCE
         --------           __ERROR__             ___INTERVAL___                     --------           __ERROR__             ___INTERVAL___

1  1963   0.0863             0.0091             0.0685 -   0.1040                     0.5012             0.0834             0.3378 -   0.6646

2  1964   0.1467             0.0134             0.1205 -   0.1729                     0.5197             0.0812             0.3607 -   0.6788

3  1965   0.0724             0.0077             0.0573 -   0.0875                     0.6675             0.1000             0.4714 -   0.8635

4  1966   0.1274             0.0096             0.1085 -   0.1463                     0.6428             0.0921             0.4624 -   0.8232

5  1967   0.0909             0.0083             0.0746 -   0.1072                     0.4607             0.0690             0.3254 -   0.5960

6  1968   0.0978             0.0087             0.0807 -   0.1150                     0.6726             0.0964             0.4836 -   0.8615

7  1969   0.1096             0.0093             0.0914 -   0.1278                     0.8125             0.2049             0.4108 -   1.2142

8  1970   0.1049             0.0102             0.0849 -   0.1248

9  1971   0.1076             0.0165             0.0753 -   0.1400

         AVERAGE            STANDARD             95% CONFIDENCE                      AVERAGE            STANDARD              95% CONFIDENCE
         ESTIMATE           ERROR___             ___INTERVAL___                      ESTIMATE           ERROR___              ___INTERVAL___

  F̄'    = 0.1049             0.0035              0.0979 -   0.1118          S̄'     = 0.6110             0.0424              0.5279 -   0.6941

                              F'''(I)                                                                  SK-1 FK
         ------------------------------------------------------------------          ------------------------------------------------------------------
I    YR  ESTIMATE           STANDARD              95% CONFIDENCE                     ESTIMATE           STANDARD              95% CONFIDENCE
         --------           __ERROR__             ___INTERVAL___                     --------           __ERROR__             ___INTERVAL___

1  1963   0.0433             0.0134             0.0170 -   0.0695                     0.0235             0.0086             0.0067 -   0.0402

2  1964   0.0894             0.0112             0.0674 -   0.1113

3  1965   0.0610             0.0080             0.0452 -   0.0768

4  1966   0.0746             0.0108             0.0534 -   0.0958

5  1967   0.0583             0.0076             0.0434 -   0.0733

6  1968   0.0613             0.0073             0.0470 -   0.0756

7  1969   0.0808             0.0077             0.0657 -   0.0959

8  1970   0.1034             0.0099             0.0839 -   0.1229

9  1971   0.0673             0.0142             0.0395 -   0.0951

         AVERAGE            STANDARD             95% CONFIDENCE
         ESTIMATE           ERROR___             ___INTERVAL___

  F̄'''  = 0.0710             0.0034              0.0643 -   0.0778
```

Example 3.5k

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H2

I        YR COVAR(F'''(I),F(I))          CORR(F'''(I),F(I))             COVAR(F'''(I),S(I))             CORR(F'''(I),S(I))
1      1963                                                                 -0.000114794                   -0.062251556
2      1964       -0.000010471              -0.056676081                    -0.000086939                   -0.081830862
3      1965       -0.000003847              -0.049178047                    -0.000051205                   -0.055976879
4      1966       -0.000006584              -0.072254022                    -0.000099861                   -0.072709400
5      1967       -0.000002827              -0.055506824                    -0.000036230                   -0.054358936
6      1968       -0.000003757              -0.056670512                    -0.000033189                   -0.051852051
7      1969       -0.000004868              -0.060706176                    -0.000063125                   -0.034231498
8      1970       -0.000005111              -0.046794813

I        YR COVAR(F'''(I+1),S(I))        CORR(F'''(I+1),S(I))           COVAR(F'''(I+1),S'(I))          CORR(F'''(I+1),S'(I))
1      1963        0.000084350               0.054695350                     0.000069015                    0.073926676
2      1964        0.000043530               0.057022187                     0.000035833                    0.054874386
3      1965        0.000093876               0.076346333                     0.000084367                    0.077975828
4      1966        0.000048864               0.050420309                     0.000039758                    0.056593216
5      1967        0.000033330               0.052220432                     0.000026214                    0.051956476
6      1968        0.000037704               0.055849972                     0.000043474                    0.058497243
7      1969        0.000107663               0.045265442                     0.000089578                    0.043964100

I        YR COVAR(F(I),S(I))             CORR(F(I),S(I))                COVAR(S(I),F(I+1))              CORR(S(I),F(I+1))
1      1963                                                                 -0.000566317                   -0.249235131
2      1964        0.000583705               0.372885923                    -0.000289453                   -0.313774667
3      1965        0.000340492               0.308022918                    -0.000538855                   -0.562594553
4      1966        0.000573209               0.535794469                    -0.000386581                   -0.456110834
5      1967        0.000286628               0.491740167                    -0.000379999                   -0.479641717
6      1968        0.000378391               0.476257935                    -0.000458447                   -0.503149047
7      1969        0.000767553               0.308389525                    -0.002009567                   -0.764648386

I        YR COVAR(F(I+1),S'(I))          CORR(F(I+1),S'(I))             COVAR(S(I),S(I+1))              CORR(S(I),S(I+1))
1      1963       -0.000463363              -0.336868308                    -0.004701938                   -0.359854127
2      1964       -0.000238274              -0.301956213                    -0.003852440                   -0.357154159
3      1965       -0.000484274              -0.574602458                    -0.008173466                   -0.566140566
4      1966       -0.000314540              -0.511951969                    -0.004953686                   -0.446678382
5      1967       -0.000298865              -0.477217070                    -0.003356640                   -0.438859686
6      1968       -0.000528613              -0.526998331                    -0.005944528                   -0.283719809
7      1969       -0.001671995              -0.742665819

I        YR COVAR(S(I),S'(I))            CORR(S(I),S'(I))               COVAR(S(I+1),S'(I))             CORR(S(I+1),S'(I))
1      1963        0.005476698               0.477006912                    -0.003847140                   -0.486381933
2      1964        0.003420540               0.444227717                    -0.003171275                   -0.343701632
3      1965        0.007543445               0.663253254                    -0.007345557                   -0.578223844
4      1966        0.006046373               0.517247460                    -0.004030544                   -0.501364979
5      1967        0.003109774               0.515808581                    -0.002639961                   -0.436641458
6      1968        0.004990235               0.590980631                    -0.006854348                   -0.297168051
7      1969        0.039337896               0.802354321

I        YR COVAR(F'(I),S'(I))           CORR(F'(I),S'(I))
1      1963       -0.000044950              -0.059560774
2      1964       -0.000108627              -0.100226310
3      1965       -0.000042712              -0.055417230
4      1966       -0.000068185              -0.076990774
5      1967       -0.000034931              -0.060945499
6      1968       -0.000056970              -0.067593315
7      1969       -0.000078765              -0.041369166

      THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.
```

Example 3.5l

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

MATRIX OF DATA VALUES -- ADULTS              H2
-------------------------------------------------
     1      10.00   13.00    6.00    1.00    1.00    3.00    1.00    2.00    0.0
     2       0.0    58.00   21.00   16.00   15.00   13.00    6.00    1.00    1.00
     3       0.0     0.0    54.00   39.00   23.00   18.00   11.00   10.00    6.00
     4       0.0     0.0     0.0    44.00   21.00   22.00    9.00    9.00    3.00
     5       0.0     0.0     0.0     0.0    55.00   39.00   23.00   11.00   12.00
     6       0.0     0.0     0.0     0.0     0.0    66.00   46.00   29.00   18.00
     7       0.0     0.0     0.0     0.0     0.0     0.0   101.00   59.00   30.00
     8       0.0     0.0     0.0     0.0     0.0     0.0     0.0    97.00   22.00
     9       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    21.00


MATRIX OF EXPECTED VALUES -- ADULTS
-------------------------------------------------
     1      10.00   10.89    5.09    3.58    2.51    2.14    1.44    0.90    0.46
     2       0.0    58.00   23.05   16.20   11.35    9.70    6.52    4.09    2.07
     3       0.0     0.0    54.00   34.71   24.32   20.78   13.97    8.77    4.44
     4       0.0     0.0     0.0    44.00   21.54   18.40   12.37    7.77    3.93
     5       0.0     0.0     0.0     0.0    55.00   36.83   24.76   15.54    7.86
     6       0.0     0.0     0.0     0.0     0.0    66.00   47.81   30.01   15.18
     7       0.0     0.0     0.0     0.0     0.0     0.0   101.00   59.10   29.90
     8       0.0     0.0     0.0     0.0     0.0     0.0     0.0    97.00   22.00
     9       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    21.00


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
-------------------------------------------------
     1       0.00    0.65    0.41   -1.37   -0.96    0.59   -0.37    1.16   -0.68
     2       0.0     0.00   -0.44   -0.05    1.09    1.07   -0.21   -1.53   -0.75
     3       0.0     0.0     0.00    0.74   -0.27   -0.62   -0.80    0.42    0.74
     4       0.0     0.0     0.0     0.00   -0.12    0.85   -0.97    0.45   -0.47
     5       0.0     0.0     0.0     0.0     0.00    0.36   -0.36   -1.16    1.48
     6       0.0     0.0     0.0     0.0     0.0     0.00   -0.27   -0.19    0.73
     7       0.0     0.0     0.0     0.0     0.0     0.0     0.00   -0.01    0.02
     8       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00    0.00
     9       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00



MATRIX OF DATA VALUES -- YOUNG               H2
-------------------------------------------------
     1      83.00   35.00   18.00   16.00    6.00    8.00    5.00    3.00    1.00
     2       0.0   103.00   21.00   13.00   11.00    8.00    6.00    6.00    0.0
     3       0.0     0.0    82.00   36.00   26.00   24.00   15.00   18.00    4.00
     4       0.0     0.0     0.0   153.00   39.00   22.00   21.00   16.00    8.00
     5       0.0     0.0     0.0     0.0   109.00   38.00   31.00   15.00    1.00
     6       0.0     0.0     0.0     0.0     0.0   113.00   64.00   29.00   22.00
     7       0.0     0.0     0.0     0.0     0.0     0.0   124.00   45.00   22.00
     8       0.0     0.0     0.0     0.0     0.0     0.0     0.0    95.00   25.00
     9       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    38.00


MATRIX OF EXPECTED VALUES -- YOUNG
-------------------------------------------------
     1      83.00   37.11   17.33   12.18    8.54    7.30    4.90    3.08    1.56
     2       0.0   103.00   20.53   14.43   10.11    8.64    5.81    3.65    1.84
     3       0.0     0.0    82.00   39.90   27.96   23.89   16.06   10.08    5.10
     4       0.0     0.0     0.0   153.00   35.67   30.48   20.49   12.86    6.51
     5       0.0     0.0     0.0     0.0   109.00   36.83   24.76   15.54    7.86
     6       0.0     0.0     0.0     0.0     0.0   113.00   59.11   37.11   18.78
     7       0.0     0.0     0.0     0.0     0.0     0.0   124.00   44.49   22.51
     8       0.0     0.0     0.0     0.0     0.0     0.0     0.0    95.00   25.00
     9       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0    38.00


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
-------------------------------------------------
     1       0.0    -0.35    0.16    1.10   -0.87    0.26    0.04   -0.04   -0.45
     2       0.0     0.0     0.11   -0.38    0.28   -0.22    0.08    1.24   -1.36
     3       0.0     0.0     0.0    -0.63   -0.38    0.02   -0.27    2.50   -0.49
     4       0.0     0.0     0.0     0.0     0.57   -1.56    0.11    0.88    0.59
     5       0.0     0.0     0.0     0.0     0.0     0.20    1.27   -0.14   -2.46
     6       0.0     0.0     0.0     0.0     0.0     0.0     0.65   -1.35    0.75
     7       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.08   -0.11
     8       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.00
     9       0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0     0.0
```

Example 3.5m

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO


THE HYPOTHESIS H3.    (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
-----------
              (2)  SURVIVAL AND RECOVERY RATES ARE AGE-DEPENDENT FOR THE FIRST TWO YEARS OF LIFE. (THIS EMBRACES
                   ASSUMPTION (3) OF H2 FOR THE TYPE OF DATA BEING ANALYSED.)

  H3 ASSUMES THE PARAMETERS S AND F ARE AGE-SPECIFIC FOR THREE AGE CLASSES (NAMELY, YOUNG, SUBADULT, AND ADULT,) BUT

ONLY TWO AGE CLASSES ARE RECOGNIZED DURING BANDING, BECAUSE SUBADULTS AND ADULTS ARE USUALLY INDISTINGUISHABLE. THUS

THE ASSUMPTIONS OF H3 GIVE RISE TO A MODEL FOR WHICH MOST OF THE PARAMETERS OF INTEREST ARE NOT ESTIMABLE, HENCE NO ESTIMATES

ARE COMPUTED HERE.  HOWEVER THE ASSUMPTIONS OF H3 ARE EXAMINED BELOW BY MEANS OF A GOODNESS OF FIT TEST AND A TEST OF H2

AGAINST THE ALTERNATIVE H3


THE HYPOTHESIS HO.
------------------

ASSUMPTIONS:  SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC BUT AGE-INDEPENDENT.
-----------

THE AGE-INDEPENDENCE ASSUMPTION OF HO IS USUALLY INAPPROPRIATE FOR BIRDS BANDED AS YOUNG AND ESTIMATION UNDER HO IS OMITTED THOUGH

TESTS RELATED TO THIS HYPOTHESIS ARE COMPUTED BELOW. IF HO IS NOT REJECTED, THE DATA SHOULD BE POOLED AND ANALYZED USING THE

MODELS THAT ASSUME PARAMETERS ARE AGE-INDEPENDENT.
```

Example 3.5n

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO


TESTS TO DESCRIMINATE BETWEEN THE MODELS UNDER H0, H1, H2, H3.  (SEE BROWNIE AND ROBSON, 1974, CORNELL BIOMETRICS
--------------------------------------------------------------   UNIT PAPER NO. BU-514-M)


H0, H1, H2 AND H3 REPRESENT A SERIES OF HYPOTHESES WITH PROGRESSIVELY MORE GENERAL ASSUMPTIONS ABOUT THE POPULATION PARA-

METERS.  TO DETERMINE WHICH ASSUMPTIONS ARE APPROPRIATE FOR A GIVEN DATA SET, A SERIES OF TESTS ARE CARRIED OUT, WHERE

EACH TEST IN THE SERIES COMPARES A GIVEN HYPOTHESIS AGAINST A MORE GENERAL ALTERNATIVE.  (THUS THE FIRST TEST IN THE

SERIES COMPARES H0 AGAINST THE MORE GENERAL H1).  IN EACH CASE, IF THE TEST RESULTS IN A 'LARGE' CHI-SQUARE VALUE,

THEN THE RESTRICTIVE HYPOTHESIS SHOULD BE REJECTED IN FAVOR OF THE MORE GENERAL ONE, WHILE A 'SMALL' CHI-SQUARE

VALUE SUGGESTS THERE IS NO REASON TO DISCARD THE SIMPLER HYPOTHESIS.



CHI-SQUARE TEST OF H0 VS H1
---------------------------


    2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC     2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                              WITH 1 DEGREE OF FREEDOM                                         WITH 1 DEGREE OF FREEDOM
    -----------------------   ----------------------------------     -----------------------   ----------------------------------

    R(I.)   N(I)-R(I.)                                               W(I)    Z(I+1)

    Q(I.)   M(I)-Q(I.)                                               Q(I,I)  Q(I.)-Q(I,I)

I= 1        37     194                                                     10      27
           175     787                   0.602                             83      92                            5.163

I= 2       131     518                                                    106     144
           168     534                   2.747                            103      65                           14.371

I= 3       161     724                                                    120     250
           205     927                   0.002                             82     123                            3.315

I= 4       108     482                                                    165     316
           259     942                   2.581                            153     106                           42.149

I= 5       140     803                                                    197     365
           194    1005                   0.714                            109      85                           26.731

I= 6       159     918                                                    261     348
           228     927                   9.632                            113     115                            3.017

I= 7       190    1060                                                    339     314
           191     940                   1.258                            124      67                           10.096

I= 8       119     819                                                    350     150
           120     786                   0.127                             95      25                            4.014

I= 9        21     291
            38     315                   3.334
|-------------------------------------------------|
|TOTAL CHI-SQUARE WITH 17 DEGREES OF FREEDOM = 129.853|
|-------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  129.85  = 0.0


THIS TEST OF THE HYPOTHESIS H0 AGAINST THE HYPOTHESIS H1 TESTS THE ASSUMPTION THAT YOUNG AND ADULTS HAVE THE
SAME SURVIVAL AND RECOVERY RATES.
```

Example 3.5o

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

CHI-SQUARE TEST OF H1 VS H2
---------------------------


     2 X 2 CONTINGENCY TABLE     CORRESPONDING CHI-SQUARE STATISTIC
                                 WITH 1 DEGREE OF FREEDOM
     -----------------------     ----------------------------------

I= 2          58      73
              48      71                       0.396

I= 3          54     107
              66     143                       0.160

I= 4          44      64
             121     252                       2.561

I= 5          55      85
             142     280                       1.467

I= 6          66      93
             195     255                       0.160

I= 7         101      89
             238     225                       0.166

I= 8          97      22
             253     128                       9.856
|------------------------------------------------------|
|TOTAL CHI-SQUARE WITH  7 DEGREES OF FREEDOM =   14.766|
|------------------------------------------------------|
PROBABILITY OF A CHI-SCUARE VALUE LARGER THAN    14.77  = 0.03912


THIS TEST OF THE HYPOTHESIS H1 AGAINST THE HYPOTHESIS H2 TESTS THE ASSUMPTION THAT RECOVERY RATES FOR NEWLY
RELEASED ADULTS ARE THE SAME AS FOR SURVIVORS OF PREVIOUSLY BANDED COHORTS.


CHI-SQUARE TEST OF H2 VS H3
---------------------------


     2 X 2 CONTINGENCY TABLE     CORRESPONDING CHI-SQUARE STATISTIC
                                 WITH 1 DEGREE OF FREEDOM
     -----------------------     ----------------------------------

I= 1          35      57
              13      14                       0.886

I= 2          21      44
              45      99                       0.023

I= 3          36      87
              85     165                       0.842

I= 4          39      67
             103     213                       0.626

I= 5          38      47
             157     208                       0.080

I= 6          64      51
             174     174                       1.105

I= 7          45      22
             208     106                       0.021
|------------------------------------------------------|
|TOTAL CHI-SQUARE WITH  7 DEGREES OF FREEDOM =    3.584|
|------------------------------------------------------|
PROBABILITY OF A CHI-SCUARE VALUE LARGER THAN     3.58  = 0.82622


REJECTION OF H2 IN FAVOR OF H3 WOULD INDICATE THAT SURVIVAL AND RECOVERY RATES ARE AGE-DEPENDENT FOR AT LEAST
THE FIRST TWO YEARS.



THE HYPOTHESES H01 AND H02 ARE MORE RESTRICTIVE THAN H1 BUT ARE NOT COMPARABLE IN THIS SENSE WITH H0, THUS H01
AND H02 DO NOT FIT INTO THE ABOVE SERIES.  IN PRACTICE THE MODELS UNDER H02 AND H1 ARE LIKELY TO BE OF MOST USE, SO A
LIKELIHOOD RATIO TEST TO DISTINGUISH BETWEEN THESE MODELS IS COMPUTED BELOW.  THIS TESTS THE ASSUMPTION THAT YOUNG AND
ADULT SURVIVAL RATES ARE CONSTANT FROM YEAR TO YEAR.

LIKELIHOOD RATIO TEST OF H02 VS H1.
-----------------------------------

CHI-SQUARE VALUE                                       =    26.94
DEGREES OF FREEDOM                                     =       14
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   26.94  = 0.01962
```

Example 3.5p

```
MALE MALLARDS BANDED PRESEASON IN THE SAN LUIS VALLEY, COLORADO

CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H3
-------------------------------------------------------

                CONTINGENCY TABLES                                        CORRESPONDING CHI-SQUARE
                                                                          STATISTICS AND DEGREES OF FREEDOM
                ------------------                                        ---------------------------------

        I= 2
                 23. 13. 16. 21.
                 15.  8. 16. 18.
                  4.  3.  1.  6.                                             3.74 WITH  6 D.F.

        I= 3
                 16. 11. 18. 23. 39.
                  6.  6.  8. 11. 13.
                  8. 12. 24. 22. 33.                                         4.62 WITH  8 D.F.

        I= 4
                  3.  9.  9. 22. 21.
                  4. 18. 15. 24. 26.
                  8. 22. 29. 50. 56.                                         3.31 WITH  8 D.F.

        I= 5
                 12. 11. 23. 39.
                  8. 16. 21. 22.
                 15. 49. 53. 96.                                             9.76 WITH  6 D.F.

        I= 6
                 18. 29. 46.
                  1. 15. 31.
                 35. 76. 97.                                                10.05 WITH  4 D.F.

        I= 7
                 30. 59.
                 22. 29.
                 54.120.                                                     2.58 WITH  2 D.F.

                                |--------------------------------------|
                                |TOTAL CHI-SQUARE     34.07 WITH 34 D.F.|
                                |--------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   34.07  = 0.46433



CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H0
-------------------------------------------------------

                                |--------------------------------------|
                                |TOTAL CHI-SQUARE    182.27 WITH 65 D.F.|
                                |--------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  182.27  = 0.0


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H1
-------------------------------------------------------

                                |--------------------------------------|
                                |TOTAL CHI-SQUARE     52.42 WITH 48 D.F.|
                                |--------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   52.42  = 0.30656


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H2
-------------------------------------------------------

                                |--------------------------------------|
                                |TOTAL CHI-SQUARE     37.66 WITH 41 D.F.|
                                |--------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   37.66  = 0.62009


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H3
-------------------------------------------------------

                                |--------------------------------------|
                                |TOTAL CHI-SQUARE     34.07 WITH 34 D.F.|
                                |--------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   34.07  = 0.46433

  FROM THE MODELS ABOVE, ONE SHOULD CHOOSE THE SIMPLEST MODEL (FEWEST PARAMETERS) THAT ADEQUATELY DESCRIBES THE DATA.

ADEQUACY MAY BE JUDGED BY EXAMINING THE RESULTS OF (1) THE GOODNESS OF FIT TESTS, AND (2) THE TESTS BETWEEN SPECIFIC

MODELS. FREQUENTLY, H02 OR H1 IS ADEQUATE.
```

Example 3.6a

```
                      YOUNG AND ADULT MALE MALLARDS BANDED PRESEASON IN SOUTHERN ONTARIO, 1965

ADULTS         INPUT MATRIX


1965    604.   43.    20.    10.    12.     3.     5.     2.     3.     0.
1966    333.    0.    36.     7.     9.     6.     4.     2.     2.     0.
1967    703.    0.     0.    40.    24.    19.     7.     5.     3.     1.
1968    497.    0.     0.     0.    34.    13.     7.     4.     7.     1.
1969    536.    0.     0.     0.     0.    46.    13.     7.     9.     1.
1970    828.    0.     0.     0.     0.     0.    56.    28.    25.     6.
1971   1353.    0.     0.     0.     0.     0.     0.    79.    42.    17.
1972   1080.    0.     0.     0.     0.     0.     0.     0.    65.    34.

YOUNG          INPUT 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.



BASIC SUBTOTALS

 I  RROW(I)  RCOL(I)  QROW(I)  QCOL(I)       T(I)      U(I)      W(I)      Z(I)
 1    98.00    43.00   257.00   132.00      98.00    257.00     43.00       0.0
 2    66.00    56.00   249.00   223.00     121.00    374.00    104.00     55.00
 3    99.00    57.00   264.00   231.00     164.00    415.00    123.00    142.00
 4    66.00    79.00   298.00   253.00     173.00    482.00    139.00    192.00
 5    76.00    87.00   299.00   286.00     170.00    528.00    180.00    218.00
 6   115.00    92.00   568.00   465.00     198.00    810.00    190.00    219.00
 7   138.00   127.00   638.00   582.00     244.00    983.00    317.00    250.00
 8    99.00   156.00   443.00   618.00     216.00    844.00    432.00    272.00
                                             60.00    226.00              185.00




THE HYPOTHESIS H2.   (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
------------
              (2)  YOUNG BIRDS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES FROM THOSE OF ADULTS.

              (3)  IN ANY YEAR, THE REPORTING RATE FOR NEW RELEASES IS DIFFERENT FROM THAT FOR SURVIVORS OF PREVIOUSLY
                   BANDED COHORTS, AND HENCE THE CORRESPONDING RECOVERY RATES ARE DIFFERENT.

   H2 IS AN EXTENSION OF H1 IN THAT THE FIRST YEAR ADULT RECOVERY RATE IN YEAR I IS DIFFERENT FROM THE RECOVERY RATE IN

YEAR I OF PREVIOUSLY BANDED ADULTS. (THE SOLICITING OF BANDS FROM HUNTERS BY CONSERVATION OFFICERS NEAR BANDING SITES

MAY GIVE RISE TO THIS SITUATION).


PARAMETERS:
-----------
F'''(I)= BAND RECOVERY RATE IN YEAR I FOR ADULTS BANDED IN YEAR I.
F(I)    = BAND RECOVERY RATE IN YEAR I FOR SURVIVORS OF COHORTS BANDED BEFORE YEAR I.
S(I)    = SURVIVAL RATE FOR ADULTS IN YEAR I.
F'(I)   = BAND RECOVERY RATE FOR YOUNG IN YEAR I.
S'(I)   = SURVIVAL RATE FOR YOUNG IN YEAR I.


STRUCTURE OF THE MODEL UNDER H2  (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
-------------------------------

                              BANDED AS ADULTS

        N(1)F'''(1)  N(1)S(1)F(2)     N(1)S(1)S(2)F(3)    N(1)S(1)S(2)S(3)F(4)
                     N(2)F'''(2)      N(2)S(2)F(3)        N(2)S(2)S(3)F(4)
                                      N(3)F'''(3)         N(3)S(3)F(4)

                              BANDED AS YOUNG

        M(1)F'(1)    M(1)S'(1)F(2)    M(1)S'(1)S(2)F(3)   M(1)S'(1)S(2)S(3)F(4)
                     M(2)F'(2)        M(2)S'(2)F(3)       M(2)S'(2)S(3)F(4
                                      M(3)F'(3)           M(3)S'(3)F(4)
```

Example 3.6b

```
ESTIMATES UNDER H2
------------------

                              F(I)                                                        S(I)
           -------------------------------------------------------   ------------------------------------------------------------
 I   YR    ESTIMATE          STANDARD          95% CONFIDENCE        ESTIMATE          STANDARD            95% CONFIDENCE
           --------          __ERROR__         ___INTERVAL___        --------          __ERROR__           ___INTERVAL___

 1  1965                                                             0.6105            0.1368              0.3423 -   0.8786
 2  1966   0.0547            0.0127            0.0298 -   0.0796     0.6509            0.1437              0.3692 -   0.9325
 3  1967   0.0524            0.0098            0.0331 -   0.0716     0.8095            0.1749              0.4666 -   1.1524
 4  1968   0.0363            0.0076            0.0214 -   0.0513     0.6526            0.1637              0.3318 -   0.9734
 5  1969   0.0397            0.0083            0.0233 -   0.0560     0.4545            0.1010              0.2566 -   0.6524
 6  1970   0.0500            0.0084            0.0335 -   0.0665     0.7594            0.1409              0.4832 -   1.0356
 7  1971   0.0487            0.0077            0.0336 -   0.0639     0.3926            0.0873              0.2216 -   0.5636
 8  1972   0.0765            0.0149            0.0473 -   0.1057
           AVERAGE           STANDARD          95% CONFIDENCE        AVERAGE           STANDARD            95% CONFIDENCE
           ESTIMATE          ERROR___          ___INTERVAL___        ESTIMATE          ERROR___            ___INTERVAL___
   _                                                           _
   F    =  0.0512            0.0039             0.0436 -   0.0588  S   = 0.6186            0.0221             0.5753 -   0.6618

                              F'(I)                                                       S'(I)
           -------------------------------------------------------   ------------------------------------------------------------
 I   YR    ESTIMATE          STANDARD          95% CONFIDENCE        ESTIMATE          STANDARD            95% CONFIDENCE
           --------          __ERROR__         ___INTERVAL___        --------          __ERROR__           ___INTERVAL___

 1  1965   0.0841            0.0070            0.0703 -   0.0978     0.5338            0.1082              0.3218 -   0.7458
 2  1966   0.1197            0.0085            0.1031 -   0.1363     0.3657            0.0646              0.2390 -   0.4924
 3  1967   0.1024            0.0076            0.0876 -   0.1172     0.5928            0.1195              0.3585 -   0.8270
 4  1968   0.1114            0.0076            0.0966 -   0.1262     0.6141            0.1268              0.3657 -   0.8626
 5  1969   0.1044            0.0071            0.0905 -   0.1184     0.4658            0.0762              0.3163 -   0.6152
 6  1970   0.1050            0.0052            0.0948 -   0.1151     0.6128            0.0937              0.4291 -   0.7964
 7  1971   0.0873            0.0042            0.0791 -   0.0956     0.4935            0.0950              0.3074 -   0.6796
 8  1972   0.0954            0.0049            0.0858 -   0.1050
           AVERAGE           STANDARD          95% CONFIDENCE        AVERAGE           STANDARD            95% CONFIDENCE
           ESTIMATE          ERROR___          ___INTERVAL___        ESTIMATE          ERROR___            ___INTERVAL___
   _                                                           _
   F'   =  0.1012            0.0024             0.0966 -   0.1058  S'  = 0.5255            0.0378             0.4515 -   0.5995

                              F'''(I)                                                     SK...SK+I-1FK+I
           -------------------------------------------------------   ------------------------------------------------------------
 I   YR    ESTIMATE          STANDARD          95% CONFIDENCE        ESTIMATE          STANDARD            95% CONFIDENCE
           --------          __ERROR__         ___INTERVAL___        --------          __ERROR__           ___INTERVAL___

 1  1965   0.0712            0.0105            0.0507 -   0.0917     0.0315            0.0053              0.0211 -   0.0419
 2  1966   0.1081            0.0170            0.0748 -   0.1415
 3  1967   0.0569            0.0087            0.0398 -   0.0740
 4  1968   0.0684            0.0113            0.0462 -   0.0906
 5  1969   0.0858            0.0121            0.0621 -   0.1095
 6  1970   0.0676            0.0087            0.0505 -   0.0847
 7  1971   0.0584            0.0064            0.0459 -   0.0709
 8  1972   0.0602            0.0072            0.0460 -   0.0744
           AVERAGE           STANDARD          95% CONFIDENCE
           ESTIMATE          ERROR___          ___INTERVAL___
   _
   F''' =  0.0721            0.0038             0.0647 -   0.0795
```

Example 3.6c

```
YOUNG AND ADULT MALE MALLARDS BANDED PRESEASON IN SOUTHERN ONTARIO, 1965

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H2


I       YR   COVAR(F'''(I),F(I))      CORR(F'''(I),F(I))      COVAR(F'''(I),S(I))      CORR(F'''(I),S(I))
1     1965                                                       -0.000071953            -0.050263757
2     1966       -0.000017757           -0.082128853             -0.000211305            -0.086413352
3     1967       -0.000004239           -0.049451699             -0.000065521            -0.042866313
4     1968       -0.000005003           -0.057858104             -0.000089832            -0.048466721
5     1969       -0.000006354           -0.062933575             -0.000072768            -0.055563022
6     1970       -0.000004083           -0.055501886             -0.000062032            -0.050443814
7     1971       -0.000002103           -0.042725413             -0.000016943            -0.030462790
8     1972       -0.000004264           -0.039589086


I       YR   COVAR(F'''(I+1),S(I))    CORR(F'''(I+1),S(I))    COVAR(F'''(I+1),S'(I))   CORR(F'''(I+1),S'(I))
1     1965        0.000198185            0.085127837              0.000173282             0.094147318
2     1966        0.000052680            0.041959024              0.000029597             0.052408655
3     1967        0.000111429            0.056245979              0.000081591             0.060283337
4     1968        0.000104494            0.052768064              0.000098331             0.064112899
5     1969        0.000037122            0.042125834              0.000038044             0.057182092
6     1970        0.000032774            0.036485280              0.000026444             0.044278074
7     1971        0.000021879            0.034649850              0.000027501             0.040021480


I       YR   COVAR(F(I),S(I))         CORR(F(I),S(I))         COVAR(S(I),F(I+1))       CORR(S(I),F(I+1))
1     1965                                                       -0.001012746            -0.582562295
2     1966        0.001079793            0.591359720             -0.000529294            -0.375399742
3     1967        0.000658314            0.383517011             -0.000860296            -0.643962160
4     1968        0.000693555            0.554897300             -0.000814945            -0.596652227
5     1969        0.000567510            0.673483256             -0.000357640            -0.420106698
6     1970        0.000597624            0.503059209             -0.000599829            -0.551351409
7     1971        0.000310095            0.460341645             -0.000855720            -0.658995815


I       YR   COVAR(F(I+1),S'(I))      CORR(F(I+1),S'(I))      COVAR(S(I),S(I+1))       CORR(S(I),S(I+1))
1     1965       -0.000885492           -0.644285909             -0.012051146            -0.612953185
2     1966       -0.000297375           -0.468890496             -0.008180991            -0.325408456
3     1967       -0.000629928           -0.690186048             -0.015446965            -0.539436014
4     1968       -0.000766877           -0.724929336             -0.009333376            -0.564697195
5     1969       -0.000366516           -0.570257835             -0.005433068            -0.381821171
6     1970       -0.000483983           -0.669112868             -0.004833199            -0.393107900
7     1971       -0.001075621           -0.761157018


I       YR   COVAR(S(I),S'(I))        CORR(S(I),S'(I))        COVAR(S(I+1),S'(I))      CORR(S(I+1),S'(I))
1     1965        0.010981679            0.742078948             -0.010536898            -0.677896799
2     1966        0.004383184            0.471876663             -0.004596349            -0.406449248
3     1967        0.014960725            0.715472751             -0.011310611            -0.578156926
4     1968        0.013491966            0.650232249             -0.008782856            -0.686103635
5     1969        0.003788974            0.492179746             -0.005567908            -0.518288443
6     1970        0.008696396            0.658718239             -0.003899758            -0.477070797
7     1971        0.006428260            0.775924154


I       YR   COVAR(F'(I),S'(I))       CORR(F'(I),S'(I))
1     1965       -0.000028583           -0.037732383
2     1966       -0.000029940           -0.054558622
3     1967       -0.000037685           -0.041738450
4     1968       -0.000039466           -0.041197019
5     1969       -0.000026321           -0.048530931
6     1970       -0.000018400           -0.037883746
7     1971       -0.000009604           -0.024000377


      THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.
```

Example 3.6d

```
YOUNG AND ADULT MALE MALLARDS BANDEC PRESEASON IN SOUTHERN ONTARIO, 1965

MATRIX OF CATA VALUES -- ADULTS
----------------------------------------
      1     43.00  20.00  10.00  12.00   3.CC   5.00   2.00   3.0C   0.0
      2      0.0   36.00   7.00   9.00   6.00   4.00   2.00   2.0C   0.0
      3      0.0    C.0   40.00  24.00  19.00   7.00   5.00   3.00   1.00
      4      0.0    0.0    0.0   34.00  13.0C   7.00   4.00   7.00   1.00
      5      0.0    0.0    0.0    0.0   46.00  13.00   7.00   9.00   1.00
      6      0.0    0.0    0.0    0.0    0.0   56.00  28.00  25.00   6.0C
      7      0.0    0.0    0.0    0.0    C.0    0.0   79.00  42.0C  17.00
      8      0.0    0.0    C.0    0.0    C.0    0.0    0.0   65.00  34.0C


MATRIX OF EXPECTED VALUES -- ADULTS
----------------------------------------
      1     43.00  20.78  13.15   7.60   5.59   3.25   2.44   1.55   0.64
      2      0.0   36.00  11.53   6.67   4.9C   2.85   2.14   1.36   0.56
      3      0.0    0.0   40.00  21.29  15.64   9.10   6.84   4.34   1.79
      4      0.0    0.0    0.0   34.00  13.28   7.72   5.81   3.68   1.51
      5      0.0    0.0    0.0    0.0   46.00  12.37   9.30   5.9C   2.43
      6      0.0    0.0    0.0    0.0    0.0   56.00  31.14  19.74   8.12
      7      0.0    C.0    0.0    0.0    0.0    0.0   79.00  41.8C  17.20
      8      0.0    0.0    0.0    0.0    C.0    0.0    0.0   65.00  34.00


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
----------------------------------------
      1      0.00  -0.17  -0.88   1.60  -1.1C   0.97  -0.28   1.17  -0.80
      2      0.0    0.00  -1.36   0.91   C.5C   0.69  -0.1C   0.55  -0.75
      3      0.0    C.0    0.00   0.60   C.86  -0.70  -0.71  -0.64  -0.59
      4      0.0    0.0    0.0    0.00  -C.08  -0.26  -0.75   1.74  -0.42
      5      0.0    0.0    0.0    0.0    C.0C   0.18  -0.76   1.28  -0.92
      6      0.0    0.0    0.0    0.0    C.0    0.00  -0.57   1.2C  -0.75
      7      0.0    0.0    0.0    0.0    0.0    0.0    0.00   0.03  -0.05
      8      0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0C   0.0C


MATRIX OF CATA VALUES -- YOUNG
----------------------------------------
      1    132.00  48.00  33.00  13.00   8.00   9.00   8.00   5.0C   1.00
      2      0.0  175.00  33.00   8.00  11.0C  10.00   4.00   7.00   1.00
      3      0.0    C.0  165.00  39.00  23.00  12.00  13.00   5.00   7.00
      4      0.0    C.0    0.0  193.00  51.0C  24.00  13.00  12.00   5.0C
      5      0.0    0.0    0.0    0.0  193.00  43.00  39.00  15.00   9.0C
      6      0.0    0.0    0.0    0.0    C.0  367.00 113.00  56.0C  32.00
      7      0.0    0.0    0.0    0.0    C.0    0.0  392.00 176.0C  70.0C
      8      0.0    C.0    0.0    0.0    C.0    0.0    0.0  342.0C 101.00


MATRIX OF EXPECTED VALUES -- YCUNG
----------------------------------------
      1    132.00  47.22  29.89  17.28  12.7C   7.38   5.55   3.52   1.45
      2      0.0  175.00  28.44  16.44  12.08   7.03   5.29   3.35   1.38
      3      0.0    0.0  165.00  35.72  26.25  15.27  11.48   7.28   3.00
      4      0.0    C.0    0.0  193.00  43.56  25.33  19.05  12.08   4.97
      5      0.0    0.0    0.0    0.0  193.00  43.70  32.87  20.84   8.58
      6      0.0    0.0    0.0    0.0    0.0  367.00 106.07  67.26  27.67
      7      0.0    C.0    C.0    0.0    C.0    0.0  392.00 174.29  71.71
      8      0.0    C.0    0.0    0.0    C.0    0.0    0.0  342.0C 101.00


MATRIX OF STANCARD NORMAL DEVIATES -- YOUNG
----------------------------------------
      1      0.0    C.11   0.57  -1.04  -1.32   0.60   1.04   0.79  -0.37
      2      0.0    0.0    0.86  -2.09  -C.31   1.12  -0.56   2.0C  -0.32
      3      0.0    0.0    0.0    0.55  -C.64  -0.84   0.45  -0.85   2.32
      4      0.0    C.0    0.0    0.0    1.14  -0.27  -1.39  -0.02   0.01
      5      0.0    C.0    0.0    0.0    0.0   -0.11   1.08  -1.29   0.14
      6      0.0    0.0    0.0    0.0    0.0    0.0    0.68  -1.39   0.83
      7      0.0    0.0    0.0    0.0    C.0    0.0    0.0    0.13  -0.2C
      8      0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0    0.0C
```

Example 3.6e

```
YOUNG AND ADULT MALE MALLARDS BANDED PRESEASON IN SOUTHERN ONTARIO, 1965   34350

CHI-SQUARE TEST OF H1 VS H2
---------------------------


     2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                               WITH 1 DEGREE OF FREEDOM
     -----------------------   ----------------------------------

I= 2         36     30
             68    112                       5.564

I= 3         40     59
             83    133                       0.112

I= 4         34     32
            105    186                       5.389

I= 5         46     30
            134    189                       9.008

I= 6         56     59
            134    191                       1.929

I= 7         79     59
            238    213                       0.851

I= 8         65     34
            367    151                       1.068
I------------------------------------------------------I
ITOTAL CHI-SQUARE WITH  7 DEGREES OF FREEDOM =  23.920I
I------------------------------------------------------I
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   23.92  = 0.00118


THIS TEST OF THE HYPOTHESIS H1 AGAINST THE HYPOTHESIS H2 TESTS THE ASSUMPTION THAT RECOVERY RATES FOR NEWLY
RELEASED ADULTS ARE THE SAME AS FOR SURVIVORS OF PREVIOUSLY BANDED COHORTS.


CHI-SQUARE TEST OF H2 VS H3
---------------------------


I------------------------------------------------------I
ITOTAL CHI-SQUARE WITH  7 DEGREES OF FREEDOM =   7.676I
I------------------------------------------------------I
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN    7.68  = 0.36205




CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H0
-----------------------------------------------------

                                    I------------------------------------I
                                    ITOTAL CHI-SQUARE    305.35 WITH 66 D.F.I
                                    I------------------------------------I

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  305.35  = 0.0


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H1
-----------------------------------------------------

                                    I------------------------------------I
                                    ITOTAL CHI-SQUARE     68.67 WITH 50 D.F.I
                                    I------------------------------------I

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   68.67  = 0.04094


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H2
-----------------------------------------------------

                                    I------------------------------------I
                                    ITOTAL CHI-SQUARE     44.75 WITH 43 D.F.I
                                    I------------------------------------I

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   44.75  = 0.39820


  FROM THE MODELS ABOVE, ONE SHOULD CHOOSE THE SIMPLEST MODEL (FEWEST PARAMETERS) THAT ADEQUATELY DESCRIBES THE DATA.

ADEQUACY MAY BE JUDGED BY EXAMINING THE RESULTS OF (1) THE GOODNESS OF FIT TESTS, AND (2) THE TESTS BETWEEN SPECIFIC

MODELS. FREQUENTLY, HO2 OR H1 IS ADEQUATE.
```

An Example

Example 3.6 serves as an illustration of a situation when the model under H_2 seems appropriate. The data are from a study on male mallards banded preseason in Southern Ontario 1965-72.

Example 3.6e shows the goodness of fit tests result in rejection of the models under both H_0 and H_1. Also, the test of H_1 vs. H_2, with a chi-square value of 23.92 and 7 df, is highly significant, with several of the individual chi-square components significantly large. Thus the H_1 model must be rejected in favor of H_2. Comparison of H_2 with H_3 does not suggest preference for H_3 so H_2 seems appropriate for these data.

Examination of the estimates of recovery rates of adults under the H_2 model indicates that first-year recovery rates (f_i''') are significantly higher than adult recovery rates in latter years (f_i). This is probably the result of band solicitation by conservation agencies in Ontario.

3.8 Summary of the Models Proposed

The models of this chapter have proved to be useful in the analysis of data from banding studies involving both young and adults of many waterfowl species. Some general conclusions can be made on the basis of these results.

The H_0 and H_{01} models are too restrictive, i.e., young and adults do have different survival and recovery rates, and recovery rates differ from year to year.

Frequently either the H_{02} or the H_1 model seems to be adequate (Examples 3.2 and 3.5). The H_1 model has the practical advantage in that estimates are easily computed with a desk calculator for small data sets.

Occasionally, as in Example 3.6, there appears to be a real difference in the recovery rate between newly released adults and that for survivors of previous releases. In this case the model under H_2 should be used although for values of N_i and M_i usually encountered in practice, the H_2 estimators will have poor precision.

In-season banding may give rise to a situation where H_2 appears to be appropriate, but in fact is not. Sometimes, however, H_2 may be a reasonable approximation when in-season banding takes place (see Example 3.4).

The H_3 model appears to be unnecessarily general for data sets we have analyzed and subadults do not have very different survival and recovery rates from those of older birds.

We conclude this chapter with a discussion which demonstrates the futility of banding only young when we have reason to believe that survival and recovery rates are different for young and adults.

3.9 Why Survival Rates Cannot be Estimated if Only Young are Banded

For many species the banding of young is much easier to accomplish than the banding of adults. For this reason there is strong temptation to concentrate the banding effort exclusively upon young; this is a pointless practice. Analysis of numerous data sets for birds has shown that the survival rate of young is typically lower than adult survival rates $(S_i' < S_i)$. Also first-year band recovery rates for young are typically higher than for adults $(f_i' > f_i)$. Consequently, an appropriate model for band recoveries from young is the H_1 model. Specifically its structure is shown below for $k = \ell = 4$.

Number banded	Year of recovery 1	2	3	4
M_1	M_1f_1'	$M_1S_1'f_2$	$M_1S_1'S_2f_3$	$M_1S_1'S_2S_3f_4$
M_2		M_2f_2'	$M_2S_2'f_3$	$M_2S_2'S_3f_4$
M_3			M_3f_3'	$M_3S_3'f_4$
M_4				M_4f_4'

From banding only young, the direct recovery rates f_i' can be estimated, but without additional information (bandings of adults) the survival rates S_i' of young can never be estimated. The impossibility of estimating S_i' persists no matter what simplifying assumptions are made about how rates vary over time. For example, assuming $S_i \equiv S$, $i = 1, \ldots, \ell - 1$ (constant adult survival rate) does not allow estimability. Assuming constant adult recovery rates $(f_i \equiv f,\ i = 2, \ldots, \ell)$ renders adult survival rates S_i estimable from young only recoveries, but leaves S_i' and f confounded. Moreover, this assumption is certainly wrong. In the extreme case of assuming no parameter varies over time the model structure depends upon only four parameters (f', f, S', S):

Number banded	Year of recovery 1	2	3	4
M_1	M_1f'	$M_1S'f$	$M_1S'Sf$	$M_1S'SSf$
M_2		M_2f'	$M_2S'f$	$M_2S'Sf$
M_3			M_3f'	$M_3S'f$
M_4				M_4f'

This is the structure for young under the H_{01} model.

Note that the parameters S' and f always occur together as a product. This is why, even in this simple model, these two parameters cannot be separately estimated; only the product $S'f$ can be estimated if no adults are banded. To illustrate this point we could define "new" parameters as $S'_o = cS'$ and $f_o = \frac{1}{c}f$ for infinitely many values of the constant c, and the product $S'f$ is not changed. Thus, the expected cell probabilities cannot be given as a product of unique parameters S' and f; for example

$$\mathbf{E}(Q_{12}) = M_1S'f = M_1(S'c)\left(\frac{1}{c}f\right) = M_1S'_of_o\,,$$

where S' and S'_o can be quite different. In statistical terms this parameterization of the band recovery model leaves S' and f nonidentifiable.

This nonidentifiability of S'_i persists in the H_1 model for banding of young only. In fact, not even S_i or f_i is estimable under the H_1 model if only young are banded. This problem may be illustrated numerically by choosing several different configurations of annual survival and recovery rates which produce exactly the same array of expected band recoveries. For $k = \ell = 4$, suppose a banding experiment on only young birds produced the following recoveries:

Number banded	Year of recovery 1	2	3	4
10,000	1,000	500	150	168
5,000		1,000	100	112
9,000			1,800	216
8,000				1,200

(symbolic expectations under H_1 were given above). The direct recovery rates (the f'_i) are unique. For this example $f'_1 = 0.1$, $f'_2 = 0.2$, $f'_3 = 0.2$, and $f'_4 = 0.15$. The remaining parameters need not be unique to produce this expected data array. In fact we used $S_2 = 0.6$ and $S_3 = 0.7$. But for the choice of S'_i and f'_i the following three parameter sets will all produce these exact same expected recovery data:

Parameter set	S'_1	S'_2	S'_3	f_2	f_3	f_4
A	0.5	0.4	0.3	0.1	0.05	0.08
B	0.25	0.2	0.15	0.2	0.1	0.16
C	0.625	0.5	0.375	0.08	0.04	0.064

For example, consider the expected recoveries from the first banded cohort in the fourth recovery year:

Parameter set	$\mathbf{E}(Q_{14}) = M_1S'_1S_2S_3f_4$
A	10,000 (.5) (.6) (.7) (.08) = 168
B	10,000 (.25) (.6) (.7) (.16) = 168
C	10,000 (.625) (.6) (.7) (.064) = 168

The reader should verify that these three parameter sets give exactly the same expected band recovery data. Consequently, it is impossible to infer from these "data" which parameter set was used to generate these data.

It is possible to test for first-year age effects in banding data for young. This is a test of whether or not Model 1 (Chapter 2) fits the data and is discussed in Section 2.6 as a test of the null hypothesis that Model 1 fits the data vs. the alternative that Model 0 fits. In fact Model 0 cannot be distinguished by any statistical test from the model for recovery data of young under H_1 (see Section 2.5, *Proper and Improper Use of Model 0)*.

For the array of expected band recoveries above, the test is based on the two contingency tables below, the elements of which come from the rows and columns of the "data" array as explained in Robson and Youngs (1971) (also see Section 2.6 for an explanation of this test):

rows	columns 2	columns 3+4
1	500	318
2	1,000	212

rows	columns 3	columns 4
1+2	250	280
3	1,800	216

The computed chi-squares, each with 1 df, are 115.8 and 474.6, respectively. We would thus conclude age-specific recovery and/or survival rates occur in the population. If these were adult bandings, we might assume Model 0 was the true model and thus proceed to get meaningful survival estimates. But when we know the recovery data came from bandings of young only, we must abandon all attempts to meaningfully estimate survival rates from these data.

It may be useful to heuristically show how S_i' can be estimated if adult recovery data are available. First

$$S_i' \rho_{i+1} = \mathbf{E}(Q_{i\cdot} - Q_{ii}) / M_i$$

is clearly estimable from band recoveries of young, while

$$\rho_{i+1} = \mathbf{E}(R_{i+1\cdot}) / N_{i+1}$$

is estimable from band recoveries of adults $(\rho_{i+1} = f_{i+1} + S_{i+1} f_{i+2} + \ldots + S_{i+1} \cdots S_{\ell-1} f_{\ell})$. An estimator of S_i' (in fact the ML estimator) is

$$\hat{S}_i' = \frac{\widehat{S_i' \rho_{i+1}}}{\hat{\rho}_{i+1}} = \frac{Q_{i\cdot} - Q_{ii}}{M_i} \Big/ \frac{R_{i+1\cdot}}{N_{i+1}} .$$

Without adult data, ρ_{i+1} cannot be estimated.

The above discussion concerning the estimation of survival rates from birds banded as young assumes the models of Chapters 2 and 3. The astute reader might ask if there exist models for band recovery data allowing the estimation of survival rates from banding only young birds. The model underlying the composite dynamic life table method is such a model (cf. Hickey 1952; Seber 1972; Anderson and Burnham 1976). Seber (1971) has considered the proper statistical analysis of data given this model structure. The composite dynamic model assumes that both recovery and survival rates are age-dependent only; i.e., these rates are unaffected by year-to-year changes in hunting regulations, habitat, weather, etc.

A second assumption of the composite dynamic model is that the age-specific recovery rate is a constant fraction of the age-specific mortality rate. For the two-age-class case (i.e., young and adults) this means we must assume

$$\frac{f_i'}{1 - S_i'} = \frac{f_i}{1 - S_i} = c \qquad , i = 1, \ldots, \ell - 1$$

where c is a constant.

Both assumptions are critical and both are demonstrably invalid for banding studies of game birds. Detailed analysis of mallard banding data has documented the untenability of both assumptions (Anderson 1975). Moreover, the estimates of age-specific survival rates computed assuming this model (i.e., the composite dynamic method) are severely biased when these assumptions fail.

It is our intention in this handbook to detail some correct models and methods for the analysis of band recovery data. Thus, we will not dwell on the composite dynamic method, nor the concomitant model under which young-only recovery data could be analyzed (e.g., Seber 1971). Firstly, the model is invalid for waterfowl (Burnham and Anderson 1979) and nongame species (Anderson et al. 1981). Secondly, serious bias occurs when these crucial assumptions are not met (Anderson et al. 1985). The interested reader can pursue this topic further by referring to reports by Anderson (1975) and Anderson and Burnham (1976).

Finally, we note that it is possible to construct yet other models that have age-specific survival rates in their structure, and are such that from banding young only these survival rates could be estimated *if the model were true.* However, such models are usually demonstrably invalid, on either theoretical and/or empirical grounds. For example, let survival be age-specific and let the recovery rate be constant, independent of age or calendar year (cf. Seber 1971). This model seems unreasonable: empirically we know f varies by calendar year and by age (young vs. adult), and theoretically it does not make sense to have age-dependent annual mortality but age-independent recovery rates (mortality is a "hidden" component of f). *We can not emphasize too strongly that, based on our current knowledge, there is no valid way to estimate age-specific survival rates from only the banding of young.*

Chapter 4. Models for Birds Banded as Young, Subadults, and Adults

4.1 Introduction

The models considered in this chapter permit estimation under the assumption that survival and recovery rates are age-dependent for the first 2 years of life. The experimental situation requires that members of the resulting three age classes be recognizable at the time of banding because the numbers banded and subsequent recoveries must be recorded separately for each age class.

We refer to the three age classes as young (birds in their first year), subadults (birds in their second year), and adults (birds over 2 years old). This terminology is commonly used in relation to certain species (e.g., geese) where 2-year-olds are not sexually mature, and it should not be confused with the use of the terms "young" and "adults" in Chapters 2 and 3.

The methods of this chapter are of limited use in relation to waterfowl studies, because three age groups are recognizable in only a few species (e.g., geese). Also, the analysis of a considerable amount of waterfowl data by the methods of Chapter 3 has indicated that the age-dependence assumption, common to the models of this chapter and to H_3 of Chapter 3, is unnecessarily general for many species. However, several sets of goose data should be collected and analyzed using the methods of this chapter to decide whether they are entirely inappropriate for waterfowl. We suggest that these methods may be of use in studies on fish populations where age may sometimes be easily determined by length, or, if necessary, by scale analysis and where the age-dependence assumption may not be too general.

The FORTRAN program BROWNIE is available to facilitate the use of these methods and sample output from the program appears in the examples. Data used in these examples are not real but have been generated stochastically.

Because the models of this chapter apply to a more specific situation and hence are likely to be less useful than their Chapter 3 counterparts, the development and treatment of these models is less extensive than those in Chapter 3. Also less detail is given in describing methods of estimation and testing procedures when analogous procedures in Chapter 3 can be referred to. (For a more detailed mathematical treatment see Brownie 1973). Similarly, the output of the FORTRAN program is much briefer than that for the two age-class models of Chapter 3. The reader should not attempt to use the models in Chapter 4 without first understanding Chapter 3.

Experimental Situation

The number banded and subsequent recoveries are recorded separately for three groups, i.e., for birds banded as young, subadults, and adults. In all other respects the experimental situation is like that of Chapter 3 and further description is omitted. We emphasize that each of the three age classes must be represented in every cohort released or parameters of interest will not be estimable.

Notation and Definitions

k = the number of years at the start of which banded birds are released.
ℓ = the number of years during which recoveries are recorded, $\ell \geqslant k$.
$s = \ell - k$ = the number of years, beyond the year of the last release, when recoveries are recorded, $s \geqslant 0$.
N_i = the number of adults banded and released in year i, $i = 1, \ldots, k$.
K_i = the number of subadults banded and released in year i, $i = 1, \ldots, k$.
M_i = the number of young banded and released in year i, $i = 1, \ldots, k$.
R_{ij} = the number of bands recovered in year j from the adults released in year i, $i = 1, \ldots, k$, $j = i, \ldots, \ell$.
Y_{ij} = the number of bands recovered in year j from the subadults released in year i, $i = 1, \ldots, k$, $j = i, \ldots, \ell$.
Q_{ij} = the number of bands recovered in year j from the young released in year i, $i = 1, \ldots, k$, $j = i, \ldots, \ell$.

The data are presented as in Chapter 3 except that three data arrays are present instead of two. This is illustrated in Table 4.1 in terms of N_i, K_i, M_i, R_{ij}, Y_{ij} and Q_{ij} for a banding study with 3 years of banding and 5 years of recovery ($k=3$, $\ell=5$, $s=2$). Some useful subtotals are indicated in the table.

Table 4.1. *Representation of the data for a banding study with $k=3$, $\ell=5$, and $s=2$ when young, subadults, and adults are banded and released each year.*

Year banded	Number banded	Year of recovery: 1	2	3	4	$\ell=5$	Row totals
			Birds banded as adults				
1	N_1	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}	$R_{1\cdot}=T_1$
2	N_2		R_{22}	R_{23}	R_{24}	R_{25}	$R_{2\cdot}$
$k=3$	N_3			R_{33}	R_{34}	R_{35}	$R_{3\cdot}$
	Column totals	$R_{\cdot 1}$	$R_{\cdot 2}$	$R_{\cdot 3}$	$R_{\cdot 4}$	$R_{\cdot 5}=T_5$	
			Birds banded as subadults				
1	K_1	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	$Y_{i\cdot}=V_1$
2	K_2		Y_{22}	Y_{23}	Y_{24}	Y_{25}	$Y_{2\cdot}$
$k=3$	K_3			Y_{33}	Y_{34}	Y_{35}	$Y_{3\cdot}$
	Column totals	$Y_{\cdot 1}$	$Y_{\cdot 2}$	$Y_{\cdot 3}$	$Y_{\cdot 4}$	$Y_{\cdot 5}=V_5$	
			Birds banded as young				
1	M_1	Q_{11}	Q_{12}	Q_{13}	Q_{14}	Q_{15}	$Q_{1\cdot}=U_1$
2	M_2		Q_{22}	Q_{23}	Q_{24}	Q_{25}	$Q_{2\cdot}$
$k=3$	M_3			Q_{33}	Q_{34}	Q_{35}	$Q_{3\cdot}$
	Column totals	$Q_{\cdot 1}$	$Q_{\cdot 2}$	$Q_{\cdot 3}$	$Q_{\cdot 4}$	$Q_{\cdot 5}=U_5$	

The arrays representing the recovery matrices are R_{ij}'s for adults, Y_{ij}'s for subadults, and Q_{ij}'s for young.

As before, row, column, and certain block totals of the recovery matrices are used in summarizing the data. The notation is similar to that of Chapter 3. Thus the row totals of the recovery matrices for adults, subadults, and young are denoted by $R_{i\cdot}$, $Y_{i\cdot}$, and $Q_{i\cdot}$, respectively, $i=1,\ldots,k$. Similarly, $R_{\cdot j}$, $Y_{\cdot j}$ and $Q_{\cdot j}$, $j=1,\ldots,\ell$ are the corresponding column totals, and

$$T_i=\begin{cases} R_{1\cdot} & ,i=1 \\ R_{i\cdot}+T_{i-1}-R_{\cdot i-1} & ,i=2,\ldots,k, \end{cases}$$

$$V_i=\begin{cases} Y_{i\cdot} & ,i=1 \\ Y_{i\cdot}+V_{i-1}-Y_{\cdot i-1} & ,i=2,\ldots,k, \end{cases}$$

$$U_i=\begin{cases} Q_{i\cdot} & ,i=1 \\ Q_{i\cdot}+U_{i-1}-Q_{\cdot i-1} & ,i=2,\ldots,k, \end{cases}$$

are corresponding block totals.

Subtotals involving elements from all three recovery matrices are

$$A_i=\begin{cases} R_{\cdot 1} & ,i=1 \\ R_{\cdot i}+Y_{\cdot i}+Q_{\cdot i}-Y_{ii}-Q_{ii}-Q_{i-1,i} & ,i=2,\ldots,k. \end{cases}$$

where A_i is the total number of recoveries from adults in year i (including recoveries from previously released young and subadults which have survived to adulthood),

$$D_i = \begin{cases} T_1 = R_{1\cdot} & , i = 1 \\ T_i + V_i + U_i - Y_{i\cdot} - Q_{i\cdot} - Q_{i-1\cdot} + Q_{i-1,i-1} & , i = 2, \ldots, k. \end{cases}$$

Synthetic data for a banding study with $k = \ell = 6$ are presented in Table 4.2 and are used to illustrate the calculation of the above subtotals.

Table 4.2 *Synthetic data for a study with* $k = \ell = 6, s = 0$.

Year banded	Number banded	Year of recovery 1	2	3	4	5	6
				Birds Banded as adults			
1	700	36	20	26	11	17	6
2	700		23	31	22	25	16
3	700			42	24	23	11
4	700				38	37	20
5	700					38	19
6	700						23
				Birds banded as subadults			
1	800	80	12	20	12	19	3
2	800		57	33	17	34	7
3	800			83	26	25	13
4	800				64	25	14
5	800					115	26
6	800						57
				Birds banded as young			
1	1,000	161	34	21	7	18	4
2	1,000		109	73	12	16	7
3	1,000			171	51	26	13
4	1,000				157	57	10
5	1,000					198	39
6	1,000						102

Subtotals

i	$\mathring{R}_{i\cdot}$	$R_{\cdot i}$	$Y_{i\cdot}$	$Y_{\cdot i}$	$Q_{i\cdot}$	$Q_{\cdot i}$	T_i	V_i	U_i	A_i	D_i
1	116	36	146	80	245	161	116	146	245	36	116
2	117	43	148	69	217	143	197	214	301	55	263
3	100	99	147	136	261	265	254	292	419	173	449
4	95	95	103	119	224	227	250	259	378	169	470
5	57	140	141	218	237	315	212	281	388	303	436
6	23	95	57	120	102	175	95	120	175	192	192

Parameters and Assumptions

The models of this chapter are closely analogous to the models in Chapter 3 and are constructed in terms of the same basic parameters: S, the annual survival rate and f, the annual recovery rate. In all three models f and S are assumed to be year-specific (indicated as usual by subscripts, e.g., f_i) and age-dependent for the three classes—young, subadult and adult (indicated by superscript primes).

Again we note that the assumption that banded birds in the population suffer independent fates is questionable for species such as geese where young and adults tend to migrate and winter in distinct family groups. This point is particularly important in this chapter because several species of geese are among the few waterfowl species in which young, subadults, and adults are distinguishable through examination in the field. Failure of this independence assumption will not bias the estimates of survival or recovery rates but will invalidate the sampling variances of estimators.

4.2 The Model Under H_4

We begin with the model which is the analogue of the H_1 model of Chapter 3, and which we call the model under H_4 (or the H_4 model). This is the simplest of the three models of this chapter and the assumptions of H_4 are:

(1) Annual survival and recovery rates are year-specific;
(2) annual survival and recovery rates are different for young, subadults, and adults; and
(3) reporting rates are not dependent on the time of release.

The parameters are:

f_i'' = recovery rate in year i for young banded in year i, $i = 1, \ldots, k$,

S_i'' = survival rate in year i for young banded in year i, $i = \begin{cases} 1, \ldots, k-1, & \text{if } \ell = k \\ 1, \ldots, k, & \text{if } \ell > k, \end{cases}$

f_i' = recovery rate in year i for subadults banded in year i or survivors of young banded in year $i-1$, $i = 1, \ldots, k$,

S_i' = survival rate in year i for subadults banded in year i or survivors of young banded in year $i-1$, $i = \begin{cases} 1, \ldots, k-1, & \text{if } \ell = k \\ 1, \ldots, k, & \text{if } \ell > k, \end{cases}$

f_i = recovery rate in year i for adults, $i = 1, \ldots, \ell$,
S_i = survival rate in year i for adults, $i = 1, \ldots, \ell - 1$.

The structure of the model is expressed by the expected numbers of band recoveries in terms of N_i, K_i, M_i, f_i, f_i', f'', S_i, S_i' and S'' as shown in Table 4.3.

Table 4.3. *Expected numbers of band recoveries under H_4 for a banding study with $k = 3$, $\ell = 5$, $s = 2$.*

Year banded	Number banded	Year of recovery				
		1	2	3	4	5
Birds banded and released as adults						
1	N_1	$N_1 f_1$	$N_1 S_1 f_2$	$N_1 S_1 S_2 f_3$	$N_1 S_1 S_2 S_3 f_4$	$N_1 S_1 S_2 S_3 S_4 f_5$
2	N_2		$N_2 f_2$	$N_2 S_2 f_3$	$N_2 S_2 S_3 f_4$	$N_2 S_2 S_3 S_4 f_5$
3	N_3			$N_3 f_3$	$N_3 S_3 f_4$	$N_3 S_3 S_4 f_5$
Birds banded and released as subadults						
1	K_1	$K_1 f_1'$	$K_1 S_1' f_2$	$K_1 S_1' S_2 f_3$	$K_1 S_1' S_2 S_3 f_4$	$K_1 S_1' S_2 S_3 S_4 f_5$
2	K_2		$K_2 f_2'$	$K_2 S_2' f_3$	$K_2 S_2' S_3 f_4$	$K_2 S_2' S_3 S_4 f_5$
3	K_3			$K_3 f_3'$	$K_3 S_3' f_4$	$K_3 S_3' S_4 f_5$
Birds banded and released as young						
1	M_1	$M_1 f_1''$	$M_1 S_1'' f_2'$	$M_1 S_1'' S_2' f_3$	$M_1 S_1'' S_2' S_3 f_4$	$M_1 S_1'' S_2' S_3 S_4 f_5$
2	M_2		$M_2 f_2''$	$M_2 S_2'' f_3'$	$M_2 S_2'' S_3' f_4$	$M_2 S_2'' S_3' S_4 f_5$
3	M_3			$M_3 f_3''$	$M_3 S_3'' f_4'$	$M_3 S_3'' S_4' f_5$

Estimation of Parameters

ML estimators of the annual recovery rates are:

$$\hat{f}_i = \frac{R_{i\cdot}}{N_i}\,\frac{A_i}{D_i} \qquad , i = 1, \ldots, k,$$

$$\hat{f}_i' = \begin{cases} \dfrac{Y_{11}}{K_1} & ,\, i=1 \\[2ex] \dfrac{Y_{i\cdot}}{K_i}\,\dfrac{Y_{ii}+Q_{i-1,i}}{Y_{i\cdot}+Q_{i-1\cdot}-Q_{i-1,i-1}} & ,\, i=2,\ldots,k, \end{cases}$$

$$\hat{f}_i'' = \frac{Q_{ii}}{M_i} \qquad ,\, i=1,\ldots,k.$$

The calculation of estimates is illustrated for the data in Table 4.2 and the results can be compared with the printout of Example 4.1b,

$$\hat{f}_1 = \frac{R_{1\cdot}\times A_1}{N_1\times D_1} = \frac{116\times 36}{700\times 116} = 0.0514,$$

$$\hat{f}_2 = \frac{R_{2\cdot}\times A_2}{N_2\times D_2} = \frac{117\times 55}{700\times 263} = 0.0350,$$

$$\hat{f}_1' = \frac{Y_{11}}{K_1} = \frac{80}{800} = 0.1000,$$

$$\hat{f}_2' = \frac{Y_{2\cdot}\times (Y_{22}+Q_{12})}{K_2\times (Y_{2\cdot}+Q_{1\cdot}-Q_{11})} = \frac{148\times 91}{800\times 232} = 0.0726,$$

$$\hat{f}_1'' = \frac{Q_{11}}{M_1} = \frac{161}{1{,}000} = 0.1610,$$

$$\hat{f}_2'' = \frac{Q_{22}}{M_2} = \frac{109}{1{,}000} = 0.1090.$$

ML estimators of annual survival rates are:

$$\hat{S}_i = \frac{R_{i\cdot}}{N_i}\,\frac{D_i-A_i}{D_i}\,\frac{N_{i+1}}{R_{i+1\cdot}} \qquad ,\, i=1,\ldots,k-1$$

$$\hat{S}_i' = \begin{cases} \dfrac{Y_{1\cdot}-Y_{11}}{K_1}\,\dfrac{N_2}{R_{2\cdot}} & ,\, i=1 \\[2ex] \dfrac{Y_{i\cdot}}{K_i}\left(1-\dfrac{Y_{ii}+Q_{i-1,i}}{Y_{i\cdot}+Q_{i-1\cdot}-Q_{i-1,i-1}}\right)\dfrac{N_{i+1}}{R_{i+1\cdot}} & ,\, i=2,\ldots,k-1, \end{cases}$$

$$\hat{S}_i'' = \frac{Q_{i\cdot}-Q_{ii}}{M_i}\,\frac{K_{i+1}}{Y_{i+1\cdot}} \qquad ,\, i=1,\ldots,k-1.$$

For example,

$$\hat{S}_1 = \frac{R_{1\cdot}\times (D_1-A_1)\times N_2}{N_1\times D_1\times R_{2\cdot}} = \frac{116\times (116-36)\times 700}{700\times 116\times 117} = 0.6838,$$

$$\hat{S}_2 = \frac{R_{2\cdot}\times (D_2-A_2)\times N_3}{N_2\times D_2\times R_{3\cdot}} = \frac{117\times (263-55)\times 700}{700\times 263\times 100} = 0.9253,$$

$$\hat{S}_1' = \frac{(Y_{1\cdot}-Y_{11})\times N_2}{K_1\times R_{2\cdot}} = \frac{66\times 700}{800\times 117} = 0.4936,$$

$$\hat{S}_2' = \frac{Y_{2\cdot}}{K_2}\left(1-\frac{Y_{22}+Q_{12}}{Y_{2\cdot}+Q_{1\cdot}-Q_{11}}\right)\frac{N_3}{R_{3\cdot}} = \frac{148}{800}\left(1-\frac{91}{232}\right)\frac{700}{100} = 0.7870,$$

$$\hat{S}_1'' = \frac{(Q_{1\cdot}-Q_{11})\times K_2}{M_1\times Y_{2\cdot}} = \frac{84\times 800}{1{,}000\times 148} = 0.4541,$$

$$\hat{S}_2'' = \frac{(Q_{2\cdot}-Q_{22})\times K_3}{M_2\times Y_{3\cdot}} = \frac{108\times 800}{1{,}000\times 147} = 0.5878.$$

It is easily seen that these calculations agree with the results in the output from program BROWNIE in Example 4.1b.

Bias-adjusted ML estimators of survival (analogous to the H_1 estimators $\tilde{S}_i$ and $\tilde{S}_i'$) are easily constructed but are not defined here. We repeat that for sample sizes which are large enough that the precision of estimators is good (and hence estimates are reliable) the difference between the adjusted and unadjusted estimators will be negligible.

If $\ell > k$, the parameters $S_k, \ldots, S_{\ell-1}, S_k', S_k'', f_{k+1}, \ldots, f_\ell$ are not separately estimable.

Sampling Variances, Standard Errors, and Confidence Intervals

As in Chapter 3, estimates of the sampling variances and 95% confidence intervals are obtained to provide an indication of the precision of the ML estimates. The notation used to define the variance estimators is like that in Chapter 3, and these are used to obtain estimates of standard errors and confidence intervals as in Chapter 3. The required variance estimators are listed below, followed by a single numerical illustration of the computational procedure.

$$\mathrm{var}(\hat{f}_i) = (\hat{f}_i)^2 \left[\frac{1}{R_{i\cdot}} - \frac{1}{N_i} + \frac{1}{A_i} - \frac{1}{D_i}\right] \qquad , i = 1, \ldots, k,$$

$$\mathrm{var}(\hat{S}_i) = (\hat{S}_i)^2 \left[\frac{1}{R_{i\cdot}} - \frac{1}{N_i} + \frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}} + \frac{1}{D_i - A_i} - \frac{1}{D_i}\right] \qquad , i = 1, \ldots, k-1,$$

$$\mathrm{var}(\hat{f}_i') = \begin{cases} \hat{f}_1'(1-\hat{f}_1')/K_1 & , i = 1 \\ f_i')^2 \left[\dfrac{1}{Y_{i\cdot}} - \dfrac{1}{K_i} + \dfrac{1}{Y_{ii} + Q_{i-1,i}} - \dfrac{1}{Y_{i\cdot} + Q_{i-1\cdot} - Q_{i-1,i-1}}\right] & , i = 2, \ldots, k, \end{cases}$$

$$\mathrm{var}(\hat{S}_i') = \begin{cases} \hat{S}_1')^2 \left[\dfrac{1}{Y_{1\cdot} - Y_{11}} - \dfrac{1}{K_1} + \dfrac{1}{R_{2\cdot}} - \dfrac{1}{N_2}\right] & , i = 1 \\ \hat{S}_i')^2 \left[\dfrac{1}{Y_{i\cdot}} - \dfrac{1}{K_i} + \dfrac{1}{R_{i+1\cdot}} - \dfrac{1}{N_{i+1}} + \dfrac{1}{Y_{i\cdot} - Y_{ii} + Q_{i-1\cdot} - Q_{i-1,i-1} - Q_{i-1,i}} - \dfrac{1}{Y_{i\cdot} + Q_{i-1\cdot} - Q_{i-1,i-1}}\right] & , i = 2, \ldots, k-1, \end{cases}$$

$$\mathrm{var}(\hat{f}_i'') = \hat{f}_i''(1-\hat{f}_i'')/M_i \qquad , i = 1, \ldots, k,$$

$$\mathrm{var}(\hat{S}_i'') = (\hat{S}_i'')^2 \left[\frac{1}{Q_{i\cdot} - Q_{ii}} - \frac{1}{M_i} + \frac{1}{Y_{i+1\cdot}} - \frac{1}{K_{i+1}}\right] \qquad , i = 1, \ldots, k-1.$$

For the data of Table 4.2,

$$\mathrm{var}(\hat{S}_2) = (\hat{S}_2)^2 \left[\frac{1}{R_{2\cdot}} - \frac{1}{N_2} + \frac{1}{R_{3\cdot}} - \frac{1}{N_3} + \frac{1}{D_2 - A_2} - \frac{1}{D_2}\right] = (0.9253)^2 \left[\frac{1}{117} - \frac{1}{700} + \frac{1}{100} - \frac{1}{700} + \frac{1}{208} - \frac{1}{263}\right] = 0.01429416,$$

$$\mathrm{se}(\hat{S}_2) = \sqrt{0.01429416} = 0.1196,$$

and $1.46 \times \mathrm{se}(\hat{S}_2) = 0.2344$, thus the 95% confidence interval for S_2 is (0.6909, 1.1597).

Comparison with the output of Example 4.1b shows that the lower bound of the confidence interval in the example is slightly different from the above result. As mentioned before, such differences are due to the greater accuracy of the calculations performed by the computer. Note also that the confidence interval for S_2 is wide and contains impossible values. In real data this would indicate an insufficient number of birds were banded. The output in Example 4.1b shows that estimated standard errors and confidence intervals are printed beside each estimate so that tedious computations can be avoided by using the FORTRAN program BROWNIE.

Sampling Covariances and Correlations

Estimates of the sampling covariances and correlations between the ML estimators are obtained as in Chapter 3, and the notation used is again similar. All the covariance estimates listed below (and the corresponding correlations) are contained in the output of the FORTRAN program (see Example 4.1c) and are presented here largely for reference purposes.

Again we point out that these are estimates of the sampling covariances and correlations between the ML estimators and do not reflect a relationship between the unknown parameters. It is important to obtain an idea of the magnitudes of the correlations, because if these are substantial, they will obscure, or be confounded with, any relationship that exists between the unknown parameters. In Example 4.1c we see that many of these correlations are estimated to be quite large. Nonzero sampling covariances are estimated by

$$\operatorname{cov}(\hat{f}_i, \hat{S}_i) = \hat{f}_i \hat{S}_i \left[\frac{1}{R_{i\cdot}} - \frac{1}{N_i} - \frac{1}{D_i}\right] \qquad , i=1,\ldots,k-1,$$

$$\operatorname{cov}(\hat{f}_{i+1}, \hat{S}_i) = -\hat{f}_{i+1} \hat{S}_i \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i=1,\ldots,k-1,$$

$$\operatorname{cov}(\hat{f}_{i+1}, \hat{S}_i') = -\hat{f}_{i+1} \hat{S}_i' \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i=1,\ldots,k-1,$$

$$\operatorname{cov}(\hat{S}_i, \hat{S}_{i+1}) = -\hat{S}_i \hat{S}_{i+1} \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i=1,\ldots,k-2,$$

$$\operatorname{cov}(\hat{S}_i, \hat{S}_i') = \hat{S}_i \hat{S}_i' \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i=1,\ldots,k-1,$$

$$\operatorname{cov}(\hat{S}_{i+1}, \hat{S}_i') = -\hat{S}_{i+1} \hat{S}_i' \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i=1,\ldots,k-2,$$

$$\operatorname{cov}(\hat{f}_i', \hat{S}_i') = \begin{cases} -\hat{f}_1' \hat{S}_1' / K_1 & , i=1 \\ \hat{f}_i' \hat{S}_i' \left[\frac{1}{Y_{i+1\cdot}} - \frac{1}{K_{i+1}}\right] & , i=2,\ldots,k-1, \end{cases}$$

$$\operatorname{cov}(\hat{f}_{i+1}', \hat{S}_i'') = -\hat{f}_{i+1}' \hat{S}_i'' \left[\frac{1}{Y_{i+1\cdot}} - \frac{1}{K_{i+1}}\right] \qquad , i=1,\ldots,k-1,$$

$$\operatorname{cov}(\hat{S}_{i+1}', \hat{S}_i'') = -\hat{S}_{i+1}' \hat{S}_i'' \left[\frac{1}{Y_{i+1\cdot}} - \frac{1}{K_{i+1}}\right] \qquad , i=1,\ldots,k-2,$$

$$\operatorname{cov}(\hat{f}_i'', \hat{S}_i'') = -\hat{f}_i'' \hat{S}_i'' / M_i \qquad , i=1,\ldots,k-1\,.$$

Goodness of Fit Test

As in Chapter 3, a goodness of fit test is computed for each model to help in judging the adequacy of each model and in choosing the best model for a given data set. The goodness of fit test of the model under H_4 is computed in program BROWNIE in a manner analogous to the computation of the goodness of fit test of the model under H_1. As before, the test statistic is chi-square distributed if the assumptions of H_4 hold, and the result is interpreted in the usual way (i.e., "large" chi-square values indicate that agreement between model and data is poor).

The test for the synthetic data of Table 4.2 (see Example 4.1h) gives a chi-square value of 26.43 with 30 df. The probability under H_4 of a value at least this large is 0.653 and hence there is no reason to suspect that the assumptions of H_4 are incorrect. (This is not surprising because the data were generated using the model under H_4 as the probability model.)

4.3 The Model Under H_6

We now consider the model which is the analog of the H_2 model of Chapter 3. The model under H_6 is the most general of the three models of this chapter. The intermediate model (the model under H_5) is discussed in the next section for reasons given there. The assumptions of H_6 are:

1. Annual survival and recovery rates are year-specific;
2. annual survival and recovery rates are different for young, subadults, and adults; and
3. in any year, the reporting rate for new releases is different from that for survivors of previous releases.

Assumptions 1 and 3 of H_6 are the same as those for H_2, and H_6 is a generalization of H_4 in the same way that H_2 is a generalization of H_1. Comments related to choosing between H_1 and H_2 (see Section 3.7) thus apply to choosing between H_4 and H_6.

Because of assumption 3, newly released adults and adult survivors from earlier releases have different recovery rates, defined as f_i'''' and f_i, respectively. Also, newly released subadults and the subadult survivors of young released the year before have different recovery rates, defined as f_i' and f_i''', respectively. All other parameters, i.e., S_i, S_i', f_i'', S_i'', are as defined for H_4.

The structure of the model is expressed in Table 4.4.

Table 4.4. *Expected numbers of band recoveries under H_6 for a banding study with $k=3$, $\ell=5$, $s=2$.*

Year banded	Number banded	Year of recovery				
		1	2	3	4	5
				Birds banded and released as adults		
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''''	$N_2S_2f_3$	$N_2S_2S_3f_4$	$N_2S_2S_3S_4f_5$
3	N_3			N_3f_3''''	$N_3S_3f_4$	$N_3S_3S_4f_5$
				Birds banded and released as subadults		
1	K_1	K_1f_1'	$K_1S_1'f_2$	$K_1S_1'S_2f_3$	$K_1S_1'S_2S_3f_4$	$K_1S_1'S_2S_3S_4f_5$
2	K_2		K_2f_2'	$K_2S_2'f_3$	$K_2S_2'S_3f_4$	$K_2S_2'S_3S_4f_5$
3	K_3			K_3f_3'	$K_3S_3'f_4$	$K_3S_3'S_4f_5$
				Birds banded and released as young		
1	M_1	M_1f_1''	$M_1S_1''f_2'''$	$M_1S_1''S_2'f_3$	$M_1S_1''S_2'S_3f_4$	$M_1S_1''S_2'S_3S_4f_5$
2	M_2		M_2f_2''	$M_2S_2''f_3'''$	$M_2S_2''S_3'f_4$	$M_2S_2''S_3'S_4f_5$
3	M_3			M_3f_3''	$M_3S_3''f_4'''$	$M_3S_3''S_4'f_5$

Estimation of Parameters

ML estimators of the individually estimable parameters are:

$$\hat{f}_i=\frac{R_{i\cdot}-R_{ii}}{N_i}\ \frac{A_i-R_{ii}}{D_i-A_i-R_{i\cdot}+R_{ii}} \qquad ,i=\begin{cases}2,\ldots,k-1 \text{ if } \ell=k\\ 2,\ldots,k \quad \text{ if } \ell>k,\end{cases}$$

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

$$\hat{f}_i'=\frac{Y_{ii}}{K_i} \qquad ,i=1,\ldots,k,$$

$$\hat{f}_i''=\frac{Q_{ii}}{M_i} \qquad ,i=1,\ldots,k,$$

$$\hat{f}_i'''=\frac{Q_{i-1,i}}{Q_{i-1\cdot}-Q_{i-1,i-1}-Q_{i-1,i}}\ \frac{Y_{i\cdot}-Y_{ii}}{K_i} \qquad ,i=\begin{cases}2,\ldots,k-1 \text{ if } \ell=k\\ 2,\ldots,k \quad \text{ if } \ell>k,\end{cases}$$

$$\hat{S}_i=\frac{R_{i\cdot}-R_{ii}}{N_i}\ \frac{N_{i+1}}{R_{i+1\cdot}-R_{i+1,i+1}}\ \frac{D_{i+1}-A_{i+1}-R_{i+1\cdot}+R_{i+1,i+1}}{D_{i+1}-R_{i+1\cdot}} \qquad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k,\end{cases}$$

$$\hat{S}_i'=\frac{Y_{i\cdot}-Y_{ii}}{K_i}\ \frac{N_{i+1}}{R_{i+1\cdot}-R_{i+1,i+1}}\ \frac{D_{i+1}-A_{i+1}-R_{i+1\cdot}+R_{i+1,i+1}}{D_{i+1}-R_{i+1\cdot}} \qquad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k,\end{cases}$$

$$\hat{S}_i''=\frac{Q_{i\cdot}-Q_{ii}-Q_{i,i+1}}{M_i}\ \frac{K_{i+1}}{Y_{i+1\cdot}-Y_{i+1,i+1}} \qquad ,i=\begin{cases}1,\ldots,k-2 \text{ if } \ell=k\\ 1,\ldots,k-1 \text{ if } \ell>k.\end{cases}$$

For example, for the data of Table 4.2,

$$\hat{S}_1 = \frac{80 \times 700 \times (263 - 55 - 117 + 23)}{700 \times 94 \times (263 - 117)} = 0.6645,$$

$$\hat{S}_1' = \frac{66 \times 700 \times (263 - 55 - 117 + 23)}{800 \times 94 \times (263 - 117)} = 0.4797,$$

$$\hat{S}_1'' = \frac{(245 - 161 - 34) \times 800}{1{,}000 \times (148 - 57)} = 0.4396.$$

The (unadjusted) ML estimators defined above are evaluated and printed out by the FORTRAN program (see Example 4.1f). Bias-adjusted estimators are not presented here for the reason given in Section 4.2.

Sampling Variances, Standard Errors, and Confidence Intervals

The output of the FORTRAN program contains an estimate of the standard error and corresponding 95% confidence interval beside each ML estimate evaluated. The standard errors and confidence intervals are estimated as in the previous section and earlier chapters using the variance estimators defined below.

$$\text{var}(\hat{f}_i) = (\hat{f}_i)^2 \left[\frac{1}{R_{i\cdot} - R_{ii}} - \frac{1}{N_i} + \frac{1}{D_i - A_i - R_{i\cdot} + R_{ii}} + \frac{1}{A_i - R_{ii}}\right] \quad , i = \begin{cases} 2,\ldots,k-1 \text{ if } \ell = k \\ 2,\ldots,k \quad \text{ if } \ell > k, \end{cases}$$

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

$$\text{var}(\hat{f}_i') = \hat{f}_i'(1 - \hat{f}_i') / K_i \quad , i = 1,\ldots,k,$$

$$\text{var}(\hat{f}_i'') = \hat{f}_i'' \, (1 - \hat{f}_i'') / M_i \quad , i = 1,\ldots,k,$$

$$\text{var}(\hat{f}_i'') = (\hat{f}_i'')^2 \left[\frac{1}{Y_{i\cdot} - Y_{ii}} - \frac{1}{K_i} + \frac{1}{Q_{i-1\cdot} - Q_{i-1,i-1} - Q_{i-1,i}} + \frac{1}{Q_{i-1,i}}\right] \quad , i = \begin{cases} 2,\ldots,k-1 \text{ if } \ell = k \\ 2,\ldots,k \quad \text{ if } \ell > k, \end{cases}$$

$$\begin{aligned} \text{var}(\hat{S}_i) = (\hat{S}_i)^2 & \left[\frac{1}{R_{i\cdot} - R_{ii}} - \frac{1}{N_i} + \frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right. \\ & \left. + \frac{1}{D_{i+1} - A_{i+1} - R_{i+1\cdot} + R_{i+1,i+1}} - \frac{1}{D_{i+1} - R_{i+1\cdot}}\right] \end{aligned} \quad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

$$\begin{aligned} \text{var}(\hat{S}_i') = (\hat{S}_i')^2 & \left[\frac{1}{Y_{i\cdot} - Y_{ii}} - \frac{1}{K_i} + \frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right. \\ & \left. + \frac{1}{D_{i+1} - A_{i+1} - R_{i+1\cdot} + R_{i+1,i+1}} - \frac{1}{D_{i+1} - R_{i+1\cdot}}\right] \end{aligned} \quad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k, \end{cases}$$

$$\text{var}(\hat{S}_i'') = (\hat{S}_i'')^2 \left[\frac{1}{Q_{i\cdot} - Q_{ii} - Q_{i,i+1}} - \frac{1}{M_i} + \frac{1}{Y_{i+1\cdot} - Y_{i+1,i+1}} - \frac{1}{K_{i+1}}\right] \quad , i = \begin{cases} 1,\ldots,k-2 \text{ if } \ell = k \\ 1,\ldots,k-1 \text{ if } \ell > k. \end{cases}$$

For a numerical illustration of the computational procedure, see Section 4.2. Use of the FORTRAN program is recommended to avoid the tedious calculations involved.

Comparison of the appropriate portions of Example 4.1 (i.e., 4.1b and 4.1f) shows that confidence intervals based on the H_6 estimators are considerably larger than those based on the corresponding H_4 estimators. This is to be expected since H_6 is the more general model with a larger number of parameters to be estimated. As discussed in a similar context in Chapters 2 and 3 this must be taken into account when deciding which model to choose for a given data set.

Sampling Covariances and Correlations

Estimates of the sampling covariances and correlations between the ML estimators are contained in the printout of program BROWNIE (see Example 4.1g). These are obtained as in Chapter 3 using the estimators defined below for all the nonzero large sample covariances. The estimation formulae are presented here mainly for reference purposes, and numerical illustrations are omitted.

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

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

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

$$\text{cov}(\hat{f}_i, \hat{f}_i'''') = -\hat{f}_i \hat{f}_i'''' / N_i \qquad , i = \begin{cases} 2, \ldots, k-1 & \text{if } \ell = k \\ 2, \ldots, k & \text{if } \ell > k, \end{cases}$$

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

$$\text{cov}(\hat{S}_i, \hat{S}_i') = \hat{S}_i \hat{S}_i' \left[\frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} + \frac{1}{D_{i+1} - A_{i+1} - R_{i+1\cdot} + R_{i+1,i+1}} - \frac{1}{D_{i+1} - R_{i+1\cdot}}\right] \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{S}_{i+1}, \hat{S}_i') = -\hat{S}_{i+1} \hat{S}_i' \left[\frac{1}{R_{i+1\cdot} - R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = \begin{cases} 1, \ldots, k-3 & \text{if } \ell = k \\ 1, \ldots, k-2 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_i', \hat{S}_i') = -\hat{f}_i' \hat{S}_i' / K_i \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_{i+1}', \hat{S}_i'') = \hat{f}_{i+1}' \hat{S}_i'' / K_{i+1} \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_i', \hat{f}_i''') = -\hat{f}_i' \hat{f}_i''' / K_i \qquad , i = \begin{cases} 2, \ldots, k-1 & \text{if } \ell = k \\ 2, \ldots, k & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{S}_{i+1}', \hat{S}_i'') = -\hat{S}_{i+1}' \hat{S}_i'' \left[\frac{1}{Y_{i+1\cdot} - Y_{i+1,i+1}} - \frac{1}{K_{i+1}}\right] \qquad , i = \begin{cases} 1, \ldots, k-3 & \text{if } \ell = k \\ 1, \ldots, k-2 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{S}_i', \hat{f}_i''') = \hat{S}_i' \hat{f}_i''' \left[\frac{1}{Y_{i\cdot} - Y_{ii}} - \frac{1}{K_i}\right] \qquad , i = \begin{cases} 2, \ldots, k-2 & \text{if } \ell = k \\ 2, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_i'', \hat{S}_i'') = -\hat{f}_i'' \hat{S}_i'' / M_i \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{S}_i'', \hat{f}_{i+1}''') = -\hat{S}_i'' \hat{f}_{i+1}''' \left[\frac{1}{Y_{i+1\cdot} - Y_{i+1,i+1}} - \frac{1}{K_{i+1}}\right] \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_i'''', \hat{S}_i) = -\hat{f}_i'''' \hat{S}_i / N_i \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_{i+1}'''', \hat{S}_i) = \hat{f}_{i+1}'''' \hat{S}_i / N_{i+1} \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_{i+1}'''', \hat{S}_i') = \hat{f}_{i+1}'''' \hat{S}_i' / N_{i+1} \qquad , i = \begin{cases} 1, \ldots, k-2 & \text{if } \ell = k \\ 1, \ldots, k-1 & \text{if } \ell > k. \end{cases}$$

Goodness of Fit Test

A goodness of fit test to H_6 is computed by the FORTRAN program in a manner analogous to that used in computing the goodness of fit test to H_3 (Section 3.6). Thus the printout, like that for H_3, consists of a series of contingency tables with resulting chi-square values and degrees of freedom. (The contingency tables are derived from the rows of the recovery matrices as described in Brownie 1973). Individual chi-square values and (separately) the degrees of freedom are summed to give a total chi-square with corresponding degrees of freedom.

In Example 4.1h, we find that for the synthetic data of Table 4.2 the total chi-square value testing goodness of fit for H_6 is 20.16 with 22 df (inadvertently not photographed for Example 4.1h). For real data, such a result would suggest that there is no indication that the assumptions of H_6 are incorrect.

4.4 The Model Under H_5

The model of this section is based on a parameterization which is intermediate in complexity between those of the models under H_4 and H_6. However, it is difficult to find a meaningful biological interpretation for this parameterization, and for this reason the model is considered last, and is included only because the related estimation and testing procedures have been coded in program BROWNIE.

To be consistent we define a hypothesis H_5, with assumptions 1 and 2 the same as those of H_4 and H_6, but with a third assumption that is artificial in that it is dictated by the parameterization and has not been derived from consideration of biological or ecological factors. The model structure, estimation formulae, etc., are presented without accompanying discussion or numerical illustrations.

The assumptions of H_5 are:

(1) Annual survival and recovery rates are year-specific;
(2) annual survival and recovery rates are different for young, subadults, and adults; and
(3) in any year, newly released subadults, and subadults that are survivors of young released the year before, have different recovery rates (namely f_i' and f_i''').

The expected numbers of recoveries under H_5 for $k=3$ and $\ell=5$ are shown in Table 4.5. Note that we cannot attribute the difference in the recovery rates for subadults to the effect of a different reporting rate for new releases, because then it would be logical to assume that the recovery rate for newly released adults is similarly affected, leading to assumption 3 of H_6.

Table 4.5. *Expected numbers of band recoveries under H_5 for a banding study with $k=3$, $\ell=5$, $s=2$.*

Year banded	Number banded	Year of recovery 1	2	3	4	5
				Birds banded and released as adults		
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	$N_2S_2f_3$	$N_2S_2S_3f_4$	$N_2S_2S_3S_4f_5$
3	N_3			N_3f_3	$N_3S_3f_4$	$N_3S_3S_4f_5$
				Birds banded and released as subadults		
1	K_1	K_1f_1'	$K_1S_1'f_2$	$K_1S_1'S_2f_3$	$K_1S_1'S_2S_3f_4$	$K_1S_1'S_2S_3S_4f_5$
2	K_2		K_2f_2'	$K_2S_2'f_3$	$K_2S_2'S_3f_4$	$K_2S_2'S_3S_4f_5$
3	K_3			K_3f_3'	$K_3S_3'f_4$	$K_3S_3'S_4f_5$
				Birds banded and released as young		
1	M_1	M_1f_1''	$M_1S_1''f_2'''$	$M_1S_1''S_2'f_3$	$M_1S_1''S_2'S_3f_4$	$M_1S_1''S_2'S_3S_4f_5$
2	M_2		M_2f_2''	$M_2S_2''f_3'''$	$M_2S_2''S_3'f_4$	$M_2S_2''S_3'S_4f_5$
3	M_3			M_3f_3''	$M_3S_3''f_4'''$	$M_3S_3''S_4'f_5$

Estimation of Parameters

ML estimators of individually estimable annual recovery and survival rates are:

$$\hat{f}_i = \frac{R_{i\cdot}}{N_i}\,\frac{A_i}{D_i} \qquad , i=1,\ldots,k,$$

$$\hat{f}_i' = \frac{Y_{ii}}{K_i} \qquad , i=1,\ldots,k,$$

$$\hat{f}_i'' = \frac{Q_{ii}}{M_i} \qquad , i=1,\ldots,k,$$

$$\hat{f}_i''' = \frac{Q_{i-1,i\cdot}}{Q_{i-1\cdot} - Q_{i-1,i-1} - Q_{i-1,i}} \; \frac{Y_{i\cdot} - Y_{ii}}{K_i} \qquad , i = \begin{cases} 2,\ldots,k-1 & \text{if } \ell = k \\ 2,\ldots,k & \text{if } \ell > k, \end{cases}$$

$$\hat{S}_i = \frac{R_{i\cdot}}{N_i} \; \frac{D_i - A_i}{D_i} \; \frac{N_{i+1}}{R_{i+1\cdot}} \qquad , i = 1,\ldots,k-1,$$

$$\hat{S}_i' = \frac{Y_{i\cdot} - Y_{ii}}{K_i} \; \frac{N_{i+1}}{R_{i+1\cdot}} \qquad , i = 1,\ldots,k-1,$$

$$\hat{S}_i'' = \frac{Q_{i\cdot} - Q_{ii} - Q_{i,i+1}}{M_i} \; \frac{K_{i+1}}{Y_{i+1\cdot} - Y_{i+1,i+1}} \qquad , i = \begin{cases} 1,\ldots,k-2 & \text{if } \ell = k \\ 1,\ldots,k-1 & \text{if } \ell > k. \end{cases}$$

Estimates obtained by evaluating the above formulae for the data of Table 4.2 are shown in Example 4.1d.

Sampling Variances, Standard Errors, and Confidence Intervals

Confidence intervals for parameters are obtained using the above ML estimators and the estimators of their sampling variances defined below,

$$\mathrm{var}(\hat{f}_i) = (\hat{f}_i)^2 \left[\frac{1}{R_{i\cdot}} - \frac{1}{N_i} + \frac{1}{A_i} - \frac{1}{D_i}\right] \qquad , i = 1,\ldots,k,$$

$$\mathrm{var}(\hat{S}_i) = (\hat{S}_i)^2 \left[\frac{1}{R_{i\cdot}} - \frac{1}{N_i} + \frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}} + \frac{1}{D_i - A_i} - \frac{1}{D_i}\right] \qquad , i = 1,\ldots,k-1,$$

$$\mathrm{var}(\hat{f}_i') = \hat{f}_i'(1 - \hat{f}_i')/K_i \qquad , i = 1,\ldots,k,$$

$$\mathrm{var}(\hat{S}_i') = (\hat{S}_i')^2 \left[\frac{1}{Y_{i\cdot} - Y_{ii}} - \frac{1}{K_i} + \frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1,\ldots,k-1,$$

$$\mathrm{var}(\hat{f}_i'') = \hat{f}_i''(1 - \hat{f}_i'')/M_i \qquad , i = 1,\ldots,k,$$

$$\mathrm{var}(\hat{S}_i'') = (\hat{S}_i'')^2 \left[\frac{1}{Q_{i\cdot} - Q_{ii} - Q_{i,i+1}} - \frac{1}{M_i} + \frac{1}{Y_{i+1\cdot} - Y_{i+1,i+1}} - \frac{1}{K_{i+1}}\right] \qquad , i = \begin{cases} 1,\ldots,k-2 & \text{if } \ell = k \\ 1,\ldots,k-1 & \text{if } \ell > k, \end{cases}$$

$$\mathrm{var}(\hat{f}_i''') = (\hat{f}_i''')^2 \left[\frac{1}{Y_{i\cdot} - Y_{ii}} - \frac{1}{K_i} + \frac{1}{Q_{i-1\cdot} - Q_{i-1,i-1} - Q_{i-1,i}} + \frac{1}{Q_{i-1,i\cdot}}\right] \qquad , i = \begin{cases} 2,\ldots,k-1 & \text{if } \ell = k \\ 2,\ldots,k & \text{if } \ell > k. \end{cases}$$

Sampling Covariances and Correlations

Estimators of the nonzero, large-sample covariances between the ML estimators are defined below. Estimates of the corresponding correlations are obtained in the usual way. Thus,

$$\mathrm{cov}(\hat{f}_i, \hat{S}_i) = \hat{f}_i \hat{S}_i \left[\frac{1}{R_{i\cdot}} - \frac{1}{N_i} - \frac{1}{D_i}\right] \qquad , i = 1,\ldots,k-1,$$

$$\mathrm{cov}(\hat{f}_{i+1}, \hat{S}_i) = -\hat{f}_{i+1} \hat{S}_i \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1,\ldots,k-1,$$

$$\mathrm{cov}(\hat{f}_{i+1}, \hat{S}_i') = \hat{f}_{i+1} \hat{S}_i' \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1,\ldots,k-1,$$

$$\mathrm{cov}(\hat{S}_i, \hat{S}_{i+1}) = -\hat{S}_i \hat{S}_{i+1} \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1,\ldots,k-2,$$

$$\mathrm{cov}(\hat{S}_{i+1}, \hat{S}_i') = -\hat{S}_{i+1} \hat{S}_i' \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1,\ldots,k-2,$$

$$\mathrm{cov}(\hat{S}_i, \hat{S}_i') = \hat{S}_i \hat{S}_i' \left[\frac{1}{R_{i+1\cdot}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1,\ldots,k-1,$$

$$\text{cov}(\hat{f}_i', \hat{S}_i') = -\hat{f}_i'\hat{S}_i' / K_i \qquad , i = 1, \ldots, k-1 ,$$

$$\text{cov}(\hat{f}_{i+1}', \hat{S}_i'') = \hat{f}_{i+1}'\hat{S}_i'' / K_{i+1} \qquad , i = \begin{cases} 1, \ldots, k-2 \text{ if } \ell = k \\ 1, \ldots, k-1 \text{ if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{f}_i', \hat{f}_i''') = -\hat{f}_i' \hat{f}_i''' / K_i \qquad , i = \begin{cases} 2, \ldots, k-1 \text{ if } \ell = k \\ 2, \ldots, k \quad \text{ if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{S}_{i+1}', \hat{S}_i'') = -\hat{S}_{i+1}'\hat{S}_i'' \left[\frac{1}{Y_{i+1\cdot} - Y_{i+1,i+1}} - \frac{1}{K_{i+1}}\right] \qquad , i = 1, \ldots, k-2 ,$$

$$\text{cov}(\hat{S}_i', \hat{f}_i''') = \hat{S}_i' \hat{f}_i''' \left[\frac{1}{Y_{i\cdot} - Y_{ii}} - \frac{1}{K_i}\right] \qquad , i = 2, \ldots, k-1 ,$$

$$\text{cov}(\hat{f}_i'', \hat{S}_i'') = -\hat{f}_i'' S_i'' / M_i \qquad , i = \begin{cases} 1, \ldots, k-2 \text{ if } \ell = k \\ 1, \ldots, k-1 \text{ if } \ell > k, \end{cases}$$

$$\text{cov}(\hat{S}_i'', \hat{f}_{i+1}''') = -\hat{S}_i'' \hat{f}_{i+1}''' \left[\frac{1}{Y_{i+1\cdot} - Y_{i+1,i+1}} - \frac{1}{K_{i+1}}\right] \qquad , i = \begin{cases} 1, \ldots, k-2 \text{ if } \ell = k \\ 1, \ldots, k-1 \text{ if } \ell > k . \end{cases}$$

Estimates of covariances and corresponding correlations obtained by evaluating the above formulae for the data of Table 4.2 are contained in Example 4.1e.

Goodness of Fit Test

A goodness of fit test to the model under H_5 is computed by the FORTRAN program in a manner similar to that used for the goodness of fit test to H_4. For Example 4.1h the chi-square value testing goodness of fit for H_5 is 22.11 with 26 df (inadvertently not photographed for Example 4.1h). Such a value is consistent with the assumptions of H_5.

4.5 Testing Between Models

As described in Chapters 2 and 3, it is important to choose the model which seems most appropriate for a given data set, in the sense that the model has sufficient parameters to provide an adequate description of the data, but not so many as to make estimation inefficient. This is accomplished by examining the results of goodness of fit tests to each model and the results of tests which compare one model with another. The latter tests are the subject of this section. As before, we always compare a simpler model with a more general alternative, and if the resulting chi-square value is significantly large, the simple model is rejected in favor of the more general one.

Because of the artificial nature of the H_5 assumptions, the important test is comparing H_4 with H_6. The corresponding test statistic is not printed out by the FORTRAN program, but is easily obtained from the test statistics for the tests of H_4 against H_5 and H_5 against H_6, as described below.

H_4 vs. H_5

The test statistic for comparing H_4 with H_5 is obtained as the sum of the single degrees of freedom chi-square statistics from each of the contingency tables

$$\begin{array}{|l:l|l} \hline Y_{ii} & Y_{i\cdot} - Y_{ii} & Y_{i\cdot} \\ \hdashline Q_{i-1,i} & Q_{i-1\cdot} - Q_{i-1,i-1} - Q_{i-1,i} & Q_{i-1\cdot} - Q_{i-1,i-1} \\ \hline & & Y_{i\cdot} + Q_{i-1\cdot} - Q_{i-1,i-1} \end{array} \qquad , i = \begin{cases} 2, \ldots, k-1 \text{ if } \ell = k \\ 2, \ldots, k \quad \text{ if } \ell > k, \end{cases}$$

The format of the printout of program BROWNIE is like that for the tests between the models of Chapter 3, as can be seen in Example 4.1h. Thus, individual contingency tables and chi-square values are printed, as well as the "TOTAL CHI-SQUARE" and corresponding degrees of freedom (which are $k-2$ if $\ell=k$ and $k-1$ if $\ell > k$).

As discussed before (see Section 3.7), information can be obtained by examining individual chi-square values as well as the total chi-square value. Each contingency table provides a test of the equality of f_i' and f_i''. Significantly large chi-square values are taken as evidence that inequality exists.

For the synthetic data of Table 4.2, which was generated using the model under H_4 as the probability model, the test results in a nonsignificant chi-square value of 4.322 with $k-2=4$ df (see Example 4.1h).

H_5 *vs.* H_6

This test is based on the contingency tables

R_{ii}	$R_{i\cdot}-R_{ii}$	$R_{i\cdot}$
A_i-R_{ii}	$D_i-R_{i\cdot}-A_i+R_{ii}$	$D_i-R_{i\cdot}$
		D_i

$$, i=\begin{cases}2,\dots,k-1 & \text{if } \ell=k\\ 2,\dots,k & \text{if } \ell>k,\end{cases}$$

and is otherwise analogous to the test described above. Similarly, the printout is like that for the test above (see Example 4.1h). The degrees of freedom for the total chi-square value are $k-2$ if $\ell=k$ and $k-1$ if $\ell>k$.

Each contingency table provides a test of the equality of f_i''' and f_i. Large chi-square values are taken as indicating that inequality exists, and that the H_6 assumptions are more appropriate than those of H_5.

For the synthetic data of Table 4.2, the total chi-square value is 1.947 with 4 df (Example 4.1h).

H_4 *vs* H_6

The test of H_4 against the alternative H_6 tests the validity of assumption 3 of H_4 against the alternative that new releases have a different reporting rate from that for survivors of earlier releases (assumption 3 of H_6). The chi-square test statistic is obtained as the sum of the test statistics for the two tests described above. The degrees of freedom, obtained by summing analogously, are $k-2+k-2=2k-4$ if $\ell=k$ and $k-1+k-1=2k-2$ if $\ell>k$. For example, using results of Example 4.1h, the chi-square value and degrees of freedom for the test of H_4 vs. H_6 for the data of Table 4.2 are $4.322+1.947=6.269$ and $4+4=8$, respectively. With such a result there is no reason to discard H_4 in favor of H_6.

Choosing between Chapter 4 models and their relationship to Chapter 3 models is discussed further in Section 4.6.

An Example

This example consists of the complete output from the FORTRAN program for analysis of the data of Table 4.2. The output is similar in format to that for the Chapter 3 models, but is not as fully documented. The output is self-explanatory with the exception of the following points.

The recovery matrices are printed out in the order – adults, young, subadults. This is different from the (more natural) way the data have been presented in Table 4.2.

The basic subtotals printed are also in a different order from those in Table 4.2. The subtotals labeled W(I), Z(I) and B(I) can be ignored. Those labeled SUMAB(I) are the same as those labeled D_i in Table 4.2 except for the k^{th} one.

An important point to note is that the program will not accept data sets with $\ell=k+1$ (i.e., $s=1$). An error message is printed if such a data set is read in, and no computations are performed.

As usual, estimates labeled F(I), S(I), etc., represent $\hat{f}_i$, $\hat{S}_i$, etc. Similarly COVAR(F(I), S(I)) represent cov$(\hat{f}_i, \hat{S}_i)$.

Estimates under H_6 labeled SK – 1 FK, S′K – 1 FK, S″K – 1 F‴K can be ignored. These are estimates of products of parameters which are of little interest.

Estimates of survival of over 100% and large confidence intervals have already been noted. Clearly, with adult recovery rates of approximately 5%, releases each year of 700-1,000 birds in each age class are not sufficient to provide reliable estimates of the different survival rates.

The results of the goodness of fit tests, as noted above, indicate that the simplest model (i.e., the model under H_4) seems adequate, and this is borne out by the tests between models. This is not surprising because, as stated earlier, the data were generated stochastically using H_4 as the probability model, with values of $f_i, S_i, f_i', S_i'; f_i'', S_i''$ which are not improbable. Comparison of the estimates (particularly the H_4 estimates) with the actual parameter values used in generating the data may be of interest; Table 4.6 gives the true parameters used in generating the data of Table 4.2. However, we emphasize that this data set represents a single sample and is not a simulation study on which we can base conclusions concerning precision and bias of estimators, power of tests, and so on.

Table 4.6. *Parameters used in generating the synthetic data used for illustrative purposes in Chapter 4.*

i	f_i	S_i	f_i'	S_i'	f_i''	S_i''
1	0.05	0.70	0.10	0.55	0.15	0.40
2	0.04	0.80	0.08	0.65	0.12	0.50
3	0.06	0.83	0.12	0.68	0.18	0.53
4	0.05	0.77	0.10	0.62	0.15	0.47
5	0.07	0.86	0.13	0.71	0.19	0.56
6	0.03	–	0.07	–	0.11	–
Averages	0.05	0.79	0.10	0.64	0.15	0.49

4.6 Summary

The methods of this chapter represent an extension of the methods of Chapters 2 and 3 to the more complex situation where young, subadults, and adults are thought to have different survival and recovery rates, and data are recorded separately for these three age groups. Similarly, analogous procedures can be developed for even more complex situations where data are recorded separately for four or more age classes and survival and recovery rates are specific to each age class. However, the limited applicability of the methods in Chapter 4 suggests that further extensions are of little value at this time, and the methods of Chapters 2 and 3 seem adequate for the analysis of most bird banding data we have encountered.

We have seen that as the complexity (or number of parameters) of the models increases, the number of birds which must be banded to obtain reliable estimates of survival also increases. This factor precludes application of the Chapter 4 methods to other larger game animals (e.g., deer) where age may be more easily determined but the tagging of thousands of animals is not feasible. Thus, in Example 4.1, where H_4 is correct, and adult recovery rates are approximately 5%, then banding 700 or more of each age class, i.e., over 2,000 birds annually, is not sufficient to yield confidence intervals of a reasonable length. If survival and recovery rates are appreciably different for subadults, then using a simpler model such as H_1 or H_2 will result in biased estimators (but shorter confidence intervals), and as usual there is a tradeoff between loss in accuracy and gain in precision. The H_1 estimates for the synthetic data of Table 4.2 are shown in Example 4.2 to illustrate this point. Comparison of the H_1 confidence intervals for adult survival rates with the true parameter values in Table 4.6 shows that only 1 of the 5 confidence intervals includes the true value.

Finally, if analysis of a given data set by Chapter 4 methods (three age classes) indicates that H_4 is preferable to H_6, it is possible that H_4 is too general and that a two-age-class model of Chapter 3 may be adequate. The data for adults and subadults should then be combined and a Chapter 3 analysis performed. (Alternatively, a Chapter 3 analysis could be carried out first and the Chapter 4 analysis used only if it is indicated). If H_1 and H_2 are rejected in favor of H_3, then the Chapter 4 methods are appropriate. However, note that H_3 and H_4 are not strictly comparable in the sense that one is more general than the other, because the H_3 parameterization allows the assumption that reporting rates are different for new releases whereas that of H_4 does not. Nevertheless, rejection of H_1 and H_2 in favor of H_3 is one indication that subadults have different survival and recovery rates and that a Chapter 4 model may be appropriate.

This is illustrated in Example 4.2 which consists of portions of a printout obtained by combining the data for adults and subadults in Example 4.1, and analyzing the combined data and the data for young using the two-age-class models of Chapter 3. The combined recovery data for adults and subadults are printed as the "INPUT MATRIX" for adults. Note that the chi-square values for H_1 vs. H_2 and for H_2 vs. H_3 are both significantly large and H_1 and H_2 are thus rejected in favor of H_3. Also, the H_3 model is the only one for which the goodness of fit test is not significant ($\chi^2 = 11.28$, df = 12). This confirms that, as indicated in Example 4.1, the H_4 model is appropriate for these data.

Example 4.1a

```
                  SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE
ADULTS        INPUT MATRIX

   1   700.   36.   20.   26.   11.   17.    6.
   2   700.    0.   23.   31.   22.   25.   16.
   3   700.    0.    0.   42.   24.   23.   11.
   4   700.    0.    0.    0.   38.   37.   20.
   5   700.    0.    0.    0.    0.   38.   19.
   6   700.    0.    0.    0.    0.    0.   23.
YOUNG         INPUT MATRIX

   1  1000.  161.   34.   21.    7.   18.    4.
   2  1000.    0.  109.   73.   12.   16.    7.
   3  1000.    0.    0.  171.   51.   26.   13.
   4  1000.    0.    0.    0.  157.   57.   10.
   5  1000.    0.    0.    0.    0.  198.   39.
   6  1000.    0.    0.    0.    0.    0.  102.
SUBADULTS     INPUT MATRIX

   1   800.   80.   12.   20.   12.   19.    3.
   2   800.    0.   57.   33.   17.   34.    7.
   3   800.    0.    0.   83.   26.   25.   13.
   4   800.    0.    0.    0.   64.   25.   14.
   5   800.    0.    0.    0.    0.  115.   26.
   6   800.    0.    0.    0.    0.    0.   57.


BASIC SUBTOTALS

 I  RROW(I)   RCOL(I)   QROW(I)   QCOL(I)      T(I)      U(I)      W(I)      Z(I)   YROW(I)   YCOL(I)      V(I)      A(I)      B(I) SUMAB(I)
 1   116.00     36.00    245.00    161.00    116.00    245.00     36.00     0.0      146.00     80.00    146.00     36.00     0.0     116.00
 2   117.00     43.00    217.00    143.00    197.00    301.00     77.00     80.00    148.00     69.00    214.00     55.00     80.00    263.00
 3   100.00     99.00    261.00    265.00    254.00    419.00    193.00    204.00    147.00    136.00    292.00    173.00    208.00    449.00
 4    95.00     95.00    224.00    227.00    250.00    378.00    165.00    219.00    103.00    119.00    259.00    169.00    276.00    470.00
 5    57.00    140.00    237.00    315.00    212.00    388.00    257.00    239.00    141.00    218.00    281.00    303.00    301.00    436.00
 6    23.00     95.00    102.00    175.00     95.00    175.00    168.00    106.00     57.00    120.00    120.00    192.00    133.00      0.0
```

Example 4.1b

```
SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE

ESTIMATES UNDER H4
-------------------


                              F(I)                                                     S(I)
       -------------------------------------------------------------      ----------------------------------------------------------
 I   YR  ESTIMATE           STANDARD           95% CONFIDENCE           ESTIMATE          STANDARD            95% CONFIDENCE
         --------           __ERROR__          ___INTERVAL___           --------          __ERROR__           ___INTERVAL___

 1    1  0.0514              0.0083          0.0351 -   0.0678           0.6838            0.0922           0.5030 -   0.8645

 2    2  0.0350              0.0051          0.0249 -   0.0450           0.9253            0.1196           0.6910 -   1.1597

 3    3  0.0550              0.0061          0.0432 -   0.0669           0.6471            0.0893           0.4719 -   0.8222

 4    4  0.0488              0.0055          0.0379 -   0.0597           1.0674            0.1735           0.7274 -   1.4073

 5    5  0.0566              0.0074          0.0421 -   0.0711           0.7560            0.1903           0.3829 -   1.1290

 6    6  0.0329              0.0067          0.0197 -   0.0461

         AVERAGE            STANDARD           95% CONFIDENCE           AVERAGE           STANDARD            95% CONFIDENCE
         ESTIMATE           ERROR___           ___INTERVAL___           ESTIMATE          ERROR___            ___INTERVAL___

   F-    = 0.0466            0.0027           0.0413 -   0.0519    S-    = 0.8159          0.0396           0.7382 -   0.8936

                              F'(I)                                                    S'(I)
       -------------------------------------------------------------      ----------------------------------------------------------
 I   YR  ESTIMATE           STANDARD           95% CONFIDENCE           ESTIMATE          STANDARD            95% CONFIDENCE
         --------           __ERROR__          ___INTERVAL___           --------          __ERROR__           ___INTERVAL___

 1    1  0.1000              0.0106          0.0792 -   0.1208           0.4936            0.0716           0.3533 -   0.6339

 2    2  0.0726              0.0080          0.0569 -   0.0883           0.7870            0.1022           0.5867 -   0.9873

 3    3  0.1124              0.0101          0.0927 -   0.1322           0.5256            0.0759           0.3770 -   0.6743

 4    4  0.0767              0.0084          0.0603 -   0.0932           0.6390            0.1147           0.4142 -   0.8638

 5    5  0.1457              0.0121          0.1221 -   0.1694           0.9284            0.2471           0.4440 -   1.4128

 6    6  0.0712              0.0091          0.0534 -   0.0891

         AVERAGE            STANDARD           95% CONFIDENCE           AVERAGE           STANDARD            95% CONFIDENCE
         ESTIMATE           ERROR___           ___INTERVAL___           ESTIMATE          ERROR___            ___INTERVAL___

   F-'   = 0.0964            0.0040           0.0886 -   0.1043    S-'   = 0.6747          0.0618           0.5536 -   0.7959

                              F''(I)                                                   S''(I)
       -------------------------------------------------------------      ----------------------------------------------------------
 I   YR  ESTIMATE           STANDARD           95% CONFIDENCE           ESTIMATE          STANDARD            95% CONFIDENCE
         --------           __ERROR__          ___INTERVAL___           --------          __ERROR__           ___INTERVAL___

 1    1  0.1610              0.0116          0.1382 -   0.1838           0.4541            0.0582           0.3400 -   0.5681

 2    2  0.1090              0.0099          0.0897 -   0.1283           0.5878            0.0691           0.4524 -   0.7231

 3    3  0.1710              0.0119          0.1477 -   0.1943           0.6990            0.0953           0.5123 -   0.8857

 4    4  0.1570              0.0115          0.1345 -   0.1795           0.3801            0.0534           0.2754 -   0.4849

 5    5  0.1980              0.0126          0.1733 -   0.2227           0.5474            0.1107           0.3303 -   0.7644

 6    6  0.1020              0.0096          0.0832 -   0.1208

         AVERAGE            STANDARD           95% CONFIDENCE           AVERAGE           STANDARD            95% CONFIDENCE
         ESTIMATE           ERROR___           ___INTERVAL___           ESTIMATE          ERROR___            ___INTERVAL___

   F-''  = 0.1497            0.0046           0.1407 -   0.1587    S-''  = 0.5337          0.0360           0.4632 -   0.6042
```

Example 4.1c

```
SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H4


 I      YR   COVAR(F(I),S(I))               CORR(F(I),S(I))               COVAR(F(I+1),S(I))             CORR(F(I+1),S(I))
 1      1         -0.000050236                   -0.065253606                 -0.000170131                   -0.359974113
 2      2          0.000107256                    0.175040771                 -0.000436564                   -0.602452058
 3      3          0.000225954                    0.417278723                 -0.000287267                   -0.580407447
 4      4          0.000363052                    0.377827136                 -0.000973392                   -0.757871878
 5      5          0.000591297                    0.419539670                 -0.001044490                   -0.814461397


 I      YR   COVAR(F(I+1),S'(I))            CORR(F(I+1),S'(I))            COVAR(S(I),S(I+1))             CORR(S(I),S(I+1))
 1      1         -0.000122813                   -0.334865286                 -0.004503824                   -0.408482207
 2      2         -0.000371326                   -0.599501246                 -0.005131979                   -0.480436117
 3      3         -0.000233370                   -0.555277654                 -0.006283324                   -0.405460569
 4      4         -0.000582744                   -0.686188617                 -0.013003692                   -0.393878819
 5      5         -0.001282716                   -0.770326529


 I      YR   COVAR(S(I),S'(I))              CORR(S(I),S'(I))              COVAR(S(I+1),S'(I))            CORR(S(I+1),S'(I))
 1      1          0.002402449                    0.364044173                 -0.003251200                   -0.379989949
 2      2          0.006242327                    0.510889461                 -0.004365079                   -0.478082769
 3      3          0.003094347                    0.456530305                 -0.005104426                   -0.387904958
 4      4          0.010991640                    0.552531049                 -0.007784970                   -0.356623619
 5      5          0.029512953                    0.627401268


 I      YR   COVAR(F'(I),S'(I))             CORR(F'(I),S'(I))             COVAR(F'(I+1),S''(I))          CORR(F'(I+1),S''(I))
 1      1         -0.000061699                   -0.081286439                 -0.000181438                   -0.389404703
 2      2          0.000068329                    0.083470093                 -0.000366871                   -0.526880039
 3      3          0.000096383                    0.126032913                 -0.000453615                   -0.567282106
 4      4          0.000160665                    0.166881758                 -0.000323680                   -0.502114632
 5      5          0.000139979                    0.046960028                 -0.000635460                   -0.630907169


 I      YR   COVAR(S'(I+1),S''(I)),         CORR(S'(I+1),S''(I)),         COVAR(F''(I),S''(I))           CORR(F''(I),S''(I))
 1      1         -0.001967903                   -0.331048945                 -0.000073103                   -0.108132806
 2      2         -0.001715530                   -0.327366480                 -0.000064065                   -0.094112224
 3      3         -0.003778400                   -0.345847340                 -0.000119534                   -0.105394170
 4      4         -0.002061862                   -0.156096394                 -0.000059682                   -0.097064273
 5      5                                                                     -0.000108379                   -0.077660210

       THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.
```

Example 4.1d

```
SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE

ESTIMATES UNDER H5
------------------

                         F(I)                                                  S(I)
 --------------------------------------------------------------   ---------------------------------------------------------
I   YR  ESTIMATE          STANDARD          95% CONFIDENCE          ESTIMATE          STANDARD          95% CONFIDENCE
        --------          __ERROR__         ___INTERVAL___          --------          __ERROR__         ___INTERVAL___

1    1  0.0514             0.0083          0.0351 -  0.0678          0.6838            0.0922          0.5030 -  0.8645

2    2  0.0350             0.0051          0.0249 -  0.0450          0.9253            0.1196          0.6910 -  1.1597

3    3  0.0550             0.0061          0.0432 -  0.0669          0.6471            0.0893          0.4719 -  0.8222

4    4  0.0488             0.0055          0.0379 -  0.0597          1.0674            0.1735          0.7274 -  1.4073

5    5  0.0566             0.0074          0.0421 -  0.0711          0.7560            0.1903          0.3829 -  1.1290

6    6  0.0329             0.0067          0.0197 -  0.0461

        AVERAGE           STANDARD          95% CONFIDENCE          AVERAGE           STANDARD          95% CONFIDENCE
        ESTIMATE          ERROR__           ___INTERVAL___          ESTIMATE          ERROR__           ___INTERVAL___

  F-     = 0.0466          0.0027           0.0413 -  0.0519   S-   = 0.8159            0.0396           0.7382 -  0.8936

                         F'(I)                                                 S'(I)
 --------------------------------------------------------------   ---------------------------------------------------------
I   YR  ESTIMATE          STANDARD          95% CONFIDENCE          ESTIMATE          STANDARD          95% CONFIDENCE
        --------          __ERROR__         ___INTERVAL___          --------          __ERROR__         ___INTERVAL___

1    1  0.1000             0.0106          0.0792 -  0.1208          0.4936            0.0716          0.3533 -  0.6339

2    2  0.0712             0.0091          0.0534 -  0.0891          0.7962            0.1077          0.5851 -  1.0074

3    3  0.1037             0.0108          0.0826 -  0.1249          0.5895            0.0903          0.4125 -  0.7665

4    4  0.0800             0.0096          0.0612 -  0.0988          0.5987            0.1205          0.3625 -  0.8348

5    5  0.1437             0.0124          0.1194 -  0.1681          0.9891            0.2785          0.4433 -  1.5349

6    6  0.0712             0.0091          0.0534 -  0.0891

        AVERAGE           STANDARD          95% CONFIDENCE          AVERAGE           STANDARD          95% CONFIDENCE
        ESTIMATE          ERROR__           ___INTERVAL___          ESTIMATE          ERROR__           ___INTERVAL___

  F-'    = 0.0950          0.0042           0.0867 -  0.1033   S-'  = 0.6934            0.0684           0.5594 -  0.8275

                         F''(I)                                                S''(I)
 --------------------------------------------------------------   ---------------------------------------------------------
I   YR  ESTIMATE          STANDARD          95% CONFIDENCE          ESTIMATE          STANDARD          95% CONFIDENCE
        --------          __ERROR__         ___INTERVAL___          --------          __ERROR__         ___INTERVAL___

1    1  0.1610             0.0116          0.1382 -  0.1838          0.4396            0.0745          0.2935 -  0.5856

2    2  0.1090             0.0099          0.0897 -  0.1283          0.4375            0.0896          0.2619 -  0.6131

3    3  0.1710             0.0119          0.1477 -  0.1943          0.8000            0.1771          0.4528 -  1.1472

4    4  0.1570             0.0115          0.1345 -  0.1795          0.3077            0.1136          0.0851 -  0.5303

5    5  0.1980             0.0126          0.1733 -  0.2227

6    6  0.1020             0.0096          0.0832 -  0.1208

        AVERAGE           STANDARD          95% CONFIDENCE          AVERAGE           STANDARD          95% CONFIDENCE
        ESTIMATE          ERROR__           ___INTERVAL___          ESTIMATE          ERROR__           ___INTERVAL___

  F-''   = 0.1497          0.0046           0.1407 -  0.1587   S-'' = 0.4962            0.0601           0.3783 -  0.6141

                         F'''(I)                                               S''K-1 F'''K
 --------------------------------------------------------------   ---------------------------------------------------------
I   YR  ESTIMATE          STANDARD          95% CONFIDENCE          ESTIMATE          STANDARD          95% CONFIDENCE
        --------          __ERROR__         ___INTERVAL___          --------          __ERROR__         ___INTERVAL___

1    1                                                               0.0390            0.0061          0.0270 -  0.0510

2    2  0.0773             0.0188          0.0405 -  0.1142

3    3  0.1669             0.0397          0.0890 -  0.2447

4    4  0.0637             0.0168          0.0308 -  0.0967

5    5  0.1853             0.0729          0.0424 -  0.3281

        AVERAGE           STANDARD          95% CONFIDENCE
        ESTIMATE          ERROR__           ___INTERVAL___

  F-'''  = 0.1233          0.0217           0.0808 -  0.1658
```

Example 4.1e

```
SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H5


I       YR  COVAR(F(I),S(I))              CORR(F(I),S(I))               COVAR(F(I+1),S(I))            CORR(F(I+1),S(I))
1        1      -0.000050236                  -0.065253596                  -0.000170131                  -0.359974129
2        2       0.000107256                   0.175040771                  -0.000436564                  -0.602452058
3        3       0.000225955                   0.417278884                  -0.000287268                  -0.580407783
4        4       0.000363052                   0.377827136                  -0.000973392                  -0.757871878
5        5       0.000591298                   0.419539808                  -0.001044490                  -0.814461310


I       YR  COVAR(F(I+1),S'(I))           CORR(F(I+1),S'(I))            COVAR(S(I),S(I+1))            CORR(S(I),S(I+1))
1        1      -0.000122813                  -0.334865325                  -0.004503824                  -0.408482036
2        2      -0.000375668                  -0.575267452                  -0.005131982                  -0.480436343
3        3      -0.000261705                  -0.523084352                  -0.006283324                  -0.405460471
4        4      -0.000545969                  -0.611926022                  -0.013003699                  -0.393878834
5        5      -0.001366614                  -0.728368381


I       YR  COVAR(S(I+1),S'(I))           CORR(S(I+1),S'(I))            COVAR(S(I),S'(I))             CORR(S(I),S'(I))
1        1      -0.003251200                  -0.379989949                   0.002402450                   0.364044263
2        2      -0.004416123                  -0.458757424                   0.006315321                   0.490237920
3        3      -0.005724199                  -0.365415640                   0.003470059                   0.430062293
4        4      -0.007293694                  -0.318028380                   0.010297995                   0.492733690
5        5                                                                   0.031443290                   0.593228061


I       YR  COVAR(F'(I),S'(I))            CORR(F'(I),S'(I))             COVAR(F'(I+1),S''(I))         CORR(F'(I+1),S''(I))
1        1      -0.000061699                  -0.081286439                   0.000039148                   0.057764651
2        2      -0.000070916                  -0.072368354                   0.000056738                   0.058733650
3        3      -0.000076447                  -0.078514990                   0.000080000                   0.047083209
4        4      -0.000059868                  -0.051801824                   0.000055288                   0.039251102
5        5      -0.000177734                  -0.051455091


I       YR  COVAR(F'(I),F'''(I))          CORR(F'(I),F'''(I))           COVAR(S'(I+1),S''(I))         CORR(S'(I+1),S''(I))
1        1                                                                  -0.003408651                  -0.424550718
2        2      -0.000006889                  -0.040264081                  -0.003707234                  -0.458119236
3        3      -0.000021639                  -0.050542093                  -0.011682007                  -0.547304148
4        4      -0.000006375                  -0.039507277                  -0.011325236                  -0.358129930
5        5      -0.000033287                  -0.036824414


I       YR  COVAR(S'(I),F'''(I))          CORR(S'(I),F'''(I))           COVAR(F''(I),S''(I))          CORR(F''(I),S''(I))
1        1                                                                  -0.000070769                  -0.081713908
2        2       0.000599824                   0.295927492                  -0.000047687                  -0.054004120
3        3       0.001413892                   0.394225527                  -0.000136800                  -0.064860548
4        4       0.000930910                   0.459240230                  -0.000048308                  -0.036976801
5        5       0.006818503                   0.335988627


I       YR  COVAR(S''(I),F'''(I+1)        CORR(S''(I),F'''(I+1)
1        1      -0.000331126                  -0.236210267
2        2      -0.001049373                  -0.294903029
3        3      -0.001243941                  -0.417408350
4        4      -0.002121057                  -0.256299700


          THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.
```

Example 4.1f

```
SYNTHETIC DATA GENERATED FCR THE THREE AGE CLASS CASE

ESTIMATES UNDER H6
------------------


                    F(I)                                                    S(I)
      ----------------------------------------------------        ----------------------------------------------------
I  YR ESTIMATE         STANDARD             95% CONFIDENCE        ESTIMATE         STANDARD             95% CONFIDENCE
      --------         __ERROR__            ___INTERVAL___        --------         __ERROR__            ___INTERVAL___

1   1                                                              0.6645           0.0990           0.4704 -  0.8586

2   2 0.0377            0.0084            0.0213 -  0.0541         1.0123           0.1656           0.6879 -  1.3368

3   3 0.0498            0.0083            C.0335 -  0.0661         0.6621           0.1209           0.4251 -  0.8991

4   4 0.0437            0.0073            C.0294 -  0.0580         0.9024           0.2446           0.4230 -  1.3817

5   5 0.0631            0.0159            0.0319 -  0.0943

      AVERAGE          STANDARD             95% CONFIDENCE        AVERAGE          STANDARD             95% CONFIDENCE
      ESTIMATE         ERRCR___             ___INTERVAL___        ESTIMATE         ERROR___             ___INTERVAL___

 F̄     = 0.0486         0.0053             0.0382 -  0.0589   S̄    = 0.8103           0.0606           0.6915 -  0.9292

                    F'(I)                                                   S'(I)
      ----------------------------------------------------        ----------------------------------------------------
I  YR ESTIMATE         STANDARD             95% CONFIDENCE        ESTIMATE         STANDARD             95% CONFIDENCE
      --------         __ERROR__            ___INTERVAL___        --------         __ERROR__            ___INTERVAL___

1   1 0.1000            0.0106            0.0792 -  0.1208         0.4797           0.0759           0.3309 -  C.6285

2   2 0.0712            0.0091            C.0534 -  0.0891         0.8575           0.1416           0.5800 -  1.1351

3   3 0.1037            0.0108            C.0826 -  0.1249         0.6393           0.1142           0.4154 -  0.8631

4   4 0.0800            0.0096            0.0612 -  0.0988         0.5402           0.1544           0.2375 -  0.8429

5   5 0.1437            0.0124            C.1194 -  0.1681

6   6 C.0712            0.0091            C.0534 -  0.0891

      AVERAGE          STANDARD             95% CONFIDENCE        AVERAGE          STANDARD             95% CONFIDENCE
      ESTIMATE         ERROR___             ___INTERVAL___        ESTIMATE         ERROR___             ___INTERVAL___

 F̄'    = 0.0950         0.0042             0.0867 -  0.1033   S̄'   = 0.6292           0.0626           0.5065 -  0.7519

                    F''(I)                                                  S''(I)
      ----------------------------------------------------        ----------------------------------------------------
I  YR ESTIMATE         STANDARD             95% CONFIDENCE        ESTIMATE         STANDARD             95% CONFIDENCE
      --------         __ERROR__            ___INTERVAL___        --------         __ERROR__            ___INTERVAL___

1   1 0.1610            0.0116            C.1382 -  0.1838         0.4396           0.0745           0.2935 -  C.5856

2   2 0.1090            0.0099            0.0897 -  0.1283         0.4375           0.0896           0.2619 -  0.6131

3   3 0.1710            0.0119            C.1477 -  0.1943         C.8000           0.1771           0.4528 -  1.1472

4   4 0.1570            0.0115            C.1345 -  0.1795         0.3077           0.1136           0.0851 -  0.5303

5   5 0.1980            0.0126            C.1733 -  0.2227

6   6 0.1020            0.0096            C.0832 -  0.1208

      AVERAGE          STANDARD             95% CONFIDENCE        AVERAGE          STANDARD             95% CONFIDENCE
      ESTIMATE         ERRCR___             ___INTERVAL___        ESTIMATE         ERROR___             ___INTERVAL___

 F̄''   = 0.1497         0.0046             0.1407 -  0.1587   S̄''  = 0.4962           0.0601           0.3783 -  0.6141

                    F'''(I)                                                 F''''(I)
      ----------------------------------------------------        ----------------------------------------------------
I  YR ESTIMATE         STANDARD             95% CONFIDENCE        ESTIMATE         STANDARD             95% CONFIDENCE
      --------         __ERROR__            ___INTERVAL___        --------         __ERROR__            ___INTERVAL___

1   1                                                              0.0514           0.0083           0.0351 -  0.0678

2   2 0.0773            0.0188            0.0405 -  0.1142         0.0329           0.0067           0.0197 -  0.0461

3   3 0.1669            0.0397            0.0890 -  0.2447         0.0600           0.0090           0.0424 -  0.0776

4   4 0.0637            0.0168            0.0308 -  0.0967         0.0543           0.0086           0.0375 -  0.0711

5   5 0.1852            0.0729            0.0424 -  0.3281         0.0543           0.0086           0.0375 -  0.0711

6   6                                                              0.0329           0.0067           0.0197 -  0.0461

      AVERAGE          STANDARD             95% CONFIDENCE        AVERAGE          STANDARD             95% CONFIDENCE
      ESTIMATE         ERRCR___             ___INTERVAL___        ESTIMATE         ERROR___             ___INTERVAL___

 F̄'''  = 0.1233         0.0217             0.0808 -  0.1658   F̄''''= 0.0476          0.0033           0.0412 -  0.0541

                    SK-1 FK                                                 S'K-1 FK
      ----------------------------------------------------        ----------------------------------------------------
I  YR ESTIMATE         STANDARD             95% CONFIDENCE        ESTIMATE         STANDARD             95% CONFIDENCE
      --------         __ERROR__            ___INTERVAL___        --------         __ERROR__            ___INTERVAL___

1 ****  0.0271          0.0061            C.0151 -  0.0392         0.0325           0.0063           0.0202 -  0.0448
                    S''K-1 F'''K
      ----------------------------------------------------        ----------------------------------------------------
I  YR ESTIMATE         STANDARD             95% CONFIDENCE        ESTIMATE         STANDARD             95% CONFIDENCE
      --------         __ERROR__            ___INTERVAL___        --------         __ERROR__            ___INTERVAL___

1 ****  0.0390          0.0061            0.0270 -  0.0510
```

Example 4.1g

```
SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H6


I       YR  COVAR(F(I),S(I))              CORR(F(I),S(I))               COVAR(F(I+1),S(I))            CORR(F(I+1),S(I))
1        1                                                                  -0.000230693                  -0.278555120
2        2        0.000351440                   0.253809794                 -0.000797047                  -0.577502687
3        3        0.000521272                   0.517080767                 -0.000466452                  -0.528731567
4        4        0.000635744                   0.356325046                 -0.002915267                  -0.748273085


I       YR  COVAR(F(I+1),S'(I))           CORR(F(I+1),S'(I))            COVAR(S(I),S(I+1))            CORR(S(I),S(I+1))
1        1       -0.000166532                  -0.262337648                 -0.006195705                  -0.377935379
2        2       -0.000675158                  -0.571871189                 -0.010598645                  -0.529396925
3        3       -0.000450368                  -0.540507844                 -0.009627990                  -0.325556066
4        4       -0.001745325                  -0.709376989


I       YR  COVAR(S(I),S'(I))             CORR(S(I),S'(I))              COVAR(S(I+1),S'(I))           CORR(S(I+1),S'(I))
1        1        0.003548755                   0.472181863                 -0.004472524                  -0.355931937
2        2        0.015222251                   0.649245835                 -0.008977845                  -0.524234652
3        3        0.007426511                   0.537706946                 -0.009295996                  -0.332807022
4        4        0.027951255                   0.740024763


I       YR  COVAR(F'(I),S'(I))            CORR(F'(I),S'(I))             COVAR(F'(I+1),S''(I))         CORR(F'(I+1),S''(I))
1        1       -0.000059963                  -0.074485214                  0.000039148                   0.057764651
2        2       -0.000076374                  -0.059296152                  0.000056738                   0.058733650
3        3       -0.000082903                  -0.067327013                  0.000080000                   0.047083209
4        4       -0.000054024                  -0.036468897                  0.000055288                   0.039251102


I       YR  COVAR(F'(I),F'''(I))          CORR(F'(I),F'''(I))           COVAR(S'(I+1),S''(I))         CORR(S'(I+1),S''(I))
1        1                                                                  -0.003671008                  -0.347862263
2        2       -0.000006889                  -0.040264081                 -0.004020289                  -0.392839777
3        3       -0.000021639                  -0.050542093                 -0.010541543                  -0.385306504
4        4       -0.000006375                  -0.039507297
5        5       -0.000033287                  -0.036824411


I       YR  COVAR(S'(I),F'''(I))          CORR(S'(I),F'''(I))           COVAR(F''(I),S''(I))          CORR(F''(I),S''(I))
1        1                                                                  -0.000070769                  -0.081713908
2        2        0.000645991                   0.242472765                 -0.000047687                  -0.054004120
3        3        0.001533288                   0.338050750                 -0.000136800                  -0.064860548
4        4        0.000840028                   0.323308882                 -0.000048308                  -0.036976801


I       YR  COVAR(S''(I),F'''(I+1))       CORR(S''(I),F'''(I+1))        COVAR(F''''(I),F(I))          CORR(F''''(I),F(I))
1        1       -0.000331126                  -0.236210267
2        2       -0.001049373                  -0.294903029                 -0.000001769                  -0.031397738
3        3       -0.001243940                  -0.417408472                 -0.000004268                  -0.057032674
4        4       -0.002121056                  -0.256299725                 -0.000003390                  -0.054265385
5        5                                                                  -0.000004893                  -0.035865676


I       YR  COVAR(F''''(I),S(I))          CORR(F''''(I),S(I))           COVAR(F''''(I+1),S(I))        CORR(F''''(I+1),S(I))
1        1       -0.000048823                  -0.059061558                  0.000031192                   0.046752769
2        2       -0.000047518                  -0.042599510                  0.000086773                   0.058391125
3        3       -0.000056750                  -0.052281857                  0.000051345                   0.049579505
4        4       -0.000069980                  -0.033412822                  0.000069980                   0.033412822


I       YR  COVAR(F''''(I+1),S'(1))       CORR(F''''(I+1),S'(1))
1        1        0.000022517                   0.044030805
2        2        0.000073503                   0.057821740
3        3        0.000049575                   0.050683783
4        4        0.000041896                   0.031675984


       THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.
```

Example 4.1h

```
SYNTHETIC DATA GENERATED FOR THE THREE AGE CLASS CASE

CHI-SQUARE TEST OF H4 VS H5
---------------------------------


     2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                               WITH 1 DEGREE OF FREEDOM
     -----------------------   ----------------------------------

I= 2         57      91
             34      50                            0.087

I= 3         83      64
             73      35                            3.247

I= 4         64      39
             51      39                            0.597

I= 5        115      26
             57      10                            0.392
|--------------------------------------------------------|
|TOTAL CHI-SQUARE WITH  4 DEGREES OF FREEDOM =     4.322|
|--------------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN     4.32  = 0.36413


CHI-SQUARE TEST OF H5 VS H6
---------------------------------


     2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                               WITH 1 DEGREE OF FREEDOM
     -----------------------   ----------------------------------

I= 2         23      94
             32     114                            0.201

I= 3         42      58
            131     218                            0.654

I= 4         38      57
            131     244                            0.845

I= 5         38      19
            265     114                            0.248
|--------------------------------------------------------|
|TOTAL CHI-SQUARE WITH  4 DEGREES OF FREEDOM =     1.947|
|--------------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN     1.95  = 0.74550



CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H6
-----------------------------------------------------------

                      CONTINGENCY TABLES                                                        CORRESPONDING CHI-SQUARE
                                                                                                STATISTICS AND DEGREES OF FREEDOM
                      ------------------                                                        ---------------------------------

            I= 1
                        6. 11. 20. 17. 26.
                        3. 12. 12. 19. 20.                                                        2.62 WITH  4 D.F.

            I= 2
                       16. 22. 25. 31.
                        7. 17. 34. 33.
                        4.  7. 18. 21.
                        9. 23. 36. 46.                                                           10.11 WITH  9 D.F.

            I= 3
                       11. 24. 23.
                       13. 26. 25.
                        7. 12. 16.
                       36. 69.113.                                                                5.13 WITH  6 D.F.

            I= 4
                       20. 37.
                       14. 25.
                       13. 26.
                       67.177.                                                                    2.30 WITH  3 D.F.



CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H4
-----------------------------------------------------------

                                          |-----------------------------------------------|
                                          |TOTAL CHI-SQUARE      26.43 WITH 30 D.F.|
                                          |-----------------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN    26.43  = 0.65322
```

Example 4.2a

```
                    SYNTHETIC DATA GENERATED FOR 3 AGE CLASSES WHERE ADULTS & SUBADULTS ARE GROUPED

ADULTS       INPUT MATRIX

  1  1500.  116.   32.   46.   23.   36.    9.
  2  1500.    0.   80.   64.   39.   59.   23.
  3  1500.    0.    0.  125.   50.   48.   24.
  4  1500.    0.    0.    0.  102.   62.   34.
  5  1500.    0.    0.    0.    0.  153.   45.
  6  1500.    0.    0.    0.    0.    0.   80.

YOUNG        INPUT MATRIX

  1  1000.  161.   34.   21.    7.   18.    4.
  2  1000.    0.  109.   73.   12.   16.    7.
  3  1000.    0.    0.  171.   51.   26.   13.
  4  1000.    0.    0.    0.  157.   57.   10.
  5  1000.    0.    0.    0.    0.  198.   39.
  6  1000.    0.    0.    0.    0.    0.  102.


THE HYPOTHESIS H1.   (SEE BROWNIE AND ROBSON, 1974.  CORNELL BIOMETRICS UNIT PAPER NO. BU-514-M)
------------------

ASSUMPTIONS:  (1)  ANNUAL SURVIVAL AND RECOVERY RATES ARE YEAR-SPECIFIC.
------------
              (2)  YOUNG BIRDS HAVE DIFFERENT SURVIVAL AND RECOVERY RATES FROM THOSE OF ADULTS.


PARAMETERS:
-----------
F(I)   = BAND RECOVERY RATE FOR ADULTS IN YEAR I.
S(I)   = SURVIVAL RATE FOR ADULTS IN YEAR I.
F'(I)  = BAND RECOVERY RATE FOR YOUNG IN YEAR I.
S'(I)  = SURVIVAL RATE FOR YOUNG IN YEAR I.


STRUCTURE OF THE MODEL UNDER H1  (IN TERMS OF EXPECTED NUMBERS OF BAND RETURNS):
--------------------------------

                              BANDED AS ADULTS

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

                              BANDED AS YOUNG

         M(1)F'(1)    M(1)S'(1)F(2)   M(1)S'(1)S(2)F(3)   M(1)S'(1)S(2)S(3)F(4)
                      M(2)F'(2)       M(2)S'(2)F(3)       M(2)S'(2)S(3)F(4
                                      M(3)F'(3)           M(3)S'(3)F(4)


ESTIMATES UNDER H1
------------------


                         F(I)                                                       S(I)
         ----------------------------------------------------------     ------------------------------------------------------
 I   YR  ESTIMATE         STANDARD          95% CONFIDENCE              ESTIMATE          STANDARD           95% CONFIDENCE
         --------         __ERROR__         ___INTERVAL___              --------          __ERROR__          ___INTERVAL___

 1    1  0.0773            0.0069         0.0638 -   0.0909              0.5492            0.0529          0.4455 -   0.6530

 2    2  0.0521            0.0046         0.0430 -   0.0612              0.7539            0.0646          0.6273 -   0.8804

 3    3  0.0770            0.0054         0.0663 -   0.0876              0.6616            0.0628          0.5385 -   0.7847

 4    4  0.0565            0.0045         0.0477 -   0.0654              0.5691            0.0566          0.4582 -   0.6801

 5    5  0.0974            0.0068         0.0840 -   0.1108              0.6419            0.0921          0.4614 -   0.8224

 6    6  0.0533            0.0058         0.0420 -   0.0647

         AVERAGE          STANDARD          95% CONFIDENCE              AVERAGE           STANDARD           95% CONFIDENCE
         ESTIMATE         ERROR___          ___INTERVAL___              ESTIMATE          ERROR___           ___INTERVAL___
 _                                                                  _
 F     = 0.0689            0.0024          0.0643 -   0.0736       S    = 0.6352             0.0202          0.5957 -   0.6746

                         F'(I)                                                      S'(I)
         ----------------------------------------------------------     ------------------------------------------------------
 I   YR  ESTIMATE         STANDARD          95% CONFIDENCE              ESTIMATE          STANDARD           95% CONFIDENCE
         --------         __ERROR__         ___INTERVAL___              --------          __ERROR__          ___INTERVAL___

 1    1  0.1610            0.0116         0.1382 -   0.1838              0.4740            0.0561          0.3640 -   0.5840

 2    2  0.1090            0.0099         0.0897 -   0.1283              0.6537            0.0705          0.5154 -   0.7919

 3    3  0.1710            0.0119         0.1477 -   0.1943              0.6788            0.0817          0.5187 -   0.8390

 4    4  0.1570            0.0115         0.1345 -   0.1795              0.5054            0.0684          0.3713 -   0.6394

 5    5  0.1980            0.0126         0.1733 -   0.2227              0.7227            0.1380          0.4522 -   0.9932

 6    6  0.1020            0.0096         0.0832 -   0.1208

         AVERAGE          STANDARD          95% CONFIDENCE              AVERAGE           STANDARD           95% CONFIDENCE
         ESTIMATE         ERROR___          ___INTERVAL___              ESTIMATE          ERROR___           ___INTERVAL___
 _                                                                  _
 F'    = 0.1497            0.0046          0.1407 -   0.1587       S'   = 0.6069             0.0393          0.5300 -   0.6839
```

Example 4.2b

```
SYNTHETIC DATA GENERATED FOR 3 AGE CLASSES WHERE ADULTS & SUBADULTS ARE GROUPED

ESTIMATED NON-ZERO COVARIANCES AND CORRELATIONS UNDER H1


I    YR  COVAR(F(I),S(I))          CORR(F(I),S(I))          COVAR(S(I),S(I+1))          CORR(S(I),S(I+1))
1     1      -0.000028316              -0.077556398              -0.001286452              -0.376411009
2     2       0.000042690               0.142444283              -0.001686778              -0.415990429
3     3       0.000099862               0.292174826              -0.001650718              -0.464165164
4     4       0.000092539               0.361385704              -0.001601589              -0.307144874
5     5       0.000176929               0.280870589


I    YR  COVAR(F(I+1),S(I))        CORR(F(I+1),S(I))        COVAR(F'(I),S'(I))          CORR(F'(I),S'(I))
1     1      -0.000088919              -0.361851499              -0.000076314              -0.117027008
2     2      -0.000196198              -0.558426777              -0.000071249              -0.102512172
3     3      -0.000163993              -0.577387812              -0.000116082              -0.119292194
4     4      -0.000242920              -0.627164979              -0.000079342              -0.100855787
5     5      -0.000405113              -0.758253173              -0.000143095              -0.082271829


I    YR  COVAR(S'(I),S(I))         CORR(S'(I),S(I))         COVAR(S'(I),S(I+1))         CORR(S'(I),S(I+1))
1     1       0.000808850               0.272324296              -0.001110225              -0.306488921
2     2       0.001666550               0.366011409              -0.001462537              -0.330182605
3     3       0.001968865               0.383562160              -0.001693755              -0.365993881
4     4       0.001260906               0.325647315              -0.001422088              -0.225829192
5     5       0.005489554               0.431890425


I    YR  COVAR(S'(I),F(I+1))       CORR(S'(I),F(I+1))
1     1      -0.000076738              -0.294634009
2     2      -0.000170115              -0.443238077
3     3      -0.000168268              -0.455269810
4     4      -0.000215694              -0.461125034
5     5      -0.000456106              -0.569585954


     THE ABOVE ARE ESTIMATES OF THE SAMPLING COVARIANCES AND CORRELATIONS BETWEEN THE PARAMETER ESTIMATORS.


         COVAR(S̄,F̄)  = -0.000023781                                    CORR(S̄,F̄)  = -0.501178877


         COVAR(S̄',F̄')= -0.000016203                                    CORR(S̄',F̄')= -0.090023922


MATRIX OF DATA VALUES -- ADULTS
-----------------------------------------------
        1     116.00   32.00   46.00   23.00   36.00    9.00
        2       0.0    80.00   64.00   39.00   59.00   23.00
        3       0.0     0.0   125.00   50.00   48.00   24.00
        4       0.0     0.0     0.0   102.00   62.00   34.00
        5       0.0     0.0     0.0     0.0   153.00   45.00
        6       0.0     0.0     0.0     0.0     0.0    80.00


MATRIX OF EXPECTED VALUES -- ADULTS
-----------------------------------------------
        1     116.00   43.06   48.11   23.49   23.12    8.23
        2       0.0    78.16   87.31   42.63   41.96   14.93
        3       0.0     0.0   115.43   56.36   55.47   19.74
        4       0.0     0.0     0.0    84.81   83.48   29.70
        5       0.0     0.0     0.0     0.0   146.04   51.96
        6       0.0     0.0     0.0     0.0     0.0    80.00


MATRIX OF STANDARD NORMAL DEVIATES -- ADULTS
-----------------------------------------------
        1       0.00   -1.71   -0.31   -0.10    2.70    0.27
        2       0.0     0.21   -2.57   -0.56    2.67    2.10
        3       0.0     0.0     0.93   -0.86   -1.02    0.97
        4       0.0     0.0     0.0     1.92   -2.42    0.80
        5       0.0     0.0     0.0     0.0     0.61   -0.98
        6       0.0     0.0     0.0     0.0     0.0     0.00


MATRIX OF DATA VALUES -- YOUNG
-----------------------------------------------
        1     161.00   34.00   21.00    7.00   18.00    4.00
        2       0.0   109.00   73.00   12.00   16.00    7.00
        3       0.0     0.0   171.00   51.00   26.00   13.00
        4       0.0     0.0     0.0   157.00   57.00   10.00
        5       0.0     0.0     0.0     0.0   198.00   39.00
        6       0.0     0.0     0.0     0.0     0.0   102.00


MATRIX OF EXPECTED VALUES -- YOUNG
-----------------------------------------------
        1     161.00   24.78   27.68   13.51   13.30    4.73
        2       0.0   109.00   50.47   24.64   24.26    8.63
        3       0.0     0.0   171.00   38.55   37.95   13.50
        4       0.0     0.0     0.0   157.00   49.42   17.58
        5       0.0     0.0     0.0     0.0   198.00   39.00
        6       0.0     0.0     0.0     0.0     0.0   102.00


MATRIX OF STANDARD NORMAL DEVIATES -- YOUNG
-----------------------------------------------
        1       0.0     1.88   -1.29   -1.78    1.30   -0.34
        2       0.0     0.0     3.25   -2.58   -1.70   -0.56
        3       0.0     0.0     0.0     2.04   -1.98   -0.14
        4       0.0     0.0     0.0     0.0     1.11   -1.82
        5       0.0     0.0     0.0     0.0     0.0     0.00
        6       0.0     0.0     0.0     0.0     0.0     0.0
```

Example 4.2c

```
SYNTHETIC DATA GENERATED FOR 3 AGE CLASSES WHERE ADULTS & SUBADULTS ARE GROUPED

CHI-SQUARE TEST OF HO VS H1
---------------------------


    2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC    2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                              WITH 1 DEGREE CF FREEDOM                                        WITH 1 DEGREE OF FREEDOM
    -----------------------   ----------------------------------    -----------------------   ----------------------------------

    R(I.)    N(I)-R(I.)                                             W(I)     Z(I+1)

    Q(I.)    M(I)-Q(I.)                                             Q(I,I)   Q(I.)-Q(I,I)

I= 1        262    1238                                                     116     146
            245     755                    18.359                           161      84                       23.480

I= 2        265    1235                                                     146     349
            217     783                     6.272                           109     108                       28.217

I= 3        247    1253                                                     329     375
            261     739                    34.390                           171      90                       26.910

I= 4        198    1302                                                     284     379
            224     776                    36.195                           157      67                       49.746

I= 5        198    1302                                                     475     169
            237     763                    46.026                           198      39                        9.200

I= 6         80    1420
            102     898                    21.053
|---------------------------------------------------|
|TOTAL CHI-SQUARE WITH 11 DEGREES OF FREEDOM = 299.847|
|---------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  299.85  = 0.0


THIS TEST OF THE HYPOTHESIS HO AGAINST THE HYPOTHESIS H1 TESTS THE ASSUMPTION THAT YOUNG AND ADULTS HAVE THE
SAME SURVIVAL AND RECOVERY RATES.


CHI-SQUARE TEST OF H1 VS H2
---------------------------


    2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                              WITH 1 DEGREE CF FREEDOM
    -----------------------   ----------------------------------

I= 2         80     185
             66     164                    0.132

I= 3        125     122
            204     253                    2.294

I= 4        102      96
            182     283                    8.685

I= 5        153      45
            322     124                    1.825
|---------------------------------------------------|
|TOTAL CHI-SQUARE WITH  4 DEGREES OF FREEDOP =  12.937|
|---------------------------------------------------|
PROBABILITY OF A CHI-SCUARE VALUE LARGER THAN   12.94  = 0.01159


THIS TEST OF THE HYPOTHESIS H1 AGAINST THE HYPOTHESIS H2 TESTS THE ASSUMPTION THAT RECOVERY RATES FOR NEWLY
RELEASED ADULTS ARE THE SAME AS FOR SURVIVORS OF PREVIOUSLY BANDED COHORTS.


CHI-SQUARE TEST OF H2 VS H3
---------------------------


    2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                              WITH 1 DEGREE OF FREEDOM
    -----------------------   ----------------------------------

I= 1         34      50
             32     114                    8.975

I= 2         73      35
            131     218                   30.151

I= 3         51      39
            131     244                   14.392

I= 4         57      10
            265     114                    6.513
|---------------------------------------------------|
|TOTAL CHI-SQUARE WITH  4 DEGREES OF FREEDOM =  60.032|
|---------------------------------------------------|
PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   60.03  = 0.0


REJECTION OF H2 IN FAVOR OF H3 WOULD INDICATE THAT SURVIVAL AND RECOVERY RATES ARE AGE-DEPENDENT FOR AT LEAST
THE FIRST TWO YEARS.



THE HYPOTHESES H01 AND H02 ARE MORE RESTRICTIVE THAN H1 BUT ARE NOT COMPARABLE IN THIS SENSE WITH HO, THUS H01
AND H02 DO NOT FIT INTO THE ABOVE SERIES.  IN PRACTICE THE MODELS UNDER H02 AND H1 ARE LIKELY TO BE OF MOST USE, SO A
LIKELIHOOD RATIO TEST TO DISTINGUISH BETWEEN THESE MODELS IS COMPUTED BELOW.  THIS TESTS THE ASSUMPTION THAT YOUNG AND
ADULT SURVIVAL RATES ARE CONSTANT FROM YEAR TO YEAR.

LIKELIHOOD RATIO TEST OF H02 VS H1.
-----------------------------------


CHI-SQUARE VALUE                                      =   11.15
DEGREES OF FREEDOM                                    =       8
PROBABILITY OF A CHI-SCUARE VALUE LARGER THAN   11.15 = 0.19347
```

Example 4.2d

```
SYNTHETIC DATA GENERATED FOR 3 AGE CLASSES WHERE ADULTS & SUBADULTS ARE GROUPED

CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H3
-----------------------------------------------------

                 CONTINGENCY TABLES                                        CORRESPONDING CHI-SQUARE
                                                                           STATISTICS AND DEGREES OF FREEDOM
                 ------------------                                        ---------------------------------

          I= 2
                 23. 39. 59. 64.
                  4.  7. 18. 21.
                  9. 23. 36. 46.                                              3.89 WITH  6 D.F.

          I= 3
                 24. 50. 48.
                  7. 12. 16.
                 36. 69.113.                                                  5.09 WITH  4 D.F.

          I= 4
                 34. 62.
                 13. 26.
                 67.177.                                                      2.29 WITH  2 D.F.

                                    |----------------------------------------|
                                    |TOTAL CHI-SQUARE      11.28 WITH 12 D.F.|
                                    |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   11.28  = 0.50497




CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H0
-----------------------------------------------------

                                    |----------------------------------------|
                                    |TOTAL CHI-SQUARE     384.10 WITH 31 D.F.|
                                    |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  384.10  = 0.0


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H1
-----------------------------------------------------

                                    |----------------------------------------|
                                    |TOTAL CHI-SQUARE      84.25 WITH 20 D.F.|
                                    |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   84.25  = 0.0


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H2
-----------------------------------------------------

                                    |----------------------------------------|
                                    |TOTAL CHI-SQUARE      71.31 WITH 16 D.F.|
                                    |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   71.31  = 0.00000


CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H3
-----------------------------------------------------

                                    |----------------------------------------|
                                    |TOTAL CHI-SQUARE      11.28 WITH 12 D.F.|
                                    |----------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN   11.28  = 0.50497

   FROM THE MODELS ABOVE, ONE SHOULD CHOOSE THE SIMPLEST MODEL (FEWEST PARAMETERS) THAT ADEQUATELY DESCRIBES THE DATA.
ADEQUACY MAY BE JUDGED BY EXAMINING THE RESULTS OF (1) THE GOODNESS OF FIT TESTS, AND (2) THE TESTS BETWEEN SPECIFIC
MODELS. FREQUENTLY, H02 OR H1 IS ADEQUATE.
```

Chapter 5. Hypotheses Tests for Pooling Band Recovery Data Sets

In Chapters 2 and 3 we dealt at length with the analysis of individual data sets. In the context of adult banding, it is straightforward that an individual data set is the band recovery data from k years of banding for ℓ years of recovery (see Section 2.1). If both adults and young are banded (Section 3.1), the two band recovery arrays are considered as one data set. In either instance, the banding is typically done in the same relatively small area each year, and the sex of the birds is often determined. Consequently, in any large-scale analysis of band recoveries we are likely to have numerous data sets representing both sexes and a variety of banding areas. In this chapter we discuss several tests developed to aid in making decisions about pooling such data sets. Basically, such tests examine the question of whether or not the data sets have the same underlying survival and recovery rate parameters.

Initially, the motivation for this type test came from the issue of combining the two data sets for adult males and females banded in the same area. This test for sex differences is mathematically the same as the test for whether two data sets from different areas (for the same sex) have the same underlying parameter values. For example, if adult male mallards had been banded each winter in northeastern Colorado and western Nebraska, we may wish to pool these data for analysis. Whereas there are only two sexes, we may have data from more than two areas. The logical extension of these tests has been developed for examining questions of pooling data sets over several areas.

In all of these tests, the null hypothesis is that the data sets are described by the same survival and recovery rate parameters and hence could be pooled. Rejection of this hypothesis is evidence that the data sets should be analyzed separately and then the estimates may be averaged, rather than pooling the data. If a test is not significant, say at the 5% level, we have evidence that the data sets may be pooled for analysis.

5.1 Testing Equality of Survival and Recovery Rates for Adult Data Sets

Testing Equality of Adult Male and Female Parameters

As far as possible, the notation of Chapter 2 will be used in this section; specifically, the reader is referred to Sections 2.1 and 2.2.

For this test to be meaningful, the data sets for males and females must come from the same banding study, more specifically, from the same area and cover the same years of banding and recovery. Under Model 1 let the parameters for the two data sets be

S_{im} = adult male survival rate in year i,
S_{if} = adult female survival rate in year i,
f_{im} = adult male recovery rate in year i,
f_{if} = adult female recovery rate in year i.

The null hypothesis is $S_{im}=S_{if}$, $i=1,\ldots,\ell-1$ and $f_{im}=f_{if}$, $i=1,\ldots,\ell$. Here the general case of k years of banding and ℓ years of recoveries is assumed (hence $\ell > k$ is possible).

The alternative hypothesis is a composite one where: (1) only recovery rates differ, (2) only survival rates differ, or (3) both recovery and survival rates differ. If the null hypothesis is rejected, we do not know where the differences in parameter values lie. The way to examine this is to analyze each data set separately by program ESTIMATE and examine the resultant parameter estimates.

From Section 2.1 the following summary statistics were defined for an individual data set:

R_i = total of row i of the data array, $i=1,\ldots,k$,
C_i = total of column i of the data array, $i=1,\ldots,\ell$,
T_i = a block total $i=1,\ldots,\ell$.

Define R_{im}, C_{im}, and T_{im} as these summary statistics for the set of male recovery data, and similarly define R_{if}, C_{if}, and T_{if} for the female recovery data. The test of the null hypothesis of no difference in male and female parameters is based on these statistics, and takes the form of $k+\ell-1$ separate chi-square test statistics, each with 1 df, computed on separate 2 by 2 contingency tables.

Defining the sample sizes for males and females as N_{im} and N_{if}, the set of contingency tables is

$$\begin{array}{|c|c|} \hline R_{im} & N_{im}-R_{im} \\ \hdashline R_{if} & N_{if}-R_{if} \\ \hline \end{array} \quad , i=1,\ldots,k$$

and

$$\begin{array}{|c|c|} \hline C_{im} & T_{im}-C_{im} \\ \hdashline C_{if} & T_{if}-C_{if} \\ \hline \end{array} \quad , i=1,\ldots,\ell-1.$$

There are $k+\ell-1$ of these tables. For each one, the usual chi-square goodness of fit statistic is computed. The first set of k, 2 by 2 tables uses the row sums and sample sizes from the data, while the second set ($\ell-1$ tables) is based on the column and block totals. For triangular data arrays, $\ell=k$.

For each of these $k+\ell-1$ tables one chi-square test statistic is computed. These values are then added to obtain one test statistic which has a chi-square distribution with $k+\ell-1$ degrees of freedom under the null hypothesis. If this test value exceeds the critical level, at the 5% significance level for instance, for a chi-square variable with $k+\ell-1$ degrees of freedom, then one rejects the null hypothesis. In this event pooling the data sets is not justified.

An Example of the Male-Female Test

This test is available as an option within the FORTRAN program BROWNIE (see Chapter 6 for details of how to input data and obtain this test).

The test of the null hypothesis is easy to use and interpret as we illustrate in Example 5.1. Data on male and female mallards banded during the winter in Illinois, 1963-70, are used to illustrate the testing procedure (the data for adult males is the same as that used in Example 2.6). The computer output, shown in Example 5.1, displays the input data, and then various summary statistics. The notation for these various subtotals (summary statistics) used by the program differs from this chapter in that RROW(I), RCOL(I), T(I) correspond to R_{im}, C_{im}, T_{im}, respectively, and QROW(I), QCOL(I), U(I) correspond to R_{if}, C_{if}, T_{if}, respectively. Finally the contingency tables and individual chi-square values are given, followed by the total chi-square which is the test statistic. The program also prints the achieved significance level of the test statistic, i.e., the probability of getting a value as large as that observed, if the null hypothesis is true.

In Example 5.1 we see that the total chi-square value is 314.17 with 18 df. The probability of a value this large, if the null hypothesis of no difference in male-female parameter values is true, is essentially zero (to 8 decimal places). Therefore we reject the null hypothesis and conclude that recovery and/or survival rates differ significantly by sex, and the data sets must be analyzed separately. From the separate analyses we can compare male and female recovery and survival rates to determine where the differences are.

The program shows the individual contingency tables for the triangular part of the data arrays under the titles MATRIX 1 and MATRIX 2. MATRIX 3 relates to the nontriangular portion of the data ($s>0$), and is not shown in detail. The output labeled MATRIX 1 gives those tables based on row totals. For example, in 1963 ($i=1$) there were 2,583 males banded. A total of 279 recoveries were recorded from this banded cohort. Similarly, for females $N_{1f}=1{,}478$ and $R_{1f}=94$. Thus the first 2 by 2 table is

$$\begin{array}{|c|c|} \hline 279 & 2{,}583-279 \\ \hdashline 94 & 1{,}478-94 \\ \hline \end{array} \equiv \begin{array}{|c|c|} \hline 279 & 2{,}304 \\ \hdashline 94 & 1{,}384 \\ \hline \end{array}$$

The individual chi-square statistic from this table is 22.232.

Example 5.1

```
                    MALE AND FEMALE MALLARDS BANDED DURING THE WINTER IN ILLINOIS, 1963-70
ADULT MALE   INPUT 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.

ADULT FEMALE INPUT MATRIX


1963  1478.   40.   31.    8.   11.    2.    0.    2.    0.    0.    0.    0.
1964  1525.    0.   72.   20.   15.    7.    5.    1.    2.    1.    0.    0.
1965   319.    0.    0.    8.    7.    3.    0.    1.    3.    1.    0.    0.
1966  1805.    0.    0.    0.   63.   27.   14.    5.    5.    2.    2.    2.
1967  1400.    0.    0.    0.    0.   39.   14.   17.   10.    7.    5.    2.
1968   900.    0.    0.    0.    0.    0.   17.    9.   15.   10.    6.    1.
1969  1400.    0.    0.    0.    0.    0.    0.   39.   21.   10.   10.    0.
1970  1789.    0.    0.    0.    0.    0.    0.    0.   63.   39.   11.    4.



BASIC SUBTOTALS

 I   RROW(I)  RCOL(I)  CROW(I)  QCOL(I)      T(I)      U(I)      W(I)      Z(I)
 1   279.00    91.00    94.00    40.00    279.00     94.00     91.00     0.0
 2   382.00   230.00   123.00   103.00    570.00    177.00    261.00    188.00
 3   129.00    96.00    23.00    36.00    469.00     97.00    124.00    363.00
 4   453.00   257.00   120.00    96.00    826.00    181.00    290.00    419.00
 5   377.00   292.00    94.00    78.00    946.00    179.00    331.00    597.00
 6   260.00   211.00    58.00    50.00    914.00    159.00    244.00    700.00
 7   291.00   307.00    80.00    74.00    994.00    189.00    342.00    771.00
 8   532.00   510.00   117.00   119.00   1219.00    232.00    566.00    761.00
                                          709.00    113.00              768.00



CONTINGENCY CHI-SQUARE TEST FOR DIFFERENCES DUE TO SEX

MATRIX  1                                                           MATRIX  2


    2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC        2 X 2 CONTINGENCY TABLE   CORRESPONDING CHI-SQUARE STATISTIC
                              WITH 1 DEGREE OF FREEDOM                                            WITH 1 DEGREE OF FREEDOM
    -----------------------   ----------------------------------        -----------------------   ----------------------------------

I= 1        279   2304                                              I= 1        91    188
             94   1384                    22.232                                40     54                           3.047

I= 2        382   2693                                              I= 2       230    340
            123   1402                    19.803                               103     74                          17.401

I= 3        129   1066                                              I= 3        96    373
             23    296                     3.583                                36     61                          12.452

I= 4        453   2965                                              I= 4       257    569
            120   1685                    52.763                                96     85                          31.348

I= 5        377   2723                                              I= 5       292    654
             94   1306                    30.535                                78    101                          11.014

I= 6        260   2140                                              I= 6       211    703
             58    842                    14.479                                50    109                           5.144

I= 7        291   2310                                              I= 7       307    687
             80   1320                    32.413                                74    115                           4.972

I= 8        532   3901                                              I= 8       510    709
            117   1672                    40.685                               119    113                           7.096



MATRIX  3

    CONTINGENCY TABLE FOR S GREATER THAN 1

          I= 9
                  110.221.378.
                    9. 34. 70.                                                    5.21 WITH  2 D.F.

                                     |-------------------------------------|
                                     |TOTAL CHI-SQUARE    314.17 WITH 18 D.F.|
                                     |-------------------------------------|

PROBABILITY OF A CHI-SQUARE VALUE LARGER THAN  314.17  = 0.0
```

Under the heading MATRIX 2, the 2 by 2 tables based on C_i and T_i are given for the triangular part of the data array. For example, $C_{1m}=91$, $T_{1m}=279$, $C_{1f}=40$, and $T_{1f}=94$. Hence for $i=1$, the table is

$$\begin{array}{|c|c|}\hline 91 & 279-91 \\ \hline 40 & 94-40 \\ \hline\end{array} \equiv \begin{array}{|c|c|}\hline 91 & 188 \\ \hline 40 & 54 \\ \hline\end{array}$$

When using this test, one should look at the individual chi-square values on 1 df, as well as the overall result (314.17, with 18 df in this example). If the overall test strongly rejects the hypothesis of no difference ($P<0.005$), one can safely conclude that the data sets should not be pooled. But in the event the result is significant at lesser levels like ($0.05<P<0.005$), one should examine the separate chi-square tests on the different contingency tables. It may well be that just one of these individual chi-square values is quite large, hence "causing" the significant overall test result. In this event it is probably the male and female recovery rates in 1 year that are different, and other parameters may be very similar by sex. Pooling the data is reasonable in this case. If, however, no one single chi-square stands out, and they all tend to be a little larger than expected under the null hypothesis, then it is probably because a slight but consistent difference exists in male and female parameters and the data should not be pooled.

This test for differences in adult male and female parameters was developed by Brownie (1973); it is the only one of its kind that is programmed for computer computation. A similar test is described below for adult and young data; also both these tests are extended to the question of pooling data for the same sex over different areas.

Testing Adult Data Sets for Pooling Over Areas

Nothing in the theory of the test described above is specific for testing for sex specificity. Although this was the motivation of the test, it is nothing more than a test of the null hypothesis that two adult band recovery data sets (having the same values of k and ℓ) have the same survival and recovery rates. It is logical to apply this type test to the question of pooling data sets across areas for the same sex, species, and banding and recovery years. For two areas, the above test suffices. For more than two areas, the test has a simple extension.

For r areas the appropriate test for pooling is based on a series of r by 2 contingency tables. As before, there will still be exactly $k+\ell-1$ tables. The null hypothesis is that Model 1 with the same annual survival and recovery rates fits all r data sets (hence they can validly be pooled).

Instead of indexing data sets by m and f, let the data and parameters have a second index for area, $a=1,\ldots,r$. For example, the row total R_i from data set a becomes R_{ia}. The set of statistics R_i, C_i, and T_i, as well as N_i, must first be tabulated for each data set. The following r by 2 contingency tables are then easily written:

$$\begin{array}{|c|c|}\hline R_{i1} & N_{i1}-R_{i1} \\ \hline R_{i2} & N_{i2}-R_{i2} \\ \hline \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \hline R_{ir} & N_{ir}-R_{ir} \\ \hline\end{array} \quad ,\, i=1,\ldots,k.$$

C_{i1}	$T_{i1}-C_{i1}$
C_{i2}	$T_{i2}-C_{i2}$
$\cdot$	$\cdot$
$\cdot$	$\cdot$
$\cdot$	$\cdot$
C_{ir}	$T_{ir}-C_{ir}$

$,i=1,\ldots,\ell-1.$

Note that each row comes from a different data set (hence different area), but in a given type of table, each row is otherwise the same. Thus the extension aspect of this test beyond the case of $r=2$ is simply to add more rows onto each basic table, each row being computed in the same way (for a given table) from a different data set.

Once these $k+\ell-1$ tables have been developed, the test is straightforward. The usual chi-square test statistic for homogeneity is computed for each table. These test statistics are each approximately chi-square distributed with $r-1$ degrees of freedom under the null hypothesis. One then adds all $k+\ell-1$ test statistics. The result is the overall test statistic for testing the null hypothesis that all data sets have the same recovery and survival rate parameters. The test statistic is distributed as chi-square with $(k+\ell-1)\ (r-1)$ degrees of freedom when the null hypothesis is true. If this test value exceeds the critical value for the chosen significance level, we reject the null hypothesis and conclude that pooling of data sets is not warranted.

5.2 Testing Adult and Young Data Sets for Pooling

We will give the details of this test for comparing r data sets; these could be on the same species and sex for different areas, or if $r=2$, they could be data sets for one area for males and females. This test is not programmed in BROWNIE and it would have to be done by hand or a program written. To compute the test, a set of summary statistics must first be computed for each data set. These are in the notation of Chapter 3,

$$\begin{array}{lll}
R_{i\cdot} & \text{(adult row totals)} & ,i=1,\ldots,k, \\
Q_{i\cdot} & \text{(young row totals)} & ,i=1,\ldots,k, \\
Q_{ii} & \text{(first year recoveries} & ,i=1,\ldots,k-1 \text{ if } \ell=k \\
 & \text{for young)} & ,i=1,\ldots,k \quad \text{if } \ell>k
\end{array}$$

$$W_i=\begin{cases} R_{\cdot i}+Q_{\cdot i}-Q_{ii} & ,i=1,\ldots,k \text{ if } s\leqslant 1 \\ R_{\cdot i}+Q_{\cdot i} & ,i=k+1,\ldots,k+s-1 \text{ if } s>1 \end{cases}$$

$$X_i=\begin{cases} T_{i}+U_i-Q_{i\cdot} & ,i=1,\ldots,k \text{ if } s\leqslant 1 \\ T_i+U_i & ,i=k+1,\ldots,k+s-1 \text{ if } s>1. \end{cases}$$

The statistics $R_{\cdot i}$, $Q_{\cdot i}$ (column totals), T_i and U_i (block totals) would have to be computed as intermediate values. Program BROWNIE computes and prints these statistics, except for X_i, for the triangular part of the data array. Thus, by analyzing each data set first, most of the computation is done by the computer.

Given these summary statistics, for each data set, one then constructs a series of r by 2 contingency tables. In each table the entries are of the same type but each row comes from a different data set.

From the statistics $R_{i\cdot a}$, construct the r by 2 contingency tables:

data set		
1	$R_{i\cdot 1}$	$N_{i1}-R_{i\cdot 1}$
2	$R_{i\cdot 2}$	$N_{i2}-R_{i\cdot 2}$
.	.	.
.	.	.
.	.	.
r	$R_{i\cdot r}$	$N_{ir}-R_{i\cdot r}$

$, i=1,\ldots,k.$

In the above, the added subscript to the notation denotes the data set from which the values came. Thus, N_{ia} is the number of adults banded in year i, and area (or sex) a. Similarly, $R_{i\cdot a}$ is the row total of recoveries from adults banded in year i in area a (or of sex a). This augmented notation will also be used below.

The additional contingency tables to construct are:

$Q_{i\cdot 1}$	$M_{i1}-Q_{i\cdot 1}$
$Q_{i\cdot 2}$	$M_{i2}-Q_{i\cdot 2}$
.	.
.	.
.	.
$Q_{i\cdot r}$	$M_{ir}-Q_{i\cdot r}$

$, \ i=1,\ldots,k$

Q_{ii1}	$Q_{i\cdot 1}-Q_{ii1}$
Q_{ii2}	$Q_{i\cdot 2}-Q_{ii2}$
.	.
.	.
.	.
Q_{iir}	$Q_{i\cdot r}-Q_{iir}$

$$,\begin{cases} i=1,\ldots,k-1 & \text{if } \ell=k \\ i=1,\ldots,k & \text{if } \ell>k \end{cases}$$

W_{i1}	$X_{i1}-W_{i1}$
W_{i2}	$X_{i2}-W_{i2}$
.	.
.	.
.	.
W_{ir}	$X_{ir}-W_{ir}$

$, \ i=1,\ldots,\ell-1.$

If $s=0$ $(\ell=k)$, there are $4k-2$ tables (each r by 2); if $s>0$, there are $4k+s-1$ tables.

For each contingency table one computes the usual chi-square test of homogeneity. Each of these individual test statistics is (approximately) distributed as chi-square with $r-1$ degrees of freedom under the null hypothesis that the H_1 model with the same survival and recovery rates fits the data from all areas. The test statistic for this null hypothesis is the sum of all these individual chi-square values. Under the null hypothesis this sum is distributed as a chi-square variable with degrees of freedom

$$(4k-2)(r-1)\,,\ \text{if}\ s=0$$
$$(4k+s-1)(r-1),\ \text{if}\ s>0\,.$$

One rejects the null hypothesis if the test statistic value exceeds the critical value for the chosen significance level.

5.3 Mathematical Background

The reader with little knowledge of mathematics can skip this section. Of the above tests, only the test for differences in adult male and female parameters has previously appeared in the literature (Brownie 1973). Consequently, this section will supply the basic mathematical background of all these tests.

Adult Data

Under Model 1 for adult data a minimal sufficient statistic (MSS) is

$$\mathcal{S}=\{R_1,\ldots,R_k,C_1,\ldots,C_{\ell-1}\}\,.$$

Because of the recursive relationship

$$T_1=R_1\,,$$
$$T_{i+1}=T_i-C_i+R_{i+1} \qquad ,\ i=1,\ldots,k-1$$

and if $s>0$,

$$T_{k+j}=T_{k+j-1}-C_{k+j-1} \qquad ,j=0,\ldots,s-1\,,$$

the T_i are implicitly part of the minimal sufficient statistic $\mathcal{S}$ (T_i becomes a shorthand notation for a function of the MSS).

As determined initially by Robson and Youngs (1971) the probability distribution of $\mathcal{S}$ is given by

$$R_i \sim B(N_i,\rho_i) \qquad ,i=1,\ldots,k$$

and

$$C_i\ \text{given}\ T_i \sim B(T_i,f_i/\rho_i) \qquad ,i=1,\ldots,\ell-1\,.$$

These distributions are mutually independent. In this notation $B(n, p)$ is the binomial distribution, and from Chapter 2

$$\rho_i=f_i+S_if_{i+1}+\cdots+S_i\cdots S_{\ell-1}f_\ell\,.$$

It is thus straightforward to write down the probability distribution of $\mathcal{S}$, $P\{\mathcal{S}\}$, under Model 1.

Let $\mathcal{S}_m$ and $\mathcal{S}_f$ be the minimal sufficient statistic for separate adult male and female recovery data under the alternative hypothesis of different parameter rates. Under the null hypothesis, H_0, of equal parameters, the MSS is

$$\mathcal{S}_0=\{R_{im}+R_{if},\ i=1,\ldots,k,\ C_{im}+C_{if},\ i=1,\ldots,\ell-1\}\,.$$

A test of H_0 can be based on the probability distribution of the MSS under the alternative hypothesis, conditional on $\mathcal{S}_0$, and given H_0 is true. Symbolically we need to find

$$\mathrm{P}_{H_0}\{\mathcal{S}_m,\mathcal{S}_f \mid \mathcal{S}_0\}\,.$$

Using the above results on the distribution of $\mathcal{S}$ and the independence of the two data sets leads to the result

$$\mathrm{P}_{H_0}\{\mathcal{S}_m,\mathcal{S}_f \mid \mathcal{S}_0\}=\prod_{i=1}^{k}\frac{\binom{N_{im}}{R_{im}}\binom{N_{if}}{R_{if}}}{\binom{N_{im}+N_{if}}{R_{im}+R_{if}}}\prod_{i=1}^{\ell-1}\frac{\binom{T_{im}}{C_{im}}\binom{T_{if}}{C_{if}}}{\binom{T_{im}+T_{if}}{C_{im}+C_{if}}}\,.$$

This is a product of $k+\ell-1$ independent hypergeometric distributions, from which the series of 2 by 2 contingency tables follow as a basis for a chi-square test of H_0. This same approach is used to derive the other tests of this chapter.

The parameters appearing in the distribution of the MSS (i.e., ρ_i and f_i/ρ_i) are a one-to-one transformation of the basic parameters S_i and f_i. We are testing the null hypotheses that $\rho_{im} \equiv \rho_{if}$, $i=1,\ldots,k$ and $(f_{im}/\rho_{im}) \equiv (f_{if}/\rho_{if})$, $i=1,\ldots,\ell-1$. Each individual equality has its own chi-square test. For each contingency table we can determine exactly what is being tested, specifically the tests based on R_{im} and R_{if} test H_0: $\rho_{im}=\rho_{if}$, while the tests based on C_{im} and C_{if} test H_0: $(f_{im}/\rho_{im})=(f_{if}/\rho_{if})$. Thus these individual tests do not relate to separate survival or recovery rate parameters; hence they have no great value by themselves.

This test could be used to compare data sets from two areas (rather than two sexes). Assume there are r areas, and we have a recovery data set from each area. Let areas be indexed by a. The null hypothesis H_0 is that there are equal annual recovery and survival rate parameters in all areas. The alternative is that there are different annual parameters for at least two areas. Under this alternative, the MSS is

$$\begin{aligned}\mathcal{S}_a &= \{\mathcal{S}_1,\ldots,\mathcal{S}_r\}\\ &= \{R_{1a},\ldots,R_{ka},C_{1a},\ldots,C_{\ell-1,a},a=1,\ldots,r\}\,.\end{aligned}$$

In this notation the second subscript denotes the area from where the basic data came. Under the null hypothesis H_0, the MSS is

$$\mathcal{S}_0=\{(R_{11}+\cdots+R_{1r}),\ldots,(R_{k1}+\cdots+R_{kr}),(C_{11}+\cdots+C_{1r}),\ldots,(C_{\ell-1,1}+\cdots+C_{\ell-1,r})\}.$$

Using independence, and the known distribution under H_0 of each $\mathcal{S}_j$, $j=1,\ldots,r$, (a product of binomials) we can determine

$$\mathrm{P}_{H_0}\{\mathcal{S}_a \mid \mathcal{S}_0\}=\prod_{i=1}^{k}\frac{\binom{N_{i1}}{R_{i1}}\cdots\binom{N_{ir}}{R_{ir}}}{\binom{N_{i1}+\cdots+N_{ir}}{R_{i1}+\cdots+R_{ir}}}\prod_{i=1}^{\ell-1}\frac{\binom{T_{i1}}{C_{i1}}\cdots\binom{T_{ir}}{C_{ir}}}{\binom{T_{i1}+\cdots+T_{ir}}{C_{i1}+\cdots+C_{ir}}}\,.$$

This distribution is a product of independent multiple hypergeometric distributions. The test of H_0 is based on the corresponding series of $k+\ell-1$ contingency tables, each r by 2 (as described in Section 5.1). The reader should be able to associate each of these r by 2 tables with the exact parametric function of survival and recovery rates which is being tested for equality (based on the distribution of the MSS).

Adult and Young Data

Next the basic theory is given for testing parameters for sex or area differences for recovery data sets from banding both adults and young. The H_1 model and the notation of Chapter 3, Sections 3.1 and 3.2 are used.

From Brownie (1973), an MSS under the H_1 model for a single data set with $s=0$ is (see also Sections 3.1 and 5.1 for definition of terms)

$$\mathcal{S}=\{R_{i\cdot}, Q_{i\cdot}, i=1,\ldots,k, Q_{ii}, i=1,\ldots,k-1, \mathrm{T}_2, T_i+U_i, i=3,\ldots,k\}\,.$$

Additional statistics, computable from $\mathcal{S}$, and necessary in the tests are

$$\begin{aligned}W_i &= R_{\cdot i}+Q_{\cdot i}-Q_{ii} &&, i=1,\ldots,k,\\ X_i &= T_i+U_i-Q_{i\cdot} &&, i=1,\ldots,k\,.\end{aligned}$$

For the case of $s>0$, the MSS is equal to $\mathcal{S}$ given above, augmented by the additional terms Q_{kk}, $R_{\cdot k+j}+Q_{\cdot k+j}$, $j=1,\ldots,s$, and W_i, X_i, $i=k+1,\ldots,k+s$, where for $i>k$ we define $W_i=R_{\cdot i}+Q_{\cdot i}$ and $X_i=T_i+U_i$.

The MSS is distributed as the product of $4k+s-1$ independent binomial random variables if $s>0$, and as $4k-2$ independent binomial random variables if $s=0$;

$$\begin{aligned}R_{i\cdot} &\sim B(N_i,\rho_i) &&, i=1,\ldots,k,\\ Q_{i\cdot} &\sim B(M_i,\rho_i') &&, i=1,\ldots,k,\\ Q_{ii}\text{ given }Q_{i\cdot} &\sim B(Q_{i\cdot},f_i'/\rho_i') &&, i=1,\ldots,k-1,\\ W_i\text{ given }X_i &\sim B(X_i,f_i/\rho_i) &&, i=1,\ldots,k+s-1\,,\end{aligned}$$

and if $s>0$ we also have

$$Q_{kk}\text{ given }Q_{k\cdot} \sim B(Q_{k\cdot},f_k'/\rho_k')\,.$$

Now let there be r data sets, indexed by $a = 1,\ldots,r$, where a becomes another index on all sample sizes, statistics, and parameters. Basically we just have r independent data sets and associated statistics. To derive the test of the null hypothesis H_0 that the underlying recovery and survival parameters (f_i, f_i', S_i, S_i', see Section 3.2 for details) are the same for all r data sets we can use the probability distribution under H_0 of $\mathcal{S}_1,\ldots,\mathcal{S}_r$ (which is the MSS under the alternative) given $\mathcal{S}_0$. Here $\mathcal{S}_0$ is the MSS if H_0 is true; it is simply the "addition" (element by element) of the individual MSS, $\mathcal{S}_1$. For example, the first k elements of $\mathcal{S}_0$ are

$$R_{i\cdot 1} + R_{i\cdot 2} + \cdots + R_{i\cdot r} = \sum_{a=1}^{r} R_{i\cdot a} \qquad , i = 1,\ldots,k .$$

Under the null hypothesis H_0, the distribution of $\mathcal{S}_0$ is known. Finally, it is easy to determine

$$P_{H_0}\{\mathcal{S}_1,\ldots,\mathcal{S}_r \mid \mathcal{S}_0\} = \frac{\prod_{a=1}^{r} P_{H_0}\{\mathcal{S}_a\}}{P_{H_0}\{\mathcal{S}_0\}} .$$

The final result for $P_{H_0}\{\mathcal{S}_1,\ldots,\mathcal{S}_r \mid \mathcal{S}_0\}$ is the product of numerous terms, it can be expressed as

$$\prod_{i=1}^{k} \frac{\binom{N_{i1}}{R_{i\cdot 1}}\cdots\binom{N_{ir}}{R_{i\cdot r}}}{\binom{N_{i1}+\cdots+N_{ir}}{R_{i\cdot 1}+\cdots+R_{i\cdot r}}} \prod_{i=1}^{k} \frac{\binom{M_{i1}}{Q_{i\cdot 1}}\cdots\binom{M_{ir}}{Q_{i\cdot r}}}{\binom{M_{i1}+\cdots+M_{ir}}{Q_{i\cdot 1}+\cdots+Q_{i\cdot r}}} \times \prod_{i=1}^{\substack{k-1 \text{ if } s=0\\ k \text{ if } s>0}} \frac{\binom{Q_{i\cdot 1}}{Q_{ii1}}\cdots\binom{Q_{i\cdot r}}{Q_{iir}}}{\binom{Q_{i\cdot 1}+\cdots+Q_{i\cdot r}}{Q_{ii1}+\cdots+Q_{iir}}} \prod_{i=1}^{k+s-1} \frac{\binom{X_{i1}}{W_{i1}}\cdots\binom{X_{ir}}{W_{ir}}}{\binom{X_{i1}+\cdots+X_{ir}}{W_{i1}+\cdots+W_{ir}}} .$$

Each term gives rise to an r by 2 contingency table from which a chi-square test statistic can be computed. For example, the tables

$R_{i\cdot 1}$	$N_{i1} - R_{i\cdot 1}$
$R_{i\cdot 2}$	$N_{i2} - R_{i\cdot 2}$
$\vdots$	$\vdots$
$R_{i\cdot r}$	$N_{ir} - R_{i\cdot r}$

$, i = 1,\ldots,k ,$

derive from the first product term of the above distribution.

Reference to the original binomial distributions from whence a table derives will show what particular parametric function of survival and recovery rates is being tested for equality over the r data sets. For the above example, $\rho_{i1} = \rho_{i2} = \cdots = \rho_{ir}$ is being tested.

Chapter 6. Comprehensive Computer Programs

The calculation of the various point estimates and their associated standard errors and covariances and tests of hypotheses is a laborious task, even for small sets of banding and recovery data. The computational requirements are large when the Maximum Likelihood equations must be solved numerically (i.e., Models 2, 3, H_{01} and H_{02}). Rounding errors can become significant and mathematical errors are nearly unavoidable if the calculations are performed by hand. To alleviate these problems, we developed two comprehensive computer programs to allow estimation and testing to be done inexpensively, accurately, and quickly. In addition, White (1983) developed program SURVIV and Conroy and Williams (1985) developed program MULT. These are extremely powerful computer programs allowing the more sophisticated user to make ML estimates and efficient tests based on more complex models. Use of the output from these programs allows the biologist to concentrate on the interpretation of the results, rather than on the tedious and often complex calculations involved.

The first program, ESTIMATE, computes estimates and test statistics for adult banding data – where parameters can be assumed to be independent of age. Point estimates of the parameters and their sampling variances and various test statistics under Models 0, 1, 2, and 3 are computed by this program.

The second program, BROWNIE, permits estimation under the various age-specific models: H_{01}, H_{02}, H_1, H_2, H_3, H_4, H_5, and H_6. Goodness of fit tests and tests between these models are also computed by BROWNIE. In addition, a chi-square test for differences between adult male and female recovery and survival rates (Section 5.1) is provided as an option.

Programs SURVIV and MULT are very general in that the user can specify the expectation of each R_{ij} cell in terms of the unknown parameters to be estimated. Arbitrary subsets of parameters can be set equal, given an *a priori* value, or constrained to a fixed or variable range. In addition, structural relations can be specified among parameters or even between parameters and auxillary variables. Advanced hypothesis testing is possible through likelihood ratio tests. Unequal intervals between times of banding can be treated, including the models given in Chapter 7. Program SURVIV allows Monte Carlo studies to be easily performed.

All four computer programs are now written in FORTRAN 77 and are designed to be run on nearly any modern computer. ESTIMATE and BROWNIE require a machine with 256k bytes of memory and a FORTRAN 77 compiler, while SURVIV and MULT require 512 k bytes of memory and a hard disk. All mainframe computers satisfy these requirements. Within the past year or two, many desktop microcomputers have achieved these capabilities (e.g., the IBM PC class and various compatible machines).

Most users will employ programs ESTIMATE and BROWNIE for the analysis of band recovery data. These programs are easy to use on a large mainframe computer. We have provided standard input formats (Figs. 6.1 and 6.2 for ESTIMATE and Figs. 6.3 and 6.4 for BROWNIE). The standard (data entry) format and appropriate instructions appear back-to-back on forms which are available. Typical data sets would be coded onto the form and entered as 5 to 25 lines for adult data, 8 to 46 lines for models H_{01}-H_3 which treat adult and young data, and 11 to 68 lines for models H_4-H_6 which treat the three-age-class case. The examples used in Chapter 2-5 were derived by the two programs. Versions of ESTIMATE and BROWNIE are available to be run interactively via a remote terminal system as part of a mainframe computer or on a desktop microcomputer. Conceptually, the information required is the same (i.e., Figs. 6.1-6.4), however, the programs ask the user for the information, accepts the data from the terminal keyboard, and then provides the analysis on a video (CRT) screen or as printed output.

Programs ESTIMATE and BROWNIE for large mainframe computers are available on 9-track, 800-BPI magnetic tapes from the Chief, EDP Section, Office of Migratory Bird Management, Patuxent Wildlife Research Center, Laurel, Maryland 20708. Generally, tapes will be loaned to the individual requesting the programs. The user will be asked to copy the tape and return the original tape so it can be loaned to others. The tape will be sent on a loan basis without charge; the user must pay only the cost of returning the tape. Data entry forms are also available upon request. Hundreds of copies of these two programs have been distributed over the past seven years. Users may wish to check nearby universities or state conservation agency offices to see if these software packages can be obtained locally.

Copies of programs SURVIV and MULT can be obtained from the original authors (see Appendix C where the papers by White (1983) and Conroy and Williams (1985) are reprinted). These are large, complex programs with many advanced features. The user must be fully acquainted with the contents of this *Handbook* or seek the help of a statistician before contemplating the use of these programs.

When using a mainframe computer, the user is usually most concerned with costs, while time is more of concern to users of microcomputers. Here we will attempt to provide some orders of magnitude for these resource requirements. The cost of compiling ESTIMATE or BROWNIE on a mainframe computer may be in the $10-15 range, depending on the user's computer facility. The cost of running a fairly large data set, say $k = \ell = 12$, is often in the

$1-2 range, again depending on the computer used and local billing rates. The cost per data set often drops substantially if several data sets are submitted as a single computer run. Programs SURVIV and MULT may cost $10-20 to compile, but run costs may be significantly higher for large data sets and complex models. We have seen a single large data set cost $40 using program SURVIV when the model is very complex. Still, these costs are negligible compared to the initial cost of obtaining the basic data in the field.

The widespread availability of microcomputers now allows many biologists the opportunity to use ESTIMATE and BROWNIE on an interactive basis. The following information will provide some run times for data sets in this *Handbook*, using an IBM PC/AT with a 80287 coprocessor (data were previously stored on a hard disk).

Program	Date Set	Page	Time (seconds)
ESTIMATE	Male Wood Ducks	21	14
	Male Mallards	25	23
	Male Mallards	45	23
BROWNIE	Young and Adult Mallards	58	18
	Synthetic	118	28
	Male and Female Mallards	146	10

These times are negligible, but do not include time required for the entry of the data (N_i and R_{ij}). A user with reasonable familiarity with a desktop computer might be able to fully analyze a typical data set in 10-20 minutes. This would include providing the answers to queries for information during an interactive session, waiting perhaps 30-200 seconds, and printing the results.

We have consulted with literally hundreds of biologists concerning the analysis of their data over the past seven years. We are willing to try to continue this service as it has provided us with impressions on how this series of models is supported by real data and showed some extensions that were needed.

Figure 6.1 Coding instructions for adult recovery data.

Instructions for Program Estimate

This form is to be used in conjunction with the FORTRAN IV computer program to estimate time-specific survival and recovery rates from banding and recovery data (see Anderson, Kimball & Fiehrer. 1974. J. Wildl. Manage. 38(2):369-370). The information on this coding form is to be punched onto a standard 80-column card. Each line represents one card. All numbers must be integers (no decimal points) and must be in the right-most columns in each field.

The first line (card) is used to identify the data set: e.g.,"PINK-FOOTED GOOSE DATA (FROM BOYD. 1956. J. ANIMAL ECOL.)." This is a Header Card and will identify the information on the computer output.

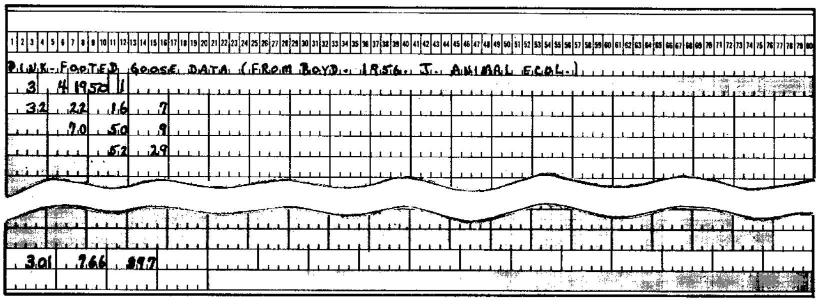
The second line (card) is a Parameter Card and contains four fields:

1. The number of years of banding must appear in columns 1-3 (maximum is 20).
2. The number of years of recovery must appear in columns 4-6 (maximum is 20).
3. The first year of the banding study must appear in columns 7-10, (e.g., 1950).
4. Column 12 must contain the number 1.

The next lines (group of cards) is for the array of recovery information. There must be a minimum of 2 years of recovery information and a maximum of 20. Each field is 4 columns wide. The number of these cards should agree with the number punched in columns 1-3 of the Parameter Card. This group of cards is termed Recovery Cards.

The final line (2 lines if more than 16 years of banding are involved) is for the number of birds banded each year. Each field is 5 columns wide. The number of birds banded the first year must appear in columns 1-5, the number banded the second year must appear in columns 6-10, etc.

An example follows:



Card #				
1	PINK-FOOTED GOOSE DATA (FROM BOYD. 1956. J. ANIMAL ECOL.)			
2	3	4	1950	1
3	32	22	16	7
4		70	50	9
5			52	29
6	301	766	897	

In this example, birds were banded in 3 years and recoveries were recorded in 4 years. Banding began in 1950 when 301 birds were banded; 32 recoveries were reported in 1950 from the cohort banded in 1950. During 1952, 16 recoveries were reported from the cohort banded in 1950, 50 recoveries were reported from the cohort banded in 1951, and 52 recoveries were reported from the cohort banded in 1952. 897 birds were banded during the final year of the experiment (1952). This data set would be punched onto 6 cards. Other data sets could follow. This example data set is discussed and analyzed by Seber (1970. Biometrika).

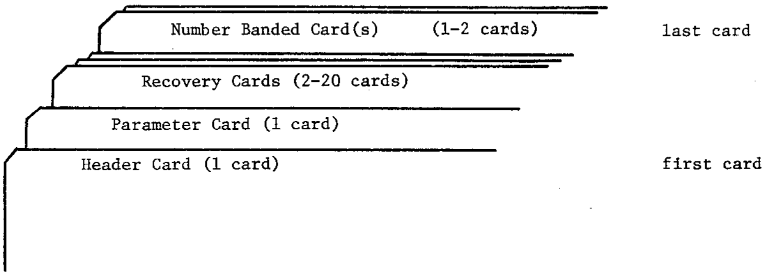
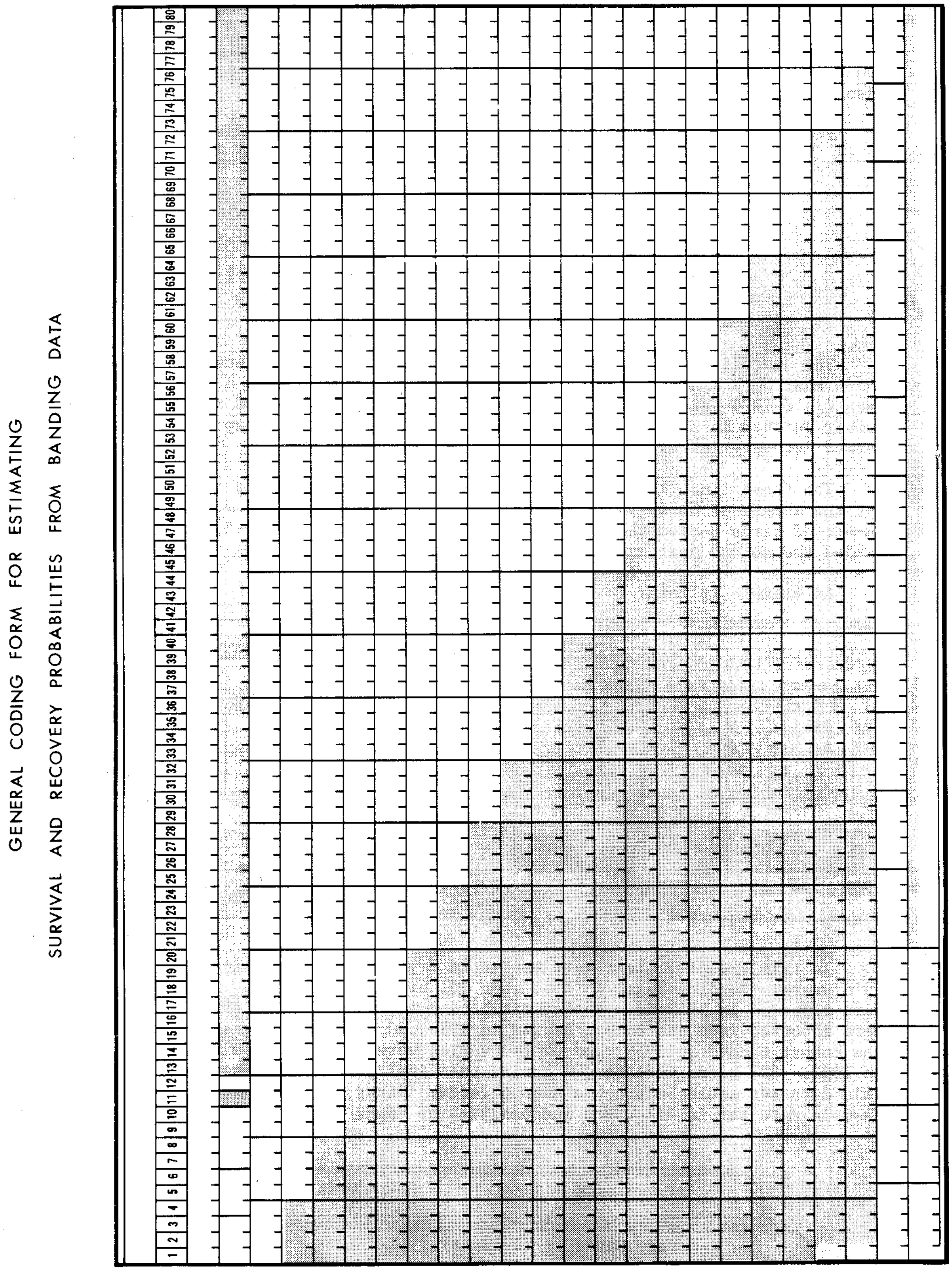


Figure 6.2 General coding form for adult recovery data.



GENERAL CODING FORM FOR ESTIMATING
SURVIVAL AND RECOVERY PROBABILITIES FROM BANDING DATA

Figure 6.3 Coding instructions for adult and young recovery data.

Instructions for Program Brownie

The form is to be used in conjunction with the FORTRAN IV computer program to estimate age-specific survival and recovery rates from banding and recovery data (see Brownie and Robson. 1974. Cornell Biometrics Unit. Paper 514-M. 27pp). The information on this coding form is to be punched onto standard 80-column cards. Each line represents one card. All numbers must be integers (no decimal points) and must be in the right-most columns in each field. Nothing should be coded in the shaded areas of the form. No blank cards should be included.

The first line (card) is used to identify the data set: e.g., YOUNG AND ADULT MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959. This is a Header Card and will identify the information on the computer output.

The second line (card) is a Parameter Card and contains 4 fields:

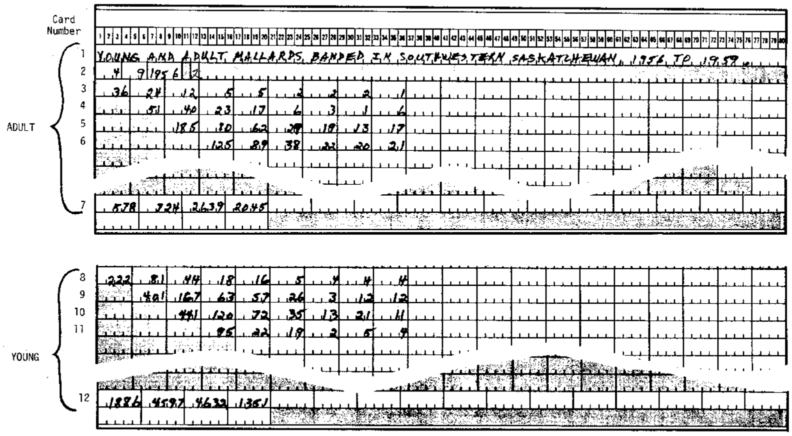
1. The number of years of banding must appear in columns 1-3 (maximum is 20).
2. The number of years of recovery must appear in columns 4-6 (maximum is 20).
3. The first year of the banding study must appear in columns 7-10(e.g., 1950).
4. A numeric code to denote the type of analysis desired must appear in column 12:
    1 a statistical test that the survival and recovery rates are equal for adult males and adult females. If this case is specified, recovery and banding for adult males and adult females are entered, rather than age classes.
    2 parameter estimation and testing for banding studies involving both young and adult birds--this is the usual case and if this code is left blank, it is the default value. The adult recovery data and banding totals are entered before the associated information for young--see example below.
    3 parameter estimation and testing for banding studies involving young, subadult, and adult birds. In this case, three recovery matrices and banding totals are entered, in the following order: adult, young, and subadult.

The next lines (group of cards) are for the array of recovery information. There must be a minimum of 2 years of recovery information and a maximum of 20. Each field is four columns wide. The number of these cards should agree with the number punched in columns 1-3 of the Parameter Card. This group of cards represents the recovery data for adults.

The next line (2 lines if more than 16 years of banding are involved) is for the number of adult birds banded each year. Each field is 5 columns wide. The number of adult birds banded the first year must appear in columns 1-5, the number banded the second year must appear in columns 6-10, etc.

Information concerning the recovery data and banding totals for young are entered on the following cards--similar to the bracketed information above. A second coding form is required for coding the recovery and banding data for young birds (however, lines 1 and 2 should be ignored). The number of years of banding and recovery years must match those of the adult data. No column or row total in the recovery arrays should be zero.

An example follows for young and adult data (information for the test of males vs. females should be entered in a similar manner):



	Card Number									
ADULT	1	YOUNG AND ADULT MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.								
	2	4	9	1956	2					
	3	36	24	12	5	5	2	2	2	1
	4		51	40	23	17	6	3	1	6
	5			185	80	62	28	19	13	17
	6				125	89	38	22	20	21
	7	578	724	2639	2045					
YOUNG	8	222	81	44	18	16	5	4	4	4
	9		401	167	63	57	26	3	12	12
	10			441	120	72	35	13	21	11
	11				95	22	19	2	5	9
	12	1886	4597	4632	1351					

In this example birds were banded in 4 years and recoveries were recorded in 9 years (line 2). Banding began in 1956 when 578 adult birds and 1886 young birds were banded (lines 7 and 12). In 1956, 36 recoveries were reported from the adults banded in 1956 (line 3). In 1958, 167 recoveries were reported from the young birds banded in 1957 (line 9). 2045 adult birds and 1351 young birds were banded during the final year of the study (lines 7 and 12). This data set would be punched onto 12 cards (two coding forms would be required). Other data sets could follow. The order of the data cards is illustrated below for the two most common cases:

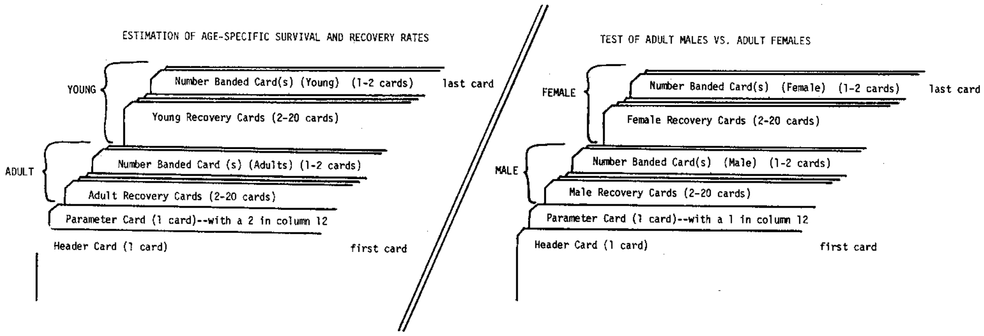
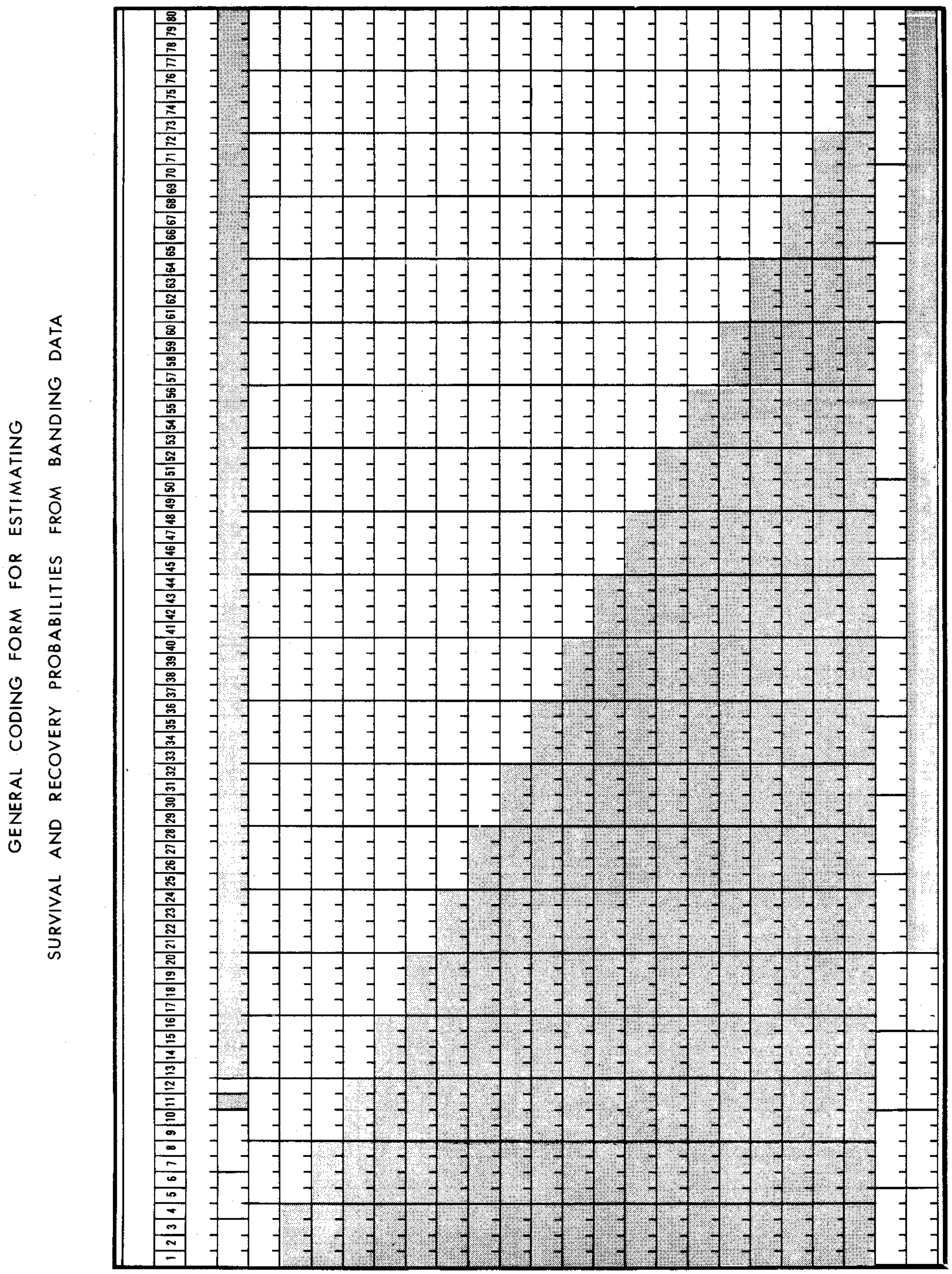


Figure 6.4 General coding form for adult and young recovery data.



Chapter 7. Analysis of Experiments Where Banding is Done Twice a Year

7.1 Introduction

The models introduced in this section relate to the specific experimental situation where adults are banded twice a year, both before and after the hunting season. These models are of interest because they show how twice-a-year banding studies are likely to provide more information about the effects of exploitation and environmental conditions on survival than do once-a-year banding studies.

These models allow a partition of the annual survival rate into a survival rate for the period between pre- and postseason bandings (which includes the hunting season), and a survival rate for the period following postseason banding (which includes the nesting season). If there is a tendency for natural mortality to be "compensatory," then survival during the postseason period should be high following a season when hunting pressure was high, and lower following a season when hunting was light. On the other hand, there may be situations when increased survival in the postseason or nesting period cannot compensate fully for the depletion of the population during the preceding hunting season, resulting in a lower annual survival rate. For example, this may occur when the population has been severely depleted by being very heavily hunted, or when environmental conditions prevailing during the nesting season are extremely adverse. The models of this section thus provide a method for obtaining information about the effects of hunting and environmental factors on survival by permitting estimation of "semi"-annual survival rates and recovery rates.

The models presented here are for data from adult birds only but analogous models could be developed for the situation of Chapter 3 when data from young birds are available. Use of the models is also restricted largely to data from resident species such as grouse, quail, and pheasant. Only resident species are considered because migratory species may migrate to very different areas so that pre- and post-season bandings may be carried out on different population segments which may be characterized by different parameters. Exceptions to this are certain species of geese which breed and winter in well-defined regions (e.g., dusky Canada goose).

Tagging twice a year (i.e., spring and fall) is a common practice in many studies on fish populations and often the data obtained usually include both live recaptures and dead recoveries. The analysis of only the dead recoveries by the methods here would not be efficient because the live recaptures would not be used. This is discussed in Section 8.2.

Data for which the models of this chapter are appropriate do not seem to be commonly available; consequently, analyses are illustrated with artificial data. However, the potential of these models for providing information about the effects of the environment and/or hunting on survival motivates their being included here.

Notation and Definitions

In the two-age-class situation of Chapter 3, recoveries are obtained from two classes of banded birds, adult or young, both banded at the same time. In this chapter we consider banding twice a year (adults only) and recoveries are obtained from two classes of banded birds depending on whether banding was done pre- or post-season. These two types of bandings and recoveries are distinguished in this chapter using a notation similar to that of Chapter 3 with the appropriate interpretation.

We consider only experiments where band recoveries do not continue beyond the year of the last or k^{th} preseason release. Thus $k = \ell$ and $s = 0$ in this chapter.

A year of the experiment is the period between consecutive preseason bandings. That is, the i^{th} year of the experiment is the period between the i^{th} and $(i+1)^{th}$ preseason bandings, $i = 1, \dots, k-1$, and the k^{th} or last year is the year following the k^{th} preseason banding. Banding should occur at the same times each year. For example, a possible program for pheasants would be preseason banding in mid-August and postseason banding in mid-December each year.

We make the following definitions:

N_i = the number of adults banded and released in the i^{th} preseason banding, $i = 1, \dots, k$.
M_i = the number of adults banded and released in the i^{th} postseason banding, $i = 1, \dots, k-1$.
R_{ij} = the number of bands returned in year j from the i^{th} preseason release, $j = i, \dots, k$, $i = 1, \dots, k$.
Q_{ij} = the number of bands returned in year j from the i^{th} postseason release, $j = i+1, \dots, k$, $i = 1, \dots, k-1$.

Note that Q_{ii} is not defined as there can be no recoveries in the i^{th} hunting season from the i^{th} postseason release, $i = 1, \dots, k$.

The data for a 4-year study (i.e., $k = 4$) are represented as in Table 7.1 below.

Table 7.1. *Representation of data for a 4-year study with banding twice a year.*

Year and time of banding		Number banded	Year of recovery 1	2	3	4	Row totals
1	preseason	N_1	T_1 R_{11}	R_{12}	R_{13}	R_{14}	R_1
	postseason	M_1		Q_{12}	Q_{13}	Q_{14}	Q_1
2	preseason	N_2		T_2 R_{22}	R_{23}	R_{24}	R_2
	postseason	M_2			Q_{23}	Q_{24}	Q_2
3	preseason	N_3			T_3 R_{33}	R_{34}	R_3
	postseason	M_3				Q_{34}	Q_3
4	preseason	N_4				R_{44}	R_4
	Column totals		C_1	C_2	C_3	$C_4=T_4$	

Subtotals which are used in calculating estimates are indicated in Table 7.1 and are defined below:

Row totals:

$$R_i=\sum_{j=i}^{k} R_{ij} \qquad ,i=1,\ldots,k,$$

$$Q_i=\sum_{j=i+1}^{k} Q_{ij} \qquad ,i=1,\ldots,k-1.$$

Column totals:

$$C_1=R_{11},$$

$$C_j=\sum_{i=1}^{j} R_{ij}+\sum_{i=1}^{j-1} Q_{ij} \qquad ,j=2,\ldots,k.$$

The outlined block totals:

$$T_1=R_1,$$

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

$$(T_k=C_k).$$

Numerical illustrations in this chapter, including calculation of the above subtotals, are all obtained using the synthetic data set shown in Table 7.2.

Table 7.2. *Synthetic data for a 5-year study ($k=5$), showing calculation of subtotals.*

Year and time of banding		Number banded	Year of recovery 1	2	3	4	5	Row totals
1	preseason	550	63	31	17	18	10	139
	postseason	350		30	19	12	6	67
2	preseason	500		48	24	24	11	107
	postseason	400			27	22	12	61
3	preseason	500			41	30	20	91
	postseason	800				70	33	103
4	preseason	400				44	21	65
	postseason	500					37	37
5	preseason	500					42	42
	Column totals C_i		63	109	128	220	192	712
	Block totals T_i		139	250	293	333	192	

The two models presented in this chapter are analogous to Models 1 and 0 of Chapter 2. As in Chapter 2, f represents the annual band recovery rate, but S, the annual survival rate, is represented as the product of two "semi-annual" survival rates. Thus $S_i = h_i n_i$, where

h_i = the survival rate during the period between the i^{th} pre- and post-season bandings, (i.e., the period including the i^{th} hunting season), $i = 1, \ldots, k-1$,

n_i = the survival rate during the period between the i^{th} postseason and the $(i+1)^{th}$ preseason bandings, $i = 1, \ldots, k-1$.

7.2 The Model Under H_7

The first model we consider is analogous to Model 1 of Chapter 2, and is called the model under H_7. The assumptions of H_7 are:

(1) Annual recovery rates (f_i) and "semi-annual" survival rates $(h_i$ and $n_i)$ are year-specific but independent of age; and

(2) reporting rates are independent of the time of release.

The parameters of the model under H_7 are h_i, n_i, and f_i, where h_i and n_i are defined above and f_i = the recovery rate in year i, for all banded adults alive at the start of year i, i.e., after the i^{th} preseason release, $i = 1, \ldots, k$.

For this model the expected or average numbers of band recoveries can be expressed in terms of N_i, M_i, f_i, h_i, and n_i and presented in the same way that the data are presented in Tables 7.1 and 7.2. Table 7.3 gives the model structure under H_7 for $k = 4$.

Table 7.3. *Expected numbers of band recoveries for a 4-year study for the model under H_7.*

Year and time of release	Number banded	Year of recovery			
		1	2	3	4
1 preseason	N_1	$N_1 f_1$	$N_1 h_1 n_1 f_2$	$N_1 h_1 n_1 h_2 n_2 f_3$	$N_1 h_1 n_1 h_2 n_2 h_3 n_3 f_4$
postseason	M_1		$M_1 n_1 f_2$	$M_1 n_1 h_2 n_2 f_3$	$M_1 n_1 h_2 n_2 h_3 n_3 f_4$
2 preseason	N_2		$N_2 f_2$	$N_2 h_2 n_2 f_3$	$N_2 h_2 n_2 h_3 n_3 f_4$
postseason	M_2			$M_2 n_2 f_3$	$M_2 n_2 h_3 n_3 f_4$
3 preseason	N_3			$N_3 f_3$	$N_3 h_3 n_3 f_4$
postseason	M_3				$M_3 n_3 f_4$
4 preseason	N_4				$N_4 f_4$

Estimation of Parameters

ML estimators of f_i, h_i and n_i are

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

$$\hat{h}_i = \frac{R_i}{N_i} \left(1 - \frac{C_i}{T_i}\right) \frac{M_i}{Q_i} = \frac{R_i}{N_i} \frac{T_i - C_i}{T_i} \frac{M_i}{Q_i} \qquad , i = 1, \ldots, k-1,$$

$$\hat{n}_i = \frac{Q_i}{M_i} \frac{N_{i+1}}{R_{i+1}} \qquad , i = 1, \ldots, k-1,$$

where R_i, Q_i, C_i, and T_i are the subtotals defined in Section 7.1.

These estimators are easily evaluated for data sets where k is not too large. The synthetic data of Table 7.2 give, for $i = 2$,

$$\hat{f}_2 = \frac{R_2 \times C_2}{N_2 \times T_2} = \frac{107 \times 109}{500 \times 250} = 0.0933,$$

$$\hat{h}_2 = \frac{R_2 \times (T_2 - C_2) \times M_2}{N_2 \times T_2 \times Q_2} = \frac{107 \times (250 - 109) \times 400}{500 \times 250 \times 61} = 0.7914\,,$$

$$\hat{n}_2 = \frac{Q_2 \times N_3}{M_2 \times R_3} = \frac{61 \times 500}{400 \times 91} = 0.8379\,,$$

and

$$\hat{S}_2 = \hat{h}_2\,\hat{n}_2 = (0.7914) \times (0.8379) = 0.6631\,.$$

A slight modification of the estimators $\hat{h}_i$ and $\hat{n}_i$ will reduce their bias, giving the bias-adjusted ML estimators

$$\tilde{h}_i = \frac{R_i}{N_i}\,\frac{(T_i - C_i)}{T_i}\,\frac{M_i + 1}{Q_i + 1} \qquad , i = 1, \ldots, k-1\,,$$

$$\tilde{n}_i = \frac{Q_i}{M_i}\,\frac{N_{i+1} + 1}{R_{i+1} + 1} \qquad , i = 1, \ldots, k-1\,.$$

For example, for $i = 2$,

$$\tilde{h}_2 = \frac{R_2 \times (T_2 - C_2) \times (M_2 + 1)}{N_2 \times T_2 \times (Q_2 + 1)} = \frac{107 \times (250 - 109) \times 401}{500 \times 250 \times 62} = 0.7806\,,$$

$$\tilde{n}_2 = \frac{Q_2 \times (N_3 + 1)}{M_2 \times (R_3 + 1)} = \frac{61 \times 501}{400 \times 92} = 0.8305\,.$$

Sampling Variances, Standard Errors, and Confidence Intervals

The procedure for obtaining confidence intervals for the above estimators is the same as that in earlier chapters. Estimators of sampling variances of the various parameter estimates are given below (e.g., the sampling variance of $\hat{f}_i$ is estimated by $\mathrm{var}(\hat{f}_i)$):

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

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

$$\mathrm{var}(\tilde{h}_i) = (\tilde{h}_i)^2 \left[\frac{1}{R_i} - \frac{1}{N_i} + \frac{1}{Q_i} - \frac{1}{M_i} + \frac{1}{T_i - C_i} - \frac{1}{T_i}\right] \qquad , i = 1, \ldots, k-1,$$

$$\mathrm{var}(\hat{n}_i) = (\hat{n}_i)^2 \left[\frac{1}{Q_i} - \frac{1}{M_i} + \frac{1}{R_{i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1, \ldots, k-1\,,$$

$$\mathrm{var}(\tilde{n}_i) = (\tilde{n}_i)^2 \left[\frac{1}{Q_i} - \frac{1}{M_i} + \frac{1}{R_{i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1, \ldots, k-1.$$

Estimates of the corresponding standard errors are given by, for example, $\mathrm{se}(\hat{f}_i) = \sqrt{\mathrm{var}(\hat{f}_i)}$.

For example, using the data of Table 7.2,

$$\mathrm{var}(\hat{f}_2) = (\hat{f}_2)^2 \left[\frac{1}{R_2} - \frac{1}{N_2} + \frac{1}{C_2} - \frac{1}{T_2}\right] = (0.0933)^2 \left[\frac{1}{107} - \frac{1}{500} + \frac{1}{109} - \frac{1}{250}\right] = 0.00010899$$

$$\mathrm{se}(\hat{f}_2) = \sqrt{0.00010899} = 0.0104\,,$$

$$1.96 \times \mathrm{se}(\hat{f}_2) = 0.0204\,,$$

and the estimated 95% confidence interval for f_2 is (0.0933-0.0204, 0.0933 + 0.0204) or (0.0729, 0.1137). As in previous chapters these are only approximate 95% confidence intervals, valid for "large" sample sizes N_i and M_i.

The data of Table 7.2 are used to evaluate all the estimators $\hat{f}_i$, $\hat{h}_i$, and $\tilde{n}_i$, and to obtain estimates of the corresponding confidence intervals as described above, and the results are presented in Table 7.4.

If required, the ML estimator of the annual survival rate is

$$\hat{S}_i = \hat{h}_i \hat{n}_i = \frac{R_i}{N_i} \; \frac{T_i - C_i}{T_i} \; \frac{N_{i+1}}{R_{i+1}} \qquad , \; i = 1, \ldots, k-1 \; ,$$

with sampling variance estimated by

$$\text{var}(\hat{S}_i) = (\hat{S}_i)^2 \left[\frac{1}{R_i} - \frac{1}{N_i} + \frac{1}{R_{i+1}} - \frac{1}{N_{i+1}} + \frac{1}{T_i - C_i} - \frac{1}{T_i}\right] .$$

We note the bias-adjusted estimator is

$$\tilde{S}_i = \frac{R_i}{N_i} \; \frac{T_i - C_i}{T_i} \; \frac{N_{i+1} + 1}{R_{i+1} + 1} \; ,$$

which is *not* the same as $\tilde{h}_i \tilde{n}_i$.

The large confidence intervals in Table 7.4 indicate that if the annual recovery rate is approximately 10% (or less), then values of N_i and M_i of about 400-500 are too small to provide reliable estimates of semiannual survival rates.

Table 7.4. *Parameter estimates from the data of Table 7.2, obtained by the model under* H_7°.

i	Estimate $\hat{f}_i$	Standard error	95% Confidence interval	Estimate $\tilde{h}_i$	Standard error	95% Confidence interval	Estimate $\tilde{n}_i$	Standard error	95% Confidence interval
1	0.1145	0.0136	0.0878-0.1412	0.7133	0.1091	0.4995-0.9271	0.8880	0.1237	0.6455-1.1305
2	0.0933	0.0104	0.0729-0.1137	0.7806	0.1218	0.5418-1.0193	0.8305	0.1256	0.5843-1.0767
3	0.0795	0.0092	0.0615-0.0975	0.7894	0.1119	0.5701-1.0087	0.7823	0.1143	0.5583-1.0063
4	0.1074	0.0129	0.0821-0.1327	0.7270	0.1521	0.4289-1.0251	0.8622	0.1866	0.4965-1.2279
5	0.0840	0.0124	0.0597-0.1083	—	—	— —	—	—	— —

Sampling Covariances and Correlations

For large N_i and M_i estimators of the nonzero or non-negligible covariances between $\hat{f}_i$, $\tilde{h}_i$ and $\tilde{n}_i$ are

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

$$\text{cov}(\tilde{h}_i, \tilde{n}_i) = -\tilde{h}_i \tilde{n}_i \left[\frac{1}{Q_i} - \frac{1}{M_i}\right] \qquad , i = 1, \ldots, k-1 \; ,$$

$$\text{cov}(\hat{f}_{i+1}, \tilde{n}_i) = -\hat{f}_{i+1} \tilde{n}_i \left[\frac{1}{R_{i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1, \ldots, k-1 \; ,$$

$$\text{cov}(\tilde{h}_{i+1}, \tilde{n}_i) = -\tilde{h}_{i+1} \tilde{n}_i \left[\frac{1}{R_{i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1, \ldots, k-2 \; .$$

Estimates of covariances involving the estimators $\hat{n}_i$ and $\hat{h}_i$ are obtained by substituting $\hat{n}_i$ and $\hat{h}_i$ for $\tilde{n}_i$ and $\tilde{h}_i$ in the above formulae.

Correlations are estimated by using the above covariance and variance estimators as described in Chapter 2. Thus, for example, the estimate of the correlation between $\hat{f}_i$ and $\tilde{h}_i$ is

$$\text{corr}(\hat{f}_i, \tilde{h}_i) = \frac{\text{cov}(\hat{f}_i, \tilde{h}_i)}{\text{se}(\hat{f}_i)\, \text{se}(\tilde{h}_i)} \qquad , i = 1, \ldots, k-1 \; .$$

For $i = 2$,

$$\text{cov}(\hat{f}_2, \tilde{h}_2) = \hat{f}_2 \tilde{h}_2 \left[\frac{1}{R_2} - \frac{1}{N_2} - \frac{1}{T_2}\right] = (0.0933)\,(0.7806) \left[\frac{1}{107} - \frac{1}{500} - \frac{1}{250}\right] = 0.000244$$

and

$$\text{corr}(\hat{f}_2,\tilde{h}_2)=\frac{\text{cov}(\hat{f}_2,\tilde{h}_2)}{\text{se}(\hat{f}_2)\,\text{se}(\tilde{h}_2)}=\frac{0.000244}{(0.0104)\,(0.1218)}=0.193\,.$$

Covariances and correlations, estimated in this way using the data of Table 7.2, are presented in Table 7.5.

Table 7.5. *Estimated covariances and correlations for the example under H_7.*

	Covariance $\text{cov}(\hat{f}_i,\tilde{h}_i)$	Correlation $\text{corr}(\hat{f}_i,\tilde{h}_i)$	Covariance $\text{cov}(\tilde{h}_i,\tilde{n}_i)$	Correlation $\text{corr}(\tilde{h}_i,\tilde{n}_i)$
$i=1$	−0.000148	−0.100	−0.007644	−0.566
2	0.000244	0.193	−0.009007	−0.589
3	0.000350	0.340	−0.005224	−0.408
4	0.000772	0.393	−0.015687	−0.553
	$\text{cov}(\hat{f}_{i+1},\tilde{n}_i)$	$\text{corr}(\hat{f}_{i+1},\tilde{n}_i)$	$\text{cov}(\tilde{h}_{i+1},\tilde{n}_i)$	$\text{corr}(\tilde{h}_{i+1},\tilde{n}_i)$
$i=1$	−0.000609	−0.473	−0.005092	−0.338
2	−0.000593	−0.514	−0.005893	−0.419
3	−0.001083	−0.735	−0.007328	−0.422
4	−0.001580	−0.683	—	—

These estimated correlations are discussed further in Section 7.5.

7.3 The Model Under H_8

The second model considered in this chapter is analogous to Model 0 of Chapter 2, and is called the model under H_8. The assumptions of H_8 are similar to those of H_7 except that the reporting rate for newly released birds is assumed to differ from that for survivors of earlier releases. Recall that in Model 0 dispersion due to migration was thought to contribute to this difference in the reporting rate for new releases, whereas the methods of this chapter are applicable mainly to resident species. Even without migration, however, dispersion after banding may be sufficient to cause a difference in the reporting rate for new releases and the model under H_8 is therefore included here.

The assumptions of H_8 are:

(1) Annual recovery rates and semiannual survival rates are year-specific but independent of age; and

(2) in any year, the reporting rate for just-released birds is different from that for survivors of previous releases (and hence the recovery rates are different also).

The parameters are: h_i and n_i $(i=1,\dots,k-1)$, as defined in Section 7.1, and

f_i^* = recovery rate in year i for banded adults released at the i^{th} preseason banding, $i=1,\dots,k$,
f_i = recovery rate in year i for banded adults alive at the start of year i, but released before the i^{th} preseason banding, $i=2,\dots,k$.

The structure of the model under H_8 is reflected in the expected numbers of band returns expressed in terms of N_i, M_i, f_i, f_i^*, h_i, and n_i as presented, for example, in Table 7.6 for a 4-year study.

Table 7.6. *Expected numbers of band recoveries for a 4-year study for the model under H_8.*

		Year of recovery			
Year and time of banding	Number banded	1	2	3	4
1 preseason	N_1	$N_1 f_1^*$	$N_1 h_1 n_1 f_2$	$N_1 h_1 n_1 h_2 n_2 f_3$	$N_1 h_1 n_1 h_2 n_2 h_3 n_3 f_4$
postseason	M_1		$M_1 n_1 f_2$	$M_1 n_1 h_2 n_2 f_3$	$M_1 n_1 h_2 n_2 h_3 n_3 f_4$
2 preseason	N_2		$N_2 f_2^*$	$N_2 h_2 n_2 f_3$	$N_2 h_2 n_2 h_3 n_3 f_4$
postseason	M_2			$M_2 n_2 f_3$	$M_2 n_2 h_3 n_3 f_4$
3 preseason	N_3			$N_3 f_3^*$	$N_3 h_3 n_3 f_4$
postseason	M_3				$M_3 n_3 f_4$
4 preseason	N_4				$N_4 f_4^*$

Note that the parameters n_3 and f_4 are not separately identifiable.

Estimation of Parameters

ML estimators of the recovery rates are

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

$$\hat{f}_i = \frac{R_i - R_{ii}}{N_i}\,\frac{C_i - R_{ii}}{T_i - R_i - C_i + R_{ii}} \qquad , i=2,\ldots,k-1.$$

Bias-adjusted ML estimators of semi annual survival rates are

$$\tilde{h}_i = \frac{R_i - R_{ii}}{N_i}\,\frac{M_i + 1}{Q_i + 1} \qquad , i=1,\ldots,k-1,$$

$$\tilde{n}_i = \frac{Q_i}{M_i}\,\frac{N_{i+1}+1}{R_{i+1} - R_{i+1,i+1} + 1}\left(\frac{T_{i+1} - R_{i+1} - C_{i+1} + R_{i+1,i+1}}{T_{i+1} - R_{i+1}}\right) \qquad , i=1,\ldots,k-2.$$

The corresponding unadjusted ML estimators are

$$\hat{h}_i = \frac{R_i - R_{ii}}{N_i}\,\frac{M_i}{Q_i}$$

$$\hat{n}_i = \frac{Q_i}{M_i}\,\frac{N_{i+1}}{R_{i+1} - R_{i+1,i+1}}\left(\frac{T_{i+1} - R_{i+1} - C_{i+1} + R_{i+1,i+1}}{T_{i+1} - R_{i+1}}\right).$$

Finally

$$\widehat{n_{k-1}f_k} = \frac{Q_{k-1}}{M_{k-1}}.$$

For example, for the data of Table 7.2,

$$\hat{f}_3^* = \frac{R_{33}}{N_3} = \frac{41}{500} = 0.0820,$$

$$\hat{f}_3 = \frac{(R_3 - R_{33}) \times (C_3 - R_{33})}{N_3 \times (T_3 - R_3 - C_3 + R_{33})} = \frac{(91-41)\times(128-41)}{500\times(293-91-128+41)} = 0.0757,$$

$$\tilde{h}_3 = \frac{(R_3 - R_{33}) \times (M_3 + 1)}{N_3 \times (Q_3 + 1)} = \frac{(91-41)\times 801}{500 \times 104} = 0.7702,$$

$$\tilde{n}_3 = \frac{Q_3 \times (N_4 + 1) \times (T_4 - R_4 - C_4 + R_{44})}{M_3 \times (R_4 - R_{44} + 1) \times (T_4 - R_4)} = \frac{103 \times 401 \times (333-65-220+44)}{800 \times (65-44+1) \times (333-65)} = 0.8056.$$

Sampling Variances, Standard Errors, and Confidence Intervals

Estimators of the sampling variances of $\hat{f}_i^*$, $\hat{f}_i$, $\tilde{h}_i$ and $\tilde{n}_i$, for large N_i and M_i are

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

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

$$\mathrm{var}(\tilde{h}_i) = (\tilde{h}_i)^2\left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i} + \frac{1}{Q_i} - \frac{1}{M_i}\right] \qquad , i=1,\ldots,k-1,$$

$$\mathrm{var}(\tilde{n}_i) = (\tilde{n}_i)^2\left[\frac{1}{Q_i} - \frac{1}{M_i} + \frac{1}{R_{i+1} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} + \frac{1}{T_{i+1} - R_{i+1} - C_{i+1} + R_{i+1,i+1}} - \frac{1}{T_{i+1} - R_{i+1}}\right] \qquad , i=1,\ldots,k-2.$$

Variance estimators for $\hat{h}_i$ and $\hat{n}_i$ are obtained by substituting $\hat{h}_i$ and $\hat{n}_i$ for $\tilde{h}_i$ and $\tilde{n}_i$ in the appropriate expressions.

These variance estimators are used to obtain estimates of the standard errors and 95% confidence intervals as described in Section 7.2. For example,

$$\mathrm{var}(\hat{f}_3) = (\hat{f}_3)^2 \left[\frac{1}{R_3 - R_{33}} - \frac{1}{N_3} + \frac{1}{T_3 - R_3 - C_3 + R_{33}} + \frac{1}{C_3 - T_{33}}\right] = (0.0757)^2 \left[\frac{1}{50} - \frac{1}{500} + \frac{1}{115} + \frac{1}{87}\right] = 0.00021885\,,$$

and

$$\mathrm{se}(\hat{f}_3) = \sqrt{0.00021885} = 0.0148\,.$$

To compute the 95% confidence interval for f_3 first compute $1.96 \times \mathrm{se}(\hat{f}_3) = 0.0290$, then the desired interval is $(0.0757 - 0.0290, 0.0757 + 0.0290)$ or $(0.0467, 0.1047)$.

Estimates of parameters, their standard errors, and confidence intervals under H_8 from the data in Table 7.2 are given in Table 7.7.

Table 7.7. *Parameter estimates from the data of Table 7.2, obtained by the model under H_8.*

i	Estimate $\hat{f}_i$	Standard error	95% Confidence interval	Estimate $\hat{f}_i^*$	Standard error	95% Confidence interval
1	–	–	–	0.1145	0.0136	0.0878-0.1412
2	0.0878	0.0183	0.0519-0.1237	0.0960	0.0132	0.0701-0.1219
3	0.0757	0.0148	0.0467-0.1047	0.0820	0.0123	0.0579-0.1061
4	0.1004	0.0249	0.0516-0.1492	0.1100	0.0156	0.0794-0.1406
5	–	–	–	0.0840	0.0124	0.0597-0.1083
	Estimate $\tilde{h}_i$	Standard error	95% Confidence interval	Estimate $\tilde{n}_i$	Standard error	95% Confidence interval
1	0.6945	0.1063	0.4862-0.9028	0.9296	0.1666	0.6031-1.2561
2	0.7632	0.1296	0.5092-1.0172	0.8529	0.1601	0.5391-1.1667
3	0.7702	0.1253	0.5246-1.0158	0.8056	0.1985	0.4165-1.1947
4	0.6922	0.1833	0.3329-1.0515	–	–	–

Comparison of Tables 7.4 and 7.7 shows that precision is lost by using the H_8 estimators, especially in the case of $\tilde{n}_i$, and the confidence intervals for the survival rates are practically useless. The discussion at the end of Section 2.5 concerning choosing between Models 1 and 0 with a goal of minimizing bias and maximizing precision applies equally to the problem of choosing between the H_7 and H_8 models, with the additional consideration that assumption 2 of H_8 is less likely to be necessary for resident species.

Sampling Covariances and Correlations

For large N_i and M_i, estimators of the nonzero or non-negligible covariances under H_8 are

$$\mathrm{cov}(\hat{f}_i^*, \hat{f}_i) = -\hat{f}_i^*\, \hat{f}_i / N_i \qquad , i = 2, \ldots, k-1\,,$$

$$\mathrm{cov}(\hat{f}_i^*, \tilde{h}_i) = -\hat{f}_i^*\, \tilde{h}_i / N_i \qquad , i = 1, \ldots, k-1\,,$$

$$\mathrm{cov}(\hat{f}_{i+1}^*, \tilde{n}_i) = \hat{f}_{i+1}^*\, \tilde{n}_i / N_{i+1} \qquad , i = 1, \ldots, k-2\,,$$

$$\mathrm{cov}(\hat{f}_i, \tilde{h}_i) = \hat{f}_i\, \tilde{h}_i \left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i}\right] \qquad , i = 2, \ldots, k-1\,,$$

$$\mathrm{cov}(\hat{f}_{i+1}, \tilde{n}_i) = -\hat{f}_{i+1}\, \tilde{n}_i \left[\frac{1}{R_{i+1} - R_{i+1,i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1, \ldots, k-2\,,$$

$$\mathrm{cov}(\tilde{h}_i,\tilde{n}_i) = -\tilde{h}_i\,\tilde{n}_i\left[\frac{1}{Q_i}-\frac{1}{M_i}\right] \qquad ,i=1,\ldots,k-2\,,$$

$$\mathrm{cov}(\tilde{h}_{i+1},\tilde{n}_i) = -\tilde{h}_{i+1}\,\tilde{n}_i\left[\frac{1}{R_{i+1}-R_{i+1,i+1}}-\frac{1}{N_{i+1}}\right] \qquad ,i=1,\ldots,k-2\,.$$

The corresponding correlations are then estimated in the usual way, for example,

$$\mathrm{cov}(\hat{f}_3,\tilde{h}_3) = \hat{f}_3\tilde{h}_3\left[\frac{1}{R_3-R_{33}}-\frac{1}{N_3}\right] = (0.0757)\times(0.7702)\left[\frac{1}{50}-\frac{1}{500}\right] = 0.001049\,,$$

$$\mathrm{corr}(\hat{f}_3,\tilde{h}_3) = \frac{\mathrm{cov}(\hat{f}_3,\tilde{h}_3)}{\mathrm{se}(\hat{f}_3)\,\mathrm{se}(\tilde{h}_3)} = \frac{0.001049}{(0.0148)\times(0.1253)} = 0.566\,.$$

Table 7.8 gives all the nonnegligible covariances and correlations for the example data of Table 7.2 under H_8.

Note that the estimators $\hat{f}_i$ and $\tilde{h}_i$ are positively correlated, and once again we emphasize that this is a property of the estimators themselves and is not indicative of a similar relationship between the parameters f_i and h_i. This should be clear in this instance, because we would expect the parameters f_i and h_i, if correlated, to be negatively so, i.e., we would expect survival during the hunting season to be low when hunting pressure is high, and vice versa.

Table 7.8. *Estimates of covariances and correlations between the H_8 estimators, evaluated from the data of Table 7.2.*

i	Covariance $\mathrm{cov}(\hat{f}_i^*,\hat{f}_i)$	Correlation $\mathrm{corr}(\hat{f}_i^*,\hat{f}_i)$	Covariance $\mathrm{cov}(\hat{f}_i^*,\tilde{h}_i)$	Correlation $\mathrm{corr}(\hat{f}_i^*,\tilde{h}_i)$
1	–	–	−0.000145	−0.100
2	−0.000017	−0.070	−0.000147	−0.086
3	−0.000012	−0.066	−0.000126	−0.082
4	−0.000028	−0.072	−0.000190	−0.066
	$\mathrm{cov}(\hat{f}_{i+1}^*,\tilde{n}_i)$	$\mathrm{corr}(\hat{f}_{i+1}^*,\tilde{n}_i)$	$\mathrm{cov}(\hat{f}_i,\tilde{h}_i)$	$\mathrm{corr}(\hat{f}_i,\tilde{h}_i)$
1	0.000178	0.081	–	–
2	0.000125	0.078	0.001002	0.422
3	0.000212	0.108	0.001049	0.566
4	0.000116	0.051	0.003136	0.687
	$\mathrm{cov}(\hat{f}_{i+1},\tilde{n}_i)$	$\mathrm{corr}(\hat{f}_{i+1},\tilde{n}_i)$	$\mathrm{cov}(\tilde{h}_i,\tilde{n}_i)$	$\mathrm{corr}(\tilde{h}_i,\tilde{n}_i)$
1	−0.001220	−0.400	−0.007286	−0.411
2	−0.001162	−0.490	−0.009044	−0.436
3	−0.003649	−0.738	−0.005248	−0.211
	$\mathrm{cov}(\tilde{h}_{i+1},\tilde{n}_i)$	$\mathrm{corr}(\tilde{h}_{i+1},\tilde{n}_i)$		
1	−0.010606	−0.491		
2	−0.011824	−0.589		
3	−0.025160	−0.691		

7.4 Tests of the Models

Goodness of Fit Tests

Goodness of fit tests to the models under H_7 and H_8 can be computed in the conventional way, as described, for example, in Section 2.2 in relation to testing fit to Model 1. For the model under H_7, the $\mathbf{E}_{ij}$'s, or estimated expected values, are obtained from entries in Table 7.3 by substituting the H_7 unadjusted ML estimates $\hat{f}_i$, $\hat{h}_i$, and $\hat{n}_i$ for the unknown parameters. These expected values are compared with the observed data values ($\mathbf{O}_{ij}$'s) and the familiar chi-square statistic is obtained as

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

If no pooling of expected values is required, this statistic has k^2-3k+2 degrees of freedom.

For the model under H_8 the expected values are obtained by substituting the H_8 unadjusted ML estimates $\hat{f}_i^*$, $\hat{f}_i$, $\hat{h}_i$, and $\hat{n}_i$ for the unknown parameters in the expectations in Table 7.6. If the assumptions of H_8 hold, the test statistic

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

is chi-square distributed with k^2-4k+4 degrees of freedom (in the absence of pooling).

An alternative method for testing fit to the models under H_7 and H_8 is described in Brownie (1973).

Testing Between the Models

A test of the model under H_7 against the alternative of the more general model under H_8 essentially tests the assumption that recovery rates are the same for new releases and survivors of earlier releases.

The test statistic is computed as the sum of the $k-2$ single degrees of freedom chi-square statistics obtained from the 2 by 2 contingency tables

C_i-R_{ii}	$T_i-R_i-C_i+R_{ii}$	T_i-R_i
R_{ii}	R_i-R_{ii}	R_i
		T_i

$i=2,\ldots,k-1.$

If the assumptions of H_7 hold, the test statistic is chi-square distributed with $k-2$ degrees of freedom. "Large" values of the statistic are taken to indicate that assumption 2 of H_7 is invalid, and the model under H_7 is rejected in favor of the model under H_8.

This is illustrated in Table 7.9 using the data of Table 7.2.

Table 7.9. *The test of* H_7 vs. H_8 *with the data of Table 7.2.*

	Contingency Table		Chi-square value
$i=2$	61	82	0.12
	48	59	
$i=3$	87	115	0.10
	41	50	
$i=4$	176	92	0.10
	44	21	

Total chi-square value with 3 df is 0.32. On the basis of this result there is no reason to reject H_7 in favor of H_8.

7.5 Summary

As stated earlier, the model under H_7 is likely to be the more useful of the two models proposed here. However, there is a need for the development of related models, particularly those that will allow for the live recaptures common in fish studies, and also models that will allow for age-dependence of parameters.

We discuss here one way in which estimates might be used to study relationships between unknown parameters, and the importance of the correlation structure between estimators in so doing.

Suppose it is desired to relate hunting pressure during the season to postseason survival, e.g., it may be postulated that these two parameters are positively correlated. The first problem is that we have no estimate of hunting pressure (H_i), but if we can assume that reporting rates are fairly constant, then variations in recovery rates (f_i) will be largely due to variations in hunting pressure. Thus, if postseason survival (n_i) is in fact positively correlated with hunting pressure, and we have accurate and precise estimates, then we may expect to detect this relationship if we plot $\tilde{n}_i$ against $\hat{f}_i$. However, the problem is not so simple, because such a plot may be influenced to a larger extent by relationships between the sampling distributions of the estimators than by a relationship between the unknown parameters. In this regard we note that the H_7 estimators $\hat{f}_i$ and $\tilde{n}_i$ are uncorrelated (in fact they are stochastically independent), but $\hat{f}_{i+1}$ and $\tilde{n}_i$ are substantially correlated (see Table 7.5). As all other relevant correlations are negligible for large N_i and M_i (see Table 7.5), a possible solution is to use pairs of estimates from every other year (e.g., $\hat{f}_1$, $\hat{n}_1$, $\hat{f}_3$, $\hat{n}_3, \ldots$) in plotting graphs. This, too, has practical problems as only rarely would enough data be available for this procedure to be useful.

We emphasize that in order to make use of such a procedure N_i and M_i must be large enough to provide estimates with high precision. Finally, the use of reward bands (Sections 2.1, 9.2) in conjunction with such a procedure may provide valuable information. Reward band studies are discussed by Henny and Burnham (1976).

Chapter 8. Discussion of Other Points

8.1 Advantages of Modern Methods

The methods developed over the past 15 years offer a number of important advantages over the various ad hoc methods and life table techniques. By "modern methods" we are referring to the development of stochastic models, based on specific and well defined assumptions, and associated inference procedures which are optimal (in some rigorously definable sense) in terms of efficiency of estimation and power of tests of hypotheses. In this section we will summarize the general advantages of modern methods presented in this handbook and in the literature.

Certain properties of a number of methods are summarized in Table 8.1. The methods presented in this handbook are grouped to allow comparison with a number of methods that are commonly used in the analysis of banding data and with various other methods in the literature. No attempt was made to include all relevant procedures. We merely wanted to suggest some general comparisons. We have also included some methods for the analysis of live recaptures in Table 8.1. This subject is discussed further in the following section.

Modern methods have several important advantages over the older approaches:

(1) They are based on a general stochastic model structure for the sampling distribution of band recoveries;
(2) the assumptions of any specific model are made explicit and often appear quite general and realistic;
(3) chi-square tests (goodness of fit tests) allow the assumptions of any given model to be checked statistically;
(4) tests between models give more specific tests of the assumptions and, in conjunction with (3), allow for choice of a "best" fitting model;
(5) estimates of annual survival and recovery rates are possible in several models;
(6) because estimators of survival and recovery rates are based on the method of Maximum Likelihood (rather than some ad hoc procedure), these estimators are asymptotically "fully efficient." That is, given the assumptions of a particular model, there are no other consistent estimators that have a smaller variance:
(7) annual changes in band reporting rates do not bias the estimates of survival rates if recovery rates are allowed to be year-dependent;
(8) estimates of sampling variances and covariances can be computed as a measure of precision; and
(9) because estimates of sampling variances and covariances are available, statistical tests of hypotheses can easily be constructed (e.g., a test of the null hypotheses that average survival in two geographic areas is equal could be made using the following test statistic,

$$z = (\tilde{\bar{\bar{S}}}_1 - \tilde{\bar{\bar{S}}}_2) \Big/ \sqrt{\operatorname{var}(\tilde{\bar{\bar{S}}}_1) + \operatorname{var}(\tilde{\bar{\bar{S}}}_2) - 2\operatorname{cov}(\tilde{\bar{\bar{S}}}_1, \tilde{\bar{\bar{S}}}_2)}$$

which is, for sufficiently large banded samples, approximately a standard normal variate under the null hypothesis). For this example, the covariance term is zero because the two data sets are totally independent.

The reader interested in further comparisons is referred to Cormack (1968), Eberhardt (1972), and Seber (1972, 1973). We agree with Eberhardt that the practice of life table analysis should be discontinued. Much better methods now exist and should be used.

8.2 Relationship to Live Recapture Data

The origins of comprehensive analysis and inference procedures for band recovery data date back only to the early 1970's (Seber 1970 and Robson and Youngs 1971). In contrast, an analysis framework for live recapture ("capture-recapture") sampling began at a sophisticated level by the early 1950's (see Jolly 1963:113). The theory for band recovery data and live recapture data was developed somewhat independently.

By 1978, the inference theory underlying band recovery data was well advanced over the comparable theory for the analysis of live recapture data. While Jolly (1965) provided a very general basis for extended theory developments few specific models and procedures were developed until the mid-1970's. Since we wrote the first edition of this *Handbook*, a great deal of theory has been published on methods for the analysis of live recapture data.

Initially, it is important to clearly distinguish between the two fundamentally different classes of sampling methods for banding or tagging studies. In the first, and the subject of this *Handbook*, data are available only on a single, terminal recovery of marked animals reported dead. These data are treated as independent, mutually exclusive outcomes (a given marked animal could be recovered dead in year 1, 2, ..., ℓ or "never") and the simple multinomial distribution is appropriate (Chapter 1). In the second, data are potentially available from multiple "encounters" of each animal through repeated captures or sightings. An example would include bird #6112 originally banded in year 1 and recaptured in years 3, 5, 6, and 8. Viewed in this way, i.e., "multiple" recaptures of bird #6112, these data are not mutually exclusive and cannot be modeled as a simple multinomial. (As we will explain, this is not the best way to consider these data.)

This second class of sampling methods has acquired the generic title "Jolly-Seber model" after the independent papers published in *Biometrika* in 1965 by G.M. Jolly and G.A.F. Seber (although one could mention J.N. Darroch and P.H. Leslie, as well as previous papers of Jolly and Seber, in the early formulation of models for this class of sample data). The Jolly-Seber model was developed for a homogeneous population segment where survival (Φ_i), population size (P_i), births (B_i) and the sampling rate (p_i) were allowed to vary for each sampling period (i). Research since the papers by Jolly (1965) and Seber (1965) on these capture-recapture models has proceeded along two divergent lines. First, some investigators (e.g., Crosbie 1979 and Jolly 1979) have explored "reduced parameter" models; these have fewer parameters, but require additional assumptions. For example, all the Φ_i could be modeled (assumed) to be a single constant Φ. Second, investigators extended the Jolly-Seber model to more general cases such as those allowing age-specific parameters (e.g., Pollock 1981a and Stokes 1984). Other examples include Jolly-Seber analogues to Model 0 of Chapter 2 (Robson 1969, Pollock 1975).

Nearly all of these restrictions and extensions were foreseen by Jolly (1965) as his basic model was extremely general. Few of the special cases of Jolly's general model have simple closed-form estimators; the Jolly-Seber model being one such exception. Iterative methods for the solution of the *ML* equations were prohibitive in the mid-1960's. The availability of computers now allows the full range of Jolly-Seber models to be considered and implemented.

While the two different classes of sampling models can be distinguished, an advanced and more careful consideration reveals many close similarities. Again, Jolly (1965) understood this, primarily through what he termed "losses on capture." Jolly (1965:244) states "...the multiple-recapture [with the number captured at time i (s_i) equalling the number released at time i (n_i)] and the single-recapture [band recovery] situations may be considered as two special cases of the general situation wherein s_i and n_i are not necessarily equal." A further fundamental link between the two classes of models is seen after a careful consideration of the conditional nature of the appropriate likelihood function for the Jolly-Seber models. In this respect, Jolly (1965:241) states "...we make use of the information that an animal is captured, and that it is later recaptured, but not the fact that it is recaptured a *second* time." For this reason he believes the phrase "multiple recapture" was misleading.

Since the first edition many research papers have appeared that extend the general theory for the Jolly-Seber class of models; these include Brownie and Robson (1983), Brownie and Pollock (In press), Buckland (1980, 82) and Clobert et al. (In press). Pollock et al. (In press) provides goodness of fit tests for the Jolly-Seber models. Very important has been the age-specific extensions to the Jolly-Seber model (Pollock 1981a, and Stokes 1984). These methods are analogues for the models we develop in Chapter 3. Crosbie (1979), Crosbie and Manly (1985), Jolly (1979) and Jolly and Dickson (1980) develop resticted Jolly-Seber models that employ few parameters, but additional assumptions.

Cormack (1979), Nichols et al. (1981), Pollock (1981b), and Seber (1982) provide an overview of new methodology in the analysis of capture-recapture sampling. Robson (1983) provides an extended numerical example for the complex multihypergeometric models and associated notation. Arnason et al. (1981) and Kremers (1984, 85) investigate tag loss in the Jolly-Seber model, while Nichols et al. (1984) study the effect of permanent trap response on the estimators.

Computer software is critically important for the analysis of capture-recapture data in a Jolly-Seber framework. Arnason and Baniuk (1978, 80), Crosbie (1979), Jolly and Dickson (1980), Nichols and Pollock, Hines and Brownie (in prep.) provide very powerful algorithms for use in the analysis of data.

The above provides an overview of the current state of development of capture-recapture analysis. We alluded to the very close link to banding analysis. Indeed, the methods are not fundamentally different. Below we present a more technical discussion of the relationship of the two methods, and a unified concept of modeling capture methods for studying survival processes.

As Jolly (1965:241) and others have recognized, capture-recapture sampling represents a series of releases and subsequent recapture. An animal can be recaptured only once after each release, because given that it has been recaptured alive, it must be re-released before it can be recaptured again. Therefore, the proper conceptualization of the recaptures is that they are *conditional* on the releases. A given animal may be captured more than once, but each recapture corresponds to a different released cohort of animals.

Consider the above case of bird #6112, first released on occasion 1, and recaptured on occasions 3, 5, 6 and 8. It was also re-released on occasions 3, 5, 6 and 8 (at which time it was *known* to be alive). Thus releases and recaptures are paired as (1, 3), (3, 5), (5, 6) and (6, 8) with no recapture after the final release at time 8. Ignoring the intervening releases, one is lead to portray the data as:

	Recaptured history								
Release time	2	3	4	5	6	7	8	9	10
1	0	1	0	1	1	0	1	0	0

Table 8.1. *A brief summary of various methods for analysis of banding and tagging data and certain of their properties. (An "X" indicates the method has the given property).*

Method/Model	Model type				Properties of method										General		
	Stochastic model	Time-specific parameters	Age-specific parameters	Multiple cohorts	Unbiased estimators (or almost unbiased)	Asymptotically fully efficient	Closed form (explicit estimators)	Estimators asymptotically normal	Unbiased w.r.t. annual change in reporting rate	Estimators for variances and covariances	Statistical testing procedures	Free of data "truncation" problems	General computer programs available[1/]	Published simulation studies	Well developed theory	Assumptions realistic[2/]	Generally a "useful" method[3/]
This Handbook																	
Adult banding																	
M_0	X	X		X	X	X	X	X	X	X	X	X	X		X	X	X
M_1	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
M_2	X	X		X	X	X		X	X	X	X	X	X	X	X	X	X
M_3	X			X	X	X		X		X	X	X	X		X		
Young and adult bandings																	
H_{01}	X		X	X	X	X		X		X	X	X	X		X		
H_{02}	X	X	X	X	X	X		X	X	X	X	X	X		X	X	X
H_1	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
H_2	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
H_3[4/]	X	X	X	X	–	–	–	–	–	–	X	X	X		X	X	
Three age classes banded																	
H_4	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
H_5	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	
H_6	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
Biannual banding of adults																	
H_7	X	X		X	X	X	X	X	X	X	X	X	X		X	X	X
H_8	X	X		X	X	X	X	X	X	X	X	X	X		X	X	X
Standard Methods Previously Used in Banding Data Analysis[9/]																	
Composite dynamic life table (Hickey 1952)			X[5/]	X			X										
Composite time-specific life table (Hickey 1952)		X		X			X										
Dynamic life table (Hickey 1952)			X[5/]														
Time-specific life table (Hickey 1952)		X					X										
Haldane (1953, 1955) (for finite $\ell = k$)	X			X	X			X		X		X					
Relative recovery rate (Ricker 1958)	X	X	X[6/]	X			X	X		X[7/]		X				X	
Williams (1963) (see Geis & Taber 1963)		X		X			X										
Other Selected Methods																	
Chapman-Robson (1960)[11/]	X			X[8/]	X	X	X	X	–[10/]	X	X	X		X	X	X	X
Cormack (1964)	X	X		X		X	X	X	–[10/]	X		X			X	X	X
Fisher-Ford (1947)	X			X	X	X		X	–[10/]			X		X			
Fordham (1970)	X	X		X		X		X	–[10/]			X	X			X	
Heinche (1913)							X							X			
Jackson (1939)							X							X			
Johnson (1974)	X	X	X	X	X	X		X	X			X	X			X	X
Jolly (1965)	X	X		X		X	X	X	–[10/]	X		X	X	X	X	X	X
Manly-Parr (1968)	X	X	X	X		X	X	X	–[10/]	X		X		X	X	X	X
Roberts (1971)	X			X		X		X	–[10/]	X		X	X				
Seber (1971)	X		X	X			X	X		X	X	X			X		

(Footnotes continued next page)

[1/] General, usually published, computer codes in FORTRAN.

[2/] Relative to bird banding experiments, primarily.

Conceived of this way, statistical modeling of the data is *very* difficult. The proper way to tabulate the data is in the manner of banding data; for this one bird the release-recapture history is:

	First recapture after release time i								
Release time i	2	3	4	5	6	7	8	9	10
1	0	1							
3			0	1					
5					1				
6						0	1		
8	(not seen again)							0	0

Tabulating all the data like this, it is easy to build up the components of the model needed as a basis for estimation of survival and recapture parameters. This approach has been used by Brownie and Robson (1983) and Brownie and Pollock (1985). The basic array of data becomes:

		First recaptured at time j given release at time i				
i	Known releases at time i	j = 2	3	4	...	k+1
1	N_1	m_{12}	m_{13}	m_{14}	...	$m_{1,k+1}$
2	N_2		m_{23}	m_{24}	...	$m_{2,k+1}$
3	N_3			m_{34}	...	$m_{3,k+1}$
.	.				.	.
.	.				.	.
.	.				.	.
k	N_k					$m_{k,k+1}$

Making the identification $R_{ij} \equiv m_{i,j-1}$ these data are exactly like band recovery data. Only now any animal recaptured at time j (= 2,..., k) may be re-released into the cohort of N_j animals released at time j (some of which may have been previously unmarked). First recaptures from each cohort are (by assumption) distributed as multinomial random variables, and cohorts are (conditionally) independent. This conditioning on release is not just a mathematical device, rather it is mandatory to do so because one knows the animals are alive at that time.

←

[3] Our judgment based on a number of considerations.

[4] Individual parameters are not estimable under H_3.

[5] Age-specific if, and only if, additional, rather unlikely, assumptions are specified.

[6] Original development was age-independent: more recently used on data for both young and adults.

[7] Estimators of variance given if $k = \ell = 2$.

[8] Only if $\ell = \infty$.

[9] Many of these methods are described throughout the literature on banding data analysis (e.g., Geis and Taber 1963, Geis 1972a, 1972b).

[10] Not relevant in live recapture experiments.

[11] Also see Lack (1943).

Under this type of recapture and re-release sampling, the expectations of the m_{ij} are:

Known releases		E (m_{ij}); m_{ij} = numbers of animals first recaptured at time j given release at time $i<j$				
i	N_i	$j = 2$	3	4	…	$k+1$
1	N_1	$N_1\phi_1 p_2$	$N_1\phi_1 q_2\phi_2 p_3$	$N_1\phi_1 q_2\phi_2 q_3\phi_3 p_4$	…	$N_1\phi_1 q_2\phi_2 q_3\phi_3 q_4 \cdots \phi_k p_{k+1}$
2	N_2		$N_2\phi_2 p_3$	$N_2\phi_2 q_3\phi_3 p_4$	…	$N_2\phi_2 q_3\phi_3 q_4 \cdots \phi_k p_{k+1}$
3	N_3			$N_3\phi_3 p_4$	…	$N_3\phi_3 q_4 \cdots \phi_p p_{k+1}$
.					.	.
.					.	.
.					.	.
k	N_k					$N_k\phi_k p_{k+1}$

Here

p_i = the capture probability on the i th occasion,
$q_i = 1-p_i$ = probability of *not* being captured on occasion i (both are conditional on the animal being alive at that time),
ϕ_i = probability of surviving in the population at risk of capture from occasion i to occasion i+1 (= the product of the physical survival rate times the probability of *not* emigrating).

Making the formal re-parameterization of

$$f_i = \phi_i\, p_{i+1} \text{ and } S_i = \phi_i q_{i+1},$$

the $E(m_{ij})$ are exactly in the form of Model M_1 in banding data. In fact, one can use ESTIMATE to analyze such a summary of recapture data and the resultant $\hat{S}_i$, $\hat{f}_i$ are fully efficient MLE's. Then the Jolly-Seber estimates of ϕ_i, p_i are found as:

$$\hat{\phi}_i = \hat{f}_i + \hat{S}_i, \quad i = 1 \ldots k-1$$
$$\hat{p}_{i+1} = \hat{f}_i/(\hat{f}_i+\hat{S}_i), \; i = 1 \ldots k,$$

$\hat{f}_k = \widehat{\phi_k\, p_{k+1}}$, cannot be decomposed; ϕ_k and p_{k+1} are not separably estimable. For more discussion of this relationship see Brownie and Pollock (in press).

We do not recommend using ESTIMATE or BROWNIE to analyze recapture data (though it can be done without loss of efficiency). It would be a nuisance to transform the parameter estimates. However, more important is that there is additional information in the recapture data (because individual animals may be seen more than twice) that is useful for testing model goodness-of-fit. That information is not in the $\{m_{ij}\}$array. See Pollock et al. (in press) regarding such goodness-of-fit testing. A component of the overall goodness-to-fit test can be computed from the $\{m_{ij}\}$array; it is exactly the type of test discussed in chapters 2, 3, 4 and 5. In particular, the tests of chapter five can be used to test equality of parameters over sexes or areas for capture-recapture data sets.

The above shows that for the purposes of parameter estimation, capture-recapture and banding data are modeled in a common framework. In fact, the only difference is the parameterization (this fact was known years ago, see e.g., Seber 1973). The difference in the parameterization is because the process of sampling released animals is so different: live recaptures vs. reported bands from hunter-killed birds. The general (abstract) statistical model for such first "re-sampling" of known releases can be written as:

$$\frac{E(m_{ij})}{N_i} = \begin{cases} \alpha_i, & j = i+1, \\ \beta_i \cdots \beta_{j-2}\, \alpha_{j-1}, & j > i+1. \end{cases}$$

Here, α_j = the probability that the animal is removed from the cohort on occasion j+1, given that it was alive at occasion j. After such a removal, it is either "lost on capture," or re-released into cohort j+1. Conversely, β_j is the probability that the animal survives in (i.e. is not removed from) the cohort from occasion j to j+1, given that it is alive at occasion j.

When sampling is by hunters harvesting birds and then reporting bands, then $\alpha_i = K_i(1-c_i)\lambda_i \equiv f_i$ and $\beta_i \equiv S_i$ = actual probability of physical survival. Here, K_i = kill rate in year i, c_i = crippling loss, and λ_i is band reporting rate. Note that S_i is also the annual survival rate for birds *not* banded. When sampling is by capture, then $\alpha_i = \phi_i p_{i+1}$; that is the animal must survive from occasion i to occasion i+1 and then be captured. Now survival in the cohort equals $\phi_i q_{i+1}$; i.e., physical survival times the probability of *not* being captured. It is the product $\beta_i = \phi_i q_{i+1}$ that is intrinsically estimable. However, the actual survival rate is ϕ_i, hence the different form of Jolly-Seber estimates vs. banding estimates of survival.

Other sampling methods correspond to still different parameterizations. For example, bands may be reported from birds found dead (common for nongame birds). This situation has been considered by Seber (1970, 71), Burnham and Anderson (1979), North and Cormack (1981), and Anderson et al. (1985). Because birds must die before being found (and then bands reported), one models $\alpha_i = (1-\phi_i)\lambda$; however we still have $\beta_i = \phi_i$. One could also use kill-trapping to sample released animals. Then, again, $\alpha_i = \phi_i p_{i+1}$ and $\beta_i = \phi_i q_{i+1}$, exactly as with live recapture; the difference is that now released cohorts are always new animals. This illustrates the irrelevance, under the Jolly-Seber model, of whether or not the animals are released after recapture.

Given this "banding framework" of modeling recapture data, one can easily get MLE's of the ϕ_i and p_i. Then the Jolly-Seber estimate of population size (P_i) is just

$$\hat{P}_i = \frac{n_i}{\hat{p}_i}, \quad i = 2, \ldots, k \text{ (with } k+1 \text{ capture times)},$$

where n_i = total number of animals captured at time i ($n_i = u_i + m_i$, $m_i = m_{1i} + \ldots + m_{i-1,i}$ = all marked animals recaptured at time i). This matter has been considered in more detail by Kremers (1984).

Now that we realize the correct modeling approach to all capture-recapture data, we see that it could all be analyzed on the basis of one unified algorithm such as SURVIV (White 1983). The key is to:

(1) Use different parameterizations for different sampling processes; generalized models (e.g., Pollock 1981b, Stokes 1984) and restricted models (e.g., Jolly 1982) can all be easily implemented.
(2) Add population size estimates, as $\hat{P} = n/\hat{p}$ for the case of recapture data.
(3) Add the additional goodness-of-fit components in the case of multiple release data (see Pollock et al. in press).

There is yet another class of studies to which these models apply: release of reared animals (e.g., hatchery raised fish) as part of experiments to study survival processes. This brings into clear focus that fact that the unifying feature here is that there are known releases, in separate, independent cohorts. These animals are then subject to forces of mortality, i.e., they experience a survival process about which we wish to make inferences. The released cohorts are sampled by some process and the resultant counts of live or dead animals, cohort by cohort, provide the basis for inferences about the survival process. The process of sampling released cohorts can be live trapping, kill trapping, hunter reporting of bands, finding dead animals, or resighting of live animals without recapturing them (see e.g., Cormack 1964, Brownie and Robson 1983).

Bearing all of these possibilities in mind, a better term than capture-recapture would be (animal) release-resampling studies. Ambitious though it is to propose a new name, it is desirable that people fully realize the common nature of all such studies of survival processes based on released cohorts of known number of animals which are then (re)sampled.

8.3 On the Interpretation of the Estimated Survival Rates

As discussed in Section 2.9 all the survival rates and their estimators in Chapters 2 and 3 pertain to the 1-year time period between successive releases of newly banded birds. Also, survival rate S is not just the probability of surviving hunting death but is the total survival rate. A better way to express this is in terms of the mortality rate $M = 1 - S$. Then M is total mortality rate from all causes, not just hunting mortality rate. We have found these two facts are sometimes confusing and difficult to understand. The tendency is to think that since all the band recoveries often come from hunting deaths, the survival rate pertains to the time periods between hunting seasons and also is just the probability of surviving a hunting death. This is not true, as we will attempt to explain below.

Because band recoveries come from dead birds, the "natural" statistical models for analysis of recovery data do not involve the numbers of survivors of banded cohorts, at subsequent banding times. However, it is possible to derive the survival rate estimators by thinking in terms of surviving birds. This approach shows that analysis of recovery data is quite similar to analysis of live recapture data, as with Jolly's (1965) model (see also Cormack 1972, 1973 for derivation of population estimators from live recapture data). This approach will be illustrated for Model 1; the notation of Chapter 2 will be used.

Consider only the first two banded cohorts. The data are

Year banded	Number banded	Year of recovery		
		1	2	3, ... ,ℓ
1	N_1	R_{11}	R_{12}	$R_{13}, \ldots, R_{1\ell}$
2	N_2		R_{22}	$R_{23}, \ldots, R_{2\ell}$.

As in Chapter 2, $C_1 \equiv R_{11}$, and row totals are R_1 and R_2.

Let N_2^* be the number of birds banded in year 1 which survive to year 2 (the time point at which the new release of N_2 birds is made). We have

$$\mathbf{E}(N_2^*) = N_1 S_1$$

where S_1 is the survival rate from year 1 to year 2 of birds in the population represented by these bandings. An estimator of S_1 would be $\hat{S}_1 = N_2^* / N_1$ if N_2^* were known. Thus, to estimate S_1, all we need to do is find an estimator of N_2^*.

Assuming, as Model 1 does, that all birds alive at year 2 have the same probabilities of subsequent survival and harvest, then on the average the same proportion of the bands of the N_2^* birds previously banded and of the N_2 birds newly banded will be subsequently recovered. This is a key point; mathematically it is

$$\mathbf{E}\left(\frac{R_2}{N_2}\right) = \mathbf{E}\left(\frac{R_{12} + R_{13} + \cdots + R_{1\ell}}{N_2^*}\right) = \mathbf{E}\left(\frac{R_1 - C_1}{N_2^*}\right).$$

From the above equation an estimator of N_2^* is derived:

$$\hat{N}_2^* = \frac{(R_1 - C_1)\, N_2}{R_2}.$$

This estimator of N_2^* is analogous to a Petersen estimator of population size.

Now an estimator of S_1 is given by

$$\hat{S}_1 = \frac{\hat{N}_2^*}{N_1} = \left(\frac{R_1 - C_1}{N_1}\right)\left(\frac{N_2}{R_2}\right).$$

This is exactly the ML estimator of S_1 under Model 1 of Chapter 2.

Several points are worth making:

(1) The estimate N_2^* relates to the time when the known number N_2 of newly banded birds are released. It is the subsequent proportions of the known N_2 bands recovered that allows estimation of N_2^*.

(2) The survival rate estimate is for the time period between years 1 and 2, because N_1 is the number of birds alive at the *start* of year 1 and N_2^* is the number of survivors of these N_1 birds at the *end* of year 1.

(3) Letting mortality $M = 1 - S$, we see that

$$\hat{M}_1 = 1 - \hat{S}_1 = \frac{N_1 - \hat{N}_2^*}{N_1} = 1 - \frac{\hat{N}_2^*}{N_1}$$

estimates total mortality, because $N_1 - \hat{N}_2^*$ estimates all deaths in year 1. In fact, we note no recoveries from year 1 are involved here, so sources of death during year 1 are irrelevant.

(4) Similarly, that the data (recoveries) came from hunting is irrelevant to the interpretation of S_1. All we require is some device for sampling the population after the N_2 new birds are released, to find the relative proportion R_2/N_2 and hence obtain the estimator of N_2^*. Hunting supplies this sampling device, but it would work just as well if birds were caught (retrapped) and live recaptures reported. In practice, hunting provides broad-coverage sampling of banded populations at a higher intensity than would otherwise be possible.

All the survival estimators of this handbook could be formulated as ratios of population size estimators between banding time periods. This would show more clearly their proper interpretation. It would, however, be very clumsy. Model 1 was used here for illustration because it is easier to deal with. For the more mathematical reader we will show how the general estimator $\hat{S}_i$ in Model 1 is derived as a ratio of population estimators.

Let banded samples of birds be released back into the population at times t_i, $i=1,\dots,k$, and let N_i^* $i=2,\dots,k$ be the total numbers of banded birds still alive just before the i^{th} release of banded birds. For example, N_2^* is the number of survivors at time t_2 of the original N_1 bandings. In general, N_i^* are survivors at time t_i of all previously banded birds (from banding times $1,\dots,i-1$).

Of the $N_i^*+N_i$ banded birds at time t_i, N_{i+1}^* survive to time t_{i+1}. Overall survival rate in this time period is S_i, hence,

$$\mathbf{E}(N_{i+1}^* \mid N_i^*, N_i) = (N_i^* + N_i)\, S_i\,.$$

Now we need an estimator of both N_i^* and N_{i+1}^*. The expected proportion of recoveries from the N_i new bandings is

$$\mathbf{E}\left(\frac{R_i}{N_i}\right).$$

Of the N_i^* survivors of previously banded birds, there are T_i-R_i total recoveries. Thus, by the assumptions of Model 1

$$\mathbf{E}\left(\frac{R_i}{N_i}\right) = \mathbf{E}\left(\frac{T_i-R_i}{N_i^*}\,\middle|\,N_i^*\right)$$

giving

$$\hat{N}_i^* = \left(\frac{T_i-R_i}{R_i}\right) N_i\,.$$

This equation holds equally well with i replaced by $i+1$ to give $\hat{N}_{i+1}^*$.

An estimator of S_i is

$$\hat{S}_i = \frac{\hat{N}_{i+1}^*}{\hat{N}_i^* + N_i},$$

Appropriate substitution gives, after simplification

$$\hat{S}_i = \frac{R_i}{N_i}\left(\frac{T_{i+1}-R_{i+1}}{T_i}\right)\frac{N_{i+1}}{R_{i+1}}$$

(see Seber 1970:319).

From Chapter 2 we have $T_i - C_i \equiv T_{i+1} - R_{i+1}$, hence

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

which is exactly the ML estimator of S_i under Model 1.

This shows $\hat{S}_i$ is the ratio of two population size estimators: $\hat{N}_{i+1}^*$ and $\hat{N}_i^*+N_i$, where N_{i+1}^* is the number of survivors at time t_{i+1} of the $N_i^*+N_i$ birds alive at t_i. This representation of $\hat{S}_i$ shows clearly that it estimates total survival between times of successive releases.

8.4 Proper Interpretation of Sampling Correlations

This handbook concentrates on the analysis of banding data with emphasis on parameter estimation, especially annual survival and recovery rates S_i and f_i, respectively. Models are presented allowing one to examine questions about whether S_i and f_i are time- and age-specific. However, *no functional relationship between true annual survival and recovery rates is assumed in any of these models.*

Biologists will usually be interested in a variety of questions concerning population dynamics of studied species. In exploited species like migratory game birds, the effect of hunting kill rate on survival rate is a fundamental question of considerable importance. Because band recovery rate is a direct index to kill rate, there will be a relationship between true annual recovery rates and true annual survival rates (approximately linear as $S_i = a + bf_i$). If the true annual parameters were known, this relationship would be easy to explore. But all we have available to us from the data analysis are estimates of these parameters. Moreover, these estimators ($\tilde{S}_i$, $\hat{f}_i$ or $\bar{\tilde{S}}$, $\hat{\bar{f}}$), whatever the model, are subject to large sampling variance and covariance. As a consequence, the relationships between parameter estimates are potentially very different from the relationship of the true parameters.

A common mistake is to regress estimates of average annual survival rate $\hat{\bar{S}}$ on estimates of average annual recovery rate $\hat{\bar{f}}$ from several different data sets, as a means of examining for the true relationship $S_i = a + bf_i$. This is not valid, primarily because there will almost always be a strong negative sampling correlation between $\hat{\bar{S}}$ and $\hat{\bar{f}}$. In Example 2.1, the estimated sampling correlation of $\hat{\bar{S}}$ and $\hat{\bar{f}}$ is -0.5710. The proper interpretation of this correlation is as follows.

If repeated samples were taken from this population of male wood ducks, all having the sample sizes and covering the same years as Example 2.1, then we could compute the usual correlation coefficient from the pairs of estimates $\hat{\bar{S}}, \hat{\bar{f}}$ and its value would be around -0.57. Because such samples would all come from the same population, the underlying true average rates $\bar{S}$ and $\bar{f}$ would be the same for every sample. Thus the sampling correlation of $\hat{\bar{S}}$ and $\hat{\bar{f}}$ tells us nothing about how true survival and recovery rates are related; it only reflects the extent to which the estimators $\hat{\bar{S}}$ and $\hat{\bar{f}}$ are correlated because they are both computed from the same data set.

As a concrete example of sampling correlation, we show the results of a simulation of band recovery data for $k = \ell = 12$ years of "banding" and "recovery" (see Anderson and Burnham 1976 for details). There were 20 independent repetitions of this "banding study." Each simulation has sample sizes (N_i) equal to 400 for all 12 years, and constant annual survival and recovery rates of 0.5 and 0.05, respectively. Thus the true model in these simulations was Model 3 wherein $S_i \equiv S$ and $f_i \equiv f$. The pairs of estimators $\hat{\bar{S}}$ and $\hat{\bar{f}}$ and all 20 independent repetitions are shown below:

$\hat{\bar{f}}$	$\hat{\bar{S}}$
0.0509	0.4561
0.0581	0.4445
0.0491	0.4851
0.0503	0.4996
0.0523	0.4858
0.0501	0.5072
0.0489	0.4792
0.0518	0.5091
0.0515	0.4702
0.0447	0.5256
0.0507	0.5177
0.0524	0.4826
0.0509	0.4929
0.0476	0.4927
0.0576	0.4775
0.0485	0.4927
0.0486	0.5105
0.0452	0.5221
0.0446	0.5236
0.0497	0.4829

Because the exact same "population" was sampled in all 20 cases, the variation among these estimates is entirely sampling variation. The actual sampling correlation of $\hat{\bar{S}}$ and $\hat{\bar{f}}$ computed from these 20 pairs is -0.712. A plot of these points reveals a very evident linear relationship of $\hat{\bar{S}}$ on $\hat{\bar{f}}$; but it reflects only the sampling variation and covariation of these estimators, and it is not valid to use such regression procedures to infer the true relationship of survival and kill (recovery) rates.

The purpose of this section is to inform the reader that this handbook does not address the question of the relationship of survival and kill rates. None of the methods presented here relate to that question. We warn the reader that the sampling correlations are to be interpreted as reflecting the extent to which the estimators are interrelated by virtue of being computed from the same sample. As such these sampling correlations themselves are a warning that it is invalid to regress, for example, $\hat{\bar{S}}$ on $\hat{\bar{f}}$ to explore the possible relationship of the true, unknown survival and recovery rates.

The fundamental problem is that of estimating the parameters a and b in the functional relationship $S_i = a + bf_i$ when the true values of S_i and f_i are not known. Instead, only estimates S_i and f_i are available. This inference problem is related to the subject of variance components and to the so-called errors-in-variables regression model (cf. Kempthorne and Folks 1971). Both these latter statistical problems are very complex.

This question of using the band recovery data to determine (test for) the effect of exploitation rate on survival rate is explored in detail in Anderson and Burnham (1976). The theoretical (data analysis) problems of validly estimating a and b are complex, but the methods do build on the models and estimation procedures presented in this handbook. A correct theory is developed and applied to the continental mallard banding data accumulated to 1971. The reader who wishes to go beyond estimation and model selection and examine the effect of exploitation on survival is referred to Nichols et al. (1984).

8.5 Some Implications of Assuming Equal Time Intervals Between Bandings

It was pointed out in Chapter 1 that in this handbook we have assumed an equal time interval between banding of each cohort. For the methods of Chapters 2, 3, and 4, this time period is 1 calendar year for bird banding studies. The bulk of banding and tagging studies do (and should) have an equal time period between bandings. However, this assumption affects the generality of some models and analysis methods we have presented. We illustrate the situation with the adult models of Chapter 2.

Models 1 and 0 are valid whether or not time periods between bandings are equal, because the survival and recovery rate parameters are allowed to be arbitrary. Consider arbitrary times of banding, say $t_1, t_2, \ldots, t_k$, and let $d_i = t_{i+1} - t_i$, $i = 1, \ldots, k-1$ be the time intervals between bandings. In Model 1 the survival rate S_i applies during time period d_i. The model structure is valid for arbitrary d_i. All the formulae for estimators of S_i and f_i are valid as are their standard error formulae. Also the goodness of fit test is valid. In fact the only aspect of the analysis based on Model 1 that would have to be modified is the computation of the mean survival rate.

Given such arbitrary time intervals between bandings, the hypothesis test of Model 1 vs. Model 0 would still make sense. The hypothesis test of Model 1 vs. Models 2 or 3 would be nonsense, however, because the hypothesis of equal survival rates per time period is not meaningful (interpretable) if time periods are themselves unequal. Models 2 and 3 would have to be modified as follows:

Let S be the survival rate for one time unit, in whatever units of time the d_i are in. Then the null hypothesis of constant survival rate becomes

$$H_0\colon S_i = S^{d_i} \quad , i = 1, \ldots, k-1 \, .$$

The appropriate structure of Model 2 to represent a time-constant survival rate is now

$$\mathbf{E}(R_{ij}) = \begin{cases} N_i f_i & , i = j \\ N_i S^{(d_i)} S^{(d_{i+1})} \cdots S^{(d_{j-1})} f_j & , i < j \, . \end{cases}$$

Appropriate estimators of S and f_i could be computed, as could the appropriate test of this model vs. Model 1 in the case of general intervals d_i to correctly test for *constant* survival rate. But these analyses would have been far more complex to devise, and for the user of the computer programs such extended methods would require inputting the time intervals $d_1, \ldots, d_{k-1}$. There is simply no great need for these more general models and methods allowing unequal time intervals.

To use the models and analysis methods presented in this handbook, or programs ESTIMATE or BROWNIE to analyze data with unequal intervals of time between banding, the following is required. First ignore all analyses using Models 2, 3, H_{01}, and H_{02}. This includes tests of these models versus any other models. Second, estimates of average survival rates cannot be computed as arithmetic averages, but must be based on the product of individual estimates of survival rates as

$$\bar{\tilde{S}}_G = \left(\prod_{i=1}^{k-1} \tilde{S}_i \right)^{1/d\cdot} , \tag{8.5.1}$$

where

$$d\cdot = d_1 + \cdots + d_{k-1} \, .$$

This estimator $\bar{\tilde{S}}_G$ then estimates average survival per time unit during the total time period t_1 to t_k (which is of length $d\cdot$).

On Averaging Survival Rate Estimators

Because most authors have considered the time intervals between banding to be arbitrary, they have recommended estimating an average annual survival rate using formula (8.5.1.). When the time periods are equal and the unit of time is taken to be $d_i \equiv d \equiv 1$, then (8.5.1) gives $\bar{\tilde{S}}_G$ as the geometric mean of the individual survival rate estimates. Yet in this equal interval case it is just as reasonable to use the arithmetic mean

$$\tilde{\tilde{S}}_A = \frac{\sum\limits_{i=1}^{k-1} \tilde{S}_i}{k-1}, \tag{8.5.2}$$

and we have done so in the computer programs.

This question of how to average the individual survival rate estimators arises only when the time intervals d_i are all equal and the S_i are allowed to vary (e.g., Models 1 or 0). Because the question is one of parameter estimation, we must first decide what is to be estimated. It is our opinion this parameter should be

$$\overline{S} = \frac{\sum\limits_{i=1}^{k-1} S_i}{k-1} .$$

Another way of approaching this question is to ask if we know the true annual survival rates $S_1,\ldots,S_{k-1}$, and if we had to compute an "average," what sort of average would we compute? If the answer is the arithmetic mean, then formula (8.5.2) should be used as the corresponding estimator.

Alternatively, conceiving of the S_i as random variables, the logical parameter of interest would be $\mathbf{E}(S_i)$ which is estimated by $\overline{S}$. Given this is the "average" we want to estimate, then the best estimator is $\tilde{\tilde{S}}_A$ as given in formula (8.5.2.) Moreover, for Model 1 (and others given here) we know

$$\mathbf{E}(\tilde{S}_i) = S_i(1-(1-\rho_{i+1})^{N_{i+1}+1}) ,$$

which means for practical purposes $\tilde{S}_i$ is unbiased. It follows $\tilde{\tilde{S}}_A$ is an unbiased estimator of $\overline{S}$. The same is not true of the geometric mean in this case. In fact it is well known that the geometric mean of a set of positive numbers is less than the arithmetic mean. It follows that $\mathbf{E}(\tilde{\tilde{S}}_G) < \overline{S}$ will be true.

Also, because $\tilde{\tilde{S}}_A$ is a linear function of the $\tilde{S}_i$, the formula for the sampling variance of $\tilde{\tilde{S}}_A$ is known exactly in terms of variances and covariances of the $\tilde{S}_i$. Because $\tilde{\tilde{S}}_G$ is a highly nonlinear function of the $\tilde{S}_i$, the formula for VAR$(\tilde{\tilde{S}}_G)$ is only an approximation. For these reasons we recommend using the simple arithmetic mean of the $\tilde{S}_i$ *when time intervals between bandings are all equal.*

8.6 Comparison of Survival Rates Over Years or Areas

Often the biologist will not be satisfied with merely estimating annual survival rates but will want to test hypotheses about variation in survival rates over time or areas. As mentioned in Section 8.4 the effect of hunting kill rate on survival rate is a fundamental question. Although proper approaches to testing this general question can be very complex, there is an easily constructed test to compare specified annual survival rates. For example, if it is known that for a given banded population hunting regulations were very restrictive in certain years and liberal in other years, then it is natural to compare average survival rates in these different years.

Let $\tilde{S}_1,\ldots,\tilde{S}_{k-1}$ be estimated annual survival rates from a given banding study. These estimates could be for either adults (as from Model 1) or for young and adults (as from H_1). In general let var$(\tilde{S}_i)$ and cov$(\tilde{S}_i,\tilde{S}_j)$, $i \neq j$, be the estimated sampling variances and covariances of these estimates. For any set of constants $c_1,\ldots,c_{k-1}$, the null hypothesis

$$H_0\colon\ c_1S_1 + c_2S_2 + \cdots + c_{k-1}S_{k-1} = 0$$

can be tested with the test statistic

$$z = \frac{c_1\tilde{S}_1 + c_2\tilde{S}_2 + \cdots + c_{k-1}\tilde{S}_{k-1}}{\sqrt{\sum\limits_{i=1}^{k-1} (c_i)^2 \mathrm{var}(\tilde{S}_i) + 2\sum\limits_{i=1}^{k-2}\ \sum\limits_{j=i+1}^{k-1} c_ic_j \,\mathrm{cov}(\tilde{S}_i,\tilde{S}_j)}}$$

The alternative hypothesis may be two-sided

$$H_a\colon\ \sum_{i=1}^{k-1} c_iS_i \neq 0 ,$$

or it may be one-sided, for example

$$H_a\colon\ \sum_{i=1}^{k-1} c_iS_i > 0 .$$

Under this null hypothesis z is asymptotically normal with mean 0 and variance 1. For such a test to be meaningful one usually requires the constants c_i to sum to zero and such a test is exactly analogous to the single degree of freedom contrast in analysis of variance (cf. Ostle 1963).

For example, if we had a data set from adults with $k=5$, analyzed under Model 1 (i.e., Model 1 was judged to be the appropriate model), then four annual survival rates are estimable. To compare the average survival rate in the first 2 years with the average survival rate in the second 2 years, one could take $c_1=c_2=1/2$ and $c_3=c_4=-1/2$. The null hypothesis is then

$$H_0\colon \frac{S_1+S_2}{2}=\frac{S_3+S_4}{2}.$$

Because this is equivalent to H_0: $S_1+S_2=S_3+S_4$, it is just as satisfactory, and mathematically simpler, to choose $c_1=c_2=1$, and $c_2=c_3=-1$. Given this choice of constants the test statistic value is

$$z=\tilde{\delta}/\sqrt{\operatorname{var}(\tilde{\delta})},$$

where $\tilde{\delta}=(\tilde{S}_1+\tilde{S}_2)-(\tilde{S}_3+\tilde{S}_4)$, and $\operatorname{var}(\tilde{\delta})=\operatorname{var}(\tilde{S}_1)+\operatorname{var}(\tilde{S}_2)+\operatorname{var}(\tilde{S}_3)+\operatorname{var}(\tilde{S}_4)+2\operatorname{cov}(\tilde{S}_1,\tilde{S}_2)-2\operatorname{cov}(\tilde{S}_2,\tilde{S}_3)+2\operatorname{cov}(\tilde{S}_3,\tilde{S}_4)$.

The appropriate alternative hypothesis in this case could be two-sided, in which case one would reject the null hypothesis if the absolute value of z were "too large." For a 5% significance level test one would reject if $|z|>1.96$. But, suppose hunting regulations were restrictive in the first 2 years of banding and liberal in the third and fourth years. Then the alternative of interest is really

$$H_a\colon \frac{S_1+S_2}{2}>\frac{S_3+S_4}{2},$$

or average survival is greater in years of restrictive regulations than in years of liberal hunting regulations. The appropriate test is then to reject the null hypothesis only if z (not $|z|$) exceeds the appropriate critical value. For a one-sided 5% significance level test we would reject if z exceeded 1.645.

For a specific example of such a test we will use the results of Example 2.6 under Model 1. This is a real data set from male mallards banded during the winter in Illinois from 1963 to 1970, inclusive. From Martin and Carney (1977) it is known that waterfowl hunting regulations were very restrictive in 1965 and 1968, and conversely they were very liberal in 1964 (cf. Anderson and Burnham 1976).

From Example 2.6f we obtain the following values

i = year	$\tilde{S}_i$	se$(\tilde{S}_i)$
1964	68.17	6.95
1965	64.65	6.26
1968	74.25	6.13

From Example 2.6 the sampling covariance of $\tilde{S}_{64}$ and $\tilde{S}_{65}$ is estimated as -29.48. (We caution the reader that the covariance of estimators are not printed on a percent basis by program ESTIMATE; hence, for comparability to results on $\tilde{S}_i$ and se$(\tilde{S}_i)$ they must be multiplied by 10,000.) The sampling covariances of $\tilde{S}_{64}$ and $\tilde{S}_{68}$, and $\tilde{S}_{65}$ and $\tilde{S}_{68}$ are both zero.

The test statistic for testing H_0: $S_{68}=S_{64}$ (restrictive vs. liberal year) vs. H_a: $S_{68}>S_{64}$ is

$$z=\frac{\tilde{S}_{68}-\tilde{S}_{64}}{\sqrt{\operatorname{var}(\tilde{S}_{68})+\operatorname{var}(\tilde{S}_{64})}}$$

For a 5% significance level we reject the null hypothesis of no difference only if z is greater than 1.645. In this example we obtain

$$z=\frac{74.25-68.17}{\sqrt{(6.13)^2+(6.95)^2}}=\frac{6.08}{9.26}=0.66\,.$$

The data do not allow rejection of the null hypothesis.

To compare average survival in the two restrictive years to the survival in the liberal year (1964) the test statistic is

$$z=\frac{\dfrac{\tilde{S}_{65}+\tilde{S}_{68}}{2}-\tilde{S}_{64}}{\sqrt{(1/2)^2\operatorname{var}(\tilde{S}_{65})+(1/2)^2\operatorname{var}(\tilde{S}_{68})+\operatorname{var}(\tilde{S}_{64})-2(1/2)\operatorname{cov}(\tilde{S}_{64},\tilde{S}_{65})}}$$

$$= \frac{\dfrac{64.65+74.25}{2}-68.17}{\sqrt{(.25)\,(6.26)^2+(.25)\,(6.13)^2+(6.95)^2+(29.48)}} = \frac{1.28}{9.85} = 0.13.$$

Again these data do not allow rejection of the null hypothesis.

Tests like these, contrasting selected survival rates, are very flexible: they can be used to compare annual survival rate estimates for different years from adult or young bandings; they can be used to compare annual (or average) survival rates of males versus females (if the tests of Chapter 5 indicate a difference and it is desired to find out where the differences are); and they can be used to compare average (or annual) survival rates over different populations (i.e., compare estimates from different data sets representing different areas). Tests like these are also suitable for comparing differences in annual, or average, recovery rates. Moreover, such tests may be made for any of the models. Hence it is impossible for us to give an exhaustive treatment of the possible uses of such contrast type tests to compare survival or recovery rates. The purpose of this brief discussion is to alert the reader to their existence and appropriateness for testing certain types of questions. We hope the reader well-trained in statistics will perceive the possible uses and correct application of this type of test. The biologist not well-trained in statistics is urged to consult a statistician when contemplating comparing survival or recovery rate estimators over years, sexes, or areas using such a z-test.

Chapter 9. Planning a Banding Study

9.1 General Remarks

In this chapter we will suggest a number of considerations in planning a banding study. The remarks made here apply to other animal marking experiments although we make specific reference to bird banding. The following material represents only guidelines because value judgements are required in planning a banding study. We can rarely give specific and rigid details because so many factors are involved.

Band recovery data represent a sample from a banded population. Banding is done so that we have marked members of the population to sample at later dates. Banding studies are merely sampling experiments. In the case of exploited species of migratory birds, we have a very large number of "samplers" over a large geographic area. Consideration must be given to banding only birds that are representative of the total population of interest. If we wish to make inferences about the adult male redhead *(Aythya americana)* population in northern Utah from sampling and analysis of banded members, we must be sure the banded sample is representative of the adult male redheads in northern Utah. This rules out, for example, the use of sick, injured, transported, or color-marked birds. We further discuss representativeness in the following section.

Although bird banding is an expensive and time-consuming effort, insufficient attention has often been given to planning the study, assessing the assumptions, and thoroughly analyzing the data. The literature emphasizes the need for detailed planning of banding operations before the field work is initiated (DeLury 1947, 1951; Orians 1958). Data gathering and data analysis should be coordinated through proper planning and a realistic evaluation of the necessary assumptions. Planning is crucial if the study is to produce accurate and precise estimates of population parameters. As DeLury (1954: 293) pointed out, "...it is an expensive impropriety to maintain that an untrustworthy estimate is better than none." Review of published material by Cormack (1979) and Seber (1972, 1982) as well as this handbook is recommended.

For the bird bander or field biologist the primary drawback in the use of the newer techniques is likely to be their complexity. The theoretical basis of most of the better methods lies deep in the field of mathematical statistics. In most situations the use of a particular method is fairly straightforward, but the mathematical theory and derivation of the various estimators may seem difficult. We have included examples illustrating the use of the methods we describe. However, it is important for the bird bander or field biologist to fully understand the assumptions and limitations of these methods as well as their strengths. This suggests that a joint effort between biologists, bird banders, and ornithologists on the one hand, and statisticians on the other, would be advantageous. This combination of talent and expertise would be particularly advantageous for comprehensive analyses of regional or flyway banding programs.

The advantages of large-scale cooperative banding programs are obvious. Entire populations can be sampled including population segments in various geographic areas. Indeed, continental, national, and regional banding programs are underway at this time for several migratory species (e.g., mallard, Canada geese, mourning dove *(Zenaidura macroura),* wood duck, and band-tailed pigeon *(Columba fasciata)*). The results of such cooperative banding programs can address major management and research needs. More comprehensive objectives and difficult questions and objectives can be approached when parameter estimates and test statistics are available from numerous banded segments of widespread populations (e.g., Anderson 1975 and Anderson and Burnham 1976). Of course, much can be learned from single banding programs aimed at one segment of the population (e.g., Canada geese wintering in southern Illinois). In any event, the computer programs described by Cowardin (1977) and Davenport (1977) may be useful.

In summary, several points could be emphasized. Bird banding studies must be developed on a scientific basis if accurate and precise results are to be expected. Planning the project and reviewing the literature should certainly precede the field work. Estimates of various population parameters are not made easily and should not be made casually. Assumptions underlying the method of analysis to be used should be fully recognized. The newer stochastic models have a number of important advantages and therefore should replace the life table approaches. Estimates of the sampling variances of parameter estimates are of great importance in bird banding studies. Estimates of population parameters, made after years of expensive field work, that have extremely wide confidence intervals represent wasted time and money.

9.2 Suggestions – Guidelines

Banding of birds may appropriately be done at various times of the year depending on objectives and availability of birds. In general, we recommend that game birds be banded just before or just after the hunting season.

Populations are usually sedentary during these periods and often large samples can be banded. Banding during migration periods is generally to be discouraged because it is nearly impossible to determine which population the estimates relate to. Banding during the breeding season is often quite disruptive to the population; however, there are many species where only breeding-season banding is feasible. Banding during the hunting season is also not recommended. However, there are instances where large samples of birds (e.g., Canada geese) can be banded only during the latter part of the hunting season. Hunting often keeps birds concentrated on refuge areas where they can be banded. The birds disperse rapidly as soon as the hunting season terminates and become difficult to band in large numbers. In some instances, Models 0 and H_2 can be used as an approximation for in-season banding data (see Sections 2.5 and 3.5). Banding programs for nongame birds should probably follow the same general periods: late summer-early fall, winter, and (for some species) during the breeding seasons.

The actual banding period should be short in relation to the period between bandings (survival period). Mortality during the banding period is undesirable. Annual banding during a 3- to 5-week period is probably satisfactory if the survival period is 12 months. In some cases mourning doves are sometimes banded over a 2- to 4-month period which is likely to be unsatisfactory if the objective relates to the estimation of population parameters. Lengthy banding periods are especially undesirable for species with a high annual mortality rate.

A very important remark concerns studies involving large banded samples of young birds without an accompanying sample of adult birds. Unless some very restrictive and unrealistic assumptions are made no satisfactory estimates of survival rates are possible (Section 3.9). In general, the use of data from banded young birds will produce seriously biased parameter estimates (similar to the large bias associated with life table methods). Unfortunately, many of the studies of gulls, terns, and other colonial nesting birds fall into this category. It is a simple matter to band young birds in the colony while still in the nest, but obtaining a sample of adults has proven very difficult and quite disruptive to the colony. Seber (1972) and Cormack (in an appendix to Fordham 1970) have developed ML methods for age-specific survival rate if it can be assumed that reporting/recovery rates are constant and independent of age and year and that survival is time-independent. These assumptions are unlikely to be supported by biological data. In addition, both authors acknowledge problems with the performance of their method. *We expect few if any new developments in methodology that will lessen the need to band a matching sample of adults.*

Differences in survival and recovery rates by sex seem to be very common. We provide a test to examine this question (Section 5.1). Generally, separate analyses should be carried out for males and females, although we have seen several sets of Canada goose data where male and female survival and recovery rates are essentially identical (this might be expected for a species that is not obviously sexually dimorphic and migrates in loose family groups). In these situations, the data for males and females (numbers banded and the recovery arrays) can be pooled before the analysis and parameter estimation takes place. This appears to be the exception rather than the rule, however, suggesting that sample sizes required for a given level of precision (Section 9.3) must be estimated separately for males and females.

Except for Section 8.2, this handbook focuses on methods for analyzing recovery data. Several very good ML methods for the analysis of live recaptures are now in the literature (Jolly 1965, Pollock 1979, Jolly 1979, Pollock 1981b, Buckland 1980, Brownie and Robson 1983, Stokes 1984, Crosbie and Manly, 1985). These methods deal **with generalized Jolly-Seber models (e.g., age-specific parameters) as well as restricted models wherein some parameters are the same over time. Also, there are some good computer programs specifically for analysis of live re-**capture data (Aranson and Baniuk 1980, Jolly and Dickson 1980, Nichols et al. in press, Crosbie and Manly 1985). It is now possible to do a fully efficient analysis of live recaptures and dead recoveries when both appear in **the same data set. Buckland (1980) discusses one way to do this, however we feel that often the parameterization for the two types of data, coded into separate [m_{ij}] arrays (See Section 8.2), must be different. While no specific software for such an analysis exist, program SURVIV (White 1983) can be used to estimate survival rates, and test if these rates differ between the data sets from live vs. dead samples. This has many important and interesting possibilities for studies such as those reported by Anderson and Sterling (1974). For these reasons, we recommend that recaptures should be recorded as part of bird banding programs in the field. That is, birds captured alive in year i from banding at the same site in previous years should be carefully recorded. Of course, with nongame birds, recapture information is often the only type of data available, and the number of recaptures are often small. Analysis of such sparse recapture data has been considered by Kreger (1973) and Mardekian and McDonald (1981).**

We have explained that the theory underlying many modern methods allows the periods between banding to be unequal, but most analyses and tests carried out by the computer programs are valid only if these periods are equal. In general, we recommend banding studies to be based on annual periods. We also question studies where birds are banded every 2 or 3 years. Although there is probably nothing fundamentally wrong with such studies, we question what objective is being satisfied by such programs.

Marking of birds (e.g., dyeing of feathers or use of patagial tags and ribbons, colored plastic leg bands, neck collars, nasal saddles, etc.) is generally done toward objectives other than estimating vital parameters. Hunters and others are much more likely to shoot and/or report a vivid pink snow goose *(Chen hyperborea)* with a colored neck collar than a normal snow goose. Therefore, if estimates of survival rates are of interest, these practices must be avoided. However, if the mark disappears within a year (e.g., molt of the dyed feathers, etc.) Models 0 or H_2 may be appropriate and thus allow the banding and recovery data to be used effectively toward two divergent objectives. Most common objectives would be best served by banding and analyzing only normal wild birds caught, banded, and released immediately, uninjured. Studies of game birds often use only recoveries reported as shot during the hunting season. This practice allows $\hat{f}_i$ to be used as an index to harvest rate. If the sole objective is to estimate survival rates, then nearly all recoveries can be used, thereby increasing sample size and precision.

The use of Reward bands (cf. Henny and Burnham 1976) has several substantial advantages in a few studies. This technique must be used with care. The advantages are primarily that an estimate of harvest rate is obtained, rather than merely an index to harvest rate. If samples of Reward and regular bands are used, an estimate of the band reporting rate can be made and thereby allow past data to be interpreted more fully. The use of Reward bands increases the number of recoveries, often by a factor of 2 or 3, given a fixed number of birds banded.

We believe banding programs should be conducted for a minimum of 5 years ($k \geq 5$). This will allow at most four estimates of survival. It takes several years to get a good-sized population of banded birds in the total population. Programs in operation for less than 5 years are not likely to produce useful estimates. Because precision is gained when $\ell > k$, the investigator may often wish to reanalyze the accumulated recovery data 2-4 years after banding is terminated to make a "final" analysis of the data. This allows the use of the additional recoveries from later years to be incorporated in the analysis.

Species having low band recovery rates are particularly troublesome. For example, adult blue-winged teal and American woodcock typically have recovery rates of about 1%. If 1,000 were banded in August of year i we would only expect to get 10 reported the first year after banding, and perhaps only 20 over the total life-span of the banded cohort. These figures are totally insufficient so we might have to envision banding 3 or 4 thousand each year to obtain estimates of annual survival that are even reasonably precise. Usually, samples of this size are not obtainable, particularly for each age and sex. One may ask if the 1,000 should be banded anyway and hope to obtain useful parameter estimates and test statistics concerning the population. Our research suggests the answer to this question is no. Johnson (1974) discusses a hypothetical example where small samples of birds were banded over a 3-year period with resulting parameter estimates having very wide confidence intervals at best. He questions the value of such a practice. The important question of how many birds should be banded annually to obtain a specified level of precision is discussed in the next section.

Needless to say, incorrect aging and/or sexing of birds in the field at the time of banding has devastating effects on the study. We have seen some examples of such problems, for example, in mourning dove data. Many songbirds are banded without identifying either the age or the sex. Data from such studies cannot be used for estimating vital statistics but perhaps some distributional data will be obtained.

The subject of "representativeness" could well occupy a chapter of its own. Consider the large-scale duck banding program conducted annually in the northern two-thirds of North America. Typically birds are banded on concentration areas (staging areas). Here, they are easily trapped and banded in sufficient numbers. It could be asked if the birds in singles or pairs or small groups on the smaller water areas might represent a different population segment and, since they are less likely to be banded, inferences from the banded sample to the entire population would not be valid. We suggest that this is not especially important because the recovery area is quite large (e.g., the major part of the Central and Mississippi Flyways) and both "types" of birds pass over the recovery (sampling) area during the fall migration. This suggests that the location of the birds during the banding period in late August will have little relevance to the sampling scheme (hunting). This conclusion might change if most hunting deaths took place on the staging areas and the birds remained on these areas during the hunting season. Neither of these conditions seems to be true, at least with migratory species.

The above example suggests we can obtain our sample by banding the birds in a representative manner, or by employing a representative recovery (sampling) procedure after a random mixing process of the birds. These considerations might be quite different with studies of resident species, either game or nongame.

In planning new banding studies of resident birds we urge that biologists consider biannual (pre- and post-hunting season) banding programs: Models H_7 and H_8 (Chapter 7). These models may allow a study of the effect of hunting on the survival process if adequate data are collected. Here again we emphasize the possible use of Reward bands in such studies. A second potential lies in data gathering efforts for Models H_4, H_5, and H_6. These models are appropriate for those species that can be identified and banded according to the three age classes at the time of banding. Furthermore, annual samples of sufficient size must be obtained for each age group. The potential for this type of study is large, however, and the data requirements are stringent, perhaps impossible except for a few species.

In summary we emphasize that banding is not a panacea. Often in migratory bird management and research the word "study" has been nearly synonymous with the word "band." Banding is an excellent tool for *some* species for *some* objectives. It is possible to obtain some basic estimates of vital statistics of the population and gain insight into some very fundamental relationships concerning population dynamics. For some studies aimed at certain objectives, banding is usually the only method we now have. In some situations, no alternative exists because we have no other means to obtain needed information. These are the situations for which the material in this handbook will be useful. Unfortunately, many species cannot easily be banded or have behavioral habits that defy banding in sufficient numbers. In these respects, banding is no different than any other tool or technique. It is useful if used intelligently in the proper situation.

9.3 Sample Size Guidelines for Adults

The General Approach

In planning a banding study one question sure to arise is how many birds to band. The variables to be specified are the annual numbers of adults to band, $N_1, \ldots, N_k$ and k, the number of years of banding. From experience in analyzing real data and knowledge of these models and methods, we suggest 300 adults per year as a minimum for any species, even those with high recovery rates (for species with low recovery rates, more than 300 birds a year must be banded for reliable results). Otherwise the hypothesis tests and confidence intervals computed will be unreliable as they require "large samples." As mentioned above we recommend at least 5 years of banding as a minimum goal when planning a banding study. Finally, it is reasonable in planning the study to aim for equal numbers of banded birds each year. Thus, in what follows we assume $N_1 = N_2 = \cdots = N_k \equiv N$.

Specific sample size considerations require that we have a criterion to be satisfied by our choice of N. Consequently, we assume it is desired to estimate the average survival $\bar{S}$, using Model 1, within a specified precision after the recoveries are available for all k years of the study. An alternate approach, which has been developed by Youngs and Robson (1975), is to determine each annual banding quota (N_i) to achieve specific precision for each annual survival rate estimator $\tilde{S}_i$. That is, their paper gives the methodology to determine N_i, given a desired variance $\mathrm{VAR}(\tilde{S}_i)$ for any (or all) years. The reader who wishes to specify each N_i separately in this way is referred to Youngs and Robson (1975).

Let $\mathrm{VAR}(\tilde{\bar{S}})$ be the sampling variance of $\bar{S}$ (see Section 2.2), and $\mathrm{SE}(\tilde{\bar{S}}) = \sqrt{\mathrm{VAR}(\tilde{\bar{S}})}$ be the corresponding standard error of $\tilde{\bar{S}}$. An approximate 95% confidence interval on $\bar{S}$ is given by $\tilde{\bar{S}} \pm 1.96\ \mathrm{SE}(\tilde{\bar{S}})$. This formula can be rewritten as $\tilde{\bar{S}}(1. \pm 1.96\ \mathrm{CV}(\tilde{\bar{S}}))$, where $\mathrm{CV}(\tilde{\bar{S}})$ is the coefficient of variation of the estimator $\tilde{\bar{S}}$, $\mathrm{CV}(\tilde{\bar{S}}) = \mathrm{SE}(\tilde{\bar{S}}) / \tilde{\bar{S}}$.

A convenient mathematical criterion for choosing the sample size N is to have $\mathrm{CV}(\tilde{\bar{S}})$ less than or equal to some specified value (e.g., 0.05 or 0.1). After k years, we can then estimate average annual survival within a preset level of precision. This is equivalent to aiming for a confidence interval on $\bar{S}$ with a specified expected width.

The true coefficient of variation of $\tilde{\bar{S}}$ depends upon the unknown annual rates S_i and f_i that will apply during the study, and upon N and k. Exact sample size determination is thus impossible. A practical course of action is to obtain the best guess of the anticipated average rates $\bar{S}$ and $\bar{f}$, then use them in place of the unknown S_i and f_i in the formula for $\mathrm{CV}(\tilde{\bar{S}})$. These anticipated average rates can usually be deduced from available data on the same or related species.

It will be shown below that $\mathrm{VAR}(\tilde{\bar{S}})$, and hence $\mathrm{CV}(\tilde{\bar{S}})$, is proportional to $1/N$. In fact once we replace S_i and f_i by $\bar{S}$ and $\bar{f}$, respectively, in the formula for $\mathrm{VAR}(\tilde{\bar{S}})$, then the expression

$$[\mathrm{CV}(\tilde{\bar{S}})]^2 = \frac{h(\bar{S}, \bar{f}, k)}{N} \tag{9.3.1}$$

results, where the function $h(\cdot)$ depends only upon $\bar{S}$, $\bar{f}$, and k.

The steps for sample size determination on adults are as follows:

(1) Specify k and the desired relative precision $\mathrm{CV}(\tilde{\bar{S}})$;

(2) supply the best known values of $\bar{S}, \bar{f}$; and

(3) compute $h(\bar{S}, \bar{f}, k)$ and from (9.3.1) obtain N as $N = h(\bar{S}, \bar{f}, k) / [\mathrm{CV}(\tilde{\bar{S}})]^2$.

We will exhibit the formula for $h(\cdot)$ below, and then give a simple way to compute it.

First a word about choosing $\mathrm{CV}(\tilde{\bar{S}})$. We recommend $\mathrm{CV}(\tilde{\bar{S}}) \leq 0.05$, or smaller. If this coefficient of variation were as large as 0.1, with for example, $\bar{S} = 0.5$, the expected 95% confidence interval on $\bar{S}$ would be (0.4, 0.6). Such a wide interval is not very useful. For the same value of $\bar{S} = 0.5$, a value of $\mathrm{CV}(\tilde{\bar{S}}) = 0.05$ gives an expected 95% confidence interval of (0.45, 0.55). If a very precise estimator of $\bar{S}$ is desired, a value of CV($\tilde{\bar{S}}$) = 0.025 or less would not be unreasonable as a goal.

The Function $h(\bar{S},\bar{f},k)$

The sampling variance of $\tilde{\bar{S}}$ is expressed, in general, as

$$\mathrm{VAR}(\tilde{\bar{S}}) = \left[\sum_{i=1}^{k-1} \mathrm{VAR}(\tilde{S}_i) + 2\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-1}\mathrm{COV}(\tilde{S}_i,\tilde{S}_j)\right] \Big/ (k-1)^2 . \tag{9.3.2}$$

From Robson and Youngs (1971), formulas for the theoretical (large sample) sampling variances and covariances needed above are

$$\mathrm{VAR}(\tilde{S}_i) = (S_i)^2 \left[\frac{1}{\mathrm{E}(R_i)} - \frac{1}{N_i} + \frac{1}{\mathrm{E}(R_{i+1})} - \frac{1}{N_{i+1}} + \frac{1}{\mathrm{E}(T_i - C_i)} - \frac{1}{\mathrm{E}(T_i)}\right] \quad , i = 1, \ldots, k-1$$

and

$$\mathrm{COV}(\tilde{S}_i, \tilde{S}_j) = \begin{cases} 0 & , j > i+1 \\ -(S_i S_{i+1}) \left[\dfrac{1}{\mathrm{E}(R_{i+1})} - \dfrac{1}{N_{i+1}}\right] & , j = i+1,\ i = 1, \ldots, k-2 . \end{cases}$$

The notation of Chapter 2 is used above; for example, R_i is the i^{th} row total of the recovery data array.

If we use the restriction of equal N_i in all years, the above variances and covariances are all proportional to $1/N$. This is because terms like $\mathbf{E}(R_i)$, $\mathbf{E}(T_i - C_i)$ and $\mathbf{E}(T_i)$ are all directly proportional to N. For example

$$\mathbf{E}(R_i) = N(f_i + S_i f_{i+1} + \cdots + S_i \cdots S_{k-1} f_k) .$$

Replacing the annual rates S_i and f_i, implicitly involved in the formulae for VAR($\tilde{S}_i$), by $\bar{S}$ and $\bar{f}$ allows us to write the approximations we must use in practice for sample size determination

$$\mathrm{VAR}(\tilde{S}_i) = \frac{(\bar{S})^2}{N} \left[\frac{1}{R_i^*} + \frac{1}{R_{i+1}^*} + \frac{1}{T_i^* - C_i^*} - \frac{1}{T_i^*} - 2\right] \tag{9.3.3}$$

$$\mathrm{COV}(\tilde{S}_i, \tilde{S}_j) = \begin{cases} 0 & , j > i+1 \\ -\dfrac{(\bar{S})^2}{N} \left[\dfrac{1}{R_{i+1}^*} - 1\right] & , j = i+1 . \end{cases} \tag{9.3.4}$$

The terms R_i^*, C_i^*, T_i^* are functions only of $\bar{S}$, $\bar{f}$ and k. They are easily computed by constructing the model structure in terms of the cell probabilities $\pi_{ij} = (\bar{S})^{j-i}\bar{f}$, treating these probabilities as data and applying the formulae for R_i, C_i and T_i to compute R_i^*, C_i^* and T_i^*. Specific formulae are

$$R_i^* = \sum_{j=i}^{k} \pi_{ij} = \bar{f}\left(\frac{1-(\bar{S})^{k-i+1}}{1-\bar{S}}\right) \quad , i = 1, \ldots, k ,$$

$$C_i^* = \sum_{i=1}^{j} \pi_{ij} = \bar{f}\left(\frac{1-(\bar{S})^{i}}{1-\bar{S}}\right) \quad , i = 1, \ldots, k ,$$

and

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

with $T_1^* \equiv R_1^*$. Explicitly

$$T_i^* = \bar{f}\left(\frac{1-(\bar{S})^{k-i+1}}{1-\bar{S}}\right)\left(\frac{1-(\bar{S})^{i}}{1-\bar{S}}\right) , i = 1, \ldots, k .$$

Using formulae (9.3.2), (9.3.3) and (9.3.4) gives, after simplification,

$$(k-1)^2\, \mathrm{VAR}(\tilde{\bar{S}}) = \frac{(\bar{S})^2}{N}\left[\sum_{i=1}^{k-1}\left(\frac{1}{T_i^* - C_i^*}\right) - \sum_{i=1}^{k-1}\left(\frac{1}{T_i^*}\right) + \frac{1}{R_1^*} + \frac{1}{R_k^*} - 2\right] .$$

From $[\mathrm{CV}(\tilde{\bar{S}})]^2 = \mathrm{VAR}(\tilde{\bar{S}}) / (\bar{S})^2$, and the above formula, we derive $h(\bar{S}, \bar{f}, k)$ of formula (9.3.1) as

$$h(\bar{S},\bar{f},k) = \frac{\left[\displaystyle\sum_{i=1}^{k-1}\left(\frac{1}{T_i^* - C_i^*}\right) - \sum_{i=1}^{k-1}\left(\frac{1}{T_i^*}\right) + \frac{1}{R_1^*} + \frac{1}{R_k^*} - 2\right]}{(k-1)^2} , \tag{9.3.5}$$

where as shown above T_i^*, C_i^*, R_i^*, and R_i^* depend only upon $\bar{S}, \bar{f}$, and k. Unfortunately, (9.3.5) does not appreciably simplify.

If the terms R_i^*, C_i^*, and T_i^* are written out as functions of $\bar{S}$ and $\bar{f}$, one sees that $h(\cdot)$ is approximately proportional to $1/\bar{f}$ times a function of only $\bar{S},k$. Consequently, to a first-order approximation the required sample size is proportional to $1/\bar{f}$. This implies a given precision could be met, for example, with half the banded sample if $\bar{f}$ could be doubled. It is worth emphasizing that average recovery rate is one of the most important controlling parameters affecting the precision of estimates of S and f from banding studies. Studies on species with very low recovery rates are usually pointless, unless huge samples can be banded annually.

Computing the Function $h(\bar{S}, \bar{f}, k)$

Given a study length k, a specified precision on $\bar{\hat{S}}$, via the CV$(\bar{\hat{S}})$, and reasonable guesses for $\bar{S}$ and $\bar{f}$, then the required value of N is computed by finding $h(\bar{S}, \bar{f}, k)$ and applying

$$N = \frac{h(\bar{S}, \bar{f}, k)}{[\text{CV}(\bar{\hat{S}})]^2} . \tag{9.3.6}$$

The function $h(\cdot)$ given in (9.3.5) is easily calculated.

First generate the triangular array of cell probabilities ("data") $\pi_{ij} = (\bar{S})^{j-i}\bar{f}$.

Year banded	Probability of band recovery in year j 1	2	3	,...,	k	Row totals
1	$\bar{f}$	$\bar{S}\bar{f}$	$\bar{S}^2\bar{f}$	,...,	$\bar{S}^{k-1}\bar{f}$	R_1^*
2		$\bar{f}$	$\bar{S}\bar{f}$	,...,	$\bar{S}^{k-2}\bar{f}$	R_2^*
3			$\bar{f}$	,...,	$\bar{S}^{k-3}\bar{f}$	R_3^*
.				.	.	.
.				.	.	.
.				.	.	.
k					$\bar{f}$	R_k^*
Column totals	C_1^*	C_2^*	C_3^*	,...,	C_k^*	

It is necessary to generate only the first row of the array; subsequent rows are merely truncated versions of row 1. Then sum the array rows and columns to obtain R_i^*, C_i^*. The mathematically inclined person can easily short-cut this process as the relationship $R_i^* = C_{k-i+1}^*$ holds.

Next the values of T_i^* are computed from the formula $T_i^* = R_i^* \, C_i^* / \bar{f}$. Finally, it is easy to compute $T_i^* - C_i^*$. It will be convenient to arrange these numbers in a brief table, computing the columns in the order shown, for $i = 1, \ldots, k-1$,

$$i \qquad R_i^* \qquad C_i^* \qquad T_i^* \qquad T_i^* - C_i^* .$$

Then compute three summary statistics from this table:

$$A = \sum_{i=1}^{k-1} \frac{1}{T_i^* - C_i^*} ,$$

$$B = \sum_{i=1}^{k-1} \frac{1}{T_i^*} ,$$

$$C = \frac{1}{R_1^*} + \frac{1}{\bar{f}} - 2 .$$

Items A and B are the sums of reciprocals of elements in columns 4 and 3, respectively, of the table.

Finally

$$h(\bar{S}, \bar{f}, k) = \frac{A - B + C}{(k-1)^2} .$$

An Example

Assume a 5-year ($k=5$) study on adult male mallards, banded preseason, is planned. From past data (e.g., Anderson 1975), $\bar{S}=0.6$ and $\bar{f}=0.07$ are reasonable values to use. Now compute the table of estimated probabilities of band recovery:

Year banded	Probability of band recovery in year j				
	1	2	3	4	5
1	0.0700	0.0420	0.0252	0.0151	0.0091
2		0.0700	0.0420	0.0252	0.0151
3			0.0700	0.0420	0.0252
4				0.0700	0.0420
5					0.0700

Only row 1 requires calculation; even there, element one is simply $\bar{f}=0.07$ and succeeding elements are just $\bar{S}=0.6$ times the previous element. For example, $0.0420=(0.6)\ (0.07)$ and $0.0252=(0.6)\ (0.0420)$.

Next, sum the rows and columns to obtain the partial table:

i	R_i^*	C_i^*
1	0.1614	0.0700
2	0.1523	0.1120
3	0.1372	0.1372
4	0.1120	0.1523

Now compute the column for $T_i^*=R_i^*C_i^*/\bar{f}$. For example,

$$T_1^*=(0.1614)\ (0.0700)\ /\ (0.0700)=0.1614,$$
$$T_2^*=(0.1523)\ (0.1120)/(0.0700)=0.2437.$$

Once the T_i^* column is written down it is easy to write the fourth column $T_i^*-C_i^*$. The completed table is

i	R_i^*	C_i^*	T_i^*	$T_i^*-C_i^*$
1	0.1614	0.0700	0.1614	0.0914
2	0.1523	0.1120	0.2437	0.1317
3	0.1372	0.1372	0.2689	0.1317
4	0.1120	0.1523	0.2437	0.0914.

Now find the sums of the reciprocals of the last two columns. On most calculators this is easy. For the example

$$A=\sum_{i=1}^{4}\frac{1}{T_i^*-C_i^*}=\frac{1}{0.0914}+\frac{1}{0.1317}+\frac{1}{0.1317}+\frac{1}{0.0914}=37.0679.$$

Similarly

$$B=\sum_{i=1}^{4}\frac{1}{T_i^*}=18.1214,$$

and

$$C=\frac{1}{R_1^*}+\frac{1}{\bar{f}}-2=\frac{1}{0.1614}+\frac{1}{0.0700}-2=18.4815.$$

Finally

$$h(0.6,0.07,5)=\frac{A-B+C}{4^2}=\frac{37.0679-18.1214+18.4815}{16}=2.3393.$$

Up to this point it has not been necessary to specify a coefficient of variation on $\bar{\bar{S}}$. Assume CV($\bar{\bar{S}}$) = 0.04 is specified; then from (9.3.6) the sample size needed (numbers to band *each* year for 5 years) is

$$N = \frac{2.3393}{[0.04]^2} = 1{,}462.$$

This is a reasonable number for a species like the mallard that is easy to trap and band. Thus, if about 1,500 adult male birds were banded each year for 5 years, then our expected 95% confidence interval around the mean survival rate (0.60) for the 5-year period would be 0.552 to 0.648 (the actual interval, being a random variable, would vary somewhat from these theoretical limits).

Other species are more difficult to capture and band; e.g., canvasback *(Aythya valisineria)* and snipe *(Capella gallinago)*. Let us now examine the situation if only 500 birds per year could be banded (assuming the same values for $\bar{S}$, $\bar{f}$, and k). Then, using the above value of $h(\cdot)$ and $N = 500$, we can compute the relative precision to be expected:

$$\mathrm{CV}(\bar{\bar{S}}) = \sqrt{\frac{h(\bar{S},\bar{f},k)}{N}} = \sqrt{\frac{2.3393}{500}} = 0.0684.$$

In this case, the expected 95% confidence interval on the mean survival (0.60) is wider, 0.520 to 0.680. Of course, the 95% confidence intervals on the *individual* annual estimates of survival would be much wider than this.

This illustrates the point that these same calculations needed for sample size determination can be used to answer the question—"If we band N birds per year for k years, with $\bar{S}$ and $\bar{f}$ as anticipated parameters, what relative precision will we achieve?" We recommend these calculations be made in any study, whether to determine N, or determine CV($\bar{\bar{S}}$) given N. If the indicated CV($\bar{\bar{S}}$) (for N given) is more than 0.1, the study is essentially worthless if survival rate estimation is the primary objective.

We emphasize that this procedure gives only guidelines for choosing the annual sample size. It is dependent upon a number of assumptions: (1) Model 1 or 2 (not Models 0 or 3) is correct, (2) the chosen values of $\bar{S}$ and $\bar{f}$ are near the values that will apply during the study, and (3) only expected confidence limits are given; the actual limits computed after the study will not be identical to the predicted ones. The assumption $N = N_1 = \cdots = N_k$ is not critical, however. If actual numbers banded vary about the target value of N by even ±10%, these sample size considerations will still by useful. Program ESTIMATE allows sample size guidelines to be computed as an option (See Chapter 6).

9.4 Sample Size Guidelines for Young

The General Approach

Adult birds should always be banded in conjunction with banding of young (e.g., see Section 3.9). Thus we are faced with determining sample sizes for adults as well as young when young are to be banded. Consequently, the reader should understand Section 9.3 before proceeding with this section.

In the problem at hand, the variables to be specified are the annual numbers of adults (N_i) and young (M_i) to band each year for k years. As with adults only, we recommend a minimum of 300 birds of each age class be banded each year. Also, we recommend 5 years of consecutive banding as a minimum goal for the study length if survival rate estimation is desired. It is reasonable in planning the study to aim for equal numbers of banded birds in each age class, each year; thus we assume $N_i \equiv N$ and $M_i \equiv M$ for $i = 1, \ldots, k$. Finally, in what follows we assume that Model H_1 will describe the band recovery data from the study, and therefore the notation of Chapter 3 will be used here.

From considerable experience with banding data we have observed that the precision of the adult survival rate estimators based on Model H_1 (adults and young) is not very different from the precision obtained with Model 1 (adults only). Consequently, we recommend that the adult sample size first be determined by the methods of Section 9.3. It will then be possible to determine an annual (target) sample size for numbers of young to band.

Let S_i' be the annual young survival rate. Assume it is desired to estimate $\bar{S}'$, the average annual young survival rate during the study, with a specified precision. Let VAR($\bar{\bar{S}}'$) be the sampling variance of the estimator $\bar{\bar{S}}'$ and SE($\bar{\bar{S}}'$) = $\sqrt{\mathrm{VAR}(\bar{\bar{S}}')}$ be the corresponding standard error of $\bar{\bar{S}}'$. An approximate 95% confidence interval on $\bar{S}'$ is given by $\bar{\bar{S}}' \pm 1.96$ SE($\bar{\bar{S}}'$). This formula can be rewritten as $\bar{\bar{S}}(1. \pm 1.96$ CV($\bar{\bar{S}}'$)), where CV($\bar{\bar{S}}'$) is the coefficient of variation of the estimator $\bar{\bar{S}}'$: CV($\bar{\bar{S}}'$) = SE($\bar{\bar{S}}'$) / $\bar{S}'$.

As with adult sample size determination, a convenient mathematical criterion for choosing the sample size M is to have CV$(\bar{S}')$ less than or equal to some specified value, such as 0.05 or 0.025. This is equivalent to choosing a sample size such as that there is a prescribed expected confidence interval width on $\bar{S}'$.

The true coefficient of variation of $\bar{\tilde{S}}'$ depends upon the unknown annual survival and recovery rates $(S_i', S_i$ f_i', and $f_i)$ for young and adults that will apply during the study. Exact sample size determination is thus impossible. As in determination of adult sample size N, a practical course of action is to obtain the best guess of the anticipated average rates, $\bar{S}'$, $\bar{S}$, $\bar{f}'$, and $\bar{f}$, then use them in place of the appropriate unknown year-specific parameters in the formula for CV$(\bar{\tilde{S}}')$. These anticipated average rates can usually be deduced from available data on the same or related species.

Given the above approach to determining sample size for young, a formula for M can be derived. Given the requisite adult sample size has been computed, it is then easy to compute M. Derivation of the formula will be given below in a separate subsection; the less mathematical reader can skip this material.

Derivation of the Formula for M

From Brownie (1973:14)

$$\text{VAR}(\tilde{S}_i') = (S_i')^2\left[\frac{1}{\mathbf{E}(Q_{i\cdot} - Q_{ii})} - \frac{1}{M_i} + \frac{1}{\mathbf{E}(R_{i+1\cdot})} - \frac{1}{N_{i+1}}\right] \tag{9.4.1.}$$

and

$$\text{COV}(\tilde{S}_i', \tilde{S}_j') = 0 \qquad i \neq j, i = 1, \ldots, k-1.$$

It follows that

$$\text{VAR}(\bar{\tilde{S}}') = \left(\frac{1}{k-1}\right)^2 \sum_{i=1}^{k-1} \text{VAR}(\tilde{S}_i').$$

Upon making the assumptions $N_i \equiv N, M_i \equiv M, S_i' \equiv \bar{S}', S_i \equiv \bar{S}, f_i' \equiv \bar{f}'$ and $f_i \equiv \bar{f}$, explicit results can be derived:

$$\text{VAR}(\bar{\tilde{S}}') = \left(\frac{1}{k-1}\right)^2 (\bar{S}')^2\left[\sum_{i=1}^{k-1}\left\{\frac{1}{\mathbf{E}(Q_{i\cdot} - Q_{ii})} + \frac{1}{\mathbf{E}(R_{i+1\cdot})}\right\} - \frac{k-1}{M} - \frac{k-1}{N}\right],$$

$$[(k-1)\text{CV}(\bar{\tilde{S}}')]^2 = \frac{1}{M}\frac{1}{\bar{S}'}\left[\sum_{i=1}^{k-1}\frac{1}{R_{i+1}^*}\right] + \frac{1}{N}\left[\sum_{i=1}^{k-1}\frac{1}{R_{i+1}^*}\right] - (k-1)\left[\frac{1}{M} + \frac{1}{N}\right]. \tag{9.4.2}$$

In formula (9.4.2) R_i^* has the same meaning as in Section 9.3; specifically it depends only upon adult parameters $\bar{S}$ and $\bar{f}$, not upon parameters for young. Because of the relationship between R_i^* and C_i^* formula (9.4.2) may be rewritten as

$$[(k-1)\text{CV}(\bar{\tilde{S}}')]^2 = \frac{1}{M}\left[\frac{A'}{\bar{S}'} - (k-1)\right] + \frac{1}{N}\left[A' - (k-1)\right],$$

where

$$A' = \sum_{i=1}^{k-1}\frac{1}{C_i^*}$$

and

$$C_i^* = \bar{f}\left(\frac{1-(\bar{S})^i}{1-\bar{S}}\right).$$

Finally the formula for M is

$$M = \frac{\left[\frac{A'}{\bar{S}'} - (k-1)\right]}{\left[(k-1)\text{CV}(\bar{\tilde{S}}')\right]^2 - \left[\frac{A' - (k-1)}{N}\right]}. \tag{9.4.3}$$

An Example and Some Considerations

The desired sample size M is given by formula (9.4.3), where the quantity A' is given by

$$A' = \sum_{i=1}^{k-1}\frac{1}{C_i^*},$$

with the C_i^* having exactly the same meaning here as in Section 9.3. Hence the C_i^* depend only upon adult parameter values, and must be computed anyway to determine the value of N which appears in formula (9.4.3).

The example of Section 9.3 will be elaborated upon here. Assume young are also to be banded, and that $\bar{S}' = 0.5$ (note $\bar{f}'$ drops out of the final formula for M, hence need not be considered). In the preceding example $S = 0.6$, $\bar{f} = 0.07$, CV($\bar{\bar{S}}$) = 0.04, and the C_i^* were computed to be $C_1^* = 0.07$, $C_2^* = 0.112$, $C_3^* = 0.1372$, and $C_4^* = 0.1523$. It is thus easy to compute

$$A' = \frac{1}{C_1^*} + \frac{1}{C_2^*} + \frac{1}{C_3^*} + \frac{1}{C_4^*} = 37.06.$$

From Section 9.3, $N = 1462$ in this example; thus from formula (9.4.3)

$$M = \frac{\left[\frac{37.06}{0.5} - 4\right]}{\left[4\,\mathrm{CV}(\bar{\bar{S}}')\right]^2 - \left[\frac{37.06-4}{1462}\right]},$$

$$M = \frac{70.12}{16\,[\mathrm{CV}(\bar{\bar{S}}')]^2 - 0.02261}. \tag{9.4.4}$$

Only now is it necessary to choose a desired value of CV($\bar{\bar{S}}'$). Because CV($\bar{\bar{S}}$) (for adults) was taken as 0.04, one might also want to use this value for young. The value of M is then computed to be

$$M = \frac{70.12}{16(0.04)^2 - 0.02261} = 23{,}452.$$

This is a surprisingly large number, and no doubt impossible to achieve. This points out that it may not be possible to achieve the same coefficient of variation for both adults and young.

The basic problem is that VAR($\bar{\bar{S}}'$) depends upon both the numbers of adults (N) and of young (M) banded. Structurally, $\hat{S}_i'$ equals $\widehat{S_i'\rho_{i+1}}/\hat{\rho}_{i+1}$, where the product $S_i'\rho_{i+1}$ is estimated from band recoveries of young only and ρ_{i+1} is estimated from adult recoveries. Thus the variance of $\hat{S}_i'$ (and of $\bar{\bar{S}}'$) has two components, one dependent upon the value of N and one dependent upon the value of M. First-year recoveries from birds banded as young are not used in the estimation of S_i' and this also explains why precision is somewhat poor. Once a value of N is chosen for adults there is a lower bound on the precision of $\bar{\bar{S}}'$ that can be achieved, even if infinitely many young birds could be banded. In this example with $N = 1{,}460$ the lower bound on CV($\bar{\bar{S}}'$) is

$$\mathrm{CV}(\bar{\bar{S}}') = \sqrt{\frac{0.02261}{N16}} = 0.03759\,.$$

In general for N given, the lower bound is

$$\mathrm{CV}(\bar{\bar{S}}') \geqslant \sqrt{\frac{A' - (k-1)}{N(k-1)^2}}\;, \tag{9.4.5}$$

(which can only be achieved if infinitely many young are banded).

For the example at hand one would realistically have to settle for a lower precision on $\bar{\bar{S}}'$, for instance CV($\bar{\bar{S}}'$) = 0.07. Then from formula (9.4.4):

$$M = \frac{70.12}{16(0.07)^2 - 0.02261} = 1{,}257\;.$$

As was pointed out in Section 9.3, an alternative use of these sample size formulae is to determine the precision that will be achieved if a given number of birds is banded (given values of k, $\bar{S}$, etc.) For the case of young, once a value of N is determined, or given, one simply solves formula (9.4.3) for CV($\bar{\bar{S}}'$) given the value of M anticipated.

As a general procedure we recommend computing the unknown terms, like A', to arrive at a reduced formula such as (9.4.4). For this example one then has

$$\mathrm{CV}(\bar{\bar{S}}')=\sqrt{\frac{70.12}{M\,16}+\frac{0.02261}{16}}=\sqrt{\frac{4.3825}{M}+0.001413}\ .$$

Any number of values of M can now be tried. If it were felt one could band twice as many young as adults, then we have $M=2(1{,}462)\approx 3{,}000$, and

$$\mathrm{CV}(\bar{\bar{S}}')=\sqrt{\frac{4.3825}{3{,}000}+0.001413}=0.0536\ .$$

Another way to use these formulae for both adults and young is to ask what would be an optimal ratio to band (young/adult) if one wants equal, but unspecified, values of $\mathrm{CV}(\bar{\bar{S}})$ and $\mathrm{CV}(\bar{\bar{S}}')$. The relevant formula is

$$\frac{M}{N}=\frac{\left[\frac{A'}{\bar{\bar{S}}'}-(k-1)\right]}{h(\bar{S},\bar{f},k)\ (k-1)^2-[A'-(k-1)]}\ .$$

Evaluating the formula for the example used throughout these two sections gives

$$\frac{M}{N}=\frac{70.12}{(2.3393)16-33.06}=16.0\ ,$$

or 16 young per adult. This is likely to be an impossible ratio when reasonable numbers of adults are banded, which fact goes back to our earlier statement that it may be unrealistic to try to achieve the same coefficient of variation for both adults and young.

In summary, the determination of sample sizes (N,M) in case both young and adults are banded is not as simple as determining N alone. It is important to bear in mind the lower limit achievable for $\mathrm{CV}(\bar{\bar{S}}')$ (once a value of N is chosen). This fact may require some adjustments in the target value of N in order to achieve suitable precision for both ages. Even if these formulae are used in no other way, they should be used to determine expected coefficients of variation for $\bar{\bar{S}}$ and $\bar{\bar{S}}'$ in a planned study with prescribed values of N and M. If this precision is too low (coefficients of variation greater than 0.1), the study is nearly worthless for survival rate estimation.

Acknowledgments

We acknowledge the following conservation agencies for their permission to use banding and recovery data as examples in this handbook: Colorado Division of Wildlife, Illinois Department of Conservation, Ontario Department of Lands and Forests, and the Saskatchewan Department of Tourism and Renewable Resources.

B. Keokosky and G. Schaar, Cornell Computer Activities Group, coded program BROWNIE. Much of the basic research leading to several of the methods presented in this handbook was funded from administrative monies of the Federal Aid program of the Fish and Wildlife Service in Washington, D.C. The Office of Migratory Bird Management paid expenses related to the development of program ESTIMATE and for the many test runs related to both programs. In addition, they have agreed to distribute magnetic tapes of the programs and to perform certain computer analyses. Ms. D. Drzal, Migratory Bird and Habitat Research Laboratory, deserves special credit for typing and coordinating the various drafts as the manuscript was being prepared. S. Talbot, Utah State University, played a major role in the final preparation of the manuscript. F. Martin, H. M. Reeves, and R. I. Smith provided essential administrative services.

Since the first edition was printed, several individuals have assisted in the further development of the computer software. Steven Cherry and Steven Lannigan made extensions and modifications to programs ESTIMATE and BROWNIE. In particular, Gary White has updated these programs to FORTRAN 77, added code to make BROWNIE interactive, and made revisions to allow both programs to be run on microcomputers. Cyndi Brooks provided the illustrations.

We appreciate the support of several of our colleagues who have allowed us to reprint their papers in Appendix C, and thank the respective editors of various journals for their cooperation in this respect.

References

In the analysis of banding or tagging studies there is little need to sharply distinguish between the single (dead) recovery case and the multiple (live) recapture case. The mathematical models and analysis methods are very similar in both cases. Consequently, we have included in this bibliography important papers that deal with estimates of survival (mortality) rates from numerous types of banding or tagging studies. We have not included papers that treat only the estimation of population size. The reader who wishes to investigate more fully the practice and analysis of mark-recaptured studies should consult the following: Cormack (1979), Nichols et al. (1984), Pollock (1981), and Seber (1982). References new to the second edition are starred.

Anderson, D.R. 1972. Bibliography on methods of analyzing bird banding data with special reference to the estimation of population size and survival. U.S. Fish Wildl. Serv., Spec. Sci. Rep.-Wildl. 156. 13 pp.

Anderson, D.R. 1975. Population ecology of the mallard: V. Temporal and geographic estimates of survival, recovery and harvest rates. U.S. Fish Wildl. Serv., Resour. Publ. 125. 110 pp.

Anderson, D.R., F.R. Fiehrer, and C.F. Kimball. 1974. A computer program for estimating survival and recovery rates. J. Wildl. Manage. 38(2):369-370.

Anderson, D.R., and K.P. Burnham. 1976. Population ecology of the mallard: VI. The effect of exploitation on survival. U.S. Fish Wildl. Serv., Resour. Publ. 128. 66 pp.

*Anderson, D.R., and K.P. Burnham. 1980. Effect of delayed reporting of band recoveries on survival estimates. J. Field Ornithol. 51(3):244-247.

Anderson, D.R., and R.T. Sterling. 1974. Population dynamics of molting pintail drakes banded in south-central Saskatchewan. J. Wildl. Manage. 38(2):266-274.

*Anderson, D.R., A.P. Wywialowski, and K.P. Burnham. 1981. Tests of the assumptions underlying life table methods for estimating parameters from cohort data. Ecology 62(4):1121-1124.

*Anderson, D.R., K.P. Burnham, and G.C. White. 1982. The effect of exploitation on annual survival of mallard ducks: An ultrastructural model. Proceed. XI Internat. Biometric Conf. Toulouse, France. 11:33-39.

*Anderson, D.R., K.P. Burnham, and G.C. White. 1985. Problems in estimating age-specific survival rates from recovery data of birds ringed as young. J. Anim. Ecol. 54:89-98.

Arnason, A.N. 1972a. Parameter estimates from mark-recapture experiments on two populations subject to migration and death. Res. Popul. Ecol. 13:97-113.

Arnason, A.N. 1972b. Prediction methods and variance estimates for the parameters of the triple catch — two population model with migration and death. Univ. of Manitoba, Dep. of Comp. Sci., Sci. Rep. 54. 31 pp.

Arnason, A.N. 1973. The estimation of population size, migration rates and survival in a stratified population. Res. Popul. Ecol. 15:1-8.

*Arnason, A.N., and L. Baniuk. 1978. POPAN-2: A data maintenance and analysis system for mark-recapture data. Charles Babbage Research Centre, Box 370, St. Pierre, Manitoba, Canada. 269 pp.

*Arnason, A.N. and L. Baniuk. 1980. A computer system for mark-recapture analysis of open populations. J. Wildl. Manage. 44(2):325-332.

Arnason, A.N., and N.S. Kreger. 1973. User's manual POPAN-1 – a program to calculate Jolly-Seber estimates from mark-recapture data with pooling methods to increase the precision. Univ. of Manitoba, Dep. of Comp. Sci., Sci. Rep. 76. 32 pp.

*Arnason, A.N., C.R. Krasey, and K.H. Mills. 1982. A computer program for predicting precision and tag-loss bias in Jolly-Seber mark-recapture estimates. Canadian Tech. Rep. of Fish. & Aquat. Sci. 1083. 42 pp.

*Arnason, A.N., and K.H. Mills. 1981. Bias and loss of precision due to tag-loss in Jolly-Seber estimates for mark-recapture experiments. Canadian J. Fish & Aquat. Sci. 38:1077-1095.

*Balser, J.P. 1981. Confidence interval estimation and tests for temporary outmigration in tag-recapture studies. PhD Thesis, Cornell Univ., Ithaca, NY. 205 pp.

*Begon, M. 1979. Investigating animal abundance: Capture-recapture for biologists. University Park Press, Baltimore, MD. 97 pp.

Beverton, R.J.H., and S.J. Holt. 1956. A review of methods for estimating mortality rates in exploited fish populations, with special reference to sources of bias in catch sampling. Pages 67-83 *in* Problems and methods of sampling fish populations. Andr. Fred. Host & Fils, Copenhagen.

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Her Majesty's Stationery Office, London. 533 pp.

Bishop, J.A., and P.M. Sheppard. 1973. An evaluation of two capture-recapture models using the technique of computer simulation. Pages 235-252 *in* M.S. Bartlett and R.W. Hiorns, eds. The mathematical theory of the dynamics of biological population. Academic Press, London.

*Blower, J.G., L.M. Cook, and J.A. Bishop. 1981. Estimating the size of animal populations. George Allen & Unwin Ltd., London. 128 pp.

Boyd, H. 1956. Statistics of the British population of the pink-footed goose. J. Anim. Ecol. 25:253-273.

Brownie, C. 1973. Stochastic models allowing age-dependent survival rates for banding experiments on exploited bird populations. PhD Thesis, Cornell University, Ithaca, NY. 114 pp.

Brownie, C. 1974a. Estimation of instantaneous mortality rates from band return data when "within-season date of kill" is also available. Biometrics Unit, Cornell University, Ithaca, NY. BU-536-M. 15 pp.

Brownie, C. 1974b. A modification of the SRY model for band-return data, allowing for differences in reporting rates near banding sites. Biometrics Unit, Cornell University, Ithaca, NY. BU-535-M. 13 pp.

Brownie, C., and D.S. Robson. 1974a. Testing for equality of the annual survival rates in the SRY model for band return data. Biometrics Unit, Cornell University, Ithaca, NY. BU-515-M. 18 pp. plus Appendix.

Brownie, C., and D.S. Robson. 1974b. Models allowing for age-dependent survival rates for band-return data. Biometrics Unit, Cornell University, Ithaca, NY. BU-514-M. 27 pp. plus tables.

Brownie, C., and D.S. Robson. 1976. Models allowing for age-dependent survival rates for band-return data. Biometrics 32(2):305-323.

*Brownie, C., and D.S. Robson. 1983. Estimation of time-specific survival rates from tag-resighting samples: A generalization of the Jolly-Seber model. Biometrics 39:437-453.

*Brownie, C., and K.H. Pollock. 1985. Analysis of multiple capture-recapture data using band-recovery methods. Biometrics 41(2):411-420.

*Buckland, S.T. 1980. A modified analysis of the Jolly-Seber capture-recapture model. Biometrics 36:419-435.

*Buckland, S.T. 1982. A mark-recapture survival analysis. J. Anim. Ecol. 51:833-847.

*Buckland, S.T., I. Rowley, and D.A. Williams. 1983. Estimation of survival from repeated sightings of tagged galahs. J. Anim. Ecol. 52:563-573.

*Burnham, K.P., and D.R. Anderson. 1979. The composite dynamic method as evidence for age-specific waterfowl mortality. J. Wildl. Manage. 43(2):356-366.

*Burnham, K.P., and J.D. Nichols. 1985. On condition bias and band-recovery data from large-scale waterfowl banding programs. Wildl. Soc. Bull. 13(3):

Carothers, A.D. 1971. An examination and extension of Leslie's test of equal catchability. Biometrics 27(3):615-630.

Carothers, A.D. 1973a. Capture-recapture methods applied to a population with known parmaters. J. Anim. Ecol. 42(1):125-146.

Carothers, A.D. 1973b. The effects of unequal catchability on Jolly-Seber estimates. Biometerics 29(1):79-100.

Chapman, D.G. 1961. Statistical problems in dynamics of exploited fisheries populations. Pages 153-168 *in* J. Neyman, ed. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley.

Chapman, D.G. 1965. The estimation of mortality and recruitment from a single-tagging experiment. Biometrics 21(3):529-542.

Chapman, D.G., and D.S. Robson. 1960. The analysis of a catch curve. Biometrics 16(3):354-368.

*Colbert, J., J.D. Lebreton, M. Clobert-Gillet, and H. Coquillart. In press. The estimation of survival in bird populations by recaptures or sightings of marked individuals. *In* B.J.T. Morgan and P. North, eds. Statistics in Ornithology. Springer Verlag, Berlin.

*Clobert, J., and J.D. Lebreton. 1985. Dependance de facteurs de milieu dans las estimations de taux de survie par capture-recapture. Biometrics 41(3). Not yet published.

*Conroy, M.J., and J.D. Nichols. 1984. Testing for variation in taxonomic extinction probabilities: A suggested methodology and some results. Paleobiology 10(3):328-337.

*Conroy, M.J., and B.K. Williams. 1984. A general methodology for maximum likelihood inference from band-recovery data. Biometrics 40(3):739-748.

*Conroy, M.J., and B.K. Williams. 1981. Sensitivity of band reporting-rate estimates to violation of assumptions. J. Wildl. Manage. 45(3):789-792.

Cormack, R.M. 1964. Estimates of survival from the sighting of marked animals. Biometrika 51(3-4):429-438.

Cormack, R.M. 1966. A test of equal catchability. Biometrics 22(2):330-342.

Cormack, R.M. 1968. The statistics of capture-recapture methods. Oceanogr. Mar. Biol. Annu. Rev. 6:455-506.

Cormack, R.M. 1972. The logic of capture-recapture estimates. Biometrics 28(2):337-343.

Cormack, R.M. 1973. Commonsense estimates from capture-recapture studies. Pages 225-234 *in* M.S. Bartlett and R.W. Hiorns, eds. The mathematical theory of the dynamics of biological populations. Academic Press, London.

*Cormack, R.M. 1979. Models for capture-recapture. Pages 217-255 *in* R.M. Cormack, G.P. Patil, and D.S. Robson, eds. Statistical ecology, Vol. 5. Sampling biological populations. International Co-operative Publishing House, Fairland, MD.

*Cormack, R.M. 1981. Loglinear models for capture-recapture experiments on open populations. Pages 217-235 *in* R.W. Hiorns and D. Cook, eds. The mathematical theory of the dynamics of biological populations II. Academic Press.

Cowardin, L.M. 1977. Analysis and machine mapping of the distribution of band recoveries. U.S. Fish Wildl. Serv., Spec. Sci. Rep.-Wildl. 198. 8 pp.

*Crosbie, S.F. 1979. The mathematical modelling of capture-mark-recapture experiments on animal populations. PhD Thesis, Univ. of Otago. Dunedin, New Zealand. 387 pp.

*Crosbie, S.F., and B.F.J. Manly. 1985. Parsimonius modelling of capture-mark-recapture studies. Biometrics 41(2):399-410.

Darroch, J.N. 1959. The multiple-recapture census. II. Estimation when there is immigration or death. Biometrika 46(3-4):336-351.

Davenport, D.A. 1977. Computerized tabulation and display of band recovery data. U.S. Fish Wildl. Serv., Spec. Sci. Rep.-Wildl. 199. 7 pp.

Davidson, R.R., and D.L. Solomon. 1974. Moment-type estimation in the exponential family. Comm. in Stat. 3(11): 1101-1108.

DeLury, D.B. 1947. On the estimation of biological populations. Biometrics 3(2):145-167.

DeLury, D.B. 1951. On the planning of experiments for the estimation of fish populations. J. Fish. Res. Board Can. 8(4):281-307.

DeLury, D.B. 1954. On the assumptions underlying estimates of mobile populations. Pages 287-293 *in* O. Kempthorne, T.A. Bancroft, J.W. Gowen, and J.L. Lush, eds. Statistics and mathematics in biology. Iowa State College Press, Cedar Falls.

Eberhardt, L.L. 1969. Population analysis. Pages 475-495 *in* R.H. Giles, ed. Wildlife management techniques. The Wildlife Society, Wash., D.C.

Eberhardt, L.L 1972. Some problems in estimating survival from banding data. Pages 153-171 *in* Population ecology of migratory birds – a symposium. Bur. Sport Fish. Wildl., Wildl. Res. Rep. 2.

Fisher, R.A., and E.B. Ford. 1947. The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. Heredity 1(2):143-174.

Fordham, R.A. 1970. Mortality and population change of dominican gulls in Wellington, New Zealand, with a statistical appendix by R.M. Cormack. J. Anim. Ecol. 39(1):13-27.

Geis, A.D. 1972a. Role of banding data in migratory bird population studies. Pages 213-228 *in* Population ecology of migratory birds. A Symposium. U.S. Fish Wildl. Serv., Wildl. Res. Rep. 2.

Geis, A.D. 1972b. Use of banding data in migratory game bird research and management. U.S. Fish Wildl. Serv., Spec. Sci. Rep.-Wildl. 154. 47 pp.

Geis, A.D., and R.D. Taber. 1963. Measuring hunting and other mortality. Pages 284-298 *in* H.S. Mosby, ed. Wildlife investigational techniques. 2nd Edition. The Wildlife Society, Wash., D.C.

Gilbert, R.O. 1973. Approximations of the bias in the Jolly-Seber capture-recapture model. Biometrics 29(3):501-526.

Gulland, J.A. 1955. On the estimation of population parameters from marked members. Biometrika 42(1-2):269-270.

Haldane, J.B.S. 1953. Some animal life tables. J. Inst. Actuaries 79(1):83-89.

Haldane, J.B.S. 1955. The calculation of mortality rates from ringing data. Pages 454-458 *in* Acta XI Congress International Ornithologists.

Hammersley, J.M. 1953. Capture-recapture analysis. Biometrika 40(3-4):265-278.

Heincke, F. 1913. Investigations on the plaice. Rapp. P.-V. Reun. Perm. Cons. Int. Explor. Mer. 17A:1(1):1-153.

Henny, C.J., and K.P. Burnham. 1976. A reward band study of mallards to estimate band reporting rates. J. Wildl. Manage. 40(1):1-14.

Hickey, J.J. 1952. Survival studies of banded birds. U.S. Fish Wildl. Serv., Spec. Sci. Rep.-Wildl. 15. 177 pp.

*Hightower, J., and R.J. Gilbert. 1984. Using the Jolly-Seber model to estimate population size, mortality, and recruitment for a reservoir fish population. Trans. Am. Fish. Soc. 113:633-641.

*Hopper, R.M., H.D. Funk, and D.R. Anderson. 1978. Age-specificity in mallards banded postseason in Eastern Colorado. J. Wildl. Manage. 42(1):263-270.

Jackson, C.H.N. 1939. The analysis of an animal population. J. Anim. Ecol. 8:238-246.

Johnson, D.H. 1974. Estimating survival rates from banding of adult and juvenile birds. J. Wildl. Manage. 38(2): 290-297.

Johnson, N.L., and S. Kotz. 1969. Discrete distributions. Houghton & Miffling, Boston. 328 pp.

Jolly, G.M. 1963. Estimates of population parameters from multiple recapture data with both death and dilution – deterministic model. Biometrika 50(1-2):113-128.

Jolly, G.M. 1965. Explicit estimates from capture-recapture data with both death and immigration – stochastic model. Biometrika 52(1-2):225-247.

*Jolly, G.M. 1979. A unified approach to mark-recapture stochastic models, exemplified by a constant survival rate model. *In* R.M. Cormack, G.P. Patil, and D.S. Robson, eds. Sampling biological populations. Statistical Ecology Series. Vol. 5. Pages 277-282. International Co-operative Publishing House, Fairland, MD.

*Jolly, G.M. 1982. Mark-recapture models with parameters constant in time. Biometrics 38:301-321.

*Jolly, G.M., and J.M. Dickson. 1980. Mark-recapture suite of programs. Pages 570-576 *in* COMPSTAT 1980: Proc. Comp. Stat, Vol 4, M.M. Barritt and D. Wishart, eds., Physica-Verlag, Vienna.

Kale, B.K. 1962. On the solution of likelihood equations by iteration processes: The multiparametric case. Biometrika 49(3):479-486.

Kempthorne, O., and L. Folks. 1971. Probability, statistics, and data analysis. Iowa State University Press, Ames. 555 pp.

Kendall, M.G., and A. Stuart. 1961. The advanced theory of statistics. Vol. 2. Hafner Publ., NY. 676 pp.

Kreger, N.S. 1973. A simulation study of Jolly's estimates for animal populations when sampling intensity is low. M.S. Thesis. University of Manitoba, Winnipeg. 171 pp.

*Kremers, W.K. 1984. The Jolly-Seber model and maximum likelihood estimation revisited. Biometrics Unit, Cornell University, Ithaca, NY. BU-847-M. 7 pp.

*Kremers, W.K. 1985. Estimation of survival rates from a Jolly-Seber model with tag loss. Biometrics Unit, Cornell University, Ithaca, NY. BU-849-M. 21 pp.

Lack, D. 1943. The age of the blackbird. Brit. Birds 36(9):166-175.

*Lakhani, K.H., and I. Newton. 1983. Estimating age-specific bird survival rates from ring recoveries – can it be done? J. Anim. Ecol. 52:83-91.

*Lebreton, J.D. 1977. Maximum likelihood estimations of survival rates from bird band returns: Some complements to age-dependent methods. Biometrie-Praximetre. 17(3-4):145-161.

*Lefebvre, L.W., D.L. Otis, and N.R. Holler. 1982. Comparison of open and closed models for cotton rat population estimates. J. Wildl. Manage. 46(1):156-163.

Leslie, P.H., and D. Chitty. 1951. The estimation of population parameters from data obtained by means of the capture-recapture method. I. The maximum likelihood equations for estimating the death-rate. Biometrika 38 (3-4):269-292.

Manly, B.F.J. 1970. A simulation study of animal population estimation using the capture-recapture method. J. Appl. Ecol. 7(1):13-39.

Manly, B.F.J. 1971a. Estimates of a marking effect with capture-recapture sampling. J. Appl. Ecol. 8(1):181-189.

Manly, B.F.J. 1971b. A simulation study of Jolly's method for analyzing capture-recapture data. Biometrics 27(2): 415-424.

Manly, B.F.J. 1972-1973. Estimating survivorship from the recaptures of animals first released at age zero. Biometrie-Praximetrie 13:1-14.

Manly, B.F.J. 1974. Estimating survival from a multi-sample single recapture census where recaptures are not made at release times. Biom. Z 16(3):185-190.

*Manly, B.F.J. 1977a. A simulation experiment on the application of the jackknife with Jolly's method for the analysis of capture-recapture data. Acta Theriologica. 22:215-223.

*Manly, B.F.J. 1977b. A note on the design of experiments to estimate survival and relative survival. Biometric J. 19:687-692.

*Manly, B.F.J. 1981. Estimation of absolute and relative survival rates from the recoveries of dead animals. New Zealand J. Ecol. 4(1):78-88.

*Manly, B.F.J. 1984. Obtaining confidence limits on parameters of the Jolly-Seber model for capture-recapture data. Biometrics 40:749-758.

Manly, B.F.J., and M.J. Parr. 1968. A new method of estimating population size, survivorship, and birth rate from capture-recapture data. Trans. Soc. Br. Entomol. 18:81-89.

Manly, B.F.J., and G.A.F. Seber. 1973. Animal life tables from capture-recapture data. Biometrics 29(3):487-500.

*Mardekian, S.Z., and L. McDonald. 1981. Simultaneous analysis of band-recovery and live-recapture data. J. Wildl. Manage. 45(2):484-488.

Martin, E.M., and S.M. Carney. 1977. Population ecology of the mallard: IV. A review of duck hunting regulations, activity and success, with special reference to the mallard. U.S. Fish Wildl. Serv., Resour. Publ. 130.

Moran, P.A.P. 1952. The estimation of death-rates from capture-mark-recapture sampling. Biometrika 39(1-2):181-188.

Murton, R.K. 1966. A statistical evaluation of the effect of wood-pigeon shooting as evidenced by the recoveries of ringed birds. The Statistician 16:183-202.

*Nelson, L.J., D.R. Anderson, and K.P. Burnham. 1980. The effect of band loss on estimates of annual survival. J. Field Ornithol. 51(1):30-38.

*Nichols, J.D., and J.E. Hines. 1983. The relationship between harvest rates and survival rates of mallards: A straightforward approach with randomly partitioned data sets. J. Wildl. Manage. 47(2):334-348.

*Nichols, J.D., M.J. Conroy, D.R. Anderson, and K.P. Burnham. 1984. Compensatory mortality in waterfowl populations: A review of the evidence and implications for research and management. Trans. N. Am. Wildl. & Natur. Resour. Conf. 49:535-554.

*Nichols, J.D., J.E. Hines, and K.H. Pollock. 1984. Effects of permanent trap response in capture probability on Jolly-Seber capture-recapture model estimates. J. Wildl. Manage. 48(1):289-294.

*Nichols, J.D., B.R. Noon, S.L. Stokes, and J.E. Hines. 1981. Remarks on the use of mark-recapture methodology in estimating avian population size. Pages 121-136 *in* C.J. Ralph and M.J. Scott, eds. Estimating numbers of terrestrial birds. Stud. Avian Biol. 6.

*Nichols, J.D., and K.H. Pollock. 1983. Estimation methodology in contemporary small mammal capture-recapture studies. J. Mammal. 64(2):253-260.

*Nichols, J.D., and K.H. Pollock. 1983. Estimating taxonomic diversity, extinction rates, and speciation rates from fossil data using capture-recapture models. Paleobiology 9(2):150-163.

*Nichols, J.D., S.L. Stokes, J.E. Hines, and M.J. Conroy. 1982. Additional comments on the assumptions of homogeneous survival rates in modern bird banding estimation models. J. Wildl. Manage. 46(4):953-962.

*North, P.M., and B.J.T. Morgan. 1979. Modelling heron survival using weather data. Biometrics 35(3):667-681.

*North, P.M., and R.M. Cormack. 1981. On Seber's method for estimating age-specific bird survival rates from ringing recoveries. Biometrics 37(1):103-112.

*North, P.M. 1979. Relating Grey Heron survival rates to winter weather conditions. Bird Study 26:23-28.

Orians, G.H. 1958. A capture-recapture analysis of a shearwater population. J. Anim. Ecol. 27(1):71-86.

Ostle, B. 1963. Statistics in research. Iowa State University Press, Ames. 585 pp.

Overton, W.S. 1969. Estimating the number of animals in wildlife populations. Pages 403-455 *in* R.H. Giles, ed. Wildlife management techniques. 2nd Edition. The Wildlife Society, Wash., D.C.

Paloheimo, J.E. 1961. Studies on estimation of mortalities. I. Comparison of a method described by Beverton and Holt and a new linear formula. J. Fish. Res. Board. Can. 18:645-662.

Parker, R.A. 1963. On the estimation of population size, mortality and recruitment. Biometrics 19(2):318-323.

Paulik, G.J. 1962. Use of the Chapman Robson survival estimate for single- and multiple-release tagging experiments. Trans. Am. Fish. Soc. 91(1):95-98.

Paulik, G.J. 1963. Estimates of mortality rates from tag recoveries. Biometrics 19(1):28-57.

Pollock, K.H. 1974. The assumption of equal catchability of animals in tag-recapture experiments. PhD Thesis, Cornell University, Ithaca, NY. 82 pp.

Pollock, K.H. 1975. A K-sample tag-recapture model allowing for unequal survival and catchability. Biometrika 62 (3):577-583.

*Pollock, K.H. 1981a. Capture-recapture models: A review of current methods, assumptions, and experimental design. Pages 426-435 *in* C.J. Ralph, and J.M. Scott, eds. Estimating the numbers of terrestrial birds. Stud. Avian., Biol. 6.

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

*Pollock, K.H. 1982. A capture-recapture design robust to unequal probability of capture. J. Wildl. Manage. 46(3): 752-757.

*Pollock, K.H., J.E. Hines, and J.D. Nichols. 1985. Goodness of fit tests for open capture-recapture models. Biometrics 41(2).

*Pollock, K.H., and R.H.K. Mann. 1983. Use of an age-dependent mark-recapture model in fisheries research. Canadian J. Fish. & Aquatic Sci. 40(9):1449-1455.

*Pollock, K.H., and D.G. Raveling. 1982. Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. J. Wildl. Manage. 46(1):88-98.

Pollock, K.H., D.L. Solomon, and D.S. Robson. 1974. Tests for mortality and recruitment in a K-sample tag-recapture experiment. Biometrics 30(1):77-87.

Pope, J.A. 1963. The design and analysis of capture-recapture experiments. Pages 342-347 *in* North Atlantic Fish Marking Symposium for the International Commission for the Northwest Atlantic Fisheries. Spec. Publ. 4.

Rafail, S.Z. 1971a. Estimation of abundance of fish populations by capture-recapture experiments. Mar. Biol. 10(1): 1-7.

Rafail, S.Z.1971b. Estimation of some parameters of large fish populations by capture-recapture experiments. Mar. Biol. 10(1):8-12.

Rao, C.R. 1965. Linear statistical inference and its application. Wiley, NY. 522 pp.

Regier, H.A., and D.S. Robson. 1967. Estimating population number and mortality rates. Pages 31-66 *in* S.D. Gerking, ed. The bioligical basis of freshwater fish production. Blackwell Scientific Publ., Oxford.

Ricker, W.E. 1958. Handbook of computations for biological statistics of fish populations. Fish. Res. Board Can. Bull. No. 119. 300 pp.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Fish. Res. Board Can. Bull. 191. 382 pp.

Roberts, J.O.L. 1971. Survival among some North American wood warblers. Bird-Banding 42(3):165-184.

Robson, D.S. 1963. Maximum likelihood estimation of a sequence of annual survival rates from a capture-recapture series. Pages 330-335 *in* North Atlantic Fish Marking Symposium. International Commission for the Northwest Atlantic Fisheries. Spec. Publ. 4.

Robson, D.S. 1969. Mark-recapture methods of population estimation. Pages 120-140 *in* N.L. Johnson and H. Smith, Jr. eds. New developments in survey sampling. Wiley Interscience, NY.

Robson, D.S. 1971a. Statistical analysis of reported tag recaptures from an exploited population: maximum likelihood estimation of instantaneous mortality rates. Biometrics Unit, Cornell University, Ithaca, NY. BU-370-M. 8 pp.

Robson, D.S. 1971b. Statistical methods for a tag-recapture experiment on a population closed to recruitment: Preliminary report. Biometrics Unit, Cornell University, Ithaca, NY. BU-374-M. 10 pp.

Robson, D.S. 1971c. A non-prejudicial method of defining class intervals for comparing frequency distribution by chi-square. Biometrics Unit, Cornell University, Ithaca, NY. BU-366-M. 12 pp.

*Robson, D.S. 1983. Mark-recapture models with survival rates specific to capture-history classes: A synthesized numerical example. Cornell Biometrics Unit Rept. BU-826-M. 11 pp.

Robson, D.S., and D.G. Chapman. 1961. Catch curves and mortality rates. Trans. Am. Fish. Soc. 90(2):181-189.

Robson, D.S., and H.A. Regier. 1968. Estimation of population number and mortality rates. Pages 124-158 *in* W.E. Ricker, ed. Methods for assessment of fish production in fresh waters. IBP Handbook No. 3, Blackwell Scientific Publ., Oxford.

Robson, D.S., and W.D. Youngs. 1971. Statistical analysis of reported tag-recaptures in the harvest from an exploited population. Biometrics Unit, Cornell University, Ithaca, NY. BU-369-M. 15 pp.

*Sandland, R.L., and P. Kirkwood. 1981. Estimation of survival in marked populations with possibly dependent sighting probabilities. Biometrika 68:531-541.

Seber, G.A.F. 1962. The multi-sample single recapture census. Biometrika 49(3-4):339-350.

Seber, G.A.F. 1965. A note on the multiple-recapture census. Biometrika 52(1-2):249-259.

Seber, G.A.F. 1966. The linear hypothesis: A general theory. Hafner Publ. Co., NY. 115 pp.

Seber, G.A.F. 1970. Estimating time-specific survival and reporting rates for adult birds from band returns. Biometrika 57(2):313-318.

Seber, G.A.F. 1971. Estimating age-specific survival rates from bird-band returns when the reporting rate is constant. Biometrika 58(3):491-497.

Seber, G.A.F. 1972. Estimating survival rates from bird-band returns. J. Wildl. Manage. 36(2):405-413.

Seber, G.A.F. 1973. Estimation of animal abundance and related parameters. Griffin, London. 506 pp.

*Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. Macmillan Publ. Co., NY. 2nd Edition. 654 pp.

Seber, G.A.F., and E.D. Le Cren. 1967. Estimating population parameters from catches large relative to the population. J. Anim. Ecol. 36(4):631-643.

*Stokes, S.L. 1984. The Jolly-Seber method applied to age-stratified populations. J. Wildl. Manage. 48(3):1053-1059. (also see the corrigendum in J. Wildl. Manage. 49(1):282)

Vaught, R.W., and L.M. Kirsch. 1966. Canada geese of the Eastern Prairie Population, with special reference to the Swan Lake flock. Mo. Dep. Conserv. Tech. Bull. 3. 91 pp.

*Weatherhead, P.J., and C.D. Ankney. 1984. Comment: A critical assumption of band-recovery models may often be violated. Wildl. Soc. Bull. 12:198-199.

*Weatherhead, P.J., and C.D. Ankney. 1985. Condition bais and band-recovery data: A reply to Burnham and Nichols. Wildl. Soc. Bull. 13(3):

White, E.G. 1971a. A computer program for capture-recapture studies of animal populations: A FORTRAN listing for the stochastic model of G.M. Jolly. Tussock Grasslands and Mountain Lands Institute. Spec. Publ. 8. 33 pp. (also a subsequent Addendum, xiii p.)

White, E.G. 1971b. A versatile FORTRAN computer program for the capture-recapture stochastic model of G.M. Jolly. J. Fish. Res. Board Can. 28:443-445.

White, E.G. 1975. Identifying population units that comply with capture-recapture assumptions in an open community of alpine grasshoppers. Res. Popul. Ecol. 16:153-187.

*White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. J. Wildl. Manage. 47(3):716-728.

*White, G.C., D.R. Anderson, K.P. Burnham, and D.L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. LA-8787-NERP. Los Alamos National Laboratory, Los Alamos, NM. 235 pp.

*Youngs, W.D., and D.S. Robson. 1975. Estimating survival rate from tag returns: Model tests and sample size determination. J. Fish. Res. Board Can. 32:2365-2371.

Appendices

As the authors pointed out in the Preface, the mathematical models and associated analysis procedures presented in this handbook are new. The first such modern model for analysis of banding data was published by Seber (1970); indeed this work spurred the developments presented here (also see Seber 1962). Robson and Youngs (1971) independently developed the model examined by Seber and extended the analysis procedures. We have referred to the Seber-Robson-Youngs model in Chapter 2 as Model 1. With the exception of Model 1, the extensive models, analyses, and testing procedures presented herein are new to the published literature. The purpose of these appendices is to document the basic references for the technical mathematical development of the models and analysis procedures presented in this handbook.

Appendices A and B are intended for the mathematical statistician who wishes to understand the theory behind the methods presented in the handbook. This theory exists in only a few published sources, the most important one being the doctoral thesis of Brownie (1973) (cf. Brownie and Robson 1976). Appendix A presents, by chapter and model, where the background theory may be found. In those instances where no literature exists on the theory, it is so stated and we sketch the underlying theory. We hope that the mathematically trained reader will be able to fill in the gaps. Appendix B presents information on several points of general methodology useful in these models and methods of analysis. For example, there is a convenient method of deriving the ML estimators of survival and recovery rates for all full-rank models which is much easier than the standard approach of taking partial derivatives of the log-likelihood. We assume the reader has read and understands Section 1.5, as well as any chapters for which greater detail is sought.

Appendix C presents 10 papers specifically related to the subject of this *Handbook* that have appeared since the first edition. They have been reprinted in an effort to allow the second edition to be up-to-date.

Appendix A: Basic Mathematical Background and References for the Specific Models Presented

Chapter 2

Models

Four models for the analysis of recovery data from individuals banded as adults are presented in Chapter 2. They are referred to here as Models 0, 1, 2, and 3 in decreasing order of generality. These models are nested in the sense that if Θ_i is the parameter space for model i, then Θ_{i+1} is a proper subspace of Θ_i, $i=0, 1, 2$.

For k years of bandings and $\ell = k+s$ years of recovery data, the cell probabilities for Model 0 are

$$\pi_{ij} = \begin{cases} f_i^* & , i=j \\ S_i \cdots S_{j-1} f_j & , i<j , \end{cases}$$

for $i=1,\ldots,k, j=i,\ldots,k+s$. The parameter space is

$$\Theta_0 = \{f_1^*, \ldots, f_k^*, f_2, \ldots, f_{k+s}, S_1, \ldots, S_{k+s-1}\} ,$$

with dimensionality $3k+2s-2$. Model 1 is obtained from Model 0 by making the assumption (restriction) that $f_i^* = f_i, i=1,\ldots,k$ Thus the parameter space for Model 1 is

$$\Theta_1 = \{f_1, \ldots, f_{k+s}, S_1, \ldots, S_{k+s-1}\} ,$$

giving $2k+2s-1$ parameters. To derive Model 2 from Model 1, assume equality of annual survival rates: $S_i = S$, $i=1,\ldots,k+s-1$.
Thus

$$\Theta_2 = \{f_1, \ldots, f_{k+s}, S\} ,$$

has $k+s+1$ parameters. Model 3 is the most restrictive possible as it assumes the restriction $f_i \equiv f, i=1,\ldots,k+s$ on Model 2. Hence, the parameter space for Model 3 is

$$\Theta_3 = \{f, S\} ,$$

with only 2 parameters.

It is seen that the parameter spaces are nested as

$$\Theta_3 \subset \Theta_2 \subset \Theta_1 \subset \Theta_0 .$$

Statistical inference under Model 1 is well documented by Seber (1970) and Robson and Youngs (1971). The latter reference (available from Cornell University, Biometrics Unit) covers the goodness of fit test for Model 1 and presents additional covariance estimators not given in Seber's paper.

Model 0 is developed in Brownie (1974b). This reference, also available from Cornell University, is the only reference to the mathematical developments of Model 0. It is comprehensive, however, covering parameter estimation, confidence intervals, and the goodness of fit for Model 0. This reference also covers the test of Model 1 vs. Model 0. As indicated above, this is a test of the null hypothesis that $f_i^* = f_i$, $i=2,\ldots,k$. Finally, Brownie (1974b) gives the relationships of the three tests: goodness of fit tests of Models 1 and 0 and the test of Model 1 vs. Model 0.

Model 2 was developed in the course of work by the authors. There is no reference to its development. Model 3 is not exactly new in that numerous authors have dealt with this case (model) of constant annual recovery and survival rates. For a review, see Seber (1973:245-246). However, no one has previously derived the exact ML estimators for this model, nor tested any of the assumptions it makes (e.g., a goodness of fit test or the test against Models 2 and 1).

The ML estimators under Models 2 and 3 do not exist in closed form. The equations to be solved are easily found in the usual manner by taking partial derivatives of the log-likelihood. For Model 2 the likelihood is (using $\underline{f} = [f_1, \ldots, f_\ell]$)

$$\mathscr{L}_2(\underline{f}, S) = \left[\prod_{i=1}^{\ell} (f_i^{C_i}) \right] S^Q \prod_{i=1}^{k} (\pi_{i,\ell+1})^{N_i - R_i} ,$$

where

$$R_i = \sum_{j=i}^{\ell} R_{ij} \qquad , i = 1, \ldots, k \text{ (row total)}$$

$$C_i = \sum_{h=1}^{\min\{i,\,k\}} R_{hi} \qquad , i = 1, \ldots, \ell \text{ (column total)}$$

$$Q = \sum_{i=1}^{k} \sum_{j=i}^{\ell} (j-i)\, R_{ij} \ ,$$

and

$$1 - \pi_{i,\ell+1} = \sum_{j=i}^{\ell} \pi_{ij} = f_i + S f_{i+1} + S^2 f_{i+2} + \cdots + S^{\ell-i} f_\ell \ .$$

From the above the reader can easily write the likelihood for Model 3 (also note it is given in Section 1.5). We assume the reader can write down the corresponding log-likelihoods and take their partial derivatives. For example, under Model 2:

$$\frac{\partial \ell n \mathscr{L}_2(\underline{\hat{f}},\underline{\hat{S}})}{\partial f_h} = \frac{C_h}{f_h} - \sum_{i=1}^{\min\{h,k\}} \left(\frac{N_i - R_i}{\pi_{i,\ell+1}}\right) S^{h-i} \qquad , h = 1, \ldots, \ell \ .$$

There is one general numerical method used in program ESTIMATE to solve the ML equations for Models 2 and 3; we outline this method in Appendix B.

In the cases where closed-form estimators do not exist the variance-covariance matrices of estimators were derived by inverting (numerically, except for Model 3) the Fisher information matrix under the respective models. We assume the reader can derive the elements of these matrices if they so desire.

Test Statistics

The goodness of fit tests for Models 2 and 3 are standard procedures which we feel are satisfactorily explained in the text. The tests between models (i.e., Model 3 vs. Model 2, Model 3 vs. Model 1, and Model 2 vs. Model 1) computed by ESTIMATE are all likelihood ratio tests. The test of the null hypothesis $S_i \equiv S$, $i = 1, \ldots, \ell - 1$ is a test of Model 2 vs. Model 1. The test statistic is

$$2\lambda_{12} = 2[\ell n \mathscr{L}_1(\underline{\hat{f}},\underline{\hat{S}}) - \ell n \mathscr{L}_2(\underline{\hat{f}},\hat{S})] \ .$$

In this notation $\mathscr{L}_1(\hat{f},\underline{\hat{S}})$ is the log-likelihood function under Model 1 evaluated at the ML estimators under Model 1. Similarly, $\mathscr{L}_2$ is the log-likelihood under Model 2 evaluated at the ML estimators computed under Model 2. If the null hypothesis is true, $-2\lambda_{12}$ is asymptotically distributed as a central chi-square statistic with $(2\ell - 1) - (\ell + 1) = \ell - 2$ degrees of freedom. The other test statistics are representable as

$$2\lambda_{23} = 2[\ell n \mathscr{L}_2(\underline{\hat{f}},\hat{S}) - \ell n \mathscr{L}_3(\hat{f},\hat{S})] \ ,$$

and $2\lambda_{13} = (2\lambda_{12} + 2\lambda_{23})$, where $2\lambda_{23}$ tests Model 2 vs. Model 3 (i.e., H_0: $f_i = f$) and $2\lambda_{13}$ tests Model 1 vs. Model 3. The interested reader should be able to write down the explicit form of $\ell n \mathscr{L}_i$.

Mean Life Span

Under all four models of Chapter 2 an estimate of adult mean life span is computed. Though there are references to this parameter in the literature (e.g., Cormack 1964), it is not very well explained. A general derivation is given here. For a bird alive at time t_0 ($\equiv 0$) its expected additional lifetime or mean life span (MLS) depends upon the survival curve for that species from t_0 on. In general this survival curve may depend upon age, sex, and calendar year. Let $S(0,t)$ be the probability of survival until time t. Without loss of generality we can write

$$S(0,t) = e^{-\int_0^t p(\tau)d\tau}$$

where $p(\tau)$ is the instantaneous mortality rate. If T = life time (a random variable), then the cdf of T satisfies

$$1 - F(t) = S(0,t) = e^{-\int_0^t p(\tau)d\tau}$$

Hence, the pdf of T is

$$f(t)=p(t)\,e^{-\int_0^t p(\tau)d\tau} \qquad ,0<t$$

and

$$\mathbf{E}(T)=\int_0^\infty t\,p(t)e^{-\int_0^t p(\tau)d\tau}\,dt\ . \tag{A1}$$

Without assuming something about $p(\tau)$ we cannot simplify this formula. The simplest case is $p(\tau)\equiv p$ which implies annual survival rate is constant: $S(0,1)=e^{-p}\equiv S$. Though this implies Model 3 or 2, it is not necessary in those models to have constant instantaneous mortality. The above assumption leads to $\mathbf{E}(T)=\text{MLS}=1/p$, and $p=-\ell n(S)$, or

$$\text{MLS}=-\frac{1}{\ell n(S)}\ . \tag{A2}$$

Formula $(A2)$ is the basis for the estimators of MLS used in program ESTIMATE. These estimators of MLS are properly considered as approximations; we feel they are useful enough to be computed, using either $\hat{S}$ from Models 2 and 3, or $\tilde{S}$ from Models 1 and 0 in place of S.

Another useful concept is "half-life" or the time period from banding of adults until we expect half the birds to be dead. This was mentioned in the example of Section 2.2. Under the assumption of constant instantaneous mortality rate $p(\tau)\equiv p$ (which implies Model 2 or 3), lifetime T is an exponential random variable with cdf

$$F(t)=1-e^{-pt}\ .$$

Hence, the time $t_{1/2}$ such that half the birds are dead and half alive, is given by

$$0.5=1-\exp(-pt_{1/2}).$$

This equation is easily solved for

$$t_{1/2}=\frac{-\ell n(0.5)}{p}=(0.69)\ \text{MLS}.$$

Thus the adult half-life is conveniently computed as 0.69 times the mean life span.

Chapter 3

All the material in Chapter 3 is new, having been developed within the past 4 years; the primary reference is the doctoral thesis by Brownie (1973).

Models

Five models are presented in Chapter 3; in order of decreasing complexity these are H_3, H_2, H_1, H_{02}, and H_{01}. These five models are nested in the same sense as those of Chapter 2. The model structure (there are two arrays, R_{ij} for adults and Q_{ij} for young) for H_3 is representable as

$$\mathbf{E}\left(\frac{R_{ij}}{N_i}\right)=\pi_{ij}=\begin{cases} f_i''' & ,j=i \\ S_i'''\,f_{i+1} & ,j=i+1 \\ S_i'''\,S_{i+1}\cdots S_{j-1}f_j & ,j=i+2,\ldots,\ell \end{cases}$$

$$\mathbf{E}\left(\frac{Q_{ij}}{M_i}\right)=\pi'_{ij}=\begin{cases} f_i' & ,j=i \\ S_i'\,f''_{i+1} & ,j=i+1 \\ S_i'\,S''_{i+1}S_{i+2}\cdots S_{j-1}f_j & ,j=i+2,\ldots,\ell \end{cases}$$

(see also Table 3.7). Because H_3 assumes three age classes, it has quite a few parameters, with the exact number difficult to write unless the value of ℓ (i.e., $\ell=k$, $k+1$, or $\ell>k+1$) is specified. For $\ell=k$ the parameter space of Model H_3 is $\theta_3=\{f_1''',\ldots,f_k''',f_2'',\ldots,f_k'',f_1',\ldots,f_k',f_2,\ldots,f_k,S_1''',\ldots,S_{k-1}''',S_2'',\ldots,S_{k-1}'',S_1',\ldots,S_{k-1}',\ S_2,\ldots,S_{k-1}\}$, which has $8(k-1)$ parameters.

Model H_2 is derived from H_3 by assuming the null hypothesis $S_i'''=S_i''=S_i$ and $f_i''=f_i$, where the range of i depends upon the value of ℓ.

Further restricting Model H_2 by assuming $f_i''' = f_i$ gives Model H_1 with the structure

$$\pi_{ij} = \begin{cases} f_i & ,j = i \\ S_i S_{i+1} \cdots S_{j-1} f_j & ,j > i \ . \end{cases}$$

$$\pi'_{ij} = \begin{cases} f_i' & ,j = i \\ S_i' S_{i+1} \cdots S_{j-1} f_j & ,j > i \ . \end{cases}$$

The parameter space for Model H_2 when $\ell = k$ is $\Theta_2 = \{f_1''', \ldots, f_k''', f_1', \ldots, f_k', f_2, \ldots, f_k, S_1', \ldots, S_{k-1}', S_1, \ldots, S_{k-1}\}$ with $5k-3$ parameters. The parameter space for Model H_1 when $\ell = k$ is $\Theta_1 = \{f_1', \ldots, f_k', f_1, \ldots, f_k, S_1', \ldots, S_{k-1}', S_1, \ldots, S_{k-1}\}$, with $4k-2$ parameters.

Model H_{02} is derived from H_1 by assuming constant annual survival rates: $S_i' = S'$, and $S_i = S$. The model structure of H_{02} is

$$\pi_{ij} = S^{j-i} f_j \qquad ,j = i, \ldots, \ell, \ i = 1, \ldots, k$$

$$\pi_{\;j} = \begin{cases} f_i' & ,i = j \\ S' S^{j-i-1} f_i & ,j \geqslant i+1 \end{cases}$$

with parameter space $\Theta_{02} = \{f_1', \ldots, f_k', f_1, \ldots, f_\ell \ , S', S\}$.

Finally, restricting H_{02} by assuming time-constant, but age-specific, recovery rates gives Model H_{01}: $f_i \equiv f, f_i' \equiv f'$. This simplest model has the following structure

$$\pi_{ij} = S^{j-i} f$$

$$\pi'_{ij} = \begin{cases} f' & ,j = i \\ S' S^{j-i-1} f & ,j \geqslant i+1 \ . \end{cases}$$

The parameter space is simply $\Theta_{01} = \{f', f, S', S\}$.

The parameter spaces of these five models are nested as $\Theta_{01} \subset \Theta_{02} \subset \Theta_1 \subset \Theta_2 \subset \Theta_3$. The best way for the reader to become familiar with these models is to write out the model structures, in detail for specific k and ℓ (e.g., $k = \ell = 4$ and $k = 3$, $\ell = 5$), and determine the exact number of parameters in each model.

There is an analogy between the one-age-class (adult) and two-age-class (adult and young) models with respect to the assumptions made about time specificity of parameters. This is diagramed below (models are presented in the order, restrictive to general):

Adults	*Adults and Young*
Model 3	H_{01}
Model 2	H_{02}
Model 1	H_1
Model 0	H_2
	H_3

Just as Model H_3 represents a three-age-class case, Model M_0 can represent a two-age-class case (see Section 2.5). Note that in program BROWNIE Model M_1 is referred to as Model H_0.

The development of Models H_1, H_2, and H_3, all estimators, covariances, tests between models, and goodness of fit tests are thoroughly dealt with by Brownie (1973) and Brownie and Robson (1976). We therefore make no further reference to these models.

Models H_{01} and H_{02} Likelihood Functions

The development of Models H_{01} and H_{02}, and associated inference procedures, is not documented anywhere. Consequently, we give the likelihood under each model (denoted $\mathscr{L}_{01}$ and $\mathscr{L}_{02}$):

$$\mathscr{L}_{01}(f', f, S', S) = (f')^A (S')^{(TQ-A)} f^{\,(TR+TQ-A)} S^B \prod_{i=1}^{k} (\pi_{i,\ell+1})^{N_i - R_i.} \prod_{i=1}^{k} (\pi'_{i,\ell+1})^{M_i - Q_i.}$$

where

$$A = \sum_{i=1}^{k} Q_{ii}$$

$$TQ = \sum_{i=1}^{k} \sum_{j=i}^{\ell} Q_{ij}$$

$$TR = \sum_{i=1}^{k} \sum_{j=i}^{\ell} R_{ij}$$

$$B = \sum_{i=1}^{k} \sum_{j=i+1}^{k} (j-i)R_{ij} + \sum_{i=1}^{k} \sum_{j=i+2}^{\ell} (j-i-1)Q_{ij}$$

$$\pi_{i,\ell+1} = 1 - f\left(\frac{1-S^{\ell-i+1}}{1-S}\right)$$

$$\pi'_{i,\ell+1} = \begin{cases} 1-f' & ,i=k=\ell \\ 1-f'-S'f\left(\frac{1-S^{\ell-i}}{1-S}\right) & ,\text{otherwise.} \end{cases}$$

Given the above, the reader can determine the log-likelihood function and its derivatives (specifically the likelihood equations). For example

$$\frac{\partial \ell n \mathcal{L}_{01}}{\partial f'} = \frac{A}{f'} - \sum_{i=1}^{k} \left(\frac{M_i - Q_{i.}}{\pi'_{i,\ell+1}}\right).$$

The likelihood under Model H_{02} is

$$\mathcal{L}_{02}(f'_1,\ldots,f'_k,f_1,\ldots,f_\ell,S',S) = \prod_{i=1}^{k} (f'_i)^{Q_{ii}} (S')^{(TQ-A)} \; f_1^{R_{11}} \prod_{j=2}^{k} (f_j)^{(R_{.j}+Q_{.j}-Q_{jj})} \prod_{j=k+1}^{\ell} (f_j)^{(R_{.j}+Q_{.j})} S^B$$
$$\prod_{i=1}^{k} (\pi_{i,\ell+1})^{N_i - R_{i.}} \prod_{i=1}^{k} (\pi'_{i,\ell+1})^{M_i - Q_{i.}}$$

In the above

$$\pi_{i,\ell+1} = 1 - (f_i + Sf_{i+1} + S^2 f_{i+2} + \cdots + S^{\ell-i} f_\ell)$$

$$\pi'_{i,\ell+1} = \begin{cases} 1-f'_k & ,\ i=k=\ell \\ 1-(f_i + S'f_{i+1} + S'Sf_{i+2} + \cdots + S'S^{\ell-i-1} f_\ell) & ,\ \text{otherwise.} \end{cases}$$

From $\mathcal{L}_{02}$ we can derive the likelihood equations (and the information matrix), for example

$$\frac{\partial \ell n \mathcal{L}_{02}}{\partial f'_h} = \frac{Q_{hh}}{f'_h} - \sum_{i=1}^{\min(h,k)} \left(\frac{M_i - Q_{i.}}{\pi'_{i,\ell+1}}\right) \qquad ,\ h=1,\ldots,k.$$

For Models H_{01} and H_{02}, the log-likelihood equations are solved for the ML estimators by program BROWNIE, using the method of scoring (see Appendix B). The tests among Models H_{01}, H_{02}, and H_1 are all likelihood ratio tests. For example, the test statistic

$$\lambda_{01,02} = -2(\ell n \mathcal{L}_{01} - \ell n \mathcal{L}_{02}) \ ,$$

with the likelihoods evaluated at the respective ML estimates, tests Model H_{01} vs. the alternative of Model H_{02}. The null hypothesis for this test is $f'_i \equiv f'$, $i=1,\ldots,k$, and $f_i \equiv f$, $i=1,\ldots,\ell$. The test statistic

$$\lambda_{02,1} = -2(\ell n \mathcal{L}_{02} - \ell n \mathcal{L}_1)$$

tests Model H_{02} vs. the alternative of Model H_1 (the reader should be able to write down the likelihood function for Model H_1). Finally, the test of the H_{01} vs. H_1 model is based on the statistic

$$\lambda_{01,1} = (\lambda_{01,02} + \lambda_{02,1}) \ .$$

Under the null hypothesis, each test statistic is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the difference in the number of parameters between the more general and less general model. For example, there are $(2k+s+2)-4 = 2k+s-2$ degrees of freedom for $\lambda_{01,02}$, and $(4k+2s-2)-(2k+s+2) = 2k+s-4$ degrees of freedom for $\lambda_{02,1}$.

The material presented in Section 3.9 is new to the literature; there are no other similar references.

Mean Life Span

No mean life span statistics are computed by program BROWNIE. Anderson (1975) gives a formula for MLS of young and applies it to mallard data. Again the formula is derived by assuming constant instantaneous mortality for adults and young, but the two age classes have different rates. It is assumed a bird is "young" for the 1 year from time of banding until the first banding anniversary. Hence

$$p(t) = \begin{cases} p' & , 0 \leq t < 1 \\ p & , 1 \leq t \end{cases} ,$$

with

$$S' = e^{-p'},$$

$$S = e^{-p}.$$

Making these substitutions in formula *(A1)* gives, for young, a MLS formula analogous to that developed for adults:

$$\mathbf{E}(T) = \int_0^1 (t\,p'\,e^{-tp'})\mathrm{d}t + \int_1^\infty (t\,p\,e^{-(p' + (t-1)p)})\mathrm{d}t = \frac{1}{-\ell n(S')} + \frac{S'}{-\ell n(S)} + \frac{S'}{\ell n(S')} .$$

Chapters 4, 5 and 7

All the material in Chapters 4 and 7, and the test of male vs. female parameters of Chapter 5, is developed in Brownie (1973). There is no other reference for any of this material. The interested reader who comprehends the material of Chapters 2 and 3 should have no trouble understanding the mathematical background of Models H_4, H_5 and H_6 (Chapter 4 of this handbook) and Models H_7 and H_8 (Chapter 7) from Brownie (1973).

The material of Chapter 5 is largely new and only the test of adult males vs. adult females appears in Brownie (1973). The other tests (males vs. females for young and adults, and extension of both these "sex" tests to r areas) were developed explicitly for this handbook. Consequently, the relevant derivations are sketched in Section 5.3. There are no other references dealing with these tests.

Chapters 8 and 9

The materials of Chapters 8 and 9 were also developed entirely in the course of this handbook; consequently, there are no references elaborating on any of this material with the one exception of Section 8.6. In recognition of this, the authors have given enough of the background for mathematical statisticians to understand these sections. Specifically, this is true of Sections 8.2, 8.3, 9.3, and 9.4. While there is no particular mathematical background to Section 8.4, the interested reader should refer to Anderson and Burnham (1976) for more elaboration on the proper interpretation of sampling correlations.

Section 8.6 deals with making direct tests of the equality of survival rates (the same sort of tests can be made for equality of recovery rates). The basic concepts are quite familiar to statisticians in terms of "contrasts" in the context of general linear hypothesis theory (cf. Seber 1966:29).

From the general theory of maximum likelihood estimation we have the parameter estimator vector $\underline{\hat{\theta}}$ is asymptotically distributed as a multivariate normal random variable. Moreover we have an estimator, $\hat{\Sigma}$ of the dispersion matrix of $\underline{\hat{\theta}}$. Consequently for any vector of constants $\underline{c}$, we have $\underline{c}'\,\underline{\hat{\theta}} = \hat{\delta}$ is asymptotically normal with mean $\underline{c}'\,\underline{\theta} = \delta$, and variance estimator

$$\mathrm{var}(\hat{\delta}) = \underline{c}'\,\hat{\Sigma}\,\underline{c} .$$

A test of the null hypothesis H_0: $\underline{c}'\,\underline{\theta} = \delta_o$, is thus derived from

$$z = \frac{\underline{c}'\,\underline{\hat{\theta}} - \delta_o}{\sqrt{\underline{c}'\,\hat{\Sigma}\,\underline{c}}} ,$$

where under this null hypothesis z is asymptotically normal (0,1). Both two-sided and one-sided alternatives can be tested.

For testing equality of annual survival rates, as from Model 1, we have $\underline{\hat{\theta}}' = (\tilde{S}_1, \ldots, \tilde{S}_{k-1})$, the elements of $\hat{\Sigma}$ are easily obtained from the computer printout, and the specific hypothesis of interest will dictate the constants $\underline{c}$.

For example, if $k-1=7$, and we know the first 3 years were ones of restrictive hunting while the last 4 were liberal regulation years, we might test the null hypothesis of equal average survival rates in these two groups of years:

$$\frac{S_1+S_2+S_3}{3}=\frac{S_4+S_5+S_6+S_7}{4}\ .$$

Hence

$$\underline{c}' = (1/3,\ 1/3,\ 1/3, -1/4, -1/4, -1/4, -1/4).$$

Such tests can be used in all the models and across data sets as well as within data sets. For example, if $\tilde{\bar{S}}_f$ and $\tilde{\bar{S}}_m$ are estimated average male and female survival rates, respectively, from the same population over the same period of years, then to test H_0: $\bar{S}_f \geq \bar{S}_m$ vs. H_a: $\bar{S}_f < \bar{S}_m$ we use

$$z=\frac{(\tilde{\bar{S}}_m-\tilde{\bar{S}}_f)}{\sqrt{\operatorname{var}(\tilde{\bar{S}}_m)+\operatorname{var}(\tilde{\bar{S}}_f)}},$$

and reject H_0 if z exceeds 1.645, for a 5% level test. When the estimators derive from different data sets, which they would in the above example, they are uncorrelated; hence, no covariance terms enter.

The only reference to this type of test for equality of survival rates using banding data is Brownie and Robson (1974b); their paper deals mainly with the optimality and power of these tests under Model 1.

Appendix B: Some General Methodology Underlying Analysis Methods Presented

All of the statistical theory applied to the various models presented here is standard (i.e., Maximum Likelihood estimation, likelihood ratio tests, chi-square goodness of fit tests). There are, however, a few general points of methodology that we wish to point out to the reader interested in the statistical theory behind these applications. These points concern primarily derivation of estimators and their sampling variances.

As was stated in Section 1.5, all models presented here are products of multinomial distributions. Moreover, all 14 models are in the exponential family of probability distributions. Two distinct situations arise with respect to estimation: the full-rank case and the less-than-full-rank case. In the full-rank case the dimensionality of the minimal sufficient statistic (MSS) is the same as that of the parameter space. The alternative is a parameter space of lower dimensionality than the MSS; in this case closed-form estimators usually do not exist. In the full-rank case there is a very useful shortcut to deriving both point estimators and their large sample variance covariance matrix (i.e., the inverse of the information matrix).

The Full-rank Case

Point Estimation

Models M_1, M_0 and H_1 through H_8 of this handbook are full rank. For these models the ML estimators may be derived by equating the observed MSS to its (vector valued) expectation and solving for the unknown parameters. By an appropriate choice of an MSS, this approach can achieve great simplicity over the "usual" approach of deriving and solving the likelihood equations (the partial derivatives of the log-likelihood). This procedure is proven by Davidson and Solomon (1974).

Symbolically, if $Z_1,\ldots,Z_r$ is the MSS from an exponential family with expectations $\mathbf{E}(Z_i)=g_i(\underline{\theta})$, where $\underline{\theta}=(\theta_1,\ldots,\theta_r)'$, then the ML estimator is that value $\hat{\underline{\theta}}$ which satisfies

$$Z_i=g_i(\hat{\underline{\theta}}) \qquad ,i=1,\ldots,r.$$

For all the applicable (full-rank) models of this handbook both the MSS and its expectation are relatively easy to find. For example, for the case of adults only, under Model M_1 (see Appendix A and Section 2.2), with $k=\ell$, a MSS is $\{R_1,\ldots,R_k,C_1,\ldots,C_{k-1}\}$. Its dimensionality is $2k-1$. The parameter space is $\{f_1,\ldots,f_k,S_1,\ldots,S_{k-1}\}$.

In terms of previously used notation we have

$$\mathbf{E}(R_i)=N_i\,[f_i+S_if_{i+1}+\cdots+S_iS_{i+1}\cdots S_{k-1}f_k]=N_i\rho_i$$
$$\mathbf{E}(C_i)=N_1S_1\cdots S_{i-1}f_i+N_2S_2\cdots S_{i-1}f_i+\cdots+N_if_i=\xi_if_i.$$

One could equate R_i and C_i to their expectations and solve for the ML estimators. A little cleverness will, however, make matters easier to the point of being simple. Note that given the above MSS we can compute the quantities $T_1,\ldots,T_k$ because $T_1\equiv R_1$, and

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

Thus, not only do the ML estimators satisfy $R_i=\mathbf{E}(R_i)=N_i\rho_i$, and $C_i=\mathbf{E}(C_i)=\xi_if_i$, but they also must satisfy $T_i=\mathbf{E}(T_i)$. With minimum difficulty we can find

$$\mathbf{E}(T_i)=\xi_i\rho_i\,.$$

Now, using this "method of expectation," we can quickly find

$$\hat{\rho}_i=\frac{R_i}{N_i}$$

$$\hat{f}_i=\frac{C_i}{T_i}\frac{R_i}{N_i}.$$

Then noting that $\rho_i = f_i + S_i\rho_{i+1}$, $i=1,\ldots,k-1$, we can solve for $\hat{S}_i$:

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

This approach is much easier than attempting to determine the $2k-1$ likelihood equations by brute force partial differentiation of the Model M_1 log-likelihood function. Admittedly it takes some trial and error to find the simplest MSS to use, but once the reader develops a familiarity with the full-rank models presented here, there should be no great difficulty deriving the ML estimators.

Because of the obvious importance of the point estimators, we will give a second example using the two-age-class model H_1 with $k=\ell$. The elements of the MSS are (see Section 3.1) $R_{i\cdot}$, $R_{\cdot i}$, $Q_{i\cdot}$, $Q_{\cdot i}$, Q_{ii} as well as derived values $T_i + U_i - Q_{i\cdot}$. Relevant expectations are

$$\begin{aligned}
\mathbf{E}(R_{i\cdot}) &= N_i\rho_i \\
\mathbf{E}(R_{\cdot i}) &= \xi_i f_i \\
\mathbf{E}(Q_{i\cdot}) &= M_i(f_i' + S_i'\rho_{i+1}) \qquad , i=1,\ldots,k-1 \\
\mathbf{E}(Q_{\cdot i}) &= \xi_i' f_i + M_i f_i' \\
\mathbf{E}(Q_{ii}) &= M_i f_i' \\
\mathbf{E}(T_i + U_i - Q_{i\cdot}) &= \xi_i\rho_i + \xi_i'\rho_i
\end{aligned}$$

It is also useful to define $W_i = R_{\cdot i} + Q_{\cdot i} - Q_{ii}$, then

$$\mathbf{E}(W_i) = \xi_i f_i + \xi_i' f_i \quad .$$

Using these various equations, it is not difficult to derive, by the method of expectation, the ML estimators:

$$\begin{aligned}
\hat{f}_i' &= \frac{Q_{ii}}{M_{ii}} \ , \\
\hat{f}_i &= \frac{R_{i\cdot}}{N_i}\left(\frac{W_i}{T_i + U_i - Q_{i\cdot}}\right) \ , \\
\hat{S}_i' &= \frac{Q_{i\cdot} - Q_{ii}}{M_i}\left(\frac{N_{i+1}}{R_{i+1\cdot}}\right) \\
\hat{S}_i &= \frac{R_{i\cdot}}{N_i}\left(\frac{T_i + U_i - Q_{i\cdot} - W_i}{T_i + U_i - Q_{i\cdot}}\right)\frac{N_{i+1}}{R_{i+1\cdot}} \quad .
\end{aligned}$$

This approach to deriving the ML estimators in the full-rank models is *much* easier than the standard approach of solving the likelihood equations (which are even hard to write down due to their complexity). It does require determining the MSS (cf. Kendall and Stuart 1961, Vol. 2: 193-195) and then for ease of solution some examination of alternative versions of the MSS is recommended.

Variance Estimation

Under Maximum Likelihood theory one usually derives the (asymptotic) variance-covariance matrix as the inverse of the Fisher information matrix. However, direct computation of that matrix requires computing mixed second partial derivatives of the log-likelihood function. Clearly this is not feasible if first partials are not even available. Seber (1970) and Brownie (1973) used the alternative Taylor series approach (or delta method—cf. Seber 1973). In the somewhat general terms used above, let the estimator be derived from the equations

$$\underline{Z} = \mathbf{E}(\underline{Z}) = \underline{g}(\underline{\theta}),$$

where $\underline{Z} = (Z_1,\ldots,Z_r)'$ and $\underline{g}(\underline{\theta}) = (g_1(\underline{\theta}),\ldots,g_r(\underline{\theta}))'$. Assuming a unique solution then we can write

$$\hat{\underline{\theta}} = \underline{g}^{-1}(\underline{Z})$$

and the asymptotic variance-covariance matrix of $\hat{\underline{\theta}}$ is given by

$$V(\hat{\underline{\theta}}) = A\,\Sigma A'$$

where Σ is the r by r variance covariance matrix of $\underline{Z}$ and

$$A = \left[\frac{\partial \underline{g}^{-1}(\underline{Z})}{\partial \underline{Z}}\bigg|_{\underline{Z}=\mathbf{E}(\underline{Z})}\right].$$

Because this method produces an asymptotically valid result, one would think it ought to be the same as, or similar to, the variance-covariance matrix of $\hat{\underset{\sim}{\theta}}$ generated as the inverse of the information matrix. In fact, $V(\hat{\underset{\sim}{\theta}})$ computed this way is *identical* to the inverse of the information matrix in this full-rank exponential family case (we do not know of a published proof of this, but we have such proof on file). It follows that not only are ML estimators given in this handbook, but their variances and covariances are identical to what would be derived by the information matrix approach.

Bias of estimators

Given closed-form formula for parameter estimators, it was possible to evaluate their expectations, hence bias. For example, under Model M_1 the ML estimator of S_i is biased:

$$\begin{aligned}\mathbf{E}(\hat{S}_i) &= \mathbf{E}\left\{\frac{R_i}{N_i}\left(1-\frac{T_i}{C_i}\right)\left(\frac{N_{i+1}}{R_{i+1}}\right)\right\}\\ &= \left\{\mathbf{E}\left(\frac{R_i}{N_i}\right)-\mathbf{E}(\hat{f}_i)\right\}\mathbf{E}\left(\frac{N_{i+1}}{R_{i+1}}\right)\\ &= S_i\rho_{i+1}\mathbf{E}\left(\frac{N_{i+1}}{R_{i+1}}\right),\end{aligned}$$

(from Robson and Youngs (1971), $\hat{f}_i$ is unbiased).

Note that $\rho_{i+1}=\mathbf{E}\left(\frac{R_{i+1}}{N_{i+1}}\right)$, hence we can write

$$\mathbf{E}(\hat{S}_i)=S_i\mathbf{E}(R_{i+1})\mathbf{E}\left(\frac{1}{R_{i+1}}\right)>S_i\ .$$

Strictly speaking $\mathbf{E}(1/R_{i+1})$ is infinite, so we interpreted it as the conditional expectation given R_{i+1} greater than zero. Because R_{i+1} is binomial (N_{i+1}, ρ_{i+1}) it is easy to investigate the bias of $\hat{S}_i$. We did so and found relative bias will run around 1% to 5% for many realistic values of N_{i+1} and ρ_{i+1}. Moreover, asymptotically the bias is of order $1/N_{i+1}$. Specifically, in the limit as $N_{i+1}\rightarrow\infty$

$$\left\{N_{i+1}\left(\frac{\mathbf{E}(\hat{S}_i)}{S_i}-1\right)\right\}\rightarrow\left(\frac{1-\rho_{i+1}}{\rho_{i+1}}\right)\ .$$

A simple modification of the ML estimator produces what we call an essentially unbiased estimator:

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

The exact expectation of $\tilde{S}_i$ is

$$\mathbf{E}(\tilde{S}_i)=S_i[1-(1-\rho_{i+1})^{N_{i+1}+1}]\ .$$

It is seen the bias of $\tilde{S}_i$ decreases exponentially; hence, it will be zero for all practical purposes for samples of banded birds that are large enough to merit analysis $(N_i\geqslant 300)$. It is for this reason we say $\tilde{S}_i$ is essentially unbiased.

The same approach was used to derive essentially unbiased estimators of survival rates (and recovery rates for some models) for the other full-rank models. For example, under Model H_1

$$\tilde{S}_i=\frac{R_{i\cdot}}{N_i}\left(\frac{T_i+U_i-Q_{i\cdot}-W_i}{T_i+U_i-Q_{i\cdot}}\right)\frac{N_{i+1}+1}{R_{i+1\cdot}+1}\ ,$$

$$\tilde{S}_i'=\left(\frac{Q_{i\cdot}-Q_{ii}}{M_i}\right)\frac{N_{i+1}+1}{R_{i+1\cdot}+1}$$

The Less-than-full-rank Case

Four of the 14 models presented here are not full-rank: M_3, M_2, H_{02}, H_{01} (the first two are for adults, the second two for bandings of adults and young). For these models the ML estimators do not exist in closed form; rather they must be found by numerical procedures. Programs ESTIMATE and BROWNIE find these estimators by the method of scoring (cf. Rao 1965:302; Kale 1962). An outline of the method of scoring is given below.

Let the likelihood function be $\mathscr{L}(\underline{\theta})$, where $\underline{\theta}=(\theta_1,\ldots,\theta_r)'$ is the vector-valued parameter. The likelihood equations are

$$g_i(\underline{\theta})=\frac{\partial \ell n \mathscr{L}(\underline{\theta})}{\partial \theta_i} \qquad ,i=1,\ldots,r.$$

The information matrix is $I(\underline{\theta})$ with elements

$$I_{hm}=\mathbf{E}\left[-\frac{\partial^2 \ell n \mathscr{L}(\underline{\theta})}{\partial \theta_h \partial \theta_m}\right] \qquad h,m=1,\ldots,r.$$

Let $\underline{g}(\underline{\theta})=(g_1(\underline{\theta}),\ldots,g_r(\underline{\theta}))'$, and let $\underline{\hat{\theta}}_0$ be an initial guess at the ML estimator $\underline{\hat{\theta}}$. The method of scoring is the iterative procedure defined by

$$\underline{\hat{\theta}}_{i+1}=\underline{\hat{\theta}}_i+\mathrm{I}^{-1}(\underline{\hat{\theta}}_i)g(\underline{\hat{\theta}}_i) \qquad ,i=0,1,2,\ldots$$

Under suitable conditions (as shown by Kale 1962) the sequence $\underline{\hat{\theta}}_i$ converges to the ML estimator of $\underline{\theta}$.

In practice one uses some stopping rule, generally based on convergence of the $\underline{\hat{\theta}}_i$. If the initial "estimate" is good, the number of iterations will be few. For use in ESTIMATE and BROWNIE, the initial estimates of parameters were obtained from an appropriate full-rank model. For example Model M_2, for $k=\ell$, has parameters, $f_1,\ldots,f_k$ and S. From Model M_1, closed-form estimates exist of the parameters $f_1,\ldots,f_k$ and $S_1,\ldots,S_{k-1}$. The only difference between the two models is that M_1 has year-specific survival rates while M_2 has a constant yearly survival. Initial estimates to start the iteration to find ML estimates under M_2 are $\hat{f}_1,\ldots,\hat{f}_k$ from Model M_1 and

$$\hat{\hat{S}}=\frac{1}{k-1}\sum_{i=1}^{k-1}\hat{S}_i\ ,$$

where each $\hat{S}_i$ derives from M_1. Similarly, for the two-age-class models H_{01}, H_{02}, the initial estimators are derived from the closed-form formulae of Model H_1. Because these give excellent starting values, convergence typically requires only 3-5 iterations. In fact we had no convergence problems, except with absolutely horrible data sets (i.e., those with almost no recoveries).

Both point estimation and variance-covariance matrix estimation are taken care of simultaneously by this iterative method because $\mathrm{I}^{-1}(\underline{\hat{\theta}})$ is also computed. When convergence occurs, one has both the ML estimator $\underline{\hat{\theta}}$ and the inverse of the information matrix available.

We leave it to the interested reader to pursue the subject in more depth, for instance by computing the likelihood equations and information matrix under some or all of the models M_3, M_2, H_{02} and H_{01}.

Appendix C: Ten Important Papers Since the First Edition.

Anderson, D.R., and K.P. Burnham. 1980. Effect of delayed reporting of band recoveries on survival estimates. J. Field Ornithol. 51(3):244-247.

Anderson, D.R., A.P. Wywialowski, and K.P. Burnham. 1981. Tests of the assumptions underlying life table methods for estimating parameters from cohort data. Ecology 62(4):1121-1124.

Anderson, D.R., K.P. Burnham, and G.C. White. 1985. Problems in estimating age-specific survival rates from recovery data of birds ringed as young. J. Anim. Ecol. 54:89-98.

Burnham, K.P., and D.R. Anderson. 1979. The composite dynamic method as evidence for age-specific waterfowl mortality. J. Wildl. Manage. 43(2):356-366.

Conroy, M.J., and B.K. Williams. 1984. A general methodology for maximum likelihood inference from band-recovery data. Biometrics 40(3):739-748.

Mardekian, S.Z., and L. McDonald. 1981. Simultaneous analysis of band-recovery and live-capture data. J. Wildl. Manage. 45(2):484-488.

Nelson, L.J., D.R. Anderson, and K.P. Burnham. 1980. The effect of band loss on estimates of annual survival. J. Field Ornithol. 51(1):30-38.

Nichols, J.D., S.L. Stokes, J.E. Hines, and M.J. Conroy. 1982. Additional comments on the assumptions of homogeneous survival rates in modern bird banding estimation models. J. Wildl. Manage. 46(4):953-962.

Pollock, K.H., and D.G. Raveling. 1982. Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. J. Wildl. Manage. 46(1):88-98.

White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. J. Wildl. Manage. 47(3):716-728.

J. Field Ornithol., 51(3): 244–247

EFFECT OF DELAYED REPORTING OF BAND RECOVERIES ON SURVIVAL ESTIMATES

By David R. Anderson and Kenneth P. Burnham

Brownie et al. (U.S. Fish and Wildl. Serv., Resource Publ. 131, 1978) presented 14 models based on an array of explicit assumptions for the study of survival in avian populations. These methods are replacing the life table methods previously used to estimate survival rates (e.g., Burnham and Anderson, *J. Wildl. Manage.,* **43:** 356–366, 1979). The new methods allow survival or recovery rates, or both, to be constant, time-specific, or time- and age-specific.

In studies to estimate survival rates for birds the data are often from recoveries of birds shot or found dead during the hunting season and reported to the Bird Banding Laboratory by sportsmen, conservation agency employees, or the general public. This note examines the bias in estimating annual survival due to a proportion of the recoveries being incorrectly reported a year late. Specifically, a few recoveries each year of, for example, adult male American Widgeon (*Anas americana*) banded in California are reported as being recovered in year i + 1 when in fact they were actually recovered the previous year i. Delayed reporting might typically be caused by people finding a band in their heavy clothing in the fall of the year and, being embarrassed about their failure to report the band when it was taken, report it a year late not mentioning the actual year of recovery. Heuristically, delayed reporting should bias estimated annual survival rates upwards because it appears from the data that the birds corresponding to the "delayed" recoveries actually lived an additional year.

METHODS

Results here are based on Seber's (*Biometrika,* **57:** 313–318, 1970) model (see Model 1 in Brownie et al., op. cit.) but with allowance for delayed reporting of a proportion q of the recoveries each year.

Let p = the probability that a band is reported in year i given that the bird also was shot in year i
$q = 1 - p$ = the probability that report of the band is delayed one year.

We assume all recoveries are either reported in year i (as all the models of Brownie et al., op. cit., assume) or one year later. That is, we assume people do not wait 2, 3, . . . , ℓ years before reporting a band recovered in year 1. In this study we allow p to have the values 0.95, 0.98, and 0.99 because these seem to represent realistic values (5%, 2%, and 1% rate of delayed reporting). Of course, if $p = 1.0$ no delayed reporting exists.

The magnitude of the bias will be influenced by the true survival rates and possibly, the true recovery rates. Therefore, we computed the the-

TABLE 1.

Symbolic representation of band recovery data assuming (a) no delayed reporting and (b) proportion $(1 - p)$ of the bands are reported to the Bird Banding Laboratory one year late.

(a) No delayed reporting, $p = 1.0$

Year banded	Number banded	Number of recoveries by year				
		1	2	3	• • •	ℓ
1	N_1	R_{11}	R_{12}	R_{13}	• • •	$R_{1\ell}$
2	N_2		R_{22}	R_{23}	• • •	$R_{2\ell}$
3	N_3			R_{33}	• • •	$R_{3\ell}$
•	•				•	•
•	•				•	•
•	•				•	•
k	N_k					$R_{k\ell}$

(b) Delayed reporting of one year, $p < 1.0$ $(q = 1 - p)$

Year banded	Number banded	Number of recoveries by year				
		1	2	3	• • •	ℓ
1	N_1	$R_{11}p$	$R_{12}p + R_{11}q$	$R_{13}p + R_{12}q$	• • •	$R_{1\ell}p + R_{1\ell-1}q$
2	N_2		$R_{22}p$	$R_{23}p + R_{22}q$	• • •	$R_{2\ell}p + R_{2\ell-1}q$
3	N_3			$R_{33}p$	• • •	$R_{3\ell}p + R_{3\ell-1}q$
•	•				•	•
•	•				•	•
•	•				•	•
k	N_k					$R_{k\ell}p + R_{k\ell-1}q$

oretical bias for three values of annual survival rate ($S_i = 0.35$, 0.60, and 0.85) and four values of recovery rate (0.01, 0.03, 0.06, 0.10). We examined the bias assuming banding was done over a 10-year period. In all, we examined 36 sets of expected recoveries generated from Table 1b: three survival rate values × four recovery rate values × three rates of delayed reporting = 36.

We computed the expected value of the estimator $\hat{S}_i$, $E(\hat{S}_i)$, for each of the 36 "data" sets using the adjusted maximum likelihood estimator for Model 1 (see Brownie et al., op. cit., p. 16). Because the estimator of S_i under Model 1 is unbiased assuming no delayed reporting, we can assess the bias of the estimator due to delayed reporting by generating data under the model structure shown in Table 1b. In addition, we made analyses under Models 0, 2 and 3 and intend to present these results qualitatively (see Brownie et al., op. cit., for details on these models). Two definitions are required for clarity:

$$\text{Bias} = E(\hat{S}_i) - S_i$$

$$\text{Percent relative bias (PRB)} = \frac{E(\hat{S}_i) - S_i}{S_i} \times 100.$$

TABLE 2.

Percent relative bias in the estimator of annual survival due to delayed reporting.

	Survival rate		
	35%	60%	85%
1%	7.31[1]	1.85	0.74
	2.91[2]	0.88	0.06
	2.40[3]	0.33	−0.21
3%	7.17	2.38	0.74
	3.06	0.83	0.35
	1.71	0.42	0.09
6%	6.60	2.30	1.04
	2.37	0.85	0.38
	0.94	0.38	0.17
10%	7.11	2.38	1.07
	2.91	0.92	0.41
	1.40	0.95	0.19

[1] $p = 0.95$ (5% delay)
[2] $p = 0.98$ (2% delay)
[3] $p = 0.99$ (1% delay)

RESULTS

We found, as might be expected, that the bias in the estimator of annual survival rate is independent of the number of birds banded. Furthermore, bias is not strongly affected by variation in recovery rates. Therefore, our results are much more general than the specific examples reported.

The percent relative bias (PRB) of the estimated average annual survival rate, $\hat{S}$ for each of the 36 cases is presented in Table 2. The bias is little affected by differing recovery rates. For survival rates of 60 and 85 percent, the PRB is less than 2.5 percent (e.g., if $S = 0.60$, then 2.5 PRB corresponds to $E(\hat{S}) = 0.615$) and is essentially negligible compared to the magnitude of the standard error commonly found in analyzing real data. The PRB was substantial only for survival rates of 35 percent where $p = 0.95$ (hence for low values of S and high values of q) ranging from 6.60 to 7.31 percent. Still, the size of the standard error is generally larger than this in most banding studies.

The PRB varied somewhat for the individual annual survival rates. Typically, the first and last estimates of annual survival were slightly more biased (e.g., years 1 and 9 in this study) than the estimates in the middle years of the study (e.g., years 3–7). This variation was slight and the estimates of PRB shown in Table 2 are indicative of what to expect for PRB on individual years.

The goodness of fit test for Model 1 (or Models 2 or 3) presented by Brownie et al. (op. cit.) will detect delayed reporting if it is substantial or if the sample size is large, or both. Of special interest is the fact that the tests of Model 1 vs. Model 0 (see Brownie et al., op. cit.) are quite sensitive to delayed reporting. This is indeed fortunate. If delayed reporting of recoveries is serious, the tests should indicate that Model 0 is appropriate and this model is less biased with respect to delayed reporting. The PRB for Model 0 ranged from −1.51 to 1.25 percent with most of the 36 cases studied having a slight negative bias. Of course, it is important to recall the fact that Model 1 is little biased by delayed reporting. In addition, we found Models 2 and 3 were also robust to delayed reporting.

A final remark concerns the direction of the bias. Except for Model 0 which is nearly unbiased, the other models have estimators that are slightly positively biased because of delayed reporting. In contrast, Nelson et al. (*J. Field Ornithol.*, **51:** 30–38, 1980) show that these same estimators were slightly negatively biased due to band loss. In nearly all sets of banding data we can expect some band loss and some delayed reporting of recoveries. Although we certainly cannot claim the two biases will cancel each other, it is at least satisfying that they do not magnify the overall bias.

CONCLUSION

The estimators of annual survival under Models 0, 1, 2 and 3 (Brownie et al., op. cit.) are generally robust to delayed reporting of band recoveries. If real data were analyzed under one of these models, the bias due to delayed reporting could probably be expected to be nearly negligible, especially considering the magnitude of the standard error of the estimates of annual or average annual survival. If banded samples are large and the proportion of recoveries reported a year late is large, then the tests should indicate that Model 0 is appropriate. This model is nearly unbiased with respect to problems in delayed reporting of recoveries. Finally, the direction of the bias in annual survival rate estimates is positive for most models due to delayed reporting and negative due to band loss.

Utah Cooperative Wildlife Research Unit, Utah State University, UMC 52, Logan UT 84322, and U.S. Fish and Wildlife Service, WELUT, Drake Creekside Bldg., 2625 Redwing Road, Ft. Collins, CO 80526. Received 8 November 1979, accepted 10 March 1980.

Ecology, 62(4), 1981, pp. 1121–1124

TESTS OF THE ASSUMPTIONS UNDERLYING LIFE TABLE METHODS FOR ESTIMATING PARAMETERS FROM COHORT DATA[1]

David R. Anderson,[2] Alice P. Wywialowski,[3] and Kenneth P. Burnham[4]

Since the late 1940's, life table methods have been used to estimate annual, age-specific mortality rates from tagged or banded animal populations. Deterministic methods in the form of a dynamic or composite dynamic life table were developed and used by Bellrose and Chase (1950) and Hickey (1952), and discussed by Geis and Tabor (1963). Seber (1971) used the same basic model, formulated it in the proper stochastic framework, and provided closed form estimators of the survival parameters based on approximations to the maximum likelihood (*ML*) estimators. Cormack (1970) found the *ML* estimates from the same stochastic model by numerical methods; however, identifiability was not explicitly addressed. Recently, North and Morgan (1979) proposed an extension of the same model by allowing the age-specific annual mortality rate to be a function of weather (i.e., they recognized time-dependent variation). Seber (1972, 1981) and Eberhardt (1972) provided further information on this basic model and its assumptions. Finally, we note that these same assumptions are often applied to capture-recapture data, which can then be analyzed by exactly these same life table methods applied to the final capture only (see, e.g., Mardekian and McDonald 1981).

Given this long-used model (i.e., the assumptions), there are a wide variety of ways to attempt to estimate its parameters. However, we question the fundamental assumptions that the model rests upon. In the general context here, there are two assumptions of primary concern:

1) Annual survival is assumed to be age specific only, hence independent of year,
2) The reporting rate λ is assumed to be a constant over all ages and years.

These assumptions are very restrictive and have been shown to be false for hunted waterfowl (Burnham and Anderson 1979). Furthermore, it can be argued that a model allowing each age to have a different survival is not parsimonious. Our objective here is to assess the validity of these assumptions for bird species that are not hunted.

Basic Model and Assumptions

The following notation will be required:

k = number of years of banding,
l = number of years during which recoveries are recorded, $l \geqslant k$,
S_i = $1 - M_i$ = annual survival rate (probability) for birds of age i,
M_i = $1 - S_i$ = annual mortality rate (probability) for birds of age i,
N_i = number of birds banded in year i, $i = 1, 2, \ldots, k$,
R_{ij} = number of banded birds recovered in year j from birds banded in year i, $i = 1, 2, \ldots, k$, $j = i, i + 1, \ldots, l$,
R_i = total recoveries from birds banded in year i, $= \sum_{j=1}^{l} R_{ij}$,
λ = band reporting rate.

The observed band recovery data can be symbolized as follows:

Number banded	Number recovered in year 1	2	3 . . . k . . . l	Total recovered
N_1	R_{11}	R_{12}	$R_{13} \ldots R_{1k} \ldots R_{1l}$	$R_{1.}$
N_2		R_{22}	$R_{23} \ldots R_{2k} \ldots R_{2l}$	$R_{2.}$
$\vdots$			$\ddots \quad \vdots \quad \vdots$	$\vdots$
N_k			$R_{kk} \ldots R_{kl}$	$R_{k.}$

Additional information on the theoretical basis for modeling banding data is contained in Brownie et al. (1978).

The model and its assumptions are expressed by the following expectations for the number of recoveries reported 1, 2, . . . , l yr after being banded as young in year i:

$$N_i(1-S_1)\lambda,\ N_iS_1(1-S_2)\lambda,\ N_iS_1S_2(1-S_3)\lambda,\ \ldots,\ N_iS_1S_2\ldots S_{l-1}(1-S_l)\lambda.$$

These expectations are obvious when one considers that for a bird to be recovered the 1st yr, it must die, its body be recovered, and its band be reported in that year. For a bird to be recovered in the 2nd yr after banding it must survive the 1st yr, die during the

2nd yr, and its body be recovered and its band reported, and so on (Cormack 1970:24). It is important to note the assumption that the reporting rate λ is constant for all years regardless of the age of the bird.

If we define $f_i \equiv (1 - S_i)\lambda$ then the expectations can be written more succinctly:

$$N_i f_1,\ N_i S_1 f_2,\ N_i S_1 S_2 f_3, \ldots,\ N_i S_1 S_2 \ldots S_{l-1} f_l.$$

In this form, Anderson and Burnham (1976:54–59) show that $f_i/(1 - S_i)$ must be constant for all ages, i. Tests of these assumptions constitute the subject of the paper.

Methods

We used the general goodness-of-fit test of Burnham and Anderson (1979:359–360), denoted as Test 1 here, to assess the overall assumptions of the method. Often the expectations for $j \gg i$ will be very small. If the expectations are <2, the chi-square approximation is of doubtful validity. Therefore, it becomes necessary to pool expected values within a row, and corresponding data, if $E(R_{ij})$ is <2. This results in a valid test, but loses 1 df for each pooled R_{ij} value (Brownie et al. 1978:20).

For the assumption that the 1st-yr recovery rates are constant, i.e., are age dependent only, hence not varying by year, we used the test of Burnham and Anderson (1979:360), denoted here as Test 2.

Selection of data for testing model assumptions.— We wished to analyze 20–40 data sets from unhunted species in an effort to assess the assumptions of the life table methods. After considerable effort, we could find only 10 data sets for critical evaluation. We searched *Bird-Banding, Bird Study,* and *North American Bird Bander* on the following key words: banding, band recovery, dynamics, life tables, mortality, populations, and survival. We sent letters to ornithologists and avian ecologists, scanned the *Literature Cited* sections of certain papers, and contacted many banders in an attempt to obtain more data sets for study. Table 1 provides detailed information on each data set.

Results

The results of the overall goodness-of-fit test (Test 1) of the 10 data sets to the life table model are presented in Table 1. For seven cases the significance levels range from .0000 to .0004 and provide strong evidence to reject the assumptions of the model except for the three sets of gull data. Significance levels for the three gull data sets ranged from .2546 to .8553, indicating a good fit.

The significance levels resulting from the application of Test 2 (Table 1) indicate a strong rejection of the null hypothesis that the 1st-yr recovery rates are independent of year for Great Blue Herons, Herring Gulls (New Brunswick), Brown Pelicans, White-Crowned Sparrows, and Chimney Swifts (significance levels varied from .0000 to .0007). Results of this test for Black-Crowned Night Herons and Grey Herons are significant at about the .05 level (.0610 and .0531, respectively). Only in Herring Gulls (Massachusetts), Dominican Gulls, and Great Horned Owls do the 1st-yr recovery rates appear to satisfy this assumption of the life table model. Overall, eight of the 10 data sets clearly do not fit the life table model; Herring Gulls (Massachusetts) and Dominican Gulls represent the exceptions.

Valid inferences must generally stem from a model employing valid assumptions, or, at least, the inferences must be robust, in practical terms, when certain underlying assumptions are violated. From the results presented here and those in Burnham and Anderson (1979) there is now conclusive evidence that the assumptions of the life table method are rarely met for bird banding data. Furthermore, the estimators of model parameters are quite sensitive to the failure of model assumptions (i.e., the estimators of survival are not robust to failure of assumptions). Further development and derivation of alternative estimation methods based on this model and its restrictive assumptions seem pointless, primarily because λ appears to be time specific and age specific (for at least the first or second age classes).

Discussion

It is important to note that the tests used to assess the structure and assumptions of the model are not dependent upon how the model parameters might be estimated. In fact, no parameters are estimated in doing the testing. The most severe problem concerns the reporting rate (λ) which is assumed to be a constant regardless of the age of the bird and the year in which it died. From the results of Test 2, λ clearly varies by year in most cases. Furthermore, because the spatial distribution of recoveries differs by age (e.g., young birds tend to be recovered nearer the banding site than adults), it seems likely that λ is also age specific. Finally, young birds are generally recovered earlier in time (in any given year) than adults.

Estimators of annual survival rate are substantially biased if λ is year and age specific. For example, $E(\hat{M}_1) \doteq M_1 (\lambda^*/\lambda)$, where λ^* and λ are the reporting rates for young and adult birds, respectively. For this reason alone, most mortality rates published in the literature are overestimated. Essentially, the age-specific variation in the reporting rate manifests itself as an overestimate of mortality rate because the assumption that λ is constant over all ages and years is not met.

TABLE 1. Results of tests of goodness of fit to the assumptions of the life table methods for 10 unhunted species of birds banded as young. Test 1 examined the overall assumptions of the method and Test 2 the assumption that 1st-yr recovery rates are dependent on age only and not on year.

Species (reference)	Banding area	Banding years	*k**	Total banded	Total recoveries	Test 1 significance level	Test 2 significance level
Black-Crowned Night Heron (Henny 1972)	North America	1946–1965	20	13471	287	.0000	.0610
Great Blue Heron (Henny 1972)	North America	1946–1965	20	5330	246	.0000	.0072
Grey Heron (Mead et al. 1979)	Great Britain and Ireland	1955–1974	20	4955	757	.0000	.0531
Herring Gull (Kadlec 1975)	Massachusetts	1967–1969	3	23066	499	.8553	.9256
Herring Gull (Paynter 1966)	New Brunswick	1934–1939	6	31694	1095	.2729	.0004
Dominican Gull (Fordham 1970)	New Zealand	1960–1965	6	7051	388	.2456	.6219
Great Horned Owl (Henny 1972)	North America	1951–1965	15	1896	163	.0001	.5920
Brown Pelican (Henny 1972)	North and South Carolina	1959–1965	7	6465	278	.0004	.0001
†White-Crowned Sparrow (L. Mewaldt, *personal communication*)	California	1973–1978	6	1568	648	.0001	.0000
†Chimney Swift (Henny 1972)	Tennessee	1946–1958	13	73047	2997	.0000	.0000

* Number of years of banding.
† Recaptures.

While it is becoming more clear that λ is age and year specific, it seems likely that the annual survival rate is also year specific. Certainly, there is year-specific variation in survival rate among many exploited species. North and Morgan (1979) found that annual survival might be a function of weather and hence, exhibits year-specific variation.

Substantial bias in parameter estimates is probably the most undesirable result of using a model based on invalid assumptions. Another problem with the life table method is that sampling variation is underestimated. This is because additional parameters are needed to account for the real variability in observed data. The life table model has too few parameters and, therefore, the sampling variance of the survival rate is underestimated. This leads to confidence intervals that have less than the nominal coverage and tests of hypotheses that are greater than α-level tests. The combination of substantial bias and underestimates of sampling variance result in a highly precise, incorrect answer.

Testing of assumptions has been largely neglected in the analysis of data from bird banding studies and, in fact, from animal tagging studies in general. If the numbers banded each year are not available, then tests will have almost no power. Even if large numbers are banded, but the recovery rate is very low, the tests will have limited power because so few data are available for evaluation. This may be the reason, in part, for failure to reject the assumptions for some of these data sets with few years of recovery data.

Conclusions

A variety of schemes has been proposed over the past three decades to estimate age-specific survival or mortality rates. A great many of these "life table methods" are based on the same model. This model assumes that annual survival varies only by age of the bird, and not by year. A second assumption is that the reporting rate is constant over all ages and years. These restrictive assumptions are not valid for most bird banding studies. This prevents valid inferences from being made using analysis of the sample data and, therefore, we recommend against the continued use of this approach.

Acknowledgments: Ms. C. Mead provided information on the number of nestlings banded to complement the data in Mead et al. 1979. Mr. L. R. Mewaldt and Mr. R. A. Paynter, Jr. sent additional information for our analysis. The comments offered by two anonymous reviewers were helpful.

Literature Cited

Anderson, D. R., and K. P. Burnham. 1976. Population ecology of the mallard: VI. The effect of exploitation on survival. Resource Publication 128, United States Fish and Wildlife Service, Washington, D.C., USA.

Bellrose, F. C., and E. B. Chase. 1950. Population losses in the mallard, black duck, and blue-winged teal. Biological Notes 22, Illinois Natural History Survey, Urbana, Illinois, USA.

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1978. Statistical inference from band recovery data: a handbook. Resource Publication 131, United States Fish and Wildlife Service, Washington, D.C., USA.

Burnham, K. P., and D. R. Anderson. 1979. The composite dynamic method as evidence for age-specific waterfowl mortality. Journal of Wildlife Management **43**(2):356–366.

Cormack, R. M. 1970. Appendix: the construction of life tables from the recovery of dead animals. Journal of Animal Ecology **39**(1):24–27.

Fordham, R. A. 1970. Mortality and population change of dominican gulls in Wellington, New Zealand. Journal of Animal Ecology **39**(1):3–27.

Geis, A. D., and R. D. Tabor. 1963. Measuring hunting and other mortality. Pages 284–298 *in* H. S. Mosby, editor. Wildlife investigation techniques. The Wildlife Society, Washington, D.C., USA.

Henny, C. J. 1972. An analysis of the population dynamics of selected avian species with special reference to changes during the modern pesticide era. Wildlife Research Report 1, United States Fish and Wildlife Service, Washington, D.C., USA.

Henny, C. J., and K. P. Burnham. 1976. A mallard reward band study to estimate band reporting rates. Journal of Wildlife Management **40**(1):1–14.

Hickey, J. J. 1952. Survival studies of banded birds. Special Scientific Report—Wildlife 15, United States Fish and Wildlife Service, Washington, D.C., USA.

Kadlec, J. A. 1975. Recovery rates and losses of aluminum, titanium, and incoloy bands on herring gulls. Bird-Banding **46**(3):230–235.

Mardekian, S. Z., and L. L. McDonald. 1981. Simultaneous analysis of band recovery and live recapture data. Journal of Wildlife Management **45**(2):484–488.

Mead, C. J., P. M. North, and B. R. Watmough. 1979. The mortality of British grey herons. Bird Study **26**(1):13–22.

North, P. M., and B. Morgan. 1979. Modelling heron survival using weather data. Biometrics **35**(3):667–681.

Paynter, R. A. 1966. A new attempt to construct life tables for Kent Island herring gulls. Bulletin of the Museum of Comparative Zoology at Harvard University **133**:489–528.

Seber, G. A. F. 1971. Estimating age-specific survival rates from bird-band returns when reporting rate is constant. Biometrika **58**(3):491–497.

———. 1972. Estimating survival rates from bird-band returns. Journal of Wildlife Management **34**(4):405–413.

———. 1981. The estimation of animal abundance and related parameters. Second edition. Charles Griffin, London, England.

[1] *Manuscript received 7 July 1980;*
revised 7 December 1980;
accepted 20 January 1981.

[2] *Utah Cooperative Wildlife Research Unit,*
Utah State University, Logan, Utah 84322 USA.

[3] *Department of Wildlife Science,*
Utah State University, Logan, Utah 84322 USA.

[4] *United States Fish and Wildlife Service,*
Fort Collins, Colorado 80526 USA.

Reprinted from

THE JOURNAL OF ANIMAL ECOLOGY
VOL. 54

BLACKWELL SCIENTIFIC PUBLICATIONS
OXFORD LONDON EDINBURGH
BOSTON MELBOURNE

Journal of Animal Ecology (1985), **54,** 89–98

PROBLEMS IN ESTIMATING AGE-SPECIFIC SURVIVAL RATES FROM RECOVERY DATA OF BIRDS RINGED AS YOUNG

BY D. R. ANDERSON,* K. P. BURNHAM† AND G. C. WHITE‡

**Utah Cooperative Wildlife Research Unit, Department of Fisheries and Wildlife, Utah State University UMC 52, Logan, Utah 84322, U.S.A., †U.S. Fish and Wildlife Service, WELUT, 2672 Redwing Road, Ft. Collins, Colorada 80526, U.S.A. and ‡Los Almos National Laboratory, Environmental Sciences LS-6 MS K495, Los Almos, New Mexico 87545, U.S.A.*

SUMMARY

(1) The life table model is frequently employed in the analysis of ringed samples of young in bird populations. The basic model is biologically unrealistic and of little use in making inferences concerning age-specific survival probabilities.

(2) This model rests on a number of restrictive assumptions, the failure of which causes serious biases. Several important assumptions are not met with real data and the estimators of age-specific survival are not robust enough to these failures.

(3) Five major problems in the use of the life table method are reviewed. Examples are provided to illustrate several of the problems involved in using this method in making inferences about survival rates and its age-specific nature.

(4) We conclude that this is an invalid procedure and it should not be used. Furthermore, ringing studies involving only young birds are pointless as regards survival estimation because no valid method exists for estimating age-specific or time-specific survival rates from such data.

(5) In our view, inferences about age-specific survival rates are possible only if both young and adult (or young, subadult and adult) age classes are ringed each year for k years ($k \geqslant 2$).

INTRODUCTION

Analysis methods for estimating age-specific survival rates from the ringing of young birds have generally been based on a single underlying model, or special case of this model. Our objective is to review this model that we will call the life table model, its assumptions, approaches to estimation of the model's unknown parameters and problems in making inference concerning age-specific survival rates. Our work was motivated, to some extent, by the recent paper by Lakhani & Newton (1983).

Basic model and notation

The following notation is required:

k = number of years of ringing,

l = numbers of years during which ring recoveries are recorded, $k \leqslant l$,

---

* Correspondence and present address: Dr D. R. Anderson, Colorado Cooperative Fish and Wildlife Research Unit, 201 Wagar Building, Colorado State University, Fort Collins, Colorado 80523, U.S.A.

† Present address: Dr K. P. Burnham, Department of Statistics, North Carolina State University, Raleigh, North Carolina 27650, U.S.A.

‡ Dr G. C. White, Department of Fishery and Wildlife Biology, Colorado State University, Fort Collins, Colorado 80523, U.S.A.

Age-specific survival rates

TABLE 1. Ringing and recovery data for young birds ringed over k years and recovered over l years ($l > k$)

Ringing year i	Number ringed N_i	Recoveries (R_{ij}) in year j						
		1	2	3	$\cdots$	k	$\cdots$	l
1	N_1	R_{11}	R_{12}	R_{13}	$\cdots$	R_{1k}	$\cdots$	R_{1l}
2	N_2		R_{22}	R_{23}	$\cdots$	R_{2k}	$\cdots$	R_{2l}
3	N_3			R_{33}	$\cdots$	R_{3k}	$\cdots$	R_{3l}
.	.				.	.		.
.	.				.	.		.
k	N_k					R_{kk}	$\cdots$	R_{kl}

ϕ_i = annual survival rate (probability) for birds of age i,
$M_i = l - \phi_i$ = annual mortality rate (probability) for birds of age i,
N_i = number of young birds ringed in year i, $i = 1, 2, \ldots, k$
R_{ij} = number of ringed birds recovered in year j from birds ringed as young in year i, $i = 1, 2, \ldots, k, j = i, i + 1, \ldots, l$,
λ = reporting rate (probability); sometimes used as a subscripted variable to indicate age-specific or time-specific parameters.

Ringing recovery data, the vector N_i and the matrix R_{ij}, are shown in Table 1. The expectations, under the life table model, as functions of the unknown parameters of interest, are found in Table 2.

Assumptions

Several assumptions are necessary in making inference from the life table model:

(1) The ringed sample is representative of the population of interest,
(2) There is no loss of rings,
(3) Survival rates are not affected by ringing,
(4) The year of recovery for those ringed birds recovered is correctly tabulated,
(5) The fate of each ringed bird is independent of the fate of other ringed birds,
(6) Annual survival ϕ_i is age-specific only, independent of year,
(7) The reporting rate λ is a constant over all age classes and years, and
(8) Every ringed bird experiences homogeneous rate parameters ϕ_i and λ (i.e., no population subgroups having heterogeneous parameters.

Assumptions 1, 2, 3, 4, 5 and 8 are not testable without additional data other than the N_i and R_{ij}. Statistical tests of assumptions 6 and 7 are possible using only the ringing data. However, if the general goodness-of-fit test shows lack of fit, it may not be clear which particular assumptions are rejected. Although assumptions 2, 6, and 7 are critical ones under this model, they are not especially critical with certain other models for the analysis of ringing data (e.g., Model H_2 in Brownie *et al.* 1978).

Estimation of parameters

Early attempts to derive estimators of the unknown parameters include Lack (1943), Franer (1945), Bellrose & Chase (1950), Hickey (1952), Haldane (1955), and Balham & Miers (1959). Of these, only Haldane explored optimal estimation, using maximum likelihood theory. He studied the special case of the life table model where $\phi_1 = \phi_2 = \ldots =$

TABLE 2. The life table model for the recovery of ringed birds banded as young ($l > k$)

Banded	Expected number of ring returns in year j from ringing in year i* 1	2	3	$\cdots$	k	$\cdots$	l
N_1	$N_1(1-\phi_1)\lambda$	$N_1\phi_1(1-\phi_2)\lambda$	$N_1\phi_1\phi_2(1-\phi_3)\lambda$	$\cdots$	$N_1\phi_1\phi_2 \cdots \phi_{k-1}(1-\phi_k)\lambda$	$\cdots$	$N_1\phi_1\phi_2 \cdots \phi_{l-1}(1-\phi_l)\lambda$
N_2		$N_2(1-\phi_1)\lambda$	$N_2\phi_1(1-\phi_2)\lambda$	$\cdots$	$N_2\phi_1\phi_2 \cdots \phi_{k-2}(1-\phi_{k-1})\lambda$	$\cdots$	$N_2\phi_1\phi_2 \cdots \phi_{l-2}(1-\phi_{l-1})\lambda$
N_3			$N_3(1-\phi_1)\lambda$	$\cdots$	$N_3\phi_1\phi_2 \cdots \phi_{k-3}(1-\phi_{k-2})\lambda$	$\cdots$	$N_3\phi_1\phi_2 \cdots \phi_{l-3}(1-\phi_{l-2})\lambda$
$\vdots$				$\ddots$	$\vdots$		$\vdots$
N_k					$N_k(1-\phi_1)\lambda$	$\cdots$	$N_k\phi_1\phi_2 \cdots \phi_{l-k}(1-\phi_{l-k+1})\lambda$

* All subscripts relate to the age of the bird; the parameters λ and ϕ_i are assumed to be independent of year.

$\phi_l = \phi$. His estimators allowed the R_{ij} matrix to be truncated; i.e., before some recoveries in later years were reported. Eberhardt (1972), Caughley (1977), and Seber (1972, 1982) provide details of these early methods.

Maximum likelihood methods for the general model (Table 2) were not fully considered until the work of Cormack (1970) and Seber (1971); also see North & Cormack (1981).

State-of-the-science estimation

Computer algorithms developed by Lebreton (1977) and White (1983) represent optimal estimation from this model and are flexible numerical procedures. Both computer programs require at least one constraint to allow identifiability of the unknown parameters (see Anderson & Burnham 1976, Brownie *et al.* 1978; Burnham & Anderson 1979; Seber 1971, 1972, 1982). Lakhani & Newton (1983) show that the imposition of such an identifiability constraint can produce substantial bias even when the true parameters deviate only slightly from the constraint. Both algorithms use maximum likelihood theory for estimation and testing of a sequence of models for arbitrary k and l without bias due to truncation. The sequence is illustrated in Table 3, depending on the constraints.

Log-likelihood ratio tests and goodness-of-fit tests are provided to allow an appropriate model to be used in making inference from a particular data set (Lebreton 1977; White 1983). These methods and their associated computer algorithms should replace all previous estimation methods for the life table model (however, in the unusual case where $l \rightarrow \infty$, closed form estimators and tests can be derived—see Botkin & Miller 1974; Seber 1972), but Lakhani & Newton (1983) have demonstrated that the resulting estimates are liable to be untrustworthy.

PROBLEMS

Although the underlying stochastic model is well defined, cf. Seber (1971), and optimal estimation algorithms have been developed (subject to an identifiability constraint), the life table model is a very poor basis for the analysis of ringing data. Support for this appears in the following sections.

Reporting rate not a constant

Estimates of ϕ_i are critically dependent on the assumption that the reporting rate λ is constant over all years and age classes. The reporting rate λ is the product of the probability of finding a ringed dead bird times the probability of reporting it. The latter probability is unlikely to be age- or year-specific. However, the probability of finding a

TABLE 3.

Model	Constraint(s)	Number of parameters
Full (Table 2)	$\phi_{l-1} = \phi_l$	l
Reduced	$\phi_{l-2} = \phi_{l-1} = \phi_l$	$l-1$
Reduced	$\phi_{l-3} = \phi_{l-2} = \phi_{l-1} = \phi_l$	$l-2$
Reduced	$\phi_{l-4} = \ldots = \phi_l$	$l-3$
.	.	.
.	.	.
.	.	.
Null	All $\phi_i = \phi$	2 (i.e., ϕ and λ)

ringed dead bird is likely to be a function of the cause of death, geographic location of the bird, and the time of the year. This can be easily shown by mapping and tabulating the spatial and temporal distribution of recoveries for the two age groups, young and adult (or young, subadult, and adult classes) (see Hopper, Funk & Anderson 1978). Biologically, we might expect λ to vary by age for, at least, young *v.* adults. Young birds often die due to different causes of mortality, at different times of the year, and in different geographical locations than adults. These basic concepts can be formulated as hypotheses and routinely and statistically tested. Typically, the reporting rate of young birds λ_l is higher than adult birds λ.

The approximate expection of the estimators of age-independent survival (where $\phi_i = \phi$ for all ages) when the reporting rate is age-specific for 2 years (λ_1 and λ_2) is

$$E(\hat{\phi}) \doteq 1 - \frac{M\lambda_1}{M\lambda_1 + \phi\lambda_2}$$

(see example 1a in Table 4). If survival is age-specific for three age classes (i.e., young, subadult, and adult, ϕ_1, ϕ_2 and ϕ_3, respectively), and reporting rate also varies for these three ages classes (λ_1, λ_2 and λ_3, respectively), the following approximations are useful:

$$E(\hat{\phi}_1) \doteq 1 - \frac{M_1\lambda_1}{M_1\lambda_1 + \phi_1 M_2\lambda_2 + \phi_1\phi_2\lambda},$$

$$E(\hat{\phi}_2) \doteq 1 - \frac{M_2\lambda_2}{M_2\lambda_2 + \phi_2\lambda}.$$

The expectations indicate a serious problem with the life table method. Implicitly, the model assumes that the same proportion of ringed birds are found and reported, regardless

TABLE 4. Examples illustrating the approximate age-specific bias in life table estimates of survival when reporting rate is also age-specific

		Reporting rate				Survival rate					
Example	Quantity	λ_1	λ_2	λ_3	λ_4	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
1a	Parameter	0·12	0·06	0·06	0·06	0·6	0·6	0·6	0·6	0·6	0·6
	Estimate*	0·086	—	—	—	0·41	0·58	0·58	0·58	0·58	0·58
	PRB†	−28	43			−32	−3	−3	−3	−3	−3
1b	Parameter	0·015	0·008	0·008	0·008	0·6	0·6	0·6	0·6	0·6	0·6
	Estimate	0·013	—	—	—	0·43	0·64	0·53	0·53	0·53	0·53
	PRB	−13	38			−28	+7	−12	−12	−12	−12
2a	Parameter	0·10	0·06	0·05	0·05	0·6	0·7	0·85	0·85	0·85	0·85
	Estimate	0·076	—	—	—	0·44	0·69	0·87	0·87	0·87	0·87
	PRB	−24	27	52		−27	−1	+2	+2	+2	+2
2b	Parameter	0·010	0·006	0·005	0·005	0·6	0·7	0·85	0·85	0·85	0·85
	Estimate	0·008	—	—	—	0·42	0·64	0·83	0·83	0·83	0·83
	PRB	−20	33	60		−30	−9	−2	−2	−2	−2
3	Parameter	0·18	0·19	0·25	0·29	0·53	0·37	0·59	0·61	0·47	0·61
	Estimate	0·17	—	—	—	0·42	0·59	0·61	0·63	0·60	0·61
	PRB	−6	−12	−32	−41	−21	+60	+3	+3	+28	0
4	Parameter	0·28	0·29	0·24	0·24	0·64	0·49	0·79	0·79	0·79	0·79
	Estimate	0·27	—	—	—	0·44	0·56	0·73	0·73	0·73	0·73
	PRB	−4	−7	13		−31	+12	−8	−8	−8	−8

* Approximate expected value of the estimator.

‡ Per cent relative bias, $(E(\phi) - \phi/\phi) \times 100$.

of the number or the proportion that die in a given year. Furthermore, this proportion is the same for all years $j = 1, 2, \ldots, l$. This is a very restrictive assumption if viewed on biological grounds. The estimators are particularly sensitive to large bias if λ varies by age (a violation of assumption 7, an assumption which is biologically unlikely). The expressions above allow the magnitude of the bias in $\hat{\phi}$ to be approximated as functions of age-specific reporting rates.

We developed a statistical test of the null hypothesis that λ was independent of year in Burnham & Anderson (1979). This test is useful for data sets where few recoveries are available. The test-statistic is distributed as chi square with $k - 1$ degrees of freedom under the null hypothesis that $\phi_1\lambda$ is constant,

$$N^2 \sum_{i=1}^{k} \frac{[R_{ii} - N_i(\frac{R}{N})]^2}{N_i R(N-R)} = \chi^2_{k-1}$$

where N is the total number banded ($N = N_1 + \ldots + N_k$) and R is the total number of first-year recoveries ($R = R_{11} + R_{22} + \ldots + R_{kk}$). For better data sets, additional tests are useful. An extensive analysis of forty-five sets of data on ringed waterfowl rejected this hypothesis in all but six cases ($P < 0{\cdot}05$). In a similar analysis of ten unhunted species we found rejection of the hypothesis in five cases (Anderson, Wywialowski & Burnham 1981).

In general, it seems that the reporting rate is usually age- and time-specific. This is a violation of the assumptions of the life table model and the estimates are quite sensitive to the failure of this assumption (also see Brownie *et al.* 1978).

Constraints

It was not until Seber (1971) that it was realized that the parameters in the life table model could not be uniquely estimated without at least one constraint on the parameters. The full model has $l + 1$ parameters, but only l unique data cells; thus, a fundamental problem is lack of identifiability. This problem is alleviated only if $l \gg k$ or if $(l - k) >$ potential longevity, which is unusual. In examining the older literature, we find the constraint $\phi_l \equiv 0$ was used often implicitly (Burnham & Anderson 1979). This constraint caused a significant age-specific bias in the estimators of survival.

Seber (1971) suggested the constraint $\phi_{l-1} \equiv \phi_l$. This is a minimum to gain identifiability of the unknown parameters. However, Lakhani & Newton (1983) found that such constraints, while seemingly biologically reasonable, could have a marked effect on the estimates of ϕ_i even if λ is constant. They conclude, for this reason alone, "... all hitherto published estimates of age-specific survival are liable to be untrustworthy, if they are based solely on recoveries of dead birds." Additional contraints to reduce the number of parameters, such as $\phi_{l-2} = \phi_{l-1} = \phi_l$, affected the estimates of the remaining parameters.

The necessity of a constraint has not always been recognized. Recently, Piper (1978) derived an estimation scheme for ϕ_i (i = age) and λ_j (j = year). The $\hat{\phi}_i$ are ML estimates found iteratively (under the assumption that λ is a constant over all years and ages) and the λ_j are found analytically (see Piper, Mundy & Ledger 1981). The ϕ_i and λ_j are impossible to estimate uniquely and we warn against the use of Piper's (1978) method.

Fit of model to data

A general goodness-of-fit test was derived by Burnham & Anderson (1979) and allows an assessment of the model using the data. The test can be computed without any constraints on the parameters and without even the need to estimate the parameters. It is a test of model fit, irrespective of the parameter estimation method. Burnham & Anderson (1979) found this model was rejected in forty-three out of forty-five sets of waterfowl data. In fact, the two data sets not rejected had the fewest and fourth fewest recoveries; therefore, probably indicating a lack of power in the test due to small sample size. A similar lack of fit was found in non-hunted bird species by Anderson, Wywialowski & Burnham (1981) where seven out of ten data sets were rejected. It is interesting to note that the remaining three were represented by ringing in only 3–6 years, hence the tests may have low power.

We tested the recovery data from Cape Vultures (*Gyps coprotheres*) ringed as nestlings (see Piper, Mundy & Ledger 1981) to determine if the first-year recovery rate (R_{ii}/N_i) of young birds was constant over all years. This hypothesis was rejected ($\chi^2 \sim 71$, d.f. $= 23$, $P \sim 0.0$). In fact, one can see from the basic data that the first-year recovery rate tends to increase during the years of study. This is a serious violation of the assumptions of the life table model. Piper's (1978) method is incorrect in that the year-specific reporting rates are not uniquely identifiable (see Seber 1971; Anderson & Burnham 1976; Burnham & Anderson 1979). We must conclude that real ringing data rarely fit the life table model.

Sampling correlations

Regardless of the specific constraint used to allow identifiability or reduce the number of parameters to estimate, the parameter estimators can be very highly correlated. For example, the case given by Lakhani & Newton (1983) has sampling correlation coefficients between the estimators of parameters of 0·97 or more. When there are such very high sampling correlations between all the estimated parameters under a given constraint, a small change in the constraint will result in a direct change in all estimated parameters. Thus, different constraints can be expected to have a marked effect on the estimates from a given data set. In fact, we believe that the presence of such very high sampling correlations is indicative that the parameter estimates will be very sensitive to the imposed constraint.

It is very difficult to make inference about age-specific survival processes when the estimators are almost perfectly correlated. Lakhani & Newton (1983) also note problems with the estimated standard errors of the estimates in that they give a "... totally false sense of accuracy and reliability ...." The danger here is that one tends to get highly precise estimates that are very biased; i.e., a precise wrong answer.

Ring loss

The effect of ring loss on estimates of survival is serious for the life table model because it allows survival only to be age-specific. ϕ_i and ring loss rates are seriously confounded and produce a strong age-specific negative bias in estimates of survival (see Ludwig 1967; Botkin & Miller 1974 for reviews). Nelson, Anderson & Burnham (1980) provide an example of this problem whereby an artificial population was analysed with each $\phi_i \equiv 0{\cdot}6$. Using the composite dynamic estimation method for the life table model (Hickey 1952) they computed estimates of survival as 0·56, 0·54, 0·52, 0·48, 0·42, 0·36, 0·25, 0, and 0 for ϕ_1, ϕ_2, ... ϕ_9, respectively. This example was similar to the data on *Larus delawarensis* given by Ludwig (1967) in his study of ring loss. Ring loss cannot be detected by goodness-of-fit tests because ring loss and mortality of ringed birds are confounded. Because of the importance of ring loss and the negative bias in estimates of survival with the life table model, we suspect much of the published literature presents untrustworthy estimates of survival of bird populations. In addition, populations inferred to have age-specific survival beyond about 3 or 4 years may be only losing their rings in later years.

Ring loss is a practical problem that causes large bias in estimators of survival in some models (such as the life table model) while causing negligible bias in other models (e.g., Models 1, 0, H_1 and H_2 of Brownie *et al.* 1978).

EXAMPLES

Six examples (Table 4) illustrate the approximate bias in estimating ϕ_i using the life table model when the reporting rate is age-specific. Expected values were calculated for each example, rounded to the nearest integer, treated as data, and then analysed using program SURVIV (White 1983). The first four examples use $N_i = 3000$, $k = 6$, and $l = 15$. Examples 1a and 1b allow all $\phi_i = \phi$ and let the reporting rate be age-specific for young (λ_1) and adults ($\lambda_2 = \lambda_3 = \lambda_4 = \lambda$). Examples 2a and 2b allow age-specific survival for young ϕ_1, subadults ϕ_2 and adults $\phi_3 = \phi_4 = \phi_5 = \phi_6 = \phi$ while allowing similar age-specificity in the reporting rate λ_1, λ_2, and λ_3. Examples 1a *v.* 1b and 2a *v.* 2b differ only in the magnitude of the reporting rates (see Table 4). Data for Example 3 and the final example were taken from Brownie *et al.* (1978).

Examples 3 and 4 allow the parameters to vary for two and three age classes, respectively. Both allow year-specific variation (another realism not allowed under the life table model). We used reasonable parameter values for many bird populations to illustrate the severe, age-specific bias in estimates from the life table model. The best model in the sequence was selected using the likelihood ratio tests provided by program SURVIV.

The information in Table 4 shows that if the reporting rate varies by age, a substantial negative bias is introduced in $\hat{\phi}_1$. Other estimators of survival are less effected, but frequently reflect negative bias. The mean life span (MLS) is estimated from the estimated survival estimates. The MLS for Experiment 1 is 1·96 years, however, the estimate, computed from the life table estimates of age-specific survival, is only 1·41 years. If ringing of young took place in a remote location where rings were not found or reported, we might expect $\lambda_1 < \lambda_2$ and then a significant positive bias would be expected in $\hat{\phi}_1$.

In is interesting to note that if the reporting rate is age-specific, a violation of the life table model, estimates of λ may be only moderately biased, but will result in much larger biases in $\hat{\phi}_1$. This shows that the estimates of ϕ_1 are very sensitive to age-specific reporting rates.

DISCUSSION

The critical underlying assumptions of the life table model are invalid and the estimators of model parameters are sensitive to these assumption failures, thus rigorous inference about the survival rates of ringed bird populations or their age-specific dynamics, is impossible. The foremost problem arises from the fact that λ varies substantially by age (e.g., young *v.* adults) as this leads to large bias in the estimator for ϕ_1. Even if the key assumptions are met for a particular data set, additional difficulties arise with even optimal estimation methods (e.g., sensitivity of the survival estimates to arbitrary constraints and possibly high sampling correlations among the estimators of ϕ_i and λ). Finally, ring loss is confounded with mortality giving increasing age-specific bias in the survival estimators. This is critical, especially for long-lived species.

Lakhani & Newton (1983) detail some of the major problems with the life table model. They were concerned mainly with the possible effects of constraints, if other model assumptions could be presumed reasonable. The model is even more biologically unreasonable and statistically inept than Lakhani & Newton (1983) document and we believe it is not possible to draw valid inferences from it. We feel it is very unlikely than an accurate and precise estimate of λ (or rather λ_i) or ϕ_1 could be obtained via independent information. For example, $\hat{\phi}_1$ using Cormack's (1964) method, as they mention, is likely to be fairly imprecise. The likely usefulness of $\hat{\lambda}_i$ or $\hat{\phi}_1$ from independent sources diminishes further in examples, such as Lakhani & Newton (1983) give, where the elements in the sampling correlation matrix all exceed 0·97. This is because with such high sampling correlations, the remaining parameters are strongly influenced by the externally estimated parameter any bias or imprecision of that estimate carries over to the other parameter estimates. Use of the life table model is likely only to introduce greater confusion, rather than gaining understanding of population dynamics. The biological and statistical problems of estimation and inference with the life table model may have led to the conclusions forwarded by Botkin & Miller (1974).

The life table method is unacceptable for the analysis of ringing studies of young birds. It is misleading, untrustworthy and gives a false sense of precision (Lakhani & Newton 1983). We cannot recommend its use. We agree with Brownie *et al.* (1978) that, "... based on our current knowledge, there is no valid way to estimate age-specific survival rates from only the banding of young."

The probable solution to problems of estimation and testing is to ring both young and adult (or young, subadult, and adult) age classes each year for $k \geqslant 2$ years. This allows a host of models to be considered that are superior to the life table (see Brownie *et al.* 1978 for twelve models allowing year-specific, age-specific, or year and age-specific survival and reporting parameters).

REFERENCES

Anderson, D. R. & Burnham, K. P. (1976). Population ecology of the mallard: VI. The effect of exploitation on survival. *United States Fish and Wildlife Service, Washington, D.C. Resource Publication,* **128,** 1–66.

Anderson, D. R., Wywialowski, A. P. & Burnham, K. P. (1981). Tests of the assumptions underlying life table methods for estimating parameters from cohort data. *Ecology,* **62,** 1121–1124.

Balham, R. W. & Miers, K. H. (1959). Mortality and survival of grey and mallard ducks banded in New Zealand. *New Zealand Department of Internal Affairs, Wildlife Publication* No. 5, 1–56.

Bellrose, F. C. & Chase, E. B. (1950). Population losses in the mallard, black duck and the blue-winged teal. *Illinois Natural History Survey, Biological Notes,* **22,** 1–27.

Botkin, D. B. & Miller, R. S. (1974). Mortality rates and survival of birds. *American Naturalist,* **108,** 181–192.

Brownie, C., Anderson, D. R., Burnham, K. P. & Robson, D. S. (1978). *Statistical inference from band recovery data: A handbook.* U.S. Fish and Wildlife Service, Resource Publication 131.

Burnham, K. P. & Anderson, D. R. (1979). The composite dynamic method as evidence for age-specific waterfowl mortality. *Journal of Wildlife Management,* **43,** 356–366.

Caughley, G. (1977). *Analysis of vertebrate populations.* John Wiley & Sons.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika,* **51,** 429–438.

Cormack, R. M. (1970). Statistical appendix. *Mortality and population change of Dominican gulls in Wellington, New Zealand* (by R. A. Fordham). *Journal of Animal Ecology,* **39,** 13–27.

Eberhardt, L. L. (1972). Some problems in estimating survival from banding data. *Population ecology of migratory birds—A symposium* pp. 278. U.S. Fish and Wildlife Research Report, 2.

Farner, D. S. (1945). Age groups and the american robin. *Wilson Bulletin,* **57,** 437–449.

Haldane, J. B. S. (1955). The calculation of mortality rates from ringing data. Proceedings of the 11th International Ornithological Congress, Basel, **11,** 454–458.

Hickey, J. J. (1952). Survival studies of banded birds. *United States Fish and Wildlife Service, Special Scientific Report (Wildlife),* **15,** 1–177.

Hopper, R. M., Funk, H. D. & Anderson, D. R. (1978). Age-specificity in mallards banded postseason in Eastern Colorado. *Journal of Wildlife Management,* **42,** 263–270.

Lack, D. (1943). The age of the blackbird. *British Birds,* **36,** 166–175.

Lakhani, K. H. & Newton, I. (1983). Estimating age-specific bird survival rates from ring recoveries—can it be done? *Journal of Animal Ecology*, **52,** 83–91.

Ludwig, J. P. (1967). Band-loss: Its effects on banding data and apparent survivorship in the ring-billed gull population of the Great Lakes. *Bird-Banding*, **38,** 309–323.

Lebreton, J. D. (1977). Maximum likelihood estimations of survival rates from bird band returns: Some complements to age-dependent methods. *Biometrie-Praximetrie*, **17,** 145–161.

Nelson, L. J., Anderson, D. R., & Burnham, K. P. (1980). The effect of band loss on estimates of annual survival. *Journal of Field Ornithology*, **51,** 30–38.

North, P. M. & Cormack, R. M. (1981). On Seber's method for estimating age-specific bird survival rates from ringing recoveries. *Biometrics*, **37,** 103-112.

Piper, S. E. (1978). *The evaluation, comparison and application (to avian population dynamics) of a technique for analysing capture-recapture experiments.* M.Sc. Thesis, University of the Witerwatersrand.

Piper, S. E., Mundy, P. J., & Ledger, J. A. (1981). Estimates of survival in the Cape Vulture, *Gyps coprotheres*. *Journal of Animal Ecology*, **50,** 815–825.

Seber, G. A. F. (1971). Estimating age-specific survival rates for birds from bird-band returns when the reporting rate is constant. *Biometrika*, **58,** 491–497.

Seber, G. A. F. (1972). Estimating survival rates from bird-band returns. *Journal of Wildlife Management*, **36,** 405–413.

Seber, G. A. F. (1982). *The estimation of animal abundance and related parameters.* Macmillan Publishing Company, New York, N.Y.

White, G. C. (1983). Numerical estimation of survival rates from band recovery and biotelementry data. *Journal of Wildlife Management*, **47,** 716–728.

(Received 5 August 1983)

THE COMPOSITE DYNAMIC METHOD AS EVIDENCE FOR AGE-SPECIFIC WATERFOWL MORTALITY

KENNETH P. BURNHAM, U.S. Fish and Wildlife Service, Fort Collins, CO 80521
DAVID R. ANDERSON, Utah Cooperative Wildlife Research Unit, Logan, UT 84322

Abstract: For the past 25 years estimation of mortality rates for waterfowl has been based almost entirely on the composite dynamic life table. We examined the specific assumptions for this method and derived a valid goodness of fit test. We performed this test on 45 data sets representing a cross section of banded samples for various waterfowl species, geographic areas, banding periods, and age/sex classes. We found that: (1) the composite dynamic method was rejected ($P < 0.001$) in 37 of the 45 data sets (in fact, 29 were rejected at $P < 0.00001$) and (2) recovery and harvest rates are year-specific (a critical violation of the necessary assumptions). We conclude that the restrictive assumptions required for the composite dynamic method to produce valid estimates of mortality rates are not met in waterfowl data. Also we demonstrate that even when the required assumptions are met, the method produces very biased estimates of age-specific mortality rates. We believe the composite dynamic method should not be used in the analysis of waterfowl banding data. Furthermore, the composite dynamic method does not provide valid evidence for age-specific mortality rates in waterfowl.

J. WILDL. MANAGE. 43(2):356–366

The composite dynamic (CD) life table has been the primary method used to estimate annual mortality rates of waterfowl populations for the past 25 years. Early descriptions and use of the method are found in Bellrose and Chase (1950) and Hickey (1952). The model underlying this estimation method, the properties of the method, and explicit computing formulas are usually not specified clearly (see Geis and Tabor 1963; Geis 1972*a,b*). Nevertheless, use of the method has been widespread (e.g., Moisan et al. 1967; Geis et al. 1971) and not at all restricted to waterfowl. Anderson and Burnham (1976), Eberhardt (1972), and Seber (1972, 1973) present recent material on the validity and usefulness of the method.

Three crucial assumptions are necessary for the composite dynamic method to be valid: (1) annual mortality rate (M_i) varies only by the age (i) of the bird, no time-specific variation of the M_i is allowed; (2) annual recovery rate is a constant fraction of annual mortality rate; and (3) virtually none of the banded birds remains alive when the data are analyzed (Geis 1972*a*:16).

A variety of implications follow from these assumptions, perhaps the most important one is that no year to year (time-specific) variation in either harvest rates, or band reporting rates is allowed. These 2 implicit assumptions are generally recognized as untrue (e.g., Anderson 1975; Henny and Burnham 1976). Assumption (3) means the recovery data from a banding study cannot be validly analyzed with the CD method for many years after banding has stopped. We believe that few biologists recognize the assumptions implicit in the method nor the degree to which the estimators of M_i may be sensitive to departures from these important assumptions.

In addition to these specific assumptions needed for the CD method, a number of general assumptions are necessary for meaningful analysis of banding data under any model (Brownie et al. 1978:6–7). In particular there must be no band loss. Band loss is indistinguishable from age-specific mortality, and is especially serious under the CD method of analysis.

A large scientific literature on various species of birds contains the results from a CD analysis of band recovery data. It has focused on 3 main areas: (1) the estimation of 1st-year mortality rate (M_1) from birds banded as young and estimation of average annual mortality rate from birds banded as adults (i.e., for adults, M_1 is interpreted as an average); (2) studying age-specificity of the mortality process; i.e., estimating M_1, M_2, M_3, . . . ; and (3) studying the effect of hunting on total annual mortality rates. We wish to address the 1st and 2nd of these as the 3rd has already been discussed by Anderson and Burnham (1976).

It is surprising that no one has reported results of any statistical tests of the assumptions required for the CD method to perform validly. Recently, Hickey (1972:264) believed, ". . . it seems particularly necessary that the basic assumptions underlying these calculations be subjected to an adequate review." That is our purpose here. The 4 objectives of this paper are: (1) to present the specific assumptions and model necessary for the CD method to be valid; (2) to present 2 statistical tests of these assumptions: (a) a goodness of fit test of the CD model to data, and (b) a test that recovery rates are year-independent; (3) to judge the validity of the CD model by analyzing a cross section of waterfowl data; and (4) to reflect on the CD method as evidence for age-specific mortality in waterfowl.

STATISTICAL CONSIDERATIONS

Assumptions and Model

To specify clearly the assumptions and the model underlying the CD method, and the method itself, we must present some mathematical background. First, we introduce our basic notation; for a more detailed discussion of some of these terms and basic background on analysis of recovery data see Anderson (1975), Anderson and Burnham (1976), and Brownie et al. (1978):

k = Number of years of banding,
l = Number of years during which recoveries are recorded, $l \geqslant k$,
$S_i = 1 - M_i$ = Annual survival rate (probability) for birds of age i,
$M_i = 1 - S_i$ = Annual mortality rate (probability) for birds of age i,
f_i = Annual band recovery rate (probability) for birds of age i,
N_i = Number of birds banded in year i, $i = 1, \ldots, k$,
R_{ij} = Number of birds recovered in year j from birds banded in year i, $i = 1, \ldots, k$, $j = 1, \ldots, l$,
$R_{i.}$ = Total number of band recoveries from the ith banded cohort (i.e., birds banded in year i) $= \sum_{j=1}^{l} R_{ij}$,
D'_i = All recoveries of bands exactly i years after banding (corresponds to i years of age for birds banded as young) $i = 1, \ldots, l$, and
TN_i = Total number of banded birds that could have contributed to the recoveries D'_i exactly i years after banding (often called "banded birds available" in year i), $i = 1, \ldots, l$.

The above 2 quantities are computed from the R_{ij} and N_i, respectively. In the case of $k = l$ they are (for $i = 1, \ldots, k$)

$$D'_i = \sum_{j=1}^{k-i+1} R_{j,j+i-1}\,,$$

$$TN_i = \sum_{j=1}^{k-i+1} N_j\,.$$

For example, for $i = 1$, $TN_1 = N_1 + N_2 + \cdots + N_k$ and $D'_1 = R_{11} + R_{22} + \cdots + R_{kk}$, while for $i = k$, $TN_k = N_1$ and $D'_k = R_{1k}$. The mathematical definitions of D'_i and TN_i for $l > k$ are complex and will not be given here. For the general computing formulas of D'_i and TN_i see Anderson and Burnham (1976:55).

Finally we define

$D_i = D_i'/TN_i$ = the proportion of recoveries of all birds aged i, $i = 1, \ldots, l$.

In the wildlife literature, the CD method is often presented only as a sequence of calculations based on the tabulated banding and recovery data. The formula representing these computations for estimating the mortality rate of birds in their ith year of life is:

$$\hat{M}_i = D_i/(D_i + D_{i+1} + \cdots + D_l) \qquad i = 1, \ldots, l-1. \quad (1)$$

Seber (1972) has given a formalization of this method and some of its properties (also see Seber 1971). Our use of the notation D'_i and D_i is the reverse of Seber's use of this same notation.

Given the estimator of formula (1), we specify the exact model necessary for this to produce valid estimates of age-specific mortality rates. It is beyond the scope of this paper to present all the needed background material on modeling band recovery data (see Brownie et al. 1978). Only key results will be given.

The CD procedure is meant to apply only in cases of age-specific survival and recovery rates. Therefore the general model structure in terms of expected band recoveries is

$$E(R_{ij}) = \begin{cases} N_1 f_1 & i = 1 \\ N_i S_1 S_2 \ldots S_{j-i} f_{j-i+1} & i > 1. \end{cases} \quad (2)$$

To complete the statistical model for the recoveries R_{ij}, we note that the recoveries from each banded cohort are multinomial random variables. Thus given N_i birds banded in year i, we have R_{ii}, $R_{i,i+1}$, . . . , R_{il}, $N_i - R_{i\cdot}$ are multinomial random variables with expectations given by (2). Note that:

$$E(R_{i\cdot}) = E(R_{ii}) + E(R_{i,i+1}) + \cdots + E(R_{i,l}).$$

Recognition of this multinomial model as the correct sampling model for the data allows rigorous consideration of estimation of parameters and tests of the assumptions (see Brownie et al. 1978).

We mention that a special case of the CD model is that of assuming constant recovery and survival rates; that is $f_i \equiv f$ and $S_i \equiv S$ for all i.

From formula (2) we derive:

$$E(D_i) = \begin{cases} f_1 & i = 1 \\ S_1 S_2 \ldots S_{i-1} f_i & i > 1. \end{cases} \quad (3)$$

A fundamental property of estimators is that they should be consistent; that is for large samples their expected value converges to the true parameter value (this is related to unbiasedness). Based on formulas (1) and (3) and the requirement of consistency, Anderson and Burnham (1976:54–55) showed that the CD method is valid only if the ratio f_i/M_i is constant for all i. Thus we must assume that:

$$\frac{f_i}{M_i} \equiv a \qquad i = 1, \ldots, l \quad (4)$$

for some constant a. This is a very restrictive assumption, and unless it is tested there is no reason to believe it is true for waterfowl (Anderson and Burnham 1976:55–56). Finally, the CD estimator of the parameter a is

$$\hat{a} = D_1 + D_2 + \cdots + D_l.$$

Estimation

The CD estimation method was developed long before the complete statistical model given above was formalized. The question of optimal inference under the CD assumptions has never been considered in the literature, nor have sampling variances been developed for $\hat{M}_i$ of formula (1). In fact the estimator of (1) can be improved upon in several ways; e.g., standard maximum likelihood theory.

Identifiability of parameters is an important subject, albeit a quite technical one (cf. Brownie et al. 1978:112). A key result in this regard is Seber's (1971) proof that the parameters $M_1, M_2, \ldots, M_l$ and a of the CD model are not even uniquely identifiable. The basic problem is that given any set of banding data or expected recoveries, numerous different sets of parameters under the CD model could generate exactly these same data. For example, let $k = l = 3$; the following parameters will produce exactly the same expected values, $E(R_{ij})$, for any numbers banded N_i:

$$a = 0.1, \quad S_1 = 0.6, \quad S_2 = 0.5, \quad S_3 = 0.4$$

and

$$a = 0.2, \quad S_1 = 0.8, \quad S_2 = 0.8125, \quad S_3 = \frac{11.2}{13} = 0.86154.$$

In order to estimate the parameters of the CD model at least 1 constraint must be imposed. The types of constraint allowable are to either (1) arbitarily specify 1 (or more) mortality rate(s), or (2) set 2 or more mortality rates equal. Seber (1971) assumed $M_{l-1} = M_l$ and derived estimators which differ from those of the CD method. A thorough statistical treatment of this model will require sequential testing for differences in age-specific mortality rates. Thus one would adopt constraints such as $M_j = M_{j+1} = \cdots = M_l$ and test the adequacy of the model for increasing values of j.

The CD method employs the first approach; arbitrarily specifying that $M_l = 1$. This implicit constraint is unrealistic. It also implies that all products such as $S_iS_{i+1} \ldots S_l$ are are zero, and this implies the usually stated condition required for validity of the CD estimators.

Using formula (3) we can show the estimator $\hat{M}_i$ is consistent for the parameter

$$M_i^* = \frac{M_i}{1 - S_iS_{i+1} \ldots S_l} \qquad i = 1, \ldots, l-1.$$

This means that the expected value of $\hat{M}_i$ is approximately M_i^*, with the approximation getting better for large samples of banded birds. Another way to say this is that $\hat{M}_i$ is a biased estimator (it overestimates M_i) unless the product $S_iS_{i+1} \ldots S_l$ is zero. This will be true only if l is much larger than i, i.e., one does not analyze the data until several years after the last cohort is banded (cf. Seber 1972). Two facts are apparent: (1) the bias in $\hat{M}_i$ is due to the implicit constraint $M_l = 1$, and (2) if a realistic constraint such as $M_{l-1} = M_l$ were used, the statistical bias would vanish as sample size increases.

Goodness of Fit Test

Let $\hat{E}(R_{ij})$ be an (asymptotically efficient) estimator of the expected value of the random variable R_{ij} under the assumptions of the CD method. Then the general goodness of fit test for this model is simply the classical chi-square test based on (observed − expected)²/expected. Symbolically the test statistic is:

$$\chi^2 = \left\{ \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{[R_{ij} - \hat{E}(R_{ij})]^2}{\hat{E}(R_{ij})} \right\} + \left\{ \sum_{i=1}^{k} \frac{[R_{i\cdot} - \hat{E}(R_{i\cdot})]^2}{[N_i - \hat{E}(R_{i\cdot})]} \right\}. \qquad (5)$$

The 1st term is based on the recoveries, while the 2nd term accounts for those bands that were never recovered. The quantity $\hat{E}(R_{i\cdot})$ is computed as

$$\hat{E}(R_{i\cdot}) = \sum_{j=i}^{l} \hat{E}(R_{ij}).$$

Under the null hypothesis that the CD model fits the data this test statistic has a chi-square distribution with

$$\frac{k(k+1)}{2} + k(l - k + 1) - l$$

degrees of freedom. Often the expectations for j much greater than i will be very small. If they are less than 2, the chi-square approximation is of doubtful validity. Therefore, it becomes necessary to pool expected values within a row, and corresponding data, if $\hat{E}(R_{ij})$ is less than 2. This results in a valid test, but loses 1 degree of freedom for each pooled R_{ij} value.

Usually the values of $\hat{E}(R_{ij})$ are found by substituting estimated parameter values in the formula for $E(R_{ij})$ (formula 2); in so doing we use $\hat{f}_i = \hat{a}\hat{M}_i$. Using the CD estimators of these parameters (and the constraint $\hat{M}_l \equiv 1$) leads to the simple formulas:

$$\hat{E}(R_{ij}) = N_i D_{j-i+1} \qquad i = 1, \ldots, k$$
$$j = i, \ldots, l.$$

For example,

$$\hat{E}(R_{ii}) = N_i \hat{f}_1 = N_i \hat{a}\hat{M}_i = N_i D_1.$$

We note that $E(R_{ij})$ is estimable without constraints needed to allow identifiability of the parameters. Therefore, the goodness of fit test (5) is not dependent upon how the parameters are estimated. In fact, this test can be computed without estimating any parameters by computing the conditional expectation of R_{ij} given the statistics $D_1, \ldots, D_l$ (many of the goodness of fit tests in Brownie et al. (1978) were constructed following this procedure). It can be shown that:

$$E(R_{ij} \mid D_1, \ldots, D_l) = N_i D_{j-i+1},$$

and mathematically a valid goodness of fit test results from using $\hat{E}(R_{ij}) = E(R_{ij} \mid D_1, \ldots, D_l)$. The validity of this test is unaffected by truncation; i.e., it is valid for all values of k and l.

This goodness of fit test provides a valid test of the 2 assumptions critical to the CD model (the M_i are age-specific only, and $f_i = aM_i$). If the test indicates that the model does not fit we are saying 1 or both of these assumptions is false. Therefore, the method cannot be used in the analysis of data and the estimates of parameters are not useful.

A more specific test can be used to assess the assumption that 1st-year recovery rates are constant, i.e., are age-dependent only, hence not varying with time. Under the assumptions of the CD method we must have $E(R_{ii}) = N_i f_1$ or the ratios R_{ii}/N_i have a constant expectation, independent of the year (i) of banding. A simple k by 2 chi-square contingency table test of this assumption can be written as:

$$\chi^2 = (N.) \sum_{i=1}^{k} \frac{\left[R_{ii} - N_i\left(\frac{R..}{N.}\right)\right]^2}{(R..)(N_i - R..)}, \qquad (6)$$

where $R..$ = total 1st-year recoveries and $N.$ = total number banded $\equiv TN_1$. This test has $k - 1$ degrees of freedom. If this test rejects the null hypothesis of constant first-year recovery rate then the CD model is invalid. This is because of time variation in the 1st-year recovery rates as would result if harvest rates vary over time.

SELECTION OF WATERFOWL DATA

An objective of this paper is to assess the validity of the CD model for waterfowl banding data in general (excepting geese and swans). Thus we want to test the null hypothesis that the CD model generally fits waterfowl data, versus it does not generally fit such data. This is not the same as testing whether the model fits a given set of data. To meet this more general objective a proper scientific approach is to select a representative sample of all banded waterfowl populations and conduct the tests of assumptions on these data sets. The inference is then not just that the model does or does not fit these given data sets but that it does or does not fit waterfowl banding data in general.

To meet this objective we started with a large tabulation of most of the duck banding data in Canada and the United States from 1955 to 1972. Then we established criteria for selecting data sets. The 2 primary criteria were that the data sets selected be representative of (banded) waterfowl populations and that the banded samples be large enough to yield a reliable test of the assumptions.

To be representative of waterfowl banding we sought to cover a variety of geographic areas, ages, sexes, species, and banding periods (i.e., preseason vs. winter). Before selecting these data sets we established some criterion for a "good" data set; by "good" we mean sufficient numbers of birds banded and of recoveries to yield a meaningful test of the CD method assumptions. Our target goals for each data set were to have at least 5 years of banding ($k \geq 5$), an average number banded per year of at least 300, and an average 1st-year recovery rate of at least 3%. In general we met these criteria, however, occasionally we used $k = 4$ and for some data sets average recovery rate was lower than 3% (but in these cases annual numbers banded were in the thousands).

Using these guidelines we selected 45 data sets relating to 9 species from the large tabulation. We emphasize that the tabulation showed only raw data and had never been analyzed by us. Thus we made our selection of data sets before ever knowing the outcome of the tests of assumptions. Once these 45 data sets were selected, we used all of them; there was no subsequent deletion or addition of data sets after the analyses. Thirty-four data sets were from preseason banding (8 adult and 26 young data sets) and 11 sets were from winter banding (all adults). Table 1 provides detailed information on each data set.

RESULTS

Year-specific Recovery Rates

The CD model allows recovery rates to be age-specific only, and these rates must be directly proportional to the age-specific mortality rates. However, waterfowl hunting regulations in North America have fluctuated greatly over the past 25 years (Martin and Carney 1977). It seems logical that recovery rates do vary by year—a violation of a necessary condition of the CD method (Anderson and Burnham 1976:13).

We tested the null hypothesis that 1st-year recovery rates are constant across years using (6); the results appear in Table 1. The test results indicate a sound rejection of the null hypothesis. We conclude that recovery rates in ducks vary significantly by years, a major violation of the CD model assumptions. The idea that year-specific variation is "averaged out" by combining data over a number of years in a "composite" is not correct. For example, this procedure produces severe overestimates of age-specific mortality rates and severe underestimates of their sampling variances. In other words, a highly precise, incorrect estimate is produced.

Goodness of Fit Test of the CD Method

Valid inferences must stem from a model employing valid assumptions. We tested the assumptions necessary for the CD method to be valid using the simple goodness of fit test (formula 5; Table 2). Information in Table 2 provides conclusive evidence that the assumptions of the

Table 1. Results of the test that 1st-year recovery rates are constant across years for 45 sets of waterfowl banding data in North America.

Banding period	Species	State or province	Age and sex	Banding years	k^a	Total banded	Total recoveries	Significance level
Preseason	Mallard	Montana	Adult ♂	1959–69	11	12,105	1,512	0.00189
		Saskatchewan	Adult ♂	1961–72	12	21,015	2,236	0.00000
		Minnesota	Adult ♀	1959–72	14	10,949	1,077	0.00000
		Wisconsin	Adult ♀	1961–72	12	16,803	1,682	0.00000
		Minnesota	Young ♂	1961–72	12	19,645	3,014	0.00000
		New York	Young ♂	1960–72	13	13,466	2,150	0.00000
		Ontario	Young ♂	1965–72	8	19,792	2,790	0.00140
		Saskatchewan	Young ♂	1955–72	18	30,627	4,263	0.00000
		Minnesota	Young ♀	1961–72	12	16,731	1,993	0.00000
		New York	Young ♀	1960–72	13	13,221	1,837	0.00000
		Ontario	Young ♀	1965–72	8	16,417	2,002	0.00019
		Saskatchewan	Young ♀	1955–72	18	28,198	2,770	0.00000
Preseason	Blue-winged teal	Iowa	Young ♂	1963–72	10	15,510	910	0.00000
		Manitoba	Young ♂	1962–72	11	17,115	691	0.00032
		Minnesota	Young ♂	1963–71	9	10,458	468	0.09759
		Ontario	Young ♂	1963–72	10	11,765	523	0.00108
		Iowa	Young ♀	1963–72	10	13,928	947	0.00000
		Manitoba	Young ♀	1962–72	11	15,723	686	0.00270
		Minnesota	Young ♀	1963–71	9	9,683	520	0.00063
		Ontario	Young ♀	1963–72	10	10,977	583	0.00295
Preseason	Pintail	California	Adult ♂	1955–72	18	34,788	3,356	0.00000
		Saskatchewan	Adult ♂	1955–58	4	22,279	1,403	0.00000
		California	Young ♂	1955–58	4	15,405	1,940	0.07258
		Alberta	Young ♀	1965–71	7	6,386	415	0.41488
		California	Young ♀	1955–58	4	10,216	1,015	0.50003
		Saskatchewan	Young ♀	1964–71	8	9,245	464	0.02668
Preseason	Wood duck	Wisconsin	Adult ♂	1962–72	11	8,126	818	0.00041
		Illinois	Young ♂	1962–72	11	11,690	1,081	0.00375
		Iowa	Young ♂	1962–72	11	9,342	1,026	0.00000
		Wisconsin	Young ♂	1960–70	11	9,197	1,067	0.00000
		Illinois	Young ♀	1962–72	11	10,335	716	0.00410
		Iowa	Young ♀	1962–72	11	8,661	855	0.00000
		Wisconsin	Young ♀	1960–70	11	7,894	816	0.00000
Preseason	Lesser scaup	Alaska	Adult ♂	1960–66	7	26,135	1,012	0.00002
Winter	Mallard	Illinois	Adult ♂	1963–72	10	27,691	2,905	0.00000
		Nebraska	Adult ♂	1965–72	8	15,272	1,562	0.00002
		Illinois	Adult ♀	1963–72	10	12,911	787	0.00723
		Nebraska	Adult ♀	1965–72	8	7,880	322	0.53884
Winter	Green-winged teal	California	Adult ♂	1955–70	16	13,731	919	0.00006
Winter	Pintail	California	Adult ♂	1955–72	18	27,727	2,263	0.00001
		California	Adult ♀	1957–72	16	17,747	764	0.01416
Winter	Redhead	New York	Adult ♂	1955–72	13	20,549	1,776	0.00000
		New York	Adult ♀	1963–72	10	4,629	371	0.00146
Winter	Canvasback	New York	Adult ♂	1955–59	5	6,123	808	0.00000
Winter	Wigeon	California	Adult ♂	1955–68	14	16,992	1,754	0.30970
				TOTALS:		685,079	62,869	

[a] Number of years of banding.

Table 2. Results of goodness of fit test for the composite dynamic model and modern methods applied to 45 sets of waterfowl-banding data in North America.

					Composite dynamic model	Modern methods[a]	
Banding period	Species	State or province	Age and sex	Banding years	Significance level	Significance level	Model
Preseason	Mallard	Montana	Adult ♂	1959–69	0.00006	0.59843	1
		Saskatchewan	Adult ♂	1961–72	0.00000	0.18843	1
		Minnesota	Adult ♀	1959–72	0.00000	0.00425	1
		Wisconsin	Adult ♀	1961–72	0.00000	0.00834	1
		Minnesota	Young ♂	1961–72	0.00000	0.17011	0[b]
		New York	Young ♂	1960–72	0.00000	0.00836	0
		Ontario	Young ♂	1965–72	0.00027	0.07432	0
		Saskatchewan	Young ♂	1955–72	0.00000	0.00631	0
		Minnesota	Young ♀	1961–72	0.00000	0.13044	0
		New York	Young ♀	1960–72	0.00000	0.28056	0
		Ontario	Young ♀	1965–72	0.00270	0.58030	0
		Sakatchewan	Young ♀	1955–72	0.00000	0.85574	0
Preseason	Blue-winged teal	Iowa	Young ♂	1963–72	0.00000	0.40978	0
		Manitoba	Young ♂	1962–72	0.00015	0.94543	0
		Minnesota	Young ♂	1963–71	0.05430	0.54376	0
		Ontario	Young ♂	1963–71	0.00042	0.01402	0
		Iowa	Young ♀	1963–72	0.00000	0.23395	0
		Manitoba	Young ♀	1962–72	0.00000	0.63155	0
		Minnesota	Young ♀	1963–71	0.00050	0.06369	0
		Ontario	Young ♀	1963–72	0.01422	0.27058	0
Preseason	Pintail	California	Adult ♂	1955–72	0.00000	0.00003	0
		Saskatchewan	Adult ♂	1955–58	0.00000	0.72319	1
		California	Young ♂	1955–58	0.00000	0.00590	0
		Alberta	Young ♀	1965–71	0.00889	0.01171	0
		California	Young ♀	1955–58	0.01128	0.00677	0
		Saskatchewan	Young ♀	1964–71	0.03752	0.30592	0
Preseason	Wood duck	Wisconsin	Adult ♂	1962–72	0.00000	0.96056	0
		Illinois	Young ♂	1962–72	0.00000	0.06970	0
		Iowa	Young ♂	1962–72	0.00001	0.94273	0
		Wisconsin	Young ♂	1960–70	0.00000	0.42157	0
		Illinois	Young ♀	1962–72	0.00014	0.22382	0
		Iowa	Young ♀	1962–72	0.00000	0.35093	0
		Wisconsin	Young ♀	1960–70	0.00000	0.02904	0
Preseason	Lesser scaup	Alaska	Adult ♂	1960–66	0.00000	0.00064	1
Winter	Mallard	Illinois	Adult ♂	1963–72	0.00000	0.42132	1
		Nebraska	Adult ♂	1965–72	0.00001	0.28064	1
		Illinois	Adult ♀	1963–72	0.00023	0.28552	1
		Nebraska	Adult ♀	1965–72	0.68993	0.61217	1
Winter	Green-winged teal	California	Adult ♂	1955–70	0.00000	0.33859	1
Winter	Pintail	California	Adult ♂	1955–72	0.00000	0.04601	1
		California	Adult ♀	1957–72	0.00000	0.02080	1
Winter	Redhead	New York	Adult ♂	1955–72	0.00000	0.01708	1
		New York	Adult ♀	1963–72	0.00489	0.64865	1
Winter	Canvasback	New York	Adult ♂	1955–59	0.00000	0.52254	1
Winter	Wigeon	California	Adult ♂	1955–68	0.00051	0.11472	0

[a] See Brownie et al. (1978).

[b] Tests of assumptions can be made for bandings of only young birds via Model 0; however, estimation is not possible unless a matching sample of adults is also available for analysis (Brownie et al. 1978:33–34).

CD method are not met (e.g., an adequate fit is rejected for 37 of the 45 data sets [82%] at the 0.001 level). This is an overall test of the assumptions necessary for the CD model. Moreover, the test results are not dependent on the particular estimation method used (the same conclusions would be reached if the constraint $M_{l-1} = M_l$ were used, rather than $M_l \equiv 1$). We conclude that the CD method is not appropriate for the analysis of waterfowl-banding data.

Goodness of Fit Tests of Modern Methods

Brownie et al. (1978) present a series of recently developed (i.e., "modern") estimation models and testing procedures to allow selection of the "best" model for a specific data set. It seems appropriate to assess the assumptions made by these modern methods. Relevant information is presented in Table 2 to allow comparisons. Only 8 data sets (18%) are rejected at the 0.01 level and only 2 data sets (4%) are rejected at the 0.001 level. Additional information of this type is given for mallards by Anderson (1975). These results generally indicate that modern methods are adequte for the analysis of waterfowl data—at the very least, they are a considerable improvement over the older methods. In addition, these modern procedures provide statistical tests (i.e., goodness of fit tests, likelihood ratio tests, and contingency-type tests) to assess the assumptions being made for a particular model. If these tests indicate that the assumptions are not met for a particular data set, the model should not be used for the analysis of that data set. Finally, none of the modern methods is affected by truncation of the recovery data and most are not affected by changes or trends in annual band reporting rates.

Bias in Age-specific Mortality Rate Estimates

We pointed out that bias will occur in the mortality rates estimates $\hat{M}_i$ unless $S_i S_{i+1} \ldots S_l$ is essentially zero. This product has not been zero in the vast majority of analyses of banding data. Truncation has occurred almost without exception (e.g., Moisan et al. 1967; Geis et al. 1971). To illustrate the effect of truncation, consider a case where mortality rate for all ages is 0.4 (i.e., $S_i \equiv S$ for all ages). Expected values of the estimator of mortality rate under the CD method can be computed using

$$E(\hat{M}_i) \doteq \frac{M}{1 - S^{k-i+1}} .$$

Although the true $M_i \equiv 0.4$, the CD method would produce the following average estimates for $i = 1, \ldots, 7$: 0.412, 0.420, 0.434, 0.460, 0.510, 0.625, and 1.000. This bias in the CD estimation procedure has frequently been incorrectly interpreted as evidence for age-specific mortality in waterfowl.

We stress that these biases occur even when the assumptions required by the CD model are true; they are an inherent property of the estimation method (primarily because of the implicit constraint $M_l = 1$). One might conclude that mortality rates increase with age when, in fact, only bias increases with age when the CD method is used.

DISCUSSION AND RECOMMENDATIONS

We contend that valid estimates and sound inference must come from a proper model and estimation procedure based on reasonable assumptions. We have found the CD model to be inappropriate as a basis for analysis of waterfowl banding data because: (1) the assumption that recovery rates do not vary by time is unrealistic; there is substantial direct information that hunting regulations influence harvest and harvest rate, and that recovery rate is closely correlated with harvest rate (Henny and Burnham 1976); (2) goodness of fit tests soundly reject the restrictive assumptions required; and (3) logical inconsistencies occur, as for example in applying the method to adults which are a mix of unknown ages when banded (cf. Anderson and Burnham 1976:55–56). Thus, our main point is that the necessary assumptions behind the CD method are logically and demonstratively invalid. Therefore, the CD method should not be used.

A 2nd point, of lesser significance, is that the estimation method itself is poor even if the necessary assumptions are true. It must be recognized that the $l+1$ parameters $M_1, \ldots, M_l$ and a are not separately estimable unless the mortality rates are subjected to at least 1 appropriate constraint (cf. Seber 1971). The CD method is implicitly based on the constraint $M_l \equiv 1$, which is simply not justified. If there are situations where the CD assumptions are tenable (e.g., recovery data from a few species of nongame birds) improved estimators should be developed for the analyses. We note, however, that the goodness of fit test (5) is valid and independent of the specific estimation method used and is not affected by truncation of the data.

It seems important to recognize that 2 results commonly found in the literature are now suspect. First, the estimates of 1st-year mortality rates of birds banded as young are severely overestimated (e.g., Jessen 1970) because the necessary condition $a = f_i/M_i$ is violated. Second, the notion that mortality increases with age is not validly supported. This notion is an artifact of the CD method due to truncation of data and the implied constraint $M_l = 1$. The concept that mortality in waterfowl is age-specific is not validly supported by analyses using the CD method.

We recommend against the use of the CD method in the analysis of waterfowl banding data. We suggest caution in interpreting published results based on CD method concerning estimate of age-specific mortality rates; differences between mortality rates by sex and geographic areas; and the effect of hunting on total annual mortality rates. We recommend use of the more modern methods. Their advantages are discussed by Anderson and Burnham (1976:18).

LITERATURE CITED

ANDERSON, D. R. 1975. Population ecology of the mallard: V. Temporal and geographic estimates of survival, recovery, and harvest rates. U.S. Fish Wildl. Serv., Resour. Publ. 125. 110pp.

———, AND K. P. BURNHAM. 1976. Population ecology of the mallard: VI. The effect of exploitation on survival. U.S. Fish Wildl. Serv., Resour. Publ. 128. 66pp.

BELLROSE, F. C., AND E. B. CHASE. 1950. Population losses in the mallard, black duck, and blue-winged teal. Ill. Nat. Hist. Surv. Biol. Notes 22. 27pp.

BROWNIE, C., D. R. ANDERSON, K. P. BURNHAM, AND D. S. ROBSON. 1978. Statistical inference from band recovery data: A handbook. U.S. Fish Wildl. Serv., Resour. Publ. 131. 212pp.

EBERHARDT, L. L. 1972. Some problems in estimating survival from banding data. Pages 153–171 *in* Population ecology of migratory birds—A symposium. U.S. Fish Wildl. Serv., Wildl. Res. Rep. 2. 278pp.

GEIS, A. D. 1972*a*. Use of banding data in migratory game bird research and management. U.S. Fish Wildl. Serv., Spec. Sci. Rep. Wildl. 154. 47pp.

———. 1972*b*. Role of banding data in migratory bird population studies. Pages 218–228 *in* Population ecology of migratory birds—A symposium. U.S. Fish Wildl. Serv., Wildl. Res. Rep. 2. 278pp.

———, R. I. SMITH, AND J. P. ROGERS. 1971. Black duck distribution, harvest characteristics, and survival. U.S. Fish Wildl. Serv., Spec. Sci. Rep. Wildl. 139. 241pp.

———, AND R. D. TABOR. 1963. Measuring hunting and other mortality. Pages 284–298 *in* H. S.

Mosby ed. Wildlife investigation techniques. The Wildlife Society, Washington, D.C. 419pp.

HENNY, C. J., AND K. P. BURNHAM. 1976. A mallard reward band study to estimate band reporting rates. J. Wildl. Manage. 40:1–14.

HICKEY, J. J. 1952. Survival studies of banded birds. U.S. Fish Wildl. Serv., Spec. Sci. Rep. Wildl. 15. 177pp.

———. 1972. Population ecology of migratory birds: Symposium summary. Pages 263–271 *in* Population ecology of migratory birds—A symposium. U.S. Fish Wildl. Serv., Wildl. Res. Rep. 2. 278pp.

JESSEN, R. L. 1970. Mallard population trends and hunting losses in Minnesota. J. Wildl. Manage. 34:93–105.

MARTIN, E. M., AND S. M. CARNEY. 1977. Population ecology of the mallard: IV. A review of duck hunting regulations, activity and success, with special reference to the mallard. U.S. Fish Wildl. Serv., Resour. Publ. 130. 137pp.

MOISAN G., R. I. SMITH, AND R. K. MARTINSON. 1967. The green-winged teal: Its distribution, migration, and population dynamics. U.S. Fish Wildl. Serv., Spec. Sci. Rep. Wildl. 100. 248pp.

SEBER, G. A. F. 1971. Estimating age-specific survival rates from bird-band returns when reporting rate is constant. Biometrika 58:491–497.

———. 1972. Estimating survival rates from bird-band returns. J. Wildl. Manage. 34:405–413.

———. 1973. The estimation of animal abundance and related parameters. Charles Griffin and Co. Ltd., London. 506pp.

Received 14 November 1977.
Accepted 27 September 1978.

BIOMETRICS 40, 739–748
September 1984

A General Methodology for Maximum Likelihood Inference from Band-Recovery Data

M. J. Conroy and B. K. Williams

U.S. Fish and Wildlife Service, Patuxent Wildlife Research Center,
Laurel, Maryland 20708, U.S.A.

SUMMARY

A numerical procedure is described for obtaining maximum likelihood estimates and associated maximum likelihood inference from band-recovery data. The procedure allows the specification of a very general form for band-recovery models, by means of which constraints on model parameters can be easily imposed by simple transformations. The method is illustrated for previously developed one-age-class band-recovery models, and is extended to new models, including the analysis with a covariate for survival rates and variable-time-period recovery models. Extensions to R-age-class band-recovery, mark–recapture models, and twice-yearly marking are discussed. A FORTRAN program, available from the authors, provides computations for the models described, and can be adapted for additional user-defined models.

1. Introduction

In band-recovery problems, birds are marked in each of several periods (for example, July–September in a series of years) and the sample consists of recoveries of bands during periods following banding. The probability that a bird banded in Year i is recovered in a subsequent Year j can be modelled as a function of time- and age-specific survival and recovery rates.

Seber (1970) and, in an unpublished report, D. S. Robson and W. D. Youngs (Cornell University Biometrics Unit Paper No. BU-360, 1971) have modelled recoveries with multinomial distributions and have developed maximum likelihood estimates (MLEs) for time-specific survival and recovery probabilities when all animals are marked as adults. Johnson (1974) and Brownie and Robson (1976) extended this methodology to allow age-specific survival and recovery probabilities for birds marked as adults and as young, and Brownie *et al.* (1978) summarized previous models, developed reduced-parameter models, and presented algorithms for estimation and hypothesis testing among a hierarchy of models.

In modelling band-recovery data, we believe that it is desirable to start with a model structure that allows specification of the sources of variability likely to occur in nature. For North American game birds these include, at a minimum, allowance for time-specificity in both survival and recovery probabilities, and where two age classes are considered, for age-specificity as well. All of these features have been incorporated into the models described by Brownie *et al.* (1978). For many well-structured data sets, and for relatively straightforward hypotheses regarding survival and recovery rates, these models have been quite adequate. However, we have encountered many examples of bird-banding experiments in which either variations in the design (for example, banding years were missing) or biological questions addressed (for example, tests of functional relationships between survival rates and exogenous variables) have suggested parameterizations which did not fit into the framework specified by Brownie *et al.* (1978).

Key words: Band recoveries; Hypothesis testing; Maximum likelihood; Multinomial models; Numerical estimation; Ring recoveries.

In the past, unique biological questions or data sets have required special developments for each case. We present a methodology which enables an investigator to specify a hypothesized relationship, H_0, among the parameters in terms of constraints on a very general band-recovery model, H_A. We also show how to obtain MLEs and associated test statistics under H_A and H_0, and we use the procedure to analyze several new models, illustrating each with a data example.

2. Estimation and Hypothesis Testing

2.1 *Definitions*

Initially we restrict our development to the one-age-class band-recovery problem, although R-age-class problems are a straightforward extension. Our notation follows Brownie *et al.* (1978):

N_i,	the number of animals marked and released in Year i, $i = 1, \ldots, k$;
R_{ij},	the number of animals that were marked in Year i and recovered in Year j, $j = i, \ldots, l$;
$R_{i,l+1}$	$\equiv N_i - \sum_{j=i}^{l} R_{ij}$, the number of animals marked in Year i and not recovered;
π_{ij},	the probability that an animal marked in Year i is recovered in Year j;
S_i,	the probability that an animal which is alive at the midpoint of the marking period of Year i survives to the midpoint of the marking period of Year $i + 1$;
f_i,	the probability that a marked animal which is alive at the midpoint of the marking period in Year i is recovered in the subsequent recovery period (for example, hunting season following banding).

We note that for North American game birds, f_i has a straightforward interpretation when the recovery period is a hunting season immediately following the banding period (i.e. 'preseason' banding). In cases in which there is a significant lag between banding and recovery, f_i implicitly incorporates a mortality component.

2.2 *Fully-Parameterized Models, H_A*

The statistical treatment of the band-recovery problem by Seber (1970, 1971) and Brownie *et al.* (1978) was based on the product of independent multinomials

$$\mathrm{pr}(\{R_{ij}\}) = \binom{N_i}{\{R_{ij}\}} \prod_{j=i}^{l+1} \pi_{ij}^{R_{ij}}$$

for i banding periods, $i = 1, \ldots, k$, and j recovery periods, $j = i, \ldots, l$, so the overall likelihood function is

$$L(\{\pi_{ij}\}) = \prod_{i=1}^{k} \mathrm{pr}(\{R_{ij}\})$$

with π_{ij} being a function of $\{f_j, S_i, \ j = 1, \ldots, l, \ i = 1, \ldots, l - 1\}$.

Initially we will consider H_A to be equivalent to Model $M1$ (Brownie *et al.*, 1978, p. 15) in which f_j and S_i are time-specific. For Model $M1$ (H_A)

$$\pi_{ij} \equiv \begin{cases} f_j, & j = i, \\ \prod_{m=i}^{j-1} S_m f_j, & i < j < l + 1, \\ 1 - \sum_{m=i}^{l} \pi_{im}, & j = l + 1. \end{cases} \qquad (1)$$

Since H_A is a fully-parameterized member of the exponential family, it can be shown (Brownie *et al.*, 1978) that for the parameter vector

$$\boldsymbol{\theta}_A \equiv (f_1, \ldots, f_k, S_1, \ldots, S_{k-1}, \alpha_1, \ldots, \alpha_{l-k})$$

with

$$\alpha_j \equiv f_{k+j} \prod_{m=k}^{k+j-1} S_m, \quad j = 1, \ldots, l - k,$$

moment-type estimators and MLEs are identical (Davidson and Solomon, 1974). This gives rise to a simple method of producing closed-form estimators of survival and recovery rates. Model H_A is of particular importance because most other one-age-class models of interest are obtained by constraining the parameters of H_A and will be reduced-parameter models, i.e. the number of parameters estimated is less than the number of elements of the minimal sufficient statistic.

2.3 *Reduced-Parameter Models, H_0*

Brownie *et al.* (1978) described two reduced-parameter models for one-age-class recovery problems. In one of the models, survival rates are constrained to be constant, whereas in the other, both survival and recovery rates are constrained thus. Moment-type estimators cannot be obtained, and closed-form solutions of the likelihood equations do not exist for either model. Since the likelihood must be redefined each time $\boldsymbol{\theta}_A$ is constrained under a particular H_0, it is understandable that only a few reduced-parameter models have been developed. We will show how this procedure can be considerably simplified by exploiting the simple multiplicative structure of the cell probabilities π_{ij}.

A commonly used procedure for obtaining MLEs when explicit solutions to the likelihood equations do not exist is the method of scoring (Kale, 1962; Rao, 1965, pp. 302–309):

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + I^{-1}(\hat{\boldsymbol{\theta}}_i)g(\hat{\boldsymbol{\theta}}_i), \tag{2}$$

where $\hat{\boldsymbol{\theta}}_i$ are the estimates of $\boldsymbol{\theta}$ at the ith iteration. The method requires a starting value $\boldsymbol{\theta}^*$, which in practice can be obtained by guesswork or by estimation from a fully-parameterized model. Also required are expressions for $g(\boldsymbol{\theta})$ and $I(\boldsymbol{\theta})$, where

$$g(\boldsymbol{\theta}) = \sum_{i=1}^{k} \sum_{j=i}^{l+1} \left(\frac{R_{ij}}{\pi_{ij}}\right)\left(\frac{\partial \pi_{ij}}{\partial \boldsymbol{\theta}}\right) \tag{3}$$

and

$$I(\boldsymbol{\theta}) = \sum_{i=1}^{k} N_i \sum_{j=i}^{l+1} \frac{1}{\pi_{ij}} \left(\frac{\partial \pi_{ij}}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial \pi_{ij}}{\partial \boldsymbol{\theta}'}\right). \tag{4}$$

Both functions can thus be obtained simply from first partial derivatives of the cell probabilities. Since the approach taken here involves constraining the parameter space of the general model H_A, we express partial derivatives $\partial\pi_{ij}/\partial\boldsymbol{\theta}_0$ by means of the chain rule

$$\frac{\partial \pi_{ij}}{\partial \boldsymbol{\theta}_0} = \left(\frac{\partial \boldsymbol{\theta}_A}{\partial \boldsymbol{\theta}_0}\right)\left(\frac{\partial \pi_{ij}}{\partial \boldsymbol{\theta}_A}\right). \tag{5}$$

The partial derivatives $\partial\pi_{ij}/\partial\boldsymbol{\theta}_A$ are, of course, constant for any Model H_0 obtained by constraining H_A, so the scoring procedure for Model H_0 is determined by a functional relationship of the form $\boldsymbol{\theta}_A = f(\boldsymbol{\theta}_0)$, with the assumption that the derivatives $\partial\boldsymbol{\theta}_A/\partial\boldsymbol{\theta}_0$ exist.

For example, if H_0 is defined by constant survivorship $S_i = S$, $i = 1, \ldots, k-1$, then

$$\frac{\partial \boldsymbol{\theta}_A}{\partial \boldsymbol{\theta}_0} = \begin{array}{c|cccc|cccc|} & S_1 & S_2 & \cdots & S_{k-1} & f_1 & f_2 & \cdots & f_k \\ \hline S & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \hline f_1 & & & & & 1 & & 0 & \\ f_2 & & & 0 & & & \ddots & & \\ \vdots & & & & & & & \ddots & \\ f_k & & & & & 0 & & & 1 \end{array}$$

After the parameters of a model have been fitted as described, the Pearson goodness-of-fit statistic (cf. Brownie *et al.*, 1978, pp. 19–20; North and Morgan, 1979) can be used to test for model adequacy. It is usually necessary to pool cells because expected counts are small, and alternatives for the classical test have been suggested (cf. North and Morgan, 1979). Models under null hypotheses (H_0) can be compared to that under an alternative (H_A) by likelihood ratio test statistics (Kendall and Stuart, 1961, Ch. 24).

2.4 *Computing Algorithm*

The expressions in §§2.2 and 2.3 have been coded in FORTRAN for an HP-3000 computer. A main program reads the data and initial values for $\boldsymbol{\theta}$, and calls subroutines to compute the statistics needed for the scoring procedure. General forms for the cell probabilities, π_{ij}, and their derivatives for the model structures described in §3 are coded in subroutines. Models under specific null hypotheses are then obtained by $\partial\boldsymbol{\theta}_A/\partial\boldsymbol{\theta}_0$, coded as additional subroutines. Coding is provided for all of the hypotheses described in §3, and for several additional models. Additional hypotheses may be constructed by the addition of brief coding for $\partial\boldsymbol{\theta}_A/\partial\boldsymbol{\theta}_0$. No further modification of the program is required provided H_0 fits into one of the previously coded general forms. Convergence ($\| \hat{\boldsymbol{\theta}}_{i+1} - \hat{\boldsymbol{\theta}}_i \| < \varepsilon$) is rapid (usually <10 iterations) for most models and data sets. Our FORTRAN coding of these algorithms is quite general and should be compatible, with minor modifications, with most FORTRAN compilers. A documented listing and detailed instructions are available upon request from the authors.

3. Specific Models

3.1 *Unequal Time Intervals Between Sampling Periods*

Background and development. Brownie *et al.* (1978, p. 179) briefly considered the case in which intervals between marking periods are of variable length. This situation often occurs with species for which marking efforts have been sporadic, resulting in gaps of missing years in the records. Survival rates, S_i, may still be estimated, but S_i is now defined as the probability of survival over an interval $[t_i, t_{i+1}]$ of length d_i, where d_i may vary with i. A parameter of interest is the single-interval survival rate, ϕ_i, where S_i and ϕ_i are related by

$$S_i = \phi_i^{d_i}. \tag{6}$$

It is assumed here that the survival rate S_i is simply a product of the survival rates ϕ_{ij} over each of the d_i intervals in $[t_i, t_{i+1}]$:

$$S_i = \prod_{j=1}^{d_i} \phi_{ij}.$$

Setting $\phi_{ij} = \phi_i$ yields (6). Estimates of ϕ_i, $i = 1, \ldots, k-1$, may be obtained from $\boldsymbol{\theta}_A$ by the relationship $\phi_i = (S_i)^{1/d_i}$. Similarly, $\text{var}(\hat{\phi}_i)$ may be obtained by application of the delta

Table 1
Band recoveries from male ring-necked ducks banded in Maryland, Virginia and North Carolina, December–February 1953–1964

Banding year	Interval (years)	Number banded	Recovery year							
			1953	1955	1956	1958	1960	1961	1963	1964
1953	2	220	9	8	1	0	0	0	0	0
1955	1	386		36	16	4	1	0	1	0
1956	2	223			16	2	1	0	1	1
1958	2	102				6	1	0	2	0
1960	1	212					11	2	3	2
1961	2	433						16	7	7
1963	1	140							4	1
1964	1	313								12

method (Seber, 1973, p. 7). However, tests of H_0: $S_i = S$ make little biological sense, because S_i depends on d_i as well as on ϕ_i, and H_0 will usually be rejected unless $d_i = d$. A more useful test is that of H_0: $\phi_i = \phi$. Reparameterization of H_A by (6) yields

$$\boldsymbol{\theta}_A = (f_1, f_2, \ldots, f_k,\ \phi_1, \phi_2, \ldots, \phi_{k-1}), \tag{7}$$

and under H_0: $\phi_i = \phi$ we have

$$\boldsymbol{\theta}_0 = (f_1^*, f_2^*, \ldots, f_k^*, \phi),$$

where

$$f_i^* = f_i, \quad S_i = \phi^{d_i}, \quad i = 1, \ldots, k.$$

The likelihood function and its derivatives follow from application of the methods in §2.3, and the relationship between $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_0$.

Example. Ring-necked ducks (*Aythya collaris*) were trapped and banded in December–February in Maryland, Virginia and North Carolina. After 1964, several contiguous years of adequate bandings enabled estimation of annual survival rates (Conroy and Eberhardt, 1983). There were several years between 1953 and 1964 in which few if any ringnecks were banded in these areas, resulting in gaps in the data (Table 1). We used the model defined in this section to estimate f_i and ϕ_i for these data (Table 2) and to test the null hypothesis that annual survival rates, ϕ_i, were temporally constant (Table 3). The models under both hypotheses fit the data ($P > .10$), and the hypothesis of constant ϕ was not rejected

Table 2
Maximum likelihood estimates of survival rates, ϕ_i, and band-recovery rates, f_i, for ring-necked-duck data in Table 1

Banding year	H_A: ϕ_i unconstrained				H_0: $\phi_i = \phi$			
	$\hat{f}_i$	$\widehat{\mathrm{SE}}$	$\hat{\phi}_i$	$\widehat{\mathrm{SE}}$	$\hat{f}_i$	$\widehat{\mathrm{SE}}$	$\hat{\phi}_i$	$\widehat{\mathrm{SE}}$
1953	0.0409	0.0134	0.5218	0.0918	0.0403	0.0132	0.6120	0.0346
1955	0.0987	0.0148	0.5477	0.1609	0.0937	0.0137	0.6120	0.0346
1956	0.0706	0.0159	0.5165	0.1191	0.0648	0.0114	0.6120	0.0346
1958	0.0529	0.0194	0.6447	0.1536	0.0413	0.0125	0.6120	0.0346
1960	0.0457	0.0132	0.5656	0.2014	0.0439	0.0119	0.6120	0.0346
1961	0.0297	0.0074	1.0530	0.2588	0.0289	0.0068	0.6120	0.0346
1963	0.0222	0.0103	0.3533	0.2028	0.0476	0.0119	0.6120	0.0346
1964	0.0383	0.0109			0.0422	0.0091	0.6120	0.0346
Average estimate:								
	0.0499	0.0048	0.6004	0.0400	0.0503	0.0042	0.6120	0.0346

Table 3

Likelihood ratio test of hypothesis that survival rates are constant for ring-necked-duck data in Table 1

Model	df	χ^2	P
H_A	4	6.655	.155
H_0	12	12.496	.407
H_0 vs H_A	6	8.543	.201

($P > .10$), suggesting a temporally constant annual survival rate of .61 with an estimated asymptotic standard error of .03.

3.2 *Survival Rate as a Function of a Covariate*

Background and development. In this model, annual survival rates, S_i, are hypothesized to be a function of an independent variable, X_i. Examples of X_i could include weather conditions, availability of habitats or food, and densities of conspecifics, competitors or predators. Previous investigations of the effects of environmental variables on S_i have generally involved estimation of S_i followed by regression of $\hat{S}_i$ on X_i. This approach makes less effective use of the data, and requires two separate analyses. Furthermore, because of autocorrelation among the $\hat{S}_i$, the usual assumptions for regression are not met. North and Morgan (1979) modelled annual survival rates of herons as a logistic function of winter temperature, but required assumptions about recovery probabilities which are unlikely to be appropriate for North American birds that are hunted. Our approach is to reparameterize H_A under the hypothesis

$$H_0\colon S_i = g(X_i), \quad i = 1, \ldots, k. \tag{8}$$

In particular, it may be reasonable to hypothesize a linear relation between S_i and X_i,

$$H_0\colon S_i = a + bX_i,$$

or between S_i' and X_i,

$$H_0\colon S_i' = a + bX_i,$$

where

$$S_i' = \ln\{S_i/(1 - S_i)\}.$$

This approach requires estimation of two parameters, a and b, in addition to f_i, as opposed to the estimation of S_i followed by the appropriate linear regression or linear logistic regression. The parameters under H_0 are

$$\boldsymbol{\theta}_0 = (f_1, f_2, \ldots, f_k, a, b),$$

and specification of the mathematical relationship between S_i and X_i in terms of a and b enables derivation of the likelihood and the maximum likelihood inference.

Example. Adult female mallards (*Anas platyrhynchos*) were banded in Manitoba, Saskatchewan and eastern Alberta in August–September 1966–1977. It was believed that annual survival rates over the interval August i–August $i + 1$ were negatively influenced by high densities of mallards per pond in May of Year $i + 1$ (Nichols, Pospahala and Hines, 1982). A plot of $\hat{S}_i$ (under the fully-parameterized model) suggested a linear relation

$$H_0\colon S_i = a + bX_i,$$

where S_i is the annual survival rate of adult females (August i–August $i + 1$), X_i is the number of mallards per pond in May $i + 1$, and a and b are parameters to be estimated.

Table 4

Band recoveries from adult female mallards banded in prairie Canada, July–September 1966–1979, *and densities of ducks per pond for May of following years*

Banding year	Number banded	Recovery year														Ducks/Pond
		1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	
1966	926	39	20	6	3	7	2	2	1	0	0	0	0	0	0	0.9787
1967	1413		53	18	17	9	8	5	4	0	0	1	0	0	0	2.2445
1968	1147			34	27	23	11	5	3	0	2	1	0	0	0	1.2638
1969	1233				57	33	21	14	5	2	1	2	0	1	0	1.0898
1970	1674					82	35	22	8	2	4	3	1	1	0	1.3363
1971	1727						71	30	13	1	8	6	2	0	2	1.3067
1972	1864							67	32	12	12	4	4	3	0	2.2601
1973	1438								46	23	21	13	5	4	7	0.6647
1974	1235									43	27	15	10	7	4	0.8903
1975	2351										67	56	26	13	10	1.3183
1976	5215											180	91	56	33	2.1163
1977	5256												167	89	61	0.8707
1978	3615													114	66	

Table 5
Maximum likelihood estimates of survival rates, S_i, and band-recovery rates, f_i, for mallard data in Table 4

Banding year	H_A: S_i unconstrained				H_{01}: $S_i = a + bX_i$		H_{02}: $S_i = S$	
	$\hat{f}_i$	$\widehat{\text{SE}}$	$\hat{S}_i$	$\widehat{\text{SE}}$	$\hat{f}_i$	$\widehat{\text{SE}}$	$\hat{f}_i$	$\widehat{\text{SE}}$
1966	0.0421	0.0066	0.5440	0.0962	0.0418	0.0066	0.4184	0.0066
1967	0.0381	0.0047	0.4686	0.0698	0.0364	0.0042	0.0371	0.0043
1968	0.0283	0.0041	0.5890	0.0767	0.0273	0.0036	0.0253	0.0033
1969	0.0430	0.0048	0.7130	0.0864	0.0415	0.0041	0.0405	0.0040
1970	0.0452	0.0043	0.6376	0.0795	0.0475	0.0038	0.0483	0.0039
1971	0.0380	0.0039	0.5428	0.0710	0.0403	0.0033	0.0409	0.0034
1972	0.0364	0.0037	0.4283	0.0578	0.0357	0.0030	0.0362	0.0030
1973	0.0348	0.0040	0.5488	0.0761	0.0328	0.0033	0.0297	0.0028
1974	0.0280	0.0036	0.7898	0.0997	0.0236	0.0026	0.0242	0.0027
1975	0.0303	0.0030	0.6211	0.0622	0.0307	0.0027	0.0325	0.0028
1976	0.0346	0.0023	0.5712	0.0489	0.0352	0.0021	0.0360	0.0022
1977	0.0309	0.0021	0.5904	0.0591	0.0329	0.0020	0.0311	0.0018
1978	0.0304	0.0025			0.0298	0.0018	0.0306	0.0019
Average estimate:	0.0355	0.0011	0.5864	0.0118	0.0351	0.0010	0.0349	0.0010
					$\hat{a}$:	0.7514		
					$\widehat{\text{SE}}$:	0.0554	$\hat{S}$:	0.5879
							$\widehat{\text{SE}}$:	0.0099
					$\hat{b}$:	−0.1170		
					$\widehat{\text{SE}}$:	0.0380		

Three models were analyzed: H_A, the fully-parameterized model with S_i and f_i varying with time; H_{01}, a reduced-parameter model with $S_i = a + bX_i$; and H_{02}, a reduced-parameter model with $S_i = S$, this model being equivalent to H_0: $b = 0$ in the linear parameterization. The relevant banding, recovery and covariate data are given in Table 4. Parameter estimates and associated standard errors for the models are shown in Table 5.

Test results indicate that all three models provide an adequate fit ($P > .50$) to the data (Table 6). The test of H_{01} versus H_A was not rejected ($P > .10$), which suggests that modelling survival rates by year-specific values, S_i, provides no more information than does the covariate model. The tests of H_{02} versus H_A and H_{01} were both rejected ($P < .01$), which suggests that survival rates are not temporally constant and that a significant linear relationship exists between S_i and X_i. A similar inference may be obtained by regression of the $\hat{S}_i$ under H_A against X_i, but this requires two analyses and the estimation of 12 additional parameters.

Table 6
Likelihood ratio tests of hypotheses that survival rates are constant, and that survival rates are a linear function of ducks/pond for the mallard data in Table 4

Model	df	χ^2	P
H_A	51	38.009	.916
H_{01}	62	52.849	.619
H_{02}	62	55.553	.517
Comparison			
H_{01} vs H_A	10	14.020	.172
H_{02} vs H_A	11	22.360	.022
H_{02} vs H_{01}	1	8.340	.004

4. Extensions

In addition to the models described in §3, we have extended this methodology to several other types of model. For example, if birds are marked twice a year, it is possible to estimate survival rates over portions of the year. In hunted species, particular interest may focus on the proportion of mortality, $1 - S_i$, that occurs during the hunting season (usually September–February) and during the remainder of the year (cf. Brownie *et al.*, 1978, p. 159).

If age classes can be identified at the time of banding, then the data and corresponding parameters can be stratified accordingly. Estimation of parameters and tests of hypotheses (for example, constant but age-specific survival rates) are a straightforward extension of the methods described in §2, and several two-age-class models have been programmed in our algorithms. Furthermore, although we assumed in §2 that the banded samples, N_i, were known and we conditioned the observations on them, our methodology could be extended to cases in which N_i are not known and the likelihood is conditioned on the total number of recoveries (Lebreton, 1977; North and Morgan, 1979).

We have also successfully employed a two-stage iterative procedure for obtaining solutions to fully-parameterized and reduced-parameter Jolly–Seber mark–recapture problems (Jolly, 1965, 1982; Seber, 1965). However, this procedure can be very sensitive to starting values and may not prove as useful for these problems as for band-recovery problems.

5. Discussion

Modelling band-recovery data in a biologically realistic manner often requires a large number of parameters. One application of our methodology is to reduce the number of parameters needed to model a population. It is important, however, that a reduction in parameters be pursued in a framework that allows objective comparison of biological hypotheses. The strength of our approach lies in the simplicity with which models under various null hypotheses can be obtained and compared by maximum likelihood methods. Once a sufficiently general model, H_A, has been specified, estimation and testing for biological hypotheses, H_0, are based on the mathematical relationship between H_A and H_0. This simple relationship avoids the necessity of developing new likelihood equations each time a new H_0 is specified. Because reduced-parameter models require numerical solutions, a procedure for testing alternative models, which requires minimum coding, has obvious advantages. In our algorithm we provide the coding for a very general model structure and for several models under specific hypotheses. Additional models, not described above, can be obtained with a minimal amount of extra coding. Users of this procedure are welcome to expand this methodology to other hypotheses, and we will provide assistance as requested.

ACKNOWLEDGEMENTS

Earlier discussions with K. Burnham motivated us to use Equations (3) and (4). We are most grateful to R. Blackwell, J. Hines and F. Fiehrer for assistance in coding the algorithm in FORTRAN. R. Wilcox provided technical assistance on the manuscript, and C. Brownie, J. Hines, J. Nichols, D. Otis and K. Pollock reviewed earlier drafts.

RÉSUMÉ

On décrit une procédure numérique permettant d'obtenir les estimations du maximum de vraisemblance et les statistiques associées adaptées à des données de reprises d'individus bagués. La procédure permet de travailler sur des modèles très généraux; des contraintes sur les paramètres du modèle peuvent être facilement imposées par des transformations simples. La méthode est illustrée sur des modèles à une classe d'âge déjà développés; elle est étendue à de nouveaux modèles incluant l'analyse avec une covariable pour des modèles de taux de survie et de reprise variable avec le temps. On discute d'extensions à la reprise d'individus bagués appartenant à R classes d'âge, de modèles de marquage–recapture et de marquage deux fois par an. Un programme FORTRAN, disponible auprès des auteurs permet de traiter les modèles décrits; il permet de traiter d'autres modèles définis par l'utilisateur.

REFERENCES

Brownie, C., Anderson, D. R., Burnham, K. P. and Robson, D. S. (1978). *Statistical Inference from Band Recovery Data—A Handbook.* U.S. Fish and Wildlife Service Resource Publication No. 131. Washington, D.C.: U.S. Department of the Interior.

Brownie, C. and Robson, D. S. (1976). Models allowing for age-dependent survival rates for band-return data. *Biometrics* **32,** 305–323.

Conroy, M. J. and Eberhardt, R. T. (1983). Variations in survival and recovery rates of ring-necked ducks. *Journal of Wildlife Management* **47,** 127–137.

Davidson, R. R. and Solomon, D. L. (1974). Moment-type estimation in the exponential family. *Communications in Statistics* **11,** 1101–1108.

Johnson, D. H. (1974). Estimating survival rates from banding of adult and juvenile birds. *Journal of Wildlife Management* **38,** 290–297.

Jolly, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration—stochastic model. *Biometrika* **52,** 113–128.

Jolly, G. M. (1982). Mark-recapture models with parameters constant in time. *Biometrics* **38,** 301–322.

Kale, B. K. (1962). On the solution of the likelihood equations by iterative processes: The multiparametric case. *Biometrika* **49,** 479–486.

Kendall, M. G. and Stuart, A. (1961). *The Advanced Theory of Statistics. Vol.* 2, *Inference and Relationship.* New York: Hafner.

Lebreton, J. D. (1977). Maximum likelihood estimations of survival rates from bird band returns: some complements to age-dependent methods. *Biometrie-Praximetrie* **117,** 145–161.

Nichols, J. D., Pospahala, R. S. and Hines, J. E. (1982). Breeding-ground habitat conditions and the survival of mallards. *Journal of Wildlife Management* **46,** 80–87.

North, P. M. and Morgan, B. J. T. (1979). Modelling heron survival using weather data. *Biometrics* **35,** 667–681.

Rao, C. R. (1965). *Linear Statistical Inference and its Application.* New York: Wiley.

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

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

Seber, G. A. F. (1971). Estimating age-specific survival rates from bird-band returns when the reporting rate is constant. *Biometrika* **58,** 491–497.

Seber, G. A. F. (1973). *The Estimation of Animal Abundance and Related Parameters.* London: Griffin.

Received February 1982; *revised May and September* 1983

SIMULTANEOUS ANALYSIS OF BAND-RECOVERY AND LIVE-RECAPTURE DATA

STEPHANIE ZOWADA MARDEKIAN[1] and LYMAN MCDONALD, Departments of Statistics and Zoology, University of Wyoming, Laramie, WY 82071.

Brownie et al. (U.S. Fish and Wildl. Serv., Resour. Publ. 131, 1978) presented 14 interrelated models for the analysis of band-recovery data. In addition, they discussed computer programs ESTIMATE (for animals banded as adults) and BROWNIE (for simultaneous banding of 2 or 3 age-groups) that fit the data to the various models and computed estimates of survival and recovery rates. The banding studies on which such models can be used involve capturing, banding, and releasing a sample of animals at the same location at regular intervals (usually 1 year). For k such banding occasions, let $N_1, N_2, \ldots, N_k$ be the numbers of adults banded and released. Further, let R_{ij} be the number of band recoveries in period j from adult animals banded on the ith occasion. In general, there are ℓ periods of recover, $\ell \geq k$. Brownie et al. (1978) expressed this recovery information conveniently in matrix form. When possible, the notation used by Brownie et al. (1978) will be followed in this paper, and it is assumed that the reader is familiar with their work.

The fate of a banded animal will be 1 of 3 mutually exclusive events. Each animal will (1) be killed or found dead and the band reported, (2) return to the banding site and be recaptured alive in a subsequent year, or (3) not be seen again and the band not reported. An animal could be recaptured alive on any number of occasions, and its ultimate fate be case 1 above. Because the models used to analyze band-recovery data are based on the multinomial distribution, we will be able to use only the *last* time an animal (band) is seen (reported); i.e., each animal is recorded only once in the summary—either as a live recapture on the occasion of its last recapture or as a band recovery, but not both.

Analysis of recovery data alone is well developed, and the same is true for recapture data. It is only the joint analysis of recovery and recapture data from a single banding program that is not well developed in the literature. Brownie et al. (1978, Section 8.2) dealt with this problem, but suggested including the intermediate recaptures and applying a recovery model (for bands from dead animals) to these recapture data. That approach keeps all the recapture data, and uses a model that is an approximation. To get an exact model in which the variance formulae for estimated survival rates are valid, this paper discards the intermediate recapture data. The proposed technique cannot be optimal because the intermediate recaptures are ignored. However, it is a simple and relatively efficient method of combining recoveries and recaptures and the existing computer programs ESTIMATE and BROWNIE. These programs can be used to examine the data.

MODELS AND ANALYSIS

To examine the method as applied to the models for animals banded as adults,

---

[1] Present address: Ortho Pharmaceutical Corp., Quality Assurance Division, U.S. Highway 202, Raritan, NJ 08869.

Table 1. Expected numbers of band recoveries under Model 1, with $k = 3$ and $\ell = 4$.

Banding occasion	Number banded	Period of recovery (between banding occasions) 1	2	3	4
1	N_1	N_1f_1	$N_1S_1f_2$	$N_1S_1S_2f_3$	$N_1S_1S_2S_3f_4$
2	N_2		N_2f_2	$N_2S_2f_3$	$N_2S_2S_3f_4$
3	N_3			N_3f_3	$N_3S_3f_4$

we will use Model 1 (Brownie et al. 1978) as the primary example. The assumptions of Model 1 are (1) age-independent survival and recovery rates and (2) time-specific survival and reporting rates independent of the date of banding. Parameters f (recovery rate) and S (survival rate) are subscripted to indicate dependence on a specific time period. For illustration, the general form of Model 1 in terms of 3 years of banding and 4 years of recovery is given (Table 1).

The corresponding expected numbers of terminal live recaptures of the banded animals is defined to be Model 1′ (Table 2). The parameter S again is subscripted to indicate dependence on a specific time period. It is assumed that all multiple recaptures have been eliminated, so the new parameter, p_i, is the probability of an animal returning to the banding site, being recaptured on the ith banding occasion, and neither the animal nor its band being reported again. In this model there are no live recaptures during the 1st period after banding, because recaptures will occur only when the banding operation is carried out again. Note that on the 3rd banding occasion the parameters S_2p_3 occur together, on the 4th occasion S_3p_4, and in general on the jth occasion ($j \geq 2$) $S_{j-1}p_j$ occur together. These parameters are estimable as a product, but not separately.

Comparing Model 1′ with Model 1, it can be seen that they are of the same form for estimation of survival rates S_1, S_2, . . . , S_{k-1} if we ignore the blanks on the diagonal and shift the columns of Model 1′ one place to the left. The term $q_{i+1} = S_ip_{i+1}$ in Model 1′ corresponds to the recovery rate, f_i, in Model 1. In general, existing computer programs for analysis of band recoveries can be used to analyze these recapture data. However, consider the addition of corresponding elements of the data matrices; that is, add the number of band recoveries in the jth period, R_{ij}, to the number of terminal live recaptures on the (j + 1)th banding occasion, R'_{ij}. Factoring out the common terms, define Model 1* (Table 3) for the expected value of $R_{ij} + R'_{ij}$, the "total" recoveries in the jth recovery period (i.e., period j, or on the [j + 1]th banding occasion) from the ith banding occasion. The survival rates, S_1, S_2, . . . , S_{k-1}, are separately estimable and can be directly

Table 2. Expected numbers of terminal live recaptures under Model 1′, with $k = 3$ and $\ell = 4$.

Banding occasion	Number banded	Banding occasion 1	2	3	4
1	N_1		$N_1S_1p_2$	$N_1S_1S_2p_3$	$N_1S_1S_2S_3p_4$
2	N_2			$N_2S_2p_3$	$N_2S_2S_3p_4$
3	N_3				$N_3S_3p_4$

Table 3. Model 1* for expected numbers of terminal live recaptures and band recoveries, with $k = 3$ and $\ell = 3$.

Banding occasion	Number banded	Recovery period 1	2	3
1	N_1	$N_1(f_1 + q_2)$	$N_1S_1(f_2 + q_3)$	$N_1S_1S_2(f_3 + q_4)$
2	N_2		$N_2(f_2 + q_3)$	$N_2S_2(f_3 + q_4)$
3	N_3			$N_3(f_3 + q_4)$

obtained from the computer program ESTIMATE. The 2nd set of parameters, $f_i + q_{i+1}$, is estimable; however, there is no clear, useful interpretation.

The other models (i.e., 0, 2, and 3) reported by Brownie et al. (1978) can be similarly transformed to, say, models 0*, 2*, and 3* for combined recoveries and recaptures. Model 2* assumes constant survival rate period to period, and thus is the same as Model 1*, except that there are no subscripts on S. Model 3* is too restrictive to be applicable in many natural cases, because it assumes that not only survival rates, but also recovery rates and probabilities of returning and being recaptured, remain constant from period to period. Model 0*, which allows newly banded animals to have different recovery rates than those previously banded, is the same as Model 1*, with the quantity $f_i' + q_{i+1}'$ replacing $f_i + q_{i+1}$ on the diagonal elements of the combined recovery and recapture matrix.

Several of the authors' colleagues and 1 of the referees have suggested that with recapture data taken during banding operations at the same (small) site, if an animal does not return exactly to that site in subsequent years, it is no longer in the sampled population, as some emigration is possible. If E represents the probability of emigration, the suggestion is that you might not estimate survival rate S, but rather $S \cdot (1 - E)$, assuming the events are independent. The models herein propose to solve this problem by embedding the term $(1 - E)$ in p. That is, consider the 3 events: (A) an animal returns to the banding site; (B) an animal is recaptured, given (A); and (C) neither the animal nor its band is reported again, given (A) and (B). If $(1 - E)$, B', and C' represent the probabilities of the 3 events, A, B and C, respectively, then $p = (1 - E) \cdot B' \cdot C'$. Under the assumption that the parameter p has the same properties as the recovery rate, f (e.g., in models 1, 1′, and 1* they are to be independent of the date of banding), then the survival rates are estimable by the maximum likelihood method developed by Brownie et al. (1978). This argument is based totally on deductive logic. It would also be advisable to conduct a statistical test of the equality of survival rates in the 2 data sets before combining them. Unfortunately such a test is not available, and development of one is beyond the scope of this note. We hope that interest in this aspect of the analysis is stimulated.

Each banded animal should be recorded only once in the combined recovery and recapture matrix for the multinomial likelihood functions to be applicable. The usual implicit assumption is made that the recovery of bands from, and live recaptures of, banded birds yield a representative sample from the overall banded population. Also, one assumes the banded population is representative of the whole population.

Analysis of the combined recoveries and recaptures under models 0*, 1*, 2*, and 3* (and choice among the models)

Table 4. Expected numbers of recoveries and terminal recaptures under $H_1{}^*$, with $k = 3$ and $\ell = 3$.

Banding occasion	Number banded	Recovery period		
		1	2	3
Birds banded and released as adults				
1	N_1	$N_1(f_1 + q_2)$	$N_1S_1(f_2 + q_3)$	$N_1S_1S_2(f_3 + q_4)$
2	N_2		$N_2(f_2 + q_3)$	$N_2S_2(f_3 + q_4)$
3	N_3			$N_3(f_3 + q_4)$
Birds banded and released as young				
1	M_1	$M_1(f_1' + q_2')$	$M_1S_1'(f_2 + q_3)$	$M_1S_1'S_2(f_3 + q_4)$
2	M_2		$M_2(f_2' + q_3')$	$M_2S_2'(f_3 + q_4)$
3	M_3			$M_3(f_3' + q_4')$

would be difficult were it not that program ESTIMATE exists for analysis of band recoveries under models 0, 1, 2, and 3, and is directly applicable. All tests of fit and comparisons of models are valid, as are the point and interval estimates of survival rates.

Brownie et al. (1978) also considered the case when there are 2 or 3 age-classes of banded animals (e.g., young, subadults, and/or adults). A similar method of combining band recoveries with live recaptures can be used in these cases. We will consider 2 age-classes and the model under H_1. Under H_1, annual survival and recovery rates (S and f) are time-specific, and young and adults have different survival and recovery rates. For the young, the symbols S' and f' are used. It is assumed that after the 1st period of life, a young animal becomes an adult.

For the expected number of terminal live recaptures, we define the model H_1'. Parameters are defined as in H_1, with 2 new parameters, p_i and p_i', representing the probability of returning to the banding site, being recaptured, and the band not being reported again, for adult and young birds, respectively. Letting $S_{i-1}p_i = q_i$ and $S_{i-1}'p_i' = q_i'$, the recoveries and recaptures can be added together as before, yielding Model $H_1{}^*$ for combined recoveries and recaptures in a given period (Table 4).

Survival rates are separately estimable,

Table 5. Recovery matrices for band recoveries and terminal live recaptures of Canada geese banded at Turbid Lake, Yellowstone National Park, Wyoming, 1967–71.

Banding occasion	Number banded	Year recovered					
		1967	1968	1969	1970	1971	1972
		Band recoveries from birds shot or found dead					
1967	198	30	14	7	8	5	
1968	157		24	11	10	8	
1969	119			10	6	9	
1970	275				34	13	
1971	249					26	
		Terminal live recaptures					
1967	198		1	29	14	4	8
1968	157			17	9	7	4
1969	119				19	12	4
1970	275					42	16
1971	249						43

Table 6. Total recoveries and terminal recaptures of Canada geese banded at Turbid Lake, Yellowstone National Park, Wyoming, 1967–71.

Banding occasion	Number banded	Recovery period				
		1	2	3	4	5
1967	198	31	43	21	12	13
1968	157		41	20	17	12
1969	119			29	18	13
1970	275				76	29
1971	249					69

whereas parameters $f_i + q_{i+1}$ and $f_i' + q_{i+1}'$ are estimable as sums. In a similar manner, the other models considered by Brownie et al. (1978) can be defined for combined recoveries and recaptures. The computer program BROWNIE can be used to fit the new models.

APPLICATION

Band-recovery and live-recapture data are available for normal wild Canada geese (*Branta canadensis*) banded at Turbid Lake in Yellowstone National Park, Wyoming, from 1967 to 1971. Recovery matrices are given for band recoveries and live recaptures (Table 5) and for the combined data (Table 6) (Mardekian, M.S. Thesis, Univ. Wyoming, Laramie, 1979). All intermediate recaptures have been eliminated.

Each of the 3 recovery matrices was analyzed with program ESTIMATE. Model 2 best described the band-recovery data, giving an estimate of the constant survival rate (estimate ± SE) of 69.41 ± 3.53%. On the live-recapture data, Model 1 fit best; the estimated mean survival rate ($\bar{x}$ ± SE) is 65.14 ± 5.35%. On the combined data set, ESTIMATE fit Model 2 (Model 2* in the new notation), with an estimate of constant survival rate of 66.45 ± 2.93%. As expected, this estimate is between the other 2, but has a smaller standard error.

SUMMARY

Models are developed for the analysis of combined band-recovery and live-recapture data. They are similar in structure to the band-recovery models reported in Brownie et al. (1978), and provide an improvement in the estimation of survival rates under the assumption that the recovery of bands from, and live recaptures of, banded birds yields a representative sample from the population. An example is given of the application of these models, using existing computer programs for the analysis.

Acknowledgments.—This work was supported by a grant from the Wyoming Game and Fish Department. We thank G. F. Wrakestraw, D. Saul, and L. Serdiuk of the Wyoming Game and Fish Department for their help and advice. Significant contributions also were made by the editor and referees.

Received 4 June 1979.
Accepted 27 May 1980.

Made in United States of America
Reprinted from JOURNAL OF FIELD ORNITHOLOGY
Vol. 51, No. 1, Winter 1980

THE EFFECT OF BAND LOSS ON ESTIMATES OF ANNUAL SURVIVAL

BY LOUIS J. NELSON, DAVID R. ANDERSON, AND KENNETH P. BURNHAM

Banding has proven to be a useful technique in the study of population dynamics of avian species. However, band loss has long been recognized as a potential problem (Hickey, 1952; Ludwig, 1967). Recently, Brownie et al. (1978) presented 14 models based on an array of explicit assumptions for the analysis of band recovery data. Various estimation models (assumption sets) allowed survival and/or recovery rates to be (a) constant, (b) time-specific, or (c) time- and age-specific. Optimal inference methods were employed and statistical tests of critical assumptions were developed and emphasized.

The methods of Brownie et al. (1978), as with all previously published methods of which we are aware, assume no loss of bands during the study. However, some band loss is certain to occur and this potentially biases the estimates of annual survival rates whatever the analysis method. A few empirical studies have estimated band loss rates (a notable exception is Ludwig, 1967); consequently, for almost all band recovery data, the exact rate of band loss is unknown. In this paper we investigate the bias in estimates of annual survival rates due to varying degrees of hypothesized band loss. Our main results are based on perhaps the most useful model, originally developed by Seber (1970), for estimation of annual survival rate. Inferences are made concerning the bias of estimated survival rates in other models because the structure of these estimators is similar.

METHODS

Our specific results are based on Seber's (1970) model (see Model 1 in Brownie et al., 1978:15–20) but with allowance for a general band retention function (Table 1). Model 1 and its assumptions regarding time-specific survival and recovery rates are specified by three sets of parameters:

N_i = Number of adult birds banded in year i,
S_i = Annual survival rate in year i (specifically, the probability that a bird alive at the beginning of year i will survive until year i + 1), and
f_i = Band recovery rate or "sampling rate" in year i (specifically, the probability that a banded bird alive at the beginning of year i will be reported in year i).

For the results presented here, a fourth set of parameters is required:

θ_i = Band retention rate to the end of the i^{th} year after banding (specifically, the probability of a banded bird retaining its band to the end of the i^{th} year following banding). θ_0 is the probability of retaining a band for a short period immediately following banding.

TABLE 1.

Expected numbers of band recoveries under Model 1 allowing for loss of some bands (only four years of banding and recovery are shown).

Year banded i	Number banded N_i	Year of recovery 1	2	3	4
1	N_1	$N_1\theta_0 f_1$	$N_1\theta_1 S_1 f_2$	$N_1\theta_2 S_1 S_2 f_3$	$N_1\theta_3 S_1 S_2 S_3 f_4$
2	N_2		$N_2\theta_0 f_2$	$N_2\theta_1 S_2 f_3$	$N_2\theta_2 S_2 S_3 f_4$
3	N_3			$N_3\theta_0 f_3$	$N_3\theta_1 S_3 f_4$
4	N_4				$N_4\theta_0 f_4$

We conceptualized four functions to express varying degrees of band retention as a function of the number of years after banding (Fig. 1). Functions A, B, and C reflect increasingly severe band loss. We believe many passerine and game species may be represented by functions A or B. The few species in which band loss is very severe are represented by function C and it would seem that this represents an extreme situation. We note that the band retention rates on Ring-billed Gulls (*Larus delawarensis*) in Table 3 of Ludwig (1967) correspond closely to our curve C. Function D is somewhat different and is an attempt to mimic raptorial species in which a proportion $1 - \theta_0$ of the birds may unfasten the band within a very short period following banding (i.e., $\theta_0 < 1$). Function D does allow for a more gradual loss of bands over time from the remaining birds that do not immediately remove their band. The four band retention functions are intended to cover the general range of conditions that seem likely.

Alternatively, we can conceptualize band loss as the band retention rate between any two successive years, given that the band was still in place at the start of year i.

$$\beta_i = \frac{\theta_i}{\theta_{i-1}}, \qquad i = 2, 3, \ldots$$

$$\beta_1 = \theta_1$$

For example, using function C (Fig. 1) we can compute that the band retention rate between the end of year 6 and the end of year 7 is 0.48 ($\beta_7 = \theta_7/\theta_6$ or $\beta_7 = 0.20/0.42$). Both representations of band loss are equivalent and either one can be computed from the other. The band retention functions in Fig. 1 are more convenient expressions for our analysis.

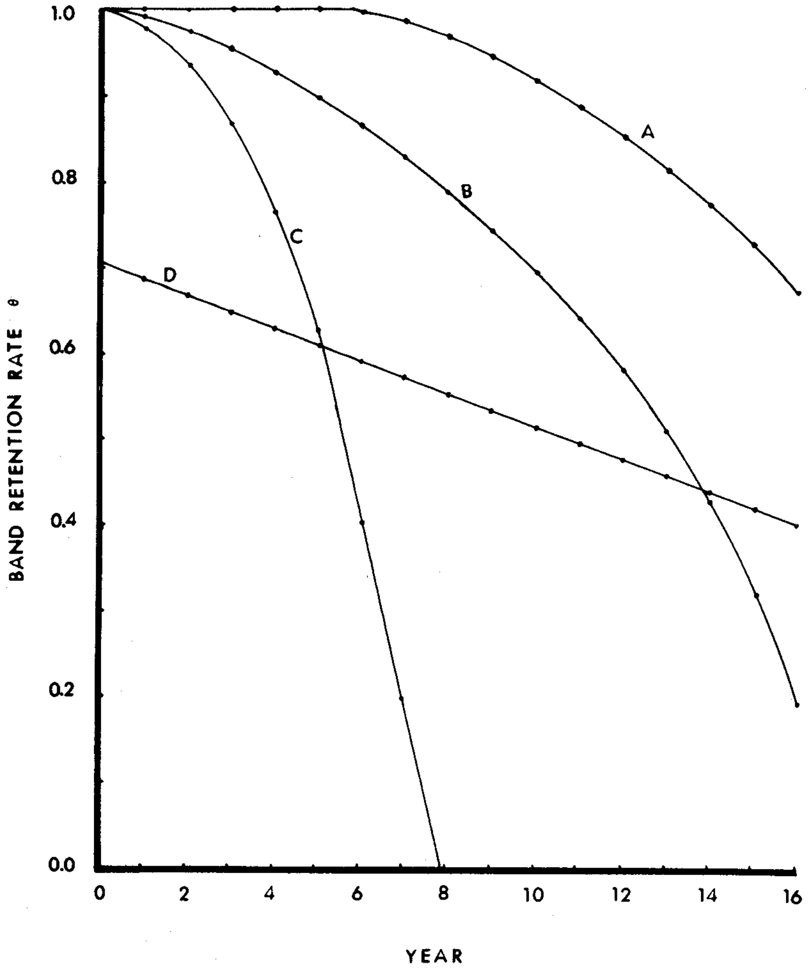


FIGURE 1. Band retention rate as a function of the number of years after banding.

The magnitude of the bias will be influenced by the true survival rate. We use the following sets of survival and recovery rate parameters to examine the effect of band loss $(1 - \theta)$ on the estimators of annual survival:

Case I	$S_i = 0.35$	$f_i = 0.10$,
Case II	$S_i = 0.60$	$f_i = 0.10$,
Case III	$S_i = 0.85$	$f_i = 0.10$,

where i = year.

TABLE 2.
Actual confidence interval coverage as a function of bias/se($\hat{S}_i$).

$\lvert\delta\rvert = b/se(\hat{S}_i)$	Coverage
0	0.950
0.1	0.949
0.2	0.945
0.25	0.943
0.3	0.939
0.4	0.932
0.5	0.921
0.6	0.908
0.7	0.892
0.75	0.884
0.8	0.874
0.9	0.853
1.0	0.830
1.5	0.677
2.0	0.484

In all cases we used $N_i = 1{,}500$ and examined bias for data sets having 16 years of banding and recovery (i.e., $i = 1, 2, \ldots, 16$). The four band retention functions (A–D) were used for each of the three cases (I–III), giving a total of 12 situations. The quantities N_i, S_i, f_i, and θ_i specify the 12 sets of expected band recovery data under the model specified in Table 1. For each of the 12 sets (where S_i is a known parameter) we computed the expected value of $\hat{S}_i$, $E(\hat{S}_i)$ using the maximum likelihood estimator for Model 1 (see Brownie et al., 1978:16). Because the estimator of S_i under Model 1 is unbiased assuming all $\theta_i = 1$ (no band loss), we can assess the bias of the estimator due to band loss by using the model structure in Table 1 which allows for band loss.

Two remaining quantities were employed in our evaluation and are defined:

$$\text{Bias} = E(\hat{S}_i) - S_i$$

$$\text{Percent relative bias (PRB)} = \frac{E(\hat{S}_i) - S_i}{S_i} \times 100,$$

where $E(\hat{S}_i)$ is computed using the maximum likelihood estimator of the parameter S_i.

The significance of the magnitude of the bias can be evaluated by comparing the bias to the standard error of the estimated annual survival rate. The maximum likelihood estimates of survival under Model 1 (i.e., $\theta_i \equiv 1$) are approximately normally distributed so a 95% confi-

TABLE 3.

Estimated annual survival rates and percent relative bias with $S_i = 0.35$ and $f_i = 0.10$ (Case I) with Model 1. The $\hat{S}_i$ are given as a percentage (i.e., $\hat{S}_i \times 100$).

	Band retention function A		Band retention function B		Band retention function C		Band retention function D	
Year (i)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)
1	34.8	-0.6	34.5	-1.4	33.6	-3.9	33.6	-3.9
5	34.4	-1.8	34.3	-2.0	32.9	-5.9	32.5	-7.2
10	34.3	-2.0	34.2	-2.3	32.9	-5.9	32.5	-7.2
15	34.9	-0.4	34.5	-1.5	33.7	-3.8	33.6	-3.9

dence interval can be computed as $\hat{S}_i \pm 1.96\ \mathrm{se}(\hat{S}_i)$. This procedure is not valid if $E(\hat{S}_i)$ is biased (e.g., by band loss). In this case $E(\hat{S}_i) = S_i + b$, where b is the bias. We can assess the coverage (the proportion of the time that this interval would include the true parameter S_i) of the usual 95% confidence interval procedure, $\hat{S}_i \pm 1.96\ \mathrm{se}(\hat{S}_i)$, by calculating the bias relative to the standard error. Therefore, let

$$\begin{aligned}\delta &= \text{bias/standard error}\\ &= b/\mathrm{se}(\hat{S}_i).\end{aligned}$$

Actual coverage can then be computed from knowledge of the absolute value of δ, $|\delta|$ (see Cochran, 1977:12–15). Of course, if $|\delta| = 0$, the coverage is 0.95. Confidence interval coverage for selected values of $|\delta|$ is given in Table 2.

The four band retention functions presented here are intended to cover the range of likely situations. Many other choices exist for band retention functions, numbers banded, survival rates, and recovery rates that may be of special concern in a given situation. These specific cases can be analyzed for the effects of band loss by the same procedures used in this paper. First, specify the parameters that define the problem (i.e., specify values for N_i, S_i, f_i, and θ_i). Second, calculate the expected recoveries using the specified parameters and the model given in Table 1. Third, compute the expected values of the maximum likelihood estimators of annual survival rates from the generated recoveries (i.e., treat it as a data set and use the formulas given by Brownie et al., 1978:16). Fourth, compute the bias, percent relative bias, and the ratio of bias to standard error with the formulas given above.

Some results on bias will be discussed for other models for adult birds (i.e., Models 2 and 3) as well as several models for young birds where

TABLE 4.

Estimated annual survival rates and percent relative bias with $S_i = 0.60$ and $f_i = 0.10$ (Case II) with Model 1. The $\hat{S}_i$ are given as a percentage (i.e., $\hat{S}_i \times 100$).

	Band retention function A		Band retention function B		Band retention function C		Band retention function D	
Year (i)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)
1	59.7	-0.5	58.8	-2.0	56.1	- 6.4	58.2	-3.0
5	59.4	-1.0	58.2	-3.0	53.9	-10.2	57.6	-4.1
10	59.4	-0.8	58.3	-2.8	53.8	-10.3	57.6	-4.0
15	59.6	-0.6	58.8	-1.9	56.9	- 5.2	58.3	-2.8

S and f are allowed to be time-specific as well as age-specific (i.e., Model H_1) (see Brownie et al., 1978 for details on these estimation models).

RESULTS

The expected annual survival rate $E(\hat{S}_i)$ and the percent relative bias (PRB) for all four band retention functions are given in Table 3 (Case I), Table 4 (Case II) and Table 5 (Case III). The PRB is small in band retention functions A and B and is larger in band retention functions C and D. However, the absolute bias is less than the standard error in nearly all instances (Table 6). The standard error of the estimate depends on the number of birds banded, the recovery rate, and the survival rate. A smaller number of birds banded or a lower recovery rate would result in a larger estimated standard error. This would indicate an even smaller ratio of bias to standard error of the estimate.

Tables 3–5 were computed with 1,500 birds banded each year and a recovery rate of 0.10. In general, fewer than 1,500 birds of a given age, sex, and species are banded each year in a particular study. Furthermore, most species have a band recovery rate considerably less than 0.10. This would suggest strongly that values of $|\delta|$ for most studies would be less than those shown in Table 6 and, therefore, the confidence interval coverage for most real data would be closer to 95% than are our examples.

We found that the bias of the estimator of annual survival rate is virtually independent of our choices of N_i, f_i, and the numbers of years of banding and recovery. Therefore, our results are much more general than the 12 specific examples reported.

The expected values of the maximum likelihood estimators for other models of banded adults (Models 0, 2, and 3; see Brownie et al., 1978) were also computed using the 12 data sets and we found them to be

TABLE 5.

Estimated annual survival rates and percent relative bias with $S_i = 0.85$ and $f_i = 0.10$ (Case III) with Model 1. The $\hat{S}_i$ are given as a percentage (i.e., $\hat{S}_i \times 100$).

	Band retention function A		Band retention function B		Band retention function C		Band retention function D	
Year (i)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)	Annual survival rate ($\hat{S}_i$)	Percent relative bias (PRB)
1	84.4	-0.8	82.1	-3.4	74.6	-12.2	82.7	-2.8
5	84.3	-0.8	81.9	-3.6	69.7	-18.0	82.3	-3.2
10	84.6	-0.5	82.2	-3.3	70.3	-17.3	82.3	-3.2
15	84.6	-0.5	82.7	-2.7	77.8	- 8.4	82.8	-2.6

generally insensitive to band loss. Bias and PRB were minimal except in severe cases with long-lived species. We did note that bias was slightly worse for the estimates of survival under Model 0. In addition, the estimators of annual survival for birds banded separately as young and adults, or young, subadults and adults (Brownie et al., 1978, Chapters 3 and 4, respectively) are also relatively insensitive to bias caused by band loss. The estimators for these models are functions of row totals of the recovery matrix and their structure is quite similar to the estimation models for adults (Models 0, 1, 2, and 3).

The effect of band loss on estimates of annual survival is quite marked in the dynamic and composite dynamic life table methods that allow *only* age-specific survival. In these models, θ_j and S_j are seriously confounded for j = age of bird or band. For example, the life table methods give

TABLE 6.

Ratio of bias to standard error ($\delta = \text{bias}/\text{se}(\hat{S}_i)$) for Cases I–III and band retention functions A–D. Refer to Table 2 for the actual coverages of 95% confidence intervals for these ratios.

	Band retention function A			Band retention function B			Band retention function C			Band retention function D		
Year (i)	Case I	Case II	Case III	Case I	Case II	Case III	Case I	Case II	Case III	Case I	Case II	Case III
1	0.1	0.1	0.2	0.1	0.3	1.0	0.3	0.8	2.7	0.3	0.3	0.6
5	0.2	0.1	0.2	0.2	0.4	1.0	0.5	1.5	4.6	0.6	0.5	0.6
10	0.2	0.1	0.1	0.2	0.4	0.8	0.5	1.5	4.2	0.6	0.5	0.6
15	0.1	0.1	0.1	0.1	0.2	0.3	0.3	0.5	0.9	0.2	0.2	0.2

the following estimates of age-specific mortality rates for Case II, function C: 43.9, 45.8, 48.1, 51.9, 57.7, 63.6, 75.0, 100.0, 100.0, etc. Because the true parameter for each age is 40%, we see that not only are the estimators quite biased, but one could easily draw the incorrect conclusion that the population exhibits a markedly age-specific mortality process. The life table methods are affected badly by band loss and have other serious deficiencies (Burnham and Anderson, 1979). We do not recommend making inference from life table analyses nor can we offer advice on how to interpret the present literature based on life table analyses. The new methods derived by Brownie et al. (1978) represent a substantial advance in the analysis of bird banding data.

The goodness of fit tests presented by Brownie et al. (1978) will detect band loss if it is substantial and if sample sizes are large. However, no such test of band loss is possible for the life table models because θ_j and S_j (j = age of bird or band) are confounded (unless a separate, specific study is made of band loss, such as Ludwig, 1967).

CONCLUSIONS

The principal conclusion from these results is that the estimates of adult annual survival rates with Model 1 are only slightly negatively biased by band loss. The effects of band loss on the estimates of annual survival are especially small for species with high mortality rates and are a significant problem only with long-lived species experiencing especially severe band loss. We again emphasize that the bias of the estimated annual survival rate is not affected by our choices of numbers banded (N_i) or recovery rates (f_i). The standard errors of $\hat{S}_i$ will be strongly affected by N_i and f_i (se($\hat{S}_i$) will decrease as either N_i and f_i increases). Most real data will have fewer birds banded than 1,500 per year and smaller recovery rates than 10%. As a result, the applicable standard error of $\hat{S}_i$ for real data will be larger (possibly much so) than we obtained here. It follows that the confidence interval coverage for S_i with real data will be closer to 95% than the results we indicate in Table 6.

The estimated recovery rates ($\hat{f}_i$) are affected primarily by θ_0 and to a much lesser degree by subsequent annual band loss rates. If the first-year band retention is 1.00, then the recovery rate estimate will be slightly inflated (generally less than 1%). If the first-year band retention is less than 1.00, then the recovery rates will be defaulted by this proportion.

SUMMARY

The effect of band loss on the estimators of annual survival rates given in Brownie et al. (1978) was examined. We examined a series of band retention functions and sets of survival rates which cover the range of real-world situations likely to be encountered. Estimates of annual survival rates were found to be only slightly negatively biased in most cases. The bias would be significant only for species with low mortality rates and severe band loss. In contrast, the bias of age-specific survival rates from the life table-type methods is quite marked.

LITERATURE CITED

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1978. Statistical inference from band recovery data: A handbook. U.S. Fish and Wildl. Serv., Resource Publ. 131.

Burnham, K. P., and D. R. Anderson. 1979. The composite dynamic method as evidence for age-specific waterfowl mortality. *J. Wildl. Manage.,* **43:** 356–366.

Cochran, W. G. 1977. Sampling Techniques. John Wiley & Sons, Inc.

Hickey, J. J. 1952. Survival studies of banded birds. U.S. Fish and Wildl. Serv., Spec. Sci. Rep.—Wildl. 15.

Ludwig, J. P. 1967. Band-loss: Its effect on banding data and apparent survivorship in the Ring-billed Gull population of the Great Lakes. *Bird-Banding,* **38:** 309–323.

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

Utah State University, Department of Wildlife Science, Logan, UT 84322. (Present address of L.J.N.: College of Forestry, Wildlife and Range Sciences, University of Idaho, Moscow, ID 83843.) Received 19 October 1978, accepted 29 June 1979.

ADDITIONAL COMMENTS ON THE ASSUMPTION OF HOMOGENEOUS SURVIVAL RATES IN MODERN BIRD BANDING ESTIMATION MODELS

JAMES D. NICHOLS, U.S. Fish and Wildlife Service, Migratory Bird and Habitat Research Laboratory, Laurel, MD 20811
S. LYNNE STOKES, U.S. Fish and Wildlife Service, Migratory Bird and Habitat Research Laboratory, Laurel, MD 20811
JAMES E. HINES, U.S. Fish and Wildlife Service, Migratory Bird and Habitat Research Laboratory, Laurel, MD 20811
MICHAEL J. CONROY, U.S. Fish and Wildlife Service, Migratory Bird and Habitat Research Laboratory, Laurel, MD 20811

Abstract: We examined the problem of heterogeneous survival and recovery rates in bird banding estimation models. We suggest that positively correlated subgroup survival and recovery probabilities may result from winter banding operations and that this situation will produce positively biased survival rate estimates. The magnitude of the survival estimate bias depends on the proportion of the population in each subgroup. Power of the suggested goodness-of-fit test to reject the inappropriate model for heterogeneous data sets was low for all situations examined and was poorest for positively related subgroup survival and recovery rates. Despite the magnitude of some of the biases reported and the relative inability to detect heterogeneity, we suggest that levels of heterogeneity normally encountered in real data sets will produce relatively small biases of average survival rates.

J. WILDL. MANAGE. 46(4):953–962

Pollock and Raveling (1982) recently emphasized the importance of considering the assumptions underlying modern bird banding estimation models (e.g., Seber 1970, Robson and Youngs 1971, Brownie and Robson 1976, Brownie et al. 1978), and devoted special attention to the assumption that all banded individuals of an identifiable demographic group (e.g., an age-sex class) have identical survival and recovery probabilities for any given year. They pointed out that in a large number of instances this assumption is biologically unrealistic, and they approximated the bias of recovery and survival rate estimators that could be expected to result from heterogeneous survival and recovery rates. Here, we present additional results that are relevant to the problem of heterogeneous survival and recovery rates.

Pollock and Raveling (1982) presented results on bias for cases in which (1) there is a negative relationship between survival and recovery probabilities of an individual and (2) individual survival probabilities differ but recovery probabilities are similar. They did not consider the situation in which survival and recovery probabilities are positively correlated. We expect this situation to occur frequently when the banding and recovery periods are separated by time spans that are not negligible with respect to mortality.

Let p_i be the conditional probability that a bird is recovered in the hunting season of year i given that it is alive at the beginning of season i. Let $S_{bh,i}$ be the probability that a bird alive at the midpoint of the banding period in year i survives until the beginning of the hunting season in year i. Let $S_{hb,i}$ be the analogous probability of surviving the interval between the beginning of the hunting season of year i and the banding period of year $i + 1$. The annual survival probability between banding periods, S_i, is then given by the product, $S_{bh,i}S_{hb,i}$. If the banding period immediately precedes the hunting season (i.e., preseason banding), then $S_{bh,i} \simeq 1.0$. If the banding period is not preseason, then $S_{bh,i} < 1.0$, with smaller $S_{bh,i}$ corresponding to longer intervals (and hence more mortality) between the banding period and the hunting season. The recovery probability of the Brownie et al. (1978) models can be defined as $f_i = S_{bh,i}p_i$.

In the case of preseason banding ($S_{bh,i} \simeq 1.0$), it seems reasonable to expect either no relationship or a negative relationship between subgroup S_i and f_i, depending on whether or not there is a negative relationship between hunting mortality (which should be reflected in p_i) and $S_{hb,i}$ and on whether or not this relationship is the basis for subgroup differences. When banding and recovery periods are separated by time spans that are not negligible with respect to mortality (i.e., when $S_{bh,i} < 1.0$), then positive, negative, or no relationship between subgroup S_i and f_i is possible. If the main difference between subgroups involves hunting mortality and $S_{hb,i}$, then a negative relationship between f_i and S_i could easily result. If the subgroups differed primarily in $S_{hb,i}$, but if this difference was not associated with differences in hunting mortality (or p_i), then no relationship would be expected between subgroup S_i and f_i. However, if the principal difference between subgroups occurs in the $S_{bh,i}$, then we would expect a positive relationship between S_i and f_i. In postseason or winter banding operations, we suspect (see Discussion) that subgroup differences in $S_{bh,i}$ will be common, and sufficiently important to produce such positive relationships between S_i and f_i.

Our objectives are to: (1) consider the magnitude and direction of bias resulting from positively correlated subgroup survival and recovery rates; (2) consider the importance of subgroup proportions in the population to the magnitude of bias, and (3) consider the power of standard goodness-of-fit tests to detect heterogeneity in survival and recovery probabilities.

We thank R. Wilcox for typing the manuscript and D. R. Anderson, K. P. Burnham, P. H. Geissler, B. R. Noon, K. H. Pollock, and an anonymous referee for providing helpful comments on an earlier draft.

METHODS

Like Pollock and Raveling (1982) we considered Model 1 of Brownie et al. (1978:15–20; see also Seber 1970, and Robson and Youngs 1971). This model was developed for adult birds and is parameterized with year-specific survival and recovery rates. The model assumes that all banded birds alive in year *i* have identical survival and recovery probabilities, S_i and f_i, respectively. We used 2 methods of approximating the bias resulting from violation of this assumption. The 1st method involved the use of Taylor series approximations and is apparently identical to that used by Pollock and Raveling (1982).

The 2nd method of approximating bias involved the use of computer simulation. For a fixed total banded sample size, we first used pseudorandom numbers to generate binomial variates corresponding to the number of birds banded in each demographic subgroup. Then, for each subgroup, we used pseudorandom numbers to generate the multinomial random vectors comprising its recovery matrix. The specific methods used to generate these matrices are similar to those described in Anderson and Burnham (1976: 66). The recovery matrix for the entire heterogeneous population is obtained by summing the subgroup matrices, and the estimators, $\hat{f}_i$ and $\tilde{S}_i$ (where $\tilde{S}_i$ is the bias-adjusted survival probability estimator), are then computed from the population matrix using the equations of Brownie et al. (1978:16). This process was repeated n times for each experimental situation considered, and bias was estimated as the difference between the mean of the n estimates and the true parameter value.

Following Pollock and Raveling (1982), we report the percent relative bias defined, for example with survival rate, as:

$$100[E(\tilde{S}_i) - S_i]/S_i$$

where $E(\tilde{S}_i)$ denotes the expected value of $\tilde{S}_i$, and S_i denotes the true survival probability. We define average percent relative bias as the arithmetic mean percent relative bias for all years of the experiment for which estimates are obtained.

We were also interested in the power of the Pearson chi-square goodness-of-fit test for Model 1 (Brownie et al. 1978: 19–20) with respect to violation of the homogeneity assumption. This was examined via computer simulation using the same basic program used to compute the bias expression. For each of n heterogeneous-model data sets, we computed the chi-square goodness-of-fit test statistic (we used the same cell pooling algorithm as in program ESTIMATE) and its associated probability level. We obtained the total number of data sets resulting in model rejection, m, and estimated power as $(m/n)_\alpha$, where power is defined as the probability of rejecting the null hypothesis when α is used as the rejection level for the chi-square test.

RESULTS

Using the Taylor series approximations, analytical expressions for the large sample bias of $\tilde{S}_i$ and $\hat{f}_i$ can be derived for any given relationship between S and f. In general, the expressions are quite complex and of little practical interest. However, in certain situations, the sign and a general feel for the magnitude of the bias can be gained by examining the expressions. For example, consider the banding experiment consisting of k banding and recovery years where there are assumed to be 2 heterogeneous subgroups within the population having proportions r and $1 - r$. These subgroups may differ in survival and/or recovery probabilities. We make the simplifying assumption that the survival and recovery rates remain constant throughout the experiment within each subgroup and we denote the parameters within each subgroup by f, S and f', S'. (The numerical results of Pollock and Raveling [1982, Table 4] are a special case of this situation.) By examining the expressions for asymptotic bias in survival and recovery rates (Appendix), we see that:

1. The bias in $\hat{f}$, although small, is always negative.
2. The bias in $\tilde{S}$ is positive if $S > S'$ and $f > f'$ or $S < S'$ and $f < f'$.
3. The bias in $\tilde{S}$ may be either positive or negative when the relationship between S and f is negative. However, for the usual magnitudes of S and f, the bias will be negative, since for a positive bias to result, ΔS (defined as $|S - S'|$) would have to be quite large relative to Δf (defined as $|f - f'|$).
4. The bias in $\tilde{S}$ is not greatly affected by the annual banded sample size, N_i.

These results are illustrated via numerical examples (Figs. 1, 2). Figure 1 contains 3 curves representing Taylor series approximations of average percent relative bias plotted as a function of the difference between subgroup survival probabilities (ΔS). The 6 plotted points corresponding to each line represent the bias as estimated from Monte Carlo simulations (each point obtained from 1,000 iterations). The simulated bias estimates correspond closely to the Taylor series approximations and indicate that the approximations are quite reasonable. We assumed 2 subgroups with survival probabilities of $S = 0.5 - \Delta S/2$, $S' = 0.5 + \Delta S/2$.

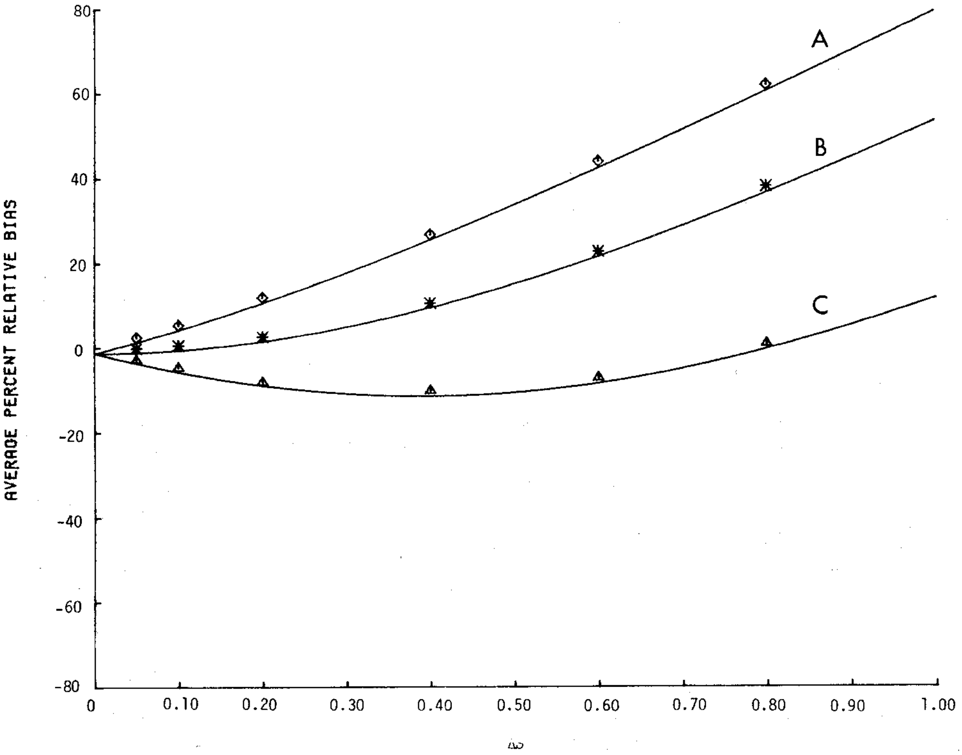


Fig. 1. Average percent relative bias of $\hat{S}_i$ plotted as a function of the difference between true subgroup survival probabilities, ΔS. In all plots, $S = 0.50 - \Delta S/2$, $S' = 0.50 + \Delta S/2$. Subgroup recovery probabilities are: plot A, $f = 0.025$, $f' = 0.075$; plot B, $f = f' = 0.05$; plot C, $f = 0.075$, $f' = 0.025$. All plots assume equal proportions (0.50) of the population and banded sample in each subgroup, and 2,000 birds banded annually for 5 years. The curves represent Taylor series approximations while the plotted points represent Monte Carlo simulation results (1,000 iterations each).

The 3 curves correspond to the following 3 sets of subgroup recovery probabilities: plot A, $f = 0.025, f' = 0.075$; plot B, $f = f' = 0.05$; plot C, $f = 0.075, f' = 0.025$. We assumed 0.50 of the population and banded sample in each subgroup. Finally, we assumed a 5-year banding program with 2,000 birds banded annually. Figure 2 assumes the same experimental situation as Fig. 1 but includes bias plotted as a function of the difference between subgroup recovery probabilities (Δf). The 3 curves in Fig. 2 correspond to the following sets of survival probabilities: plot A, $S = 0.40$, $S' = 0.60$; plot B, $S = S' = 0.50$; plot C, $S = 0.60$, $S' = 0.40$. Both Figs. 1 and 2 demonstrate the importance of the relationship between survival and recovery probabilities to the magnitude and direction of bias. Although a negative relationship does generally result in negative bias, as reported by Pollock and Raveling (1982), a positive relationship between the probabilities can result in large positive bias (Figs. 1, 2).

Most of the numerical examples of Pollock and Raveling (1982) assume equal sizes for all subgroups. However, the bias resulting from survival rate heterogeneity is a function of the proportions in each subgroup (see bias expression in Appen-

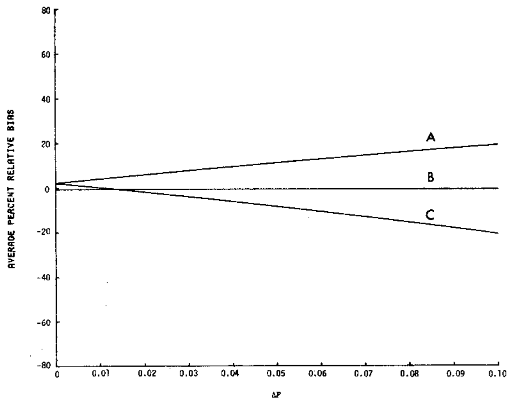


Fig. 2. Average percent relative bias of $\tilde{S}_i$ plotted as a function of the difference between subgroup recovery probabilities (Δf). In all plots, $f = 0.05 - \Delta f/2$, $f' = 0.05 + \Delta f/2$. Subgroup survival probabilities are: plot A, $S = 0.40$, $S' = 0.60$; plot B, $S = S' = 0.50$; plot C, $S = 0.60$, $S' = 0.40$. All plots assume equal proportions (0.50) of the population and banded sample in each subgroup, and 2,000 birds banded annually for 5 years. The curves represent Taylor series approximations.

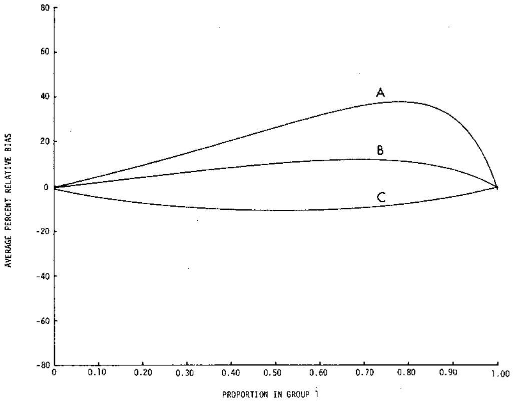


Fig. 3. Average percent relative bias of $\tilde{S}_i$ plotted as a function of the proportion of the population and banded sample in subgroup 1 (the subgroup characterized by S and f). Subgroup survival and recovery probabilities are: plot A, $S = 0.30$, $f = 0.025$, $S' = 0.70$, $f' = 0.075$; plot B, $S = 0.30$, $f = f' = 0.05$, $S' = 0.70$; plot C, $S = 0.30$, $f = 0.075$, $S' = 0.70$, $f' = 0.025$. All plots assume 2,000 birds banded annually for 5 years. The curves represent Taylor series approximations.

dix). We plotted Taylor series approximations of average percent relative bias as a function of the proportion of the population in subgroup 1 (characterized by S and f), for the following parameter values: plot A, $S = 0.30$, $S' = 0.70$, $f = 0.025$, $f' = 0.075$; plot B, $S = 0.30$, $S' = 0.70$, $f = f' = 0.05$; plot C, $S = 0.30$, $S' = 0.70$, $f = 0.075$, $f' = 0.025$ (Fig. 3). We assumed a 5-year banding program with 2,000 total birds banded annually. Monte Carlo simulations with extreme subgroup proportions (e.g., 0.05, 0.95) again indicated that the Taylor series approximations were quite reasonable. Plot C (negative relationship between f and S) is fairly symmetrical with the greatest bias occurring around equal subgroup proportions of 0.5. However, in plots A and B (positive relationship between S and f, and no relationship), bias is greater when a larger proportion of individuals is in the low survival probability subgroup (and, in the case of plot A, the low recovery probability subgroup). This skewness with positive relationship and no relationship between S and f is even more pronounced with greater differences in subgroup survival probabilities (greater ΔS).

Brownie et al. (1978) provided goodness-of-fit tests to determine whether or not a given model seems to fit a particular band recovery data set. If these tests are likely to detect violations of the homogeneous survival and recovery probability assumption, then biased estimators will not present as big a problem as when the violations are not detectable. We plotted empirical power estimates ($\alpha = 0.05$) for the Model 1 goodness-of-fit test (Brownie et al. 1978:19–20) (Fig. 4). Plots A, B, and C represent the same sets of subgroup recovery probabilities as in Fig. 1. We again assumed 5 years of banding, equal proportions in each subgroup, and 2,000 total birds banded annually. The estimated power of the goodness-of-fit test to reject the inappropriate Model 1 was low for lower values of ΔS in all plots

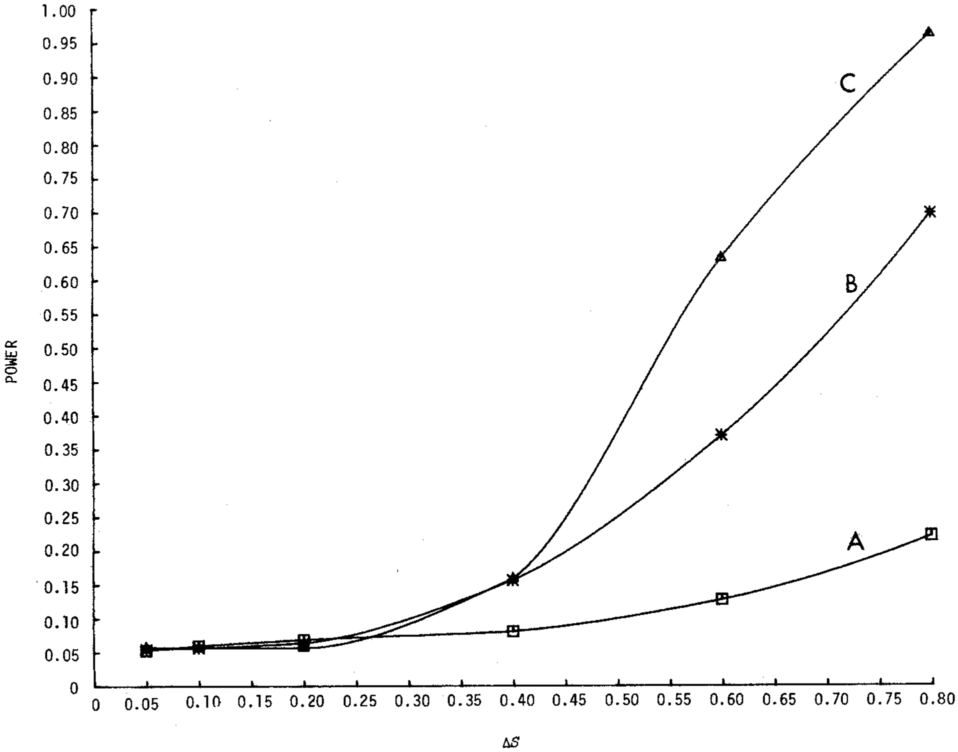


Fig. 4. Estimated power ($\alpha = 0.05$) of the goodness-of-fit test to reject Model 1, plotted as a function of the difference between subgroup survival probabilities (ΔS). In all plots $S = 0.50 - \Delta S/2$, $S' = 0.50 + \Delta S/2$. Subgroup recovery probabilities are: plot A, $f = 0.025$, $f' = 0.075$; plot B, $f = f' = 0.05$; plot C, $f = 0.075$, $f' = 0.025$. All plotted points assume equal proportions (0.50) of the population and banded sample in each subgroup, and 2,000 birds banded annually for 5 years. The plotted points represent Monte Carlo simulation results (1,000 iterations each).

(Fig. 4). These lower ΔS values also correspond to the smaller biases of survival rate estimates (Fig. 1). At larger values of ΔS (e.g., $\Delta S > 0.40$), estimated power differed substantially for the 3 kinds of subgroups considered. The test was most powerful in the case of a negative relationship between subgroup survival and recovery probabilities and least powerful when a positive relationship existed.

We also plotted estimated power of the goodness-of-fit test as a function of banded sample size (Fig. 5). Plots A, B, and C correspond to the same sets of subgroup survival and recovery probabilities as in Fig. 3. We again assumed 5 years of banding. In all 3 plots, the estimated power of the test increased with banded sample size. The 3 plots differed substantially, with the greatest power corresponding to the subgroup with negatively related S and f.

DISCUSSION

When considering the implications of our results for actual banding studies it is important to keep in mind the kinds of heterogeneity that might exist in migratory bird populations. Pollock and Raveling (1982) provided an example of Canada geese (*Branta canadensis maxima*) banded at Marshy Point on Lake Mani-

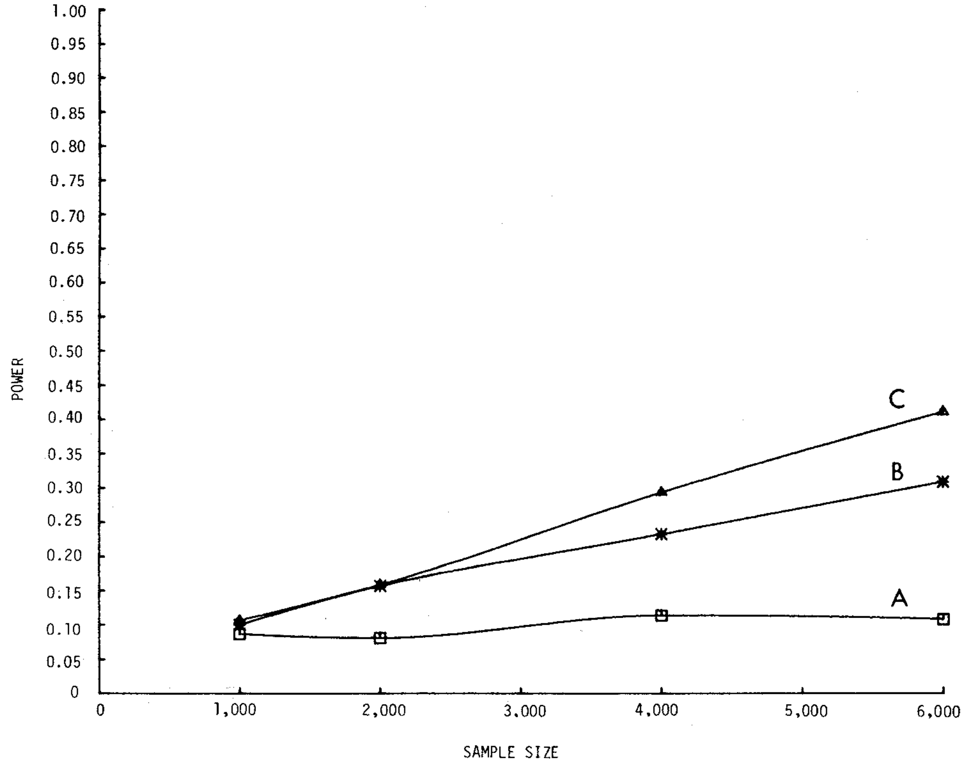


Fig. 5. Estimated power ($\alpha = 0.05$) of the goodness-of-fit test to reject Model 1, plotted as a function of total annual banded sample size. Subgroup survival and recovery probabilities are: plot A, $S = 0.30$, $f = 0.025$, $S' = 0.70$, $f' = 0.075$; plot B, $S = 0.30$, $f = f' = 0.05$, $S' = 0.70$; plot C, $S = 0.30$, $f = 0.075$, $S' = 0.70$, $f' = 0.025$. All plotted points assume equal proportions (0.50) of the population and banded sample in each subgroup, and a 5-year banding experiment. The plotted points represent Monte Carlo simulation results (1,000 iterations each).

toba, Canada, which represented 2 population segments that wintered in different areas. In most duck species, it is not unusual to find birds banded in specific breeding areas and recovered in a variety of different wintering areas. Similarly, birds from a particular winter-banded sample are often recovered in different breeding and staging areas in the fall. In banding analyses it is common to combine different banding stations within a relatively large geographic area into banding reference areas. While such areas are usually defined based on similarity of recovery distribution of birds banded at the included stations, it is certainly possible that the use of reference areas might result in the combination of population segments that differ to some degree in migration pathways. In any case, differences in migration pathways, breeding grounds and wintering grounds are expected to result in differences in environmental conditions and hunting pressure, and thus in possible differences in survival and recovery probabilities.

If reproduction increases mortality risks of birds (Johnson and Sargeant 1977, Bailey 1981) then breeding and nonbreeding segments of a population would be expected to exhibit different survival probabilities. Age may represent another source of heterogeneity in survival and recovery probabilities. Winter-banded birds of many species are not separated by age-classes when banded. Although survival and recovery rates of first year vs. older winter-banded mallards (*Anas platyrhynchos*) from Colorado were similar (Hopper et al. 1978), it is certainly possible that this is not true of all other waterfowl species (e.g., diving ducks seem to exhibit greater age-specific dimorphism in winter). Preseason-banded birds are generally assigned to either of 2 age-classes (adult or young), but test results of Anderson (1975:18–19) suggest that in some areas subadult mallards may exhibit different survival and recovery probabilities than either adults or young.

In many of the previous examples of sources of heterogeneity, it is probably most useful to think of a population or banded sample as consisting of discrete population segments. However, when some sources of heterogeneity are considered it may be more useful to think of survival rates as differing from individual to individual and following some probability distribution (see model M_h of Otis et al. 1978). For example, phenotypic variation in the ability to withstand periods of severe winter stress and even variation in tendency to follow different migration pathways might be assumed to vary from individual to individual in a population, resulting in a specific survival and recovery probability being associated with each individual. Pollock and Raveling (1982) modeled this situation by allowing individual survival rates to follow a series of beta distributions, but their conclusions about estimator bias were qualitatively similar to those they obtained with the discrete population segment model. Because of the qualitatively similar results of the 2 approaches and because of our interest in sources of heterogeneity that are likely to result in population segments, we only considered this approach in the present work.

Our results have emphasized the importance of the relationship between individual survival and recovery probabilities to the bias resulting from such heterogeneity. In general, presence of a negative relationship between S and f tends to result in negatively biased survival estimates. Uncorrelated and positively correlated S and f tend to produce positively biased survival estimates, often of greater magnitude. We expect winter banding operations (or any banding operation separated from the recovery period by a substantial amount of time) to often result in positively related S and f, while preseason-banded birds may exhibit either no relationship or a negative relationship between S and f. This observation may explain the tendency in some species (mallards, J. Nichols, unpubl. data; black ducks, *Anas rubripes*, W. Blandin, unpubl. data) for survival estimates of winter-banded birds to be somewhat higher than those of preseason-banded birds. The proportion of individuals in each population subgroup also influences the magnitude of the resulting bias of survival rate estimates. In the cases we examined (which included a number of situations in addition to those presented in Fig. 3), negatively related S and f tended to exhibit the greatest biases when subgroup proportions were similar (~0.50). However, with uncorrelated and positively correlated S and f, the largest biases occurred when a larger proportion of the population was in the subgroup with the lower S.

Our empirical power estimates of the Brownie et al. (1978) goodness-of-fit test were low for small to moderate differences between subgroup survival probabilities (ΔS), but biases were also relatively small for such differences. In the case of positively correlated S and f, the power of the test was low for all ΔS examined. Therefore, among the situations we examined, the power of the goodness-of-fit test was lowest for the situation (positively related S and f) producing the largest relative bias. Power of the test increased with increasing banded sample sizes, as expected.

Our suspicion that winter banding programs of heterogeneous populations tend to produce a positive correlation between subgroup survival and recovery probabilities leads us to suggest that preseason banding programs should offer somewhat more protection from survival estimate bias resulting from heterogeneity. The absolute value of relative bias of the $\hat{S}_i$ resulting from negatively correlated and uncorrelated S and f (as expected from preseason programs) is smaller than for the situation of a positive correlation. In addition, the goodness-of-fit test was more powerful in the situations expected to accompany preseason banding, although even then this test is not expected to be of much use in detecting heterogeneity at expected levels with normal sample sizes. Finally, we emphasize that portions of our figures represent situations that are unrealistic. For example, we allowed the difference between subgroup survival probabilities (ΔS) to range between 0.0 and 1.0 in Figs. 1 and 4. When considering the possible biological sources of heterogeneity that we have discussed, it seems that most situations would probably involve $\Delta S \leq 0.20$, with $\Delta S \leq 0.10$ being most common. For ΔS in this range, relative biases of the $\hat{S}_i$ are small, especially when considered in relation to their standard errors (estimated standard errors of the annual survival estimates generally would be expected to range from approximately 0.05 to 0.10 for the situations considered in Figs. 1–4, and would be larger with smaller banded sample sizes). Larger ΔS may exist in some instances (Pollock and Raveling 1982), but we suspect they will represent rare situations.

LITERATURE CITED

ANDERSON, D. R. 1975. Population ecology of the mallard. V. Temporal and geographic estimates of survival, recovery and harvest rates. U.S. Dep. Inter., Fish and Wildl. Serv. Resour. Publ. 125. 110pp.

———, AND K. P. BURNHAM. 1976. Population ecology of the mallard. VI. The effect of exploitation on survival. U.S. Dep. Inter., Fish and Wildl. Serv. Resour. Publ. 128. 66pp.

BAILEY, R. O. 1981. A theoretical approach to problems in waterfowl management. Trans. North Am. Wildl. and Nat. Resour. Conf. 46: 58–71.

BROWNIE, C., AND D. S. ROBSON. 1976. Models allowing for age-dependent survival rates for band return data. Biometrics 32:305–323.

———, D. R. ANDERSON, K. P. BURNHAM, AND D. S. ROBSON. 1978. Statistical inference from band recovery data: a handbook. U.S. Dep. Inter., Fish and Wildl. Serv. Resour. Publ. 130. 212pp.

HOPPER, R. M., H. D. FUNK, AND D. R. ANDERSON. 1978. Age specificity in mallards banded postseason in eastern Colorado. J. Wildl. Manage. 42:263–270.

JOHNSON, D. H., AND A. B. SARGEANT. 1977. Impact of red fox predation on the sex ratio of prairie mallards. U.S. Dep. Inter., Fish and Wildl. Serv. Wildl. Res. Rep. 6. 56pp.

OTIS, D. L., K. P. BURNHAM, G. C. WHITE, AND D. R. ANDERSON. 1978. Statistical inference from capture data on closed animal populations. Wildl. Monogr. 62. 135pp.

POLLOCK, K. H., AND D. G. RAVELING. 1982. Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. J. Wildl. Manage. 46:88–98.

ROBSON, D. S., AND W. D. YOUNGS. 1971. Statistical analysis of reported tag-recaptures in the harvest from an exploited population. Cornell Univ., Ithaca, N.Y., Biometrics Unit, BU-369-M. 15pp.

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

Received 2 June 1981.
Accepted 31 December 1981.

APPENDIX

An analytical expression for the large sample bias of $\hat{f}_i$ and $\tilde{S}_i$ is given for the situation described in the text. We assume that

1. There are 2 heterogeneous subgroups within the population which differ in recovery and/or survival rate. Let r and $1 - r$ denote the proportions of the population in subgroups 1 and 2, respectively.
2. There are k banding and recovery seasons within the banding experiments.
3. The recovery and survival rates are constant within each subgroup over the course of the experiment.

We define $\bar{f} = rf + (1 - r)f'$ and $\bar{S} = rS + (1 - r)S'$ to be the "average" recovery and survival rates in the population. Let

$$\delta_i = ff'r(1-r)\sum_{j=1}^{k-i}(S^j - S'^j),$$

$$\gamma_i = \sum_{j=1}^{i-1} N_j(S^{i-j} - S'^{i-j}),$$

and

$$\beta_i = r(1-r)(S - S')\sum_{j=1}^{k-i} \cdot (S^{j-1}f - S'^{j-1}f').$$

Then for large samples we have

$$E(\hat{f}_i) \approx \bar{f} - \delta_i\gamma_i/E(T_i),$$

and $E(\tilde{S}_i) \approx \bar{S} + [\delta_i\gamma_i/E(T_i) + \beta_i]/\rho_{i+1}$,

where $\rho_i = \sum_{j=0}^{k-i}[rS^jf + (1-r)S'^jf']$

is the probability of ever recovering a bird banded in year i, and T_i is the ith block total of the recovery matrix (see Brownie et al. 1978:13).

Then Bias $(\hat{f}_i) = -\delta_i\gamma_i/E(T_i) < 0$,

and Bias $(\tilde{S}_i) = [\delta_i\gamma_i/E(T_i) + \beta_i]/\rho_{i+1} > 0$

if $S > S'$ and $f \geq f'$
or $S < S'$ and $f \leq f'$.

ASSUMPTIONS OF MODERN BAND-RECOVERY MODELS, WITH EMPHASIS ON HETEROGENEOUS SURVIVAL RATES

KENNETH H. POLLOCK,[1] Division of Statistics, University of California, Davis, CA 95616
DENNIS G. RAVELING, Division of Wildlife and Fisheries Biology, University of California, Davis, CA 95616

Abstract: The assumptions inherent in modern band-recovery models are reviewed with particular attention to homogeneity of survival and band-recovery rates for all individuals in the population. If this assumption fails, the following implications emerge: (1) the models only enable estimation of average annual survival and band-recovery rates; (2) estimators of these averages probably give underestimates; (3) the degree of underestimation in practical studies is difficult to assess, but may sometimes be important for survival estimates; (4) if sampling is nonrandom and heterogeneity of survival and recovery rates is present in the population, then any estimates could be misleading; and (5) if survival rates are homogeneous but recovery rates are heterogeneous (due perhaps to geographical variation in hunting pressure and reporting rates), then there is no bias in survival estimates. An example where data from neck-collared birds showed heterogeneity in segments of a Canada goose (*Branta canadensis*) population is discussed. We believe that the practical limitations of bird-banding studies deserve careful review by population biologists and managers.

J. WILDL. MANAGE. 46(1):88–98

Band-recovery data have long been used to estimate mortality rates in exploited migratory bird populations. However, it is only in recent years that methods of analysis of banding data have been rigorously considered. By making certain assumptions, Seber (1970, 1971, 1972, 1973) formulated a rigorous statistical model that gave rise to explicit maximum-likelihood estimators of survival- and recovery-rate parameters. Extensions allowing for age dependence of survival and recovery rates were made by Johnson (1974), Brownie and Robson (1976), and Brownie et al. (1978), who, in a comprehensive monograph, detailed a variety of age-dependent models.

Development of modern methods clearly identified the weaknesses of some of the older procedures such as the composite dynamic life table. Burnham and Anderson (1979) analyzed 45 substantial data sets from migratory waterfowl band recoveries and demonstrated that the composite dynamic life table should no longer be used, as only 2 data sets (4%) fit the model at $\alpha = 0.05$.

Burnham and Anderson (1979) also found that modern methods are not necessarily always adequate for analyzing banding data, as 14 of 45 data sets (30%) did not fit. In this paper we further examine the assumptions of the modern methods.

An important assumption of all recovery models is that all banded individuals of an identifiable class (e.g., by species, age, sex) in the sample have the same annual survival and recovery rates (Brownie et al. 1978:6). We consider what happens when this assumption fails, and the importance of random sampling in conjunction with homogeneity of survival and recovery rates.

First, we consider a specific model (Model 1: Adults Only Banded) when the homogeneity assumption holds. We shall refer to this as the Homogeneous Population Model. This will be followed by the same model when the homogeneity assumption fails, the Heterogeneous Population Model. We discuss a popula-

---

[1] Present address: Department of Statistics, North Carolina State University, Box 5457, Raleigh, NC 27650.

tion of Canada geese for which segments with heterogeneous survival and recovery rates are known to exist from the use of individually identifiable neck-collared birds (Raveling 1978). Finally, we provide a general discussion of all the assumptions underlying modern band-recovery models, with an emphasis on the practical implications of assumption failure.

We thank D. R. Anderson, D. W. Anderson, K. P. Burnham, D. S. Gilmer, D. H. Johnson, and R. M. McLandress for helpful comments on an earlier version of the manuscript.

HOMOGENEOUS POPULATION MODEL

Here we consider the situation where all birds are banded as adults and all birds have the same survival and recovery rates in a particular year (Model 1 of Brownie et al. 1978:15). For simplicity, we will also assume the number of years of recoveries is the same as the number of years of banding.

The data matrix for observed recoveries is given in Table 1. Under this model, each row of recoveries follows a multinomial distribution, and the expected recoveries take the form given in Table 2.

The maximum-likelihood estimators of the survival- and recovery-rate parameters are given by the following equations:

$$\hat{S}_i = \left(\frac{R_i}{N_i} - \hat{f}_i\right) \bigg/ \left(\frac{R_{i+1}}{N_{i+1}}\right) \qquad (1)$$

$$\hat{f}_i = \frac{R_i C_i}{N_i T_i}. \qquad (2)$$

If we assume the model is correct, then Eq. (2) gives unbiased estimators of the recovery rates. The survival-rate estimators (Eq. [1]) are biased, but the bias is small for banding studies of reasonable size. The maximum-likelihood estimator adjusted for bias is

$$\hat{S}_i = \left(\frac{R_i}{N_i} - \hat{f}_i\right) \bigg/ \left(\frac{R_{i+1} + 1}{N_{i+1} + 1}\right). \qquad (3)$$

Table 1. The algebraic representation of band recoveries with *k* banding and recovery years.

Year banded	Number banded		Recoveries by hunting season 1	2	3	...	k	Row totals
1	N_1	T_1[a]	R_{11}	R_{12}	R_{13}	...	R_{1k}	R_1
2	N_2			R_{22}	R_{23}	...	R_{2k}	R_2
3	N_3		T_2		R_{33}	...	R_{3k}	R_3
⋮	⋮			T_3		⋮	⋮	⋮
k	N_k						R_{kk}	R_k
Column totals			C_1	C_2	C_3	...	C_k	

[a] T_i is a rectangular block of data with $T_1 = R_1$, $T_i = R_i + T_{i-1} - C_{i-1}$ for $i = 2, \ldots, k$, where R_i is the ith-row total and C_i is the ith-column total for $i = 1, \ldots, k$.

HETEROGENEOUS POPULATION MODEL

This is the same model, except now each bird has its own specific survival and recovery probability for a given year. Also, all survival and recovery rates for the birds banded are a random sample from a multivariate probability distribution. Under this model, it is easy to show (Seber 1973:316) that each row of recoveries is still multinomially distributed as before, but now the expected numbers of band recoveries differ (Table 3).

All expectations $E(\cdot)$ are with respect to the appropriate multivariate probability distribution. Suppose, for example, we consider $E(S_1 S_2 f_3)$, then

$$E(S_1 S_2 f_3) = \int_0^1 \int_0^1 \int_0^1 S_1 S_2 f_3 f(S_1, S_2, f_3)\, dS_1\, dS_2\, df_3$$

with $f(S_1, S_2, f_3)$ the joint distribution of S_1, S_2, and f_3 for the whole population of animals. Put in simple terms, $E(S_1 S_2 f_3)$

Table 2. The expected number of band recoveries under the Homogeneous Population Model.

Year banded	Number banded	Recoveries by hunting season				
		1	2	3	...	k
1	N_1	N_1f_1[a]	$N_1S_1f_2$	$N_1S_1S_2f_3$[b]	...	$N_1S_1 \ldots S_{k-1}f_k$
2	N_2		N_2f_2	$N_2S_2f_3$	...	$N_2S_2 \ldots S_{k-1}f_k$
3	N_3			N_3f_3	...	$N_3S_3 \ldots S_{k-1}f_k$
⋮	⋮				⋮	⋮
k	N_k					N_kf_k

[a] f_i is the recovery rate for year i, i.e., the probability of a bird being killed, retrieved by hunter, and its band reported for $i = 1, \ldots, k$.
[b] S_i is the survival rate from year i to year $i + 1$, for $i = 1, \ldots, k - 1$.

is the average value of the product $S_1S_2f_3$ over the whole population. If results like

$$E(S_1S_2f_3) = E(S_1)E(S_2)E(f_3) \tag{4}$$

were true for all the products, then Table 3 would have the same form as Table 2, with average values replacing the 1 constant value for all animals. In this instance, the estimators (Eqs. [1] and [2]) would have the same properties as before, except they would be estimating average survival and recovery rates. Actually, the result is only true if animals' survival rates and band-recovery rates are completely independent of each other. This is extremely unlikely in practice. It is likely that survival rates from year to year for the same animal will have a strong positive relationship. Also, the relationship between survival and recovery rates is likely to be negative, because recovery typically occurs for those animals killed by hunters.

Using Taylor Series approximations (Seber 1973:7), approximate expectations can be derived for the estimators $\tilde{S}_i$ and $\hat{f}_i$ in Eqs. (3) and (2). We find

$$E(\tilde{S}_i) \simeq \left[E\left(\frac{R_i}{N_i}\right) - E(\hat{f}_i)\right] \bigg/ E\left(\frac{R_{i+1}+1}{N_{i+1}+1}\right)$$

and

$$E(\hat{f}_i) \simeq E\left(\frac{R_i}{N_i}\right)\frac{E(C_i)}{E(T_i)}, \tag{5}$$

which can be expressed in terms of the expected values (Table 3).

$$E(\tilde{S}_i) \simeq \frac{\begin{gathered} E(f_i) - E(\hat{f}_i) + E(S_if_{i+1}) \\ + E(S_iS_{i+1}f_{i+2}) + \ldots \\ + E(S_i \ldots S_{k-1}f_k) \end{gathered}}{\begin{gathered} E(f_{i+1}) + E(S_{i+1}f_{i+2}) + \ldots \\ + E(S_{i+1} \ldots S_{k-1}f_k) \end{gathered}}. \tag{6}$$

The expected value of $\tilde{S}_i$ given in Eq. (6) is a very good approximation because of the independence of numerator and denominator in the estimator (Eq. [3]).

To find the $E(\hat{f}_i)$ in Eq. (5), it is only necessary to give expressions for $E(R_i/N_i)$, $E(C_i)$, and $E(T_i)$, which are summations of the appropriate cells (Table 3). Bias expressions are then calculated as follows:

$$B(\tilde{S}_i) = E(\tilde{S}_i) - E(S_i) \tag{7}$$

$$B(\hat{f}_i) = E(\hat{f}_i) - E(f_i). \tag{8}$$

The percent relative bias is determined by dividing the bias by the true value of the parameter and expressing as a percentage.

NUMERICAL ANALYSIS OF BIAS

The dynamics of exploited populations are poorly understood. Although it is likely (from density-dependent factors influencing survival) that hunting mortality is partially compensated for by a decrease in natural mortality (Anderson and

Table 3. The expected number of band recoveries under the Heterogeneous Population Model.

Year banded	Number banded	Recoveries by hunting season				
		1	2	3	...	k
1	N_1	$N_1E(f_1)$[a]	$N_1E(S_1f_2)$	$N_1E(S_1S_2f_3)$	...	$N_1E(S_1 \ldots S_{k-1}f_k)$
2	N_2		$N_2E(f_2)$	$N_2E(S_2f_3)$	...	$N_2E(S_2 \ldots S_{k-1}f_k)$
3	N_3			$N_3E(f_3)$	...	$N_3E(S_3 \ldots S_{k-1}f_k)$
$\vdots$	$\vdots$				$\vdots$	$\vdots$
k	N_k					$N_kE(f_k)$

[a] $E(\cdot)$ are expectations.

Burnham 1976, 1978), the exact relationship is unknown. This uncertainty makes analysis of the bias due to the existence of heterogeneity of individuals with respect to survival and recovery extremely difficult.

First, let us consider the situation of 2 or 3 population segments of equal size with a wide range of heterogeneous survival rates such that the average annual survival rate of the whole population is held constant at either 0.3 or 0.6. For each case, we consider a wide range of heterogeneous recovery rates with a constant average of 0.05. The averages of the percent relative bias values for the yearly survival-rate estimators over a 3- or 5-year banding study are presented in Tables 4 and 5. These results show that for a given amount of heterogeneity in the survival rates, the bias moves from positive to negative as the recovery rates of the segments become more different. The values for 2 population segments (Table 4) are more extreme than those for 3 population segments (Table 5).

Now we consider the more realistic situation of a heterogeneous population where individual survival rates follow a series of beta distributions (Johnson and Kotz 1970), with the average annual survival rate held constant at 0.33, 0.5, or 0.67. For each case, we consider 2 very different relationships for the recovery rates. First, we assume that the individual's recovery rate is independent of its survival rate and that the average annual recovery rate over the population is 0.05. Second, we assume that the individual's recovery rate is a constant proportion of its mortality rate (1 minus the survival rate), so $f = a(1 - s)$, and that the average

Table 4. Average percent relative bias of survival-rate estimators for a heterogeneous population with 2 population segments for 3 and 5 years of banding data.

	Expected survival rates					
	$E(S) = 0.3$			$E(S) = 0.6$		
	Survival rates of segments			Survival rates of segments		
	0.35, 0.25	0.4, 0.2	0.45, 0.15	0.7, 0.5	0.8, 0.4	0.9, 0.3
I[a]	+ 0.5 (+ 1.3)[b]	+ 2.3 (+ 5.0)	+ 5.0 (+10.9)	+ 0.8 (+ 2.3)	+ 3.3 (+ 8.7)	+ 7.5 (+18.5)
II	− 3.7 (− 2.9)	− 6.2 (− 3.5)	− 7.5 (− 1.5)	− 3.3 (− 2.0)	− 5.1 (+ 0.5)	− 5.0 (+ 7.2)
III	− 8.0 (− 7.3)	−14.8 (−12.6)	−20.8 (−15.6)	− 7.7 (− 6.5)	−13.9 (− 8.8)	−18.5 (− 6.8)
IV	−12.3 (−11.8)	−23.8 (−22.6)	−35.0 (−31.5)	−12.1 (−11.3)	−23.3 (−19.9)	−33.4 (−24.7)
V	−16.7 (−16.7)	−33.3 (−33.3)	−50.0 (−50.0)	−16.7 (−16.7)	−33.3 (−33.3)	−50.0 (−50.0)

[a] The recovery rates of the 2 population segments are I: 0.05, 0.05; II: 0.0375, 0.0625; III: 0.025, 0.075; IV: 0.0125, 0.0875; and V: 0, 0.1. In all cases $E(f) = 0.05$.

[b] Figures in parentheses are 5-year values.

Table 5. Average percent relative bias of survival-rate estimators for a heterogeneous population with 3 population segments for 3 and 5 years of banding data.

	Expected survival rates					
	$E(S) = 0.3$			$E(S) = 0.6$		
	Survival rates of segments			Survival rates of segments		
	0.35–0.25[a]	0.4–0.2	0.45–0.15	0.7–0.5	0.8–0.4	0.9–0.3
I[b]	+ 0.4 (+ 0.8)[c]	+ 1.5 (+ 3.3)	+ 3.4 (+ 7.4)	+ 0.6 (+ 1.5)	+ 2.2 (+ 5.9)	+ 5.0 (+12.8)
II	− 2.4 (− 2.0)	− 4.1 (− 2.3)	− 5.0 (− 1.0)	− 2.2 (− 1.3)	− 3.3 (+ 0.4)	− 3.3 (+ 4.8)
III	− 5.2 (− 4.8)	− 9.8 (− 8.2)	−13.7 (−10.1)	− 5.1 (− 4.2)	− 9.1 (− 5.7)	−12.1 (− 4.4)
IV	− 8.1 (− 7.8)	−15.7 (−14.5)	−22.7 (−19.9)	− 8.0 (− 7.3)	−15.1 (−12.4)	−21.4 (−15.2)
V	−11.0 (−10.8)	−21.7 (−21.1)	−32.1 (−30.7)	−10.9 (−10.6)	−21.4 (−20.0)	−31.3 (−28.2)

[a] There are 3 population segments with survival rates of 0.35, 0.3 and 0.25. In all cases the middle survival rate is the same as $E(S)$.

[b] The recovery rates of the 3 population segments are I: 0.05, 0.05, 0.05; II: 0.0375, 0.05, 0.0625; III: 0.025, 0.05, 0.075; IV: 0.0125, 0.05, 0.0875; and V: 0, 0.05, 0.1.

[c] Figures in parentheses are 5-year values.

annual recovery rate over the population is 0.05. The average percent relative bias of survival-rate estimators over a 3- or 5-year study varies (Table 6). These results show that the bias in the survival-rate estimators is very dependent on which relationship between recovery rate (f) and survival rate (s) is considered. If the recovery rate is independent of the survival rate, then there is a positive bias that is most severe in longer studies and when the average annual survivial rate is low. If the relationship is $f = a(1 - s)$, then there is a negative bias that is most severe in shorter studies and when the average annual survival rate is high.

DISCUSSION

It is difficult to make strong quantitative statements about the nature of the bias of survival- and recovery-rate estimators in banding studies. First, the exact nature of the relationship between survival and recovery rates is unknown. Second, few data are available documenting the nature of heterogeneity for

Table 6. Average percent relative bias of survival-rate estimators for a heterogeneous population where individual survival rates follow a beta distribution for 3 and 5 years of banding data.

$E(S) = 0.33$				
Parameter values	$\alpha = 1, \beta = 2$	$\alpha = 3, \beta = 6$	$\alpha = 6, \beta = 12$	$\alpha = 25, \beta = 50$
SE of beta distributions	0.236	0.149	0.108	0.053
f independent of s	+10.8 (+26.0)[a]	+4.4 (+10.2)	+2.3 (+5.3)	+0.6 (+1.3)
$f = 0.075(1 - s)$	−17.0 (− 6.2)	−6.1 (− 1.1)	−3.1 (−0.3)	−0.7 (0)
$E(S) = 0.5$				
Parameter values	$\alpha = 1, \beta = 1$	$\alpha = 5, \beta = 5$	$\alpha = 10, \beta = 10$	$\alpha = 40, \beta = 40$
SE of beta distributions	0.290	0.151	0.109	0.056
f independent of s	+ 9.1 (+21.7)	+2.5 (+6.3)	+1.3 (+3.3)	+0.3 (+0.9)
$f = 0.1(1 - s)$	−26.1 (−16.0)	−6.7 (−3.2)	−3.5 (−1.5)	−0.9 (−0.4)
$E(S) = 0.67$				
Parameter values	$\alpha = 2, \beta = 1$	$\alpha = 6, \beta = 3$	$\alpha = 12, \beta = 6$	$\alpha = 50, \beta = 25$
SE of beta distributions	0.236	0.149	0.108	0.053
f independent of s	+ 4.0 (+ 9.7)	+1.6 (+4.2)	+0.8 (+2.3)	+0.2 (+0.6)
$f = 0.15(1 - s)$	−20.9 (−14.9)	−8.3 (−5.7)	−4.4 (−2.8)	−1.1 (−0.7)

[a] Figures in parentheses are 5-year values.

most species. However, that recoveries frequently occur in widely different geographic areas from banded samples of migratory waterfowl demonstrates that different paths (population segments) exist (Bellrose 1976). In fact, brood mates raised on the same pond may undertake drastically different migrations (Martinson and Hawkins 1968). Therefore, individuals within banded samples are exposed to different environmental conditions including hunting pressure and cooperation in reporting bands (Henny and Burnham 1976). Thus, it is possible to provide some useful new qualitative statements about the nature of possible bias of survival- and recovery-rate estimators when heterogeneity is present in the population. Readers should also keep in mind that bias is of relatively more practical importance in precise studies where the standard errors are small.

Recovery Estimators.—(1) If population segments have the same survival rates but different recovery rates, then there is no bias in the estimator of the average recovery rate. (2) If population segments have heterogeneous survival and recovery rates, then the recovery-rate estimators at the beginning and end of a study always have no bias and those in the middle have a negative bias. The bias is more severe for species where the average survival is high and for studies of longer duration. The degree of bias is much less than for the survival-rate estimators and can probably be ignored for most practical purposes.

Survival Estimators.—(1) If population segments have the same survival rates but different recovery rates, then there is no bias in the survival estimators. This could occur if there were reporting rate changes in different geographical areas (Henny and Burnham 1976). (2) If survival probabilities are heterogeneous over the population, there is likely to be a strong positive relationship between the survival probabilities of an individual from year to year. There is also likely to be a negative relationship between survival and recovery probabilities for an individual. In this situation, survival-rate estimators will generally have a negative bias. The negative bias will be more severe for species where the average survival rate is high and for studies that are of short duration. The seriousness of this negative bias to a banding study will depend on the species under study and will be hard to evaluate in general. (3) It is theoretically possible for the survival-rate estimators to have a positive bias. This could occur if there were segments with markedly different survival rates but similar recovery rates. This implies that the difference in survival of the segments would have to be mostly due to differences in natural mortality. This could occur if drastically different environmental conditions were encountered by the segments (e.g., disease level, food supply, weather). All of these factors can vary on a local or regional scale.

A CANADA GOOSE POPULATION

There has been little documentation of the existence of populations of migratory birds with heterogeneous survival and recovery rates. A major reason is that most studies on the dynamics of migratory bird populations are based on data from band recoveries from dead birds. These data are not suited to a thorough examination of the nature of possible heterogeneity.

The analysis of the geographic distribution of band recoveries does indicate the segments of a population of banded individuals may take different migratory routes. A good example comes from the

Table 7. Characteristics of the Marshy Point adult Canada goose population viewed as 2 distinct segments.

Characteristic	Segment		Average annual population values
	Rochester	Remainder	
Proportion of population	0.75[a]	0.25[a]	
Survival rate	0.83[b]	0.51[b]	0.75[c]
Hunting mortality rate	0.11[d]	0.30[d]	
Recovery rate	0.03[e]	0.08[e]	0.04[e]

[a] From Raveling (1978).

[b] Weighted averages from Raveling (1978)—uses Kirwin, Kansas, data as representative of the remainder. These estimates of survival are based on those birds alive and that returned to Marshy Point. They have approximate standard errors of 0.02 and 0.07, respectively.

[c] This is a weighted average of segment values. The average survival rate is given by $[(3 \times 0.83) + (1 \times 0.51)]/4 = 0.75$.

[d] From Raveling (1978). The Rochester rate is for Manitoba plus Minnesota. The remainder rate was the data from Manitoba plus Kansas plus other Central Flyway states.

[e] These values are obtained by assuming 0.85 of the birds are retrieved and 0.3 are reported (Raveling 1978). For example, the recovery rate for the Rochester segment is: $0.11 \times 0.85 \times 0.3 = 0.03$.

study of Anderson and Sterling (1974) on pintail (*Anas acuta*) drakes banded in Saskatchewan: "Banded pintail drakes were taken by hunters in Canada, all four Flyways in the United States, in Mexico and in Russia, with a single report from Cuba." We believe that it is likely that the pintails undertaking such different migrations had different survival and recovery rates because of different hunting pressures and environmental conditions. Six of the 8 pintail data sets analyzed by Burnham and Anderson (1979) did not fit modern methods of analysis, however, we emphasize that other explanations are possible.

An alternative type of data, not without its own problems, is repeated observations of individually identifiable birds on both the breeding and wintering grounds. This approach was taken by Raveling (1978), who described a population of giant Canada geese (*B. c. maxima*) neck-collared at Marshy Point Goose Sanctuary on the southeastern shore of Lake Manitoba, Canada, during the summers of 1968–70.

From 1968 to 1976, the marked geese were observed on breeding and staging areas at Marshy Point and on wintering grounds at Rochester, Minnesota; Kirwin, Kansas; and Swan Lake, Missouri. Raveling (1978) concluded that this population of geese consisted of segments with differing hunting mortalities and overall survival rates. About 75% of the birds migrated to Rochester, Minnesota, and were subjected to an annual hunting mortality rate of about 19% and total annual mortality of 25%. The remainder of the birds wintered in several areas, with most (15% of the total population) going to Kirwin, Kansas. They had an annual hunting mortality of about 48% and a total annual mortality of 58% (including mortality on breeding grounds in Manitoba). It should be emphasized that these estimates are subject to sampling error and perhaps some bias due to collar loss.

To obtain an approximation of the degree of bias in a traditional banding study on this population, let us assume we have a population of *adult* birds divided into 2 segments (Table 7). These characteristics describe the population based on direct observation of 90% of the marked individuals (Raveling 1978). However, this is admittedly still a simplification of the real population. Based on 3 years of data, the expected value of the survival estimator from Eq. (6) is approximately 0.69, as compared with the average population value of 0.75. This gives a relative bias of about −8%. Over a longer study the bias would be less.

More important to the biologist than any negative bias of the average survival estimate is that a traditional banding study would have revealed almost nothing about the characteristics of the segments. The segment not going to Rochester (25%) was being subjected to a high mortality rate clearly beyond the reproductive potential of geese. Such differential mortality rates for segments could cause rapid distributional changes and have important implications for management (Raveling 1978).

A REVIEW OF THE ASSUMPTIONS OF MODERN BANDING MODELS

Considering the attention given in the literature to the assumptions of capture-recapture sampling (Cormack 1968, Seber 1973), band-recovery models should be subjected to the same scrutiny. Here we consider all of the assumptions of Brownie et al. (1978:6), with an emphasis on the practical implications of assumption failure.

The Sample is Representative of the Target Population

This assumption is important, particularly if heterogeneity of survival and band-recovery rates occurs. Let us take an extreme example for illustration, where we have a population composed 90% of birds with an annual survival rate of 0.8, and 10% of birds with an annual survival rate of 0.4. The average survival rate for the population is 0.76, but suppose nonrandom sampling makes our population appear equally divided into 2 segments. The apparent average survival rate is 0.6, and in a 5-year banding study there would also be a negative bias on survival estimators. The expected value of the survival estimators would be approximately 0.55 if we assume recovery rates of 0.025 and 0.075 for the 2 segments (Eq. [6]).

The following are some possible causes of nonrepresentative samples. (1) Segments may use different parts of a staging or winter area due to behavioral characteristics of the species. This has been demonstrated in the stratification of Canada goose population segments at a migration stopover and winter location (Raveling 1979). If this heterogeneous spatial distribution on a small scale is ignored and only easily accessible areas are sampled rather than using a random sample, large biases could result. (2) Banding at different times on a staging area could be of segments with different timing of migration or different wintering grounds and, hence, possible different survival rates. Canada geese banded at the same trap sites in July, August, and September had different migration patterns, and thus different proportions of segments were sampled in the different time periods (Raveling 1978). Dispersal away from natal sites and concentration at staging areas before autumn migration are common features of many duck populations (Bellrose 1976). It is likely that trapping efforts capture varying components of dispersing birds, which represent different characteristics in terms of their molt and behavior patterns that could be reflected in different migration patterns and vulnerability to hunting. (3) Banding is likely to be nonrandom with respect to age- and sex classes in a population (Raveling 1966). This could be important if age and sex categories cannot be accurately established in the field. It is common during winter banding of several duck species to consider all birds adults, because the molt of juveniles has progressed to make them nearly indistinguishable from adults when plumage criteria are used to separate age-classes. It is well known that different age and sex categories may have different survival rates.

Age and Sex of Individuals are Correctly Determined

This is related to (3) above, and will clearly depend on the species being studied. If the assumption is false, one will, in effect, have segments of different age and sex categories, which is likely to cause heterogeneity of survival and recovery rates.

There is No Band Loss

This is rarely completely true, but it is likely that for recent studies the loss rate is low. Nelson et al. (1980) examined this assumption and found that there is a negative bias on survival estimators that is worse for species with high survival rates. They concluded that, unless band loss was high, this bias would be of little significance. However, it does occur on occasion, especially for sea birds and raptors. Ludwig (1967) documented severe band loss in a study on ring-billed gulls (*Larus delawarensis*) in the Great Lakes.

Survival Rates are Not Influenced by Banding

This assumption is important because if there was substantial mortality due to banding, the survival estimators would not apply to the nonbanded birds. Direct damage will depend on the species and the method of capture, and should usually be detectable by experienced biologists. Detection may not always be easy, however. For example, Lensink (1964:12) concluded that serious injuries to ducklings caught by dogs were frequently undetected, as band-recovery rates from those birds were much lower than from birds caught in traps. Indirect damage by dispersing family groups and other social relationships is more subtle, but may be more common than assumed.

The Year (Hunting Season) of Band Recoveries is Correctly Tabulated

Errors here are probably rather infrequent and should not cause major difficulties in most practical applications of the models (Anderson and Burnham 1980). However, errors do occur and usually involve a hunter reporting a band taken in a later season than when it actually was taken. Raveling (unpubl. data) has 9 records of instances in which a band was sent in with the wrong season identified (of 327 recoveries, or a minimum of 2.8% error). Reasons for this occurring include fears by hunters of government apprehension when sending a band in after a hunting season had closed; sending in bands as taken in a year when a hunter's interest was stimulated by friends or game department officials, but after the banded bird(s) had been obtained; fear of reporting bands of birds killed outside the legal hunting season; and hunters reporting band numbers from birds previously taken by friends or acquaintances as being taken by themselves in the current year. To the extent that such instances occur, they operate to produce a positive bias on survival estimators.

The Fate of Each Banded Individual is Independent of the Fate of Other Banded Individuals

This assumption is probably violated in many practical applications of bird-banding models. Birds are not independent entities in terms of survival or other characteristics (Sulzbach and Cooke 1978). This will not bias any estimators, but will mean that true sampling variances are larger than those given by the statistical models. Thus, any calculated confidence intervals will be narrower than they should be.

A simple example for illustration is to consider a population composed of independent pairs of birds that behave as though they are a single individual. A sample of N individuals from this population is effectively only one-half of N and, hence, any sampling variances will be much larger than those for the models that assume the sample is N independent individuals. The actual situation in real populations is much more complex, with many partially dependent members, but the effective sample size will still be less than the actual sample size. This would impact the confidence intervals of analyses of survival for geese and swans because of their strong pair and family associations.

The Fate of a Given Banded Animal is a Multinomial Random Variable

This is not really an assumption. It follows from the assumption of independence.

All Banded Individuals of an Identifiable Class Have the Same Annual Survival Rates and Recovery Rates

The effects of failure of this assumption have been discussed in detail elsewhere in this paper. We feel it is likely that populations with heterogeneous survival and band-recovery rates exist in practice because: (1) Migration is often on more than 1 flyway, and even on a single flyway there is often a series of wintering points. (2) Hunting pressure varies drastically from state to state and even in different regions of the same state, especially because of the existence of refuges. (3) Habitat on the breeding and wintering grounds is variable.

CONCLUSIONS

We feel that wildlife scientists involved with banding studies should give careful consideration to the model assumptions we have discussed. We believe that particular attention should be given to whether the sampling is random and to whether heterogeneity of survival and band-recovery rates is pronounced. The question of the effect of the banding process itself either directly or indirectly on survival is also important.

If population segments exist and are of interest in their own right, other forms of data such as repeated observations of color-marked, live birds on the breeding and wintering grounds should be considered if possible. Clearly, alternative methods of data collection will not be without their own difficulties.

LITERATURE CITED

ANDERSON, D. R., AND K. P. BURNHAM. 1976. Population ecology of the mallard: VI. The effect of exploitation on survival. U.S. Fish and Wildl. Serv., Resour. Publ. 128. 66pp.

———, AND ———. 1978. Effect of restrictive and liberal hunting regulations on annual survival rates of the mallard in North America. Trans. North Am. Wildl. and Nat. Resour. Conf. 43:181–186.

———, AND ———. 1980. Effect of delayed reporting of band recoveries on survival estimates. J. Field Ornithol. 51:244–247.

———, AND R. T. STERLING. 1974. Population dynamics of molting pintail drakes banded in south-central Saskatchewan. J. Wildl. Manage. 38:266–274.

BELLROSE, F. C. 1976. Ducks, geese and swans of North America. Stackpole Books, Harrisburg, Pa. 544pp.

BROWNIE, C., AND D. S. ROBSON. 1976. Models allowing for age-dependent survival rates for band return data. Biometrics 32:305–323.

———, D. R. ANDERSON, K. P. BURNHAM, AND D. S. ROBSON. 1978. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildl. Serv., Resour. Publ. 121. 212pp.

BURNHAM, K. P., AND D. R. ANDERSON. 1979. The composite dynamic method as evidence for age-specific waterfowl mortality. J. Wildl. Manage. 43:356–366.

CORMACK, R. M. 1968. The statistics of capture-recapture methods. Oceanogr. Mar. Biol. Annu. Rev. 6:455–506.

HENNY, C. J., AND K. P. BURNHAM. 1976. A reward band study of mallards to estimate band reporting rates. J. Wildl. Manage. 40:1–14.

JOHNSON, D. H. 1974. Estimating survival rates from banding of adult and juvenile birds. J. Wildl. Manage. 38:290–297.

JOHNSON, N. L., AND S. KOTZ. 1970. Continuous univariate distributions—2. Houghton Mifflin, Boston, Mass. 306pp.

LENSINK, C. J. 1964. Distribution of recoveries from bandings of ducklings. U.S. Fish and

Wildl. Serv. Spec. Sci. Rep. 89, Washington, D.C. 146pp.

LUDWIG, J. P. 1967. Band-loss: its effect on banding data and apparent survivorship in the ring-billed gull population of the Great Lakes. Bird-Banding 38:309–323.

MARTINSON, R. K., AND A. S. HAWKINS. 1968. Lack of association among duck brood mates during migration and wintering. Auk 85:684–686.

NELSON, L. J., D. R. ANDERSON, AND K. P. BURNHAM. 1980. The effect of band loss on estimates of annual survival. J. Field Ornithol. 51:32–38.

RAVELING, D. G. 1966. Factors affecting age ratios of samples of Canada geese caught with cannon-nets. J. Wildl. Manage. 30:682–691.

———. 1978. Dynamics of distribution of Canada geese in winter. Trans. North Am. Wildl. and Nat. Resour. Conf. 43:206–225.

———. 1979. Traditional use of migration and winter roost sites by Canada geese. J. Wildl. Manage. 43:229–235.

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

———. 1971. Estimating age-specific survival rates from bird-band returns when the reporting rate is constant. Biometrika 58:491–497.

———. 1972. Estimating survival rates from bird-band returns. J. Wildl. Manage. 36:405–413.

———. 1973. Estimation of animal abundance and related parameters. Griffin, London, England. 506pp.

SULZBACH, D., AND F. COOKE. 1978. Elements of nonrandomness in mass-captured samples of snow geese. J. Wildl. Manage. 42:437–441.

Received 16 June 1980.
Accepted 5 April 1981.

NUMERICAL ESTIMATION OF SURVIVAL RATES FROM BAND-RECOVERY AND BIOTELEMETRY DATA

GARY C. WHITE, Environmental Science, MS-K495, Los Alamos National Laboratory, Los Alamos, NM 87545[1]

Abstract: The estimation of survival rates from tagging or banding data has been well developed by Brownie et al. (1978). However, problems occur when sparse data sets result in undefined estimates, when survival estimates exceed unity, when a hypothesis about the data cannot be tested by any of the available models, and when constraints on model estimators are desired. This paper presents a general analysis method whereby the models of Brownie et al. (1978) and many other methods described in the literature are merely special cases. Models are specified algebraically as cell probabilities consisting of functions of the survival rates and other parameters to be estimated. These algebraic expressions and the observed cell values are input to the computer program SURVIV to provide maximum-likelihood estimates of the unknown parameters and perform hypothesis tests on the data. The generality of the model specification also allows estimation of survival rates from biotelemetry data.

J. WILDL. MANAGE. 47(3):716–728

Estimates of the critical parameters of a wildlife population are required for its sound management. Estimates of recruitment and survival are needed to harvest a population in an optimal fashion. In this paper, numerical methods of maximum-likelihood estimation of survival rates from biotelemetry and banding data are described and a general computer code to perform the estimation is presented.

Maximum-likelihood (ML) estimation is a rigorous statistical method of obtaining estimators for unknown parameters. ML estimators fall into 2 classes: (1) closed form and (2) iterative or numerical. In either case, a likelihood function $L(\theta|X)$ involving the parameters of interest (the vector θ) is formed, given the data (X). In the context here, θ usually represents survival and recovery rates. In all cases, the ML estimators can be found as the solutions of

$$\frac{\partial \ln L(\theta|X)}{\partial \theta_i} = 0.$$

That is, the partial derivative of the natural logarithm of the likelihood function with respect to the ith parameter is set equal to zero and solved for θ_i. From calculus, this procedure is known to find the maximum of the log-likelihood function (hence the name "maximum likelihood"). If θ is a vector of 9 parameters, then there are 9 equations to be solved (usually nonlinear simultaneous equations). Often, these equations can be solved algebraically, such as the ML estimators presented by Seber (1970).

More frequently, it is impossible, or at least difficult, to solve these equations algebraically. In this case, ML estimators of each parameter θ_i can be found by iteratively (using numerical techniques) maximizing the likelihood function $L(\theta|X)$. Systems that can be solved algebraically can also be solved numerically, but not vice versa.

Numerous authors (cf. Robson 1963; Murton 1966; Cormack 1970; Seber 1970, 1971, 1972; Johnson 1974; Youngs and Robson 1975; Brownie et al. 1978; North and Cormack 1981) have developed estimators of survival rates from band- or tag-recovery data, making a wide variety of specific assumptions. In each of the references cited, maximum-likelihood methods were used to derive the estimators and their sampling variances and covariances.

---

[1] **Present address: Wagar Bldg., Colorado State University, Ft. Collins, CO 80523.**

Jolly (1965:228–230) demonstrated how to build a general model (likelihood function) for capture–recapture sampling for open populations. He then derived estimators for only a few special cases of this model due to the difficulties of algebraically solving the likelihood equations. Numerical solution of complex likelihood equations such as those presented by Jolly (1965) provides a means of obtaining estimates for any particular model, including the most general model, from a complex likelihood function. The biologist or biometrician is no longer constrained by the algebraic complexity of solving the likelihood equations from using a model that provides a realistic representation of the biological situation.

Brownie et al. (1978:9–12) demonstrated how a conceptual model of band-recovery data is used to construct a likelihood function and derive ML estimators. All of the references cited use the technique they describe. The computer program presented here also uses the same technique to perform ML estimation of survival and recovery parameters. Any of the models mentioned in the references cited can be handled by the program, because they are merely special cases of the more general theory. The results of Brownie et al. (1978) are the most recent and comprehensive; thus, comparisons with their work will be made through the remainder of the paper.

Band- or tag-recovery models are more general than needed for biotelemetry data in that the fate of radio-marked animals is usually known, but legbanded birds are often not recovered. Whereas the recovery rate (*f*) of bands must be estimated, the recovery rate of radios is usually known exactly (i.e., $f = 1$). Maximum-likelihood estimation is also useful to obtain estimates of survival rates from biotelemetry data.

Herein are described procedures and a computer program that treat survival-estimation methods as specific cases of a more general model. Methods of estimating parameters and testing general hypotheses about the model parameters are presented. The trade-offs between numerical vs. analytical maximum-likelihood estimation of survival rates are discussed and the necessity of an easy-to-use program illustrated. A Monte Carlo simulation of the effects of bounding parameters on their bias is presented and an example of survival-rate estimation for an elk (*Cervus elaphus nelsoni*) biotelemetry study is provided to illustrate the practical application of the approach described in this paper.

I thank D. R. Anderson for many helpful comments and suggestions on the manuscript and K. P. Burnham for discussions on the numerical estimation of survival rates from banding experiments. This work was funded by Contract W-7405-ENG.36 from the U.S. Department of Energy to Los Alamos National Laboratory.

PARAMETER ESTIMATION BY NUMERICAL METHODS

The general model for estimating survival rates from biotelemetry data is structured identically to that of Brownie et al. (1978:9–12), consisting of a set of multinomial distributions tied together by common parameters to describe the expected cell probabilities, with the numbers of animals banded or tagged treated as fixed observable numbers. The model is best illustrated by considering a simple example. Assume that 40 and 50 animals are radiocollared in years 1 and 2, respectively. The fate of these 90 radio collars is known (Table 1), based on a 3-year battery life. Thus, the "disappearance" of the animal after 3 years is attributed to bat-

Table 1. Simulated and expected recovery history of 90 radio collars placed on elk in years 1 and 2. The true survival rates for years 1–4 were 0.6, 0.7, 0.5, and 0.6, respectively.

Year collared	Number collared	Year collars returned 1	2	3	4	Battery failure
			Simulated Recoveries			
1	40	12	8	13		7
2	50		13	17	7	13
			Expected Recoveries			
1	N_1	$N_1(1 - S_1)$	$N_1S_1(1 - S_2)$	$N_1S_1S_2(1 - S_3)$		$N_1S_1S_2S_3$
2	N_2		$N_2(1 - S_2)$	$N_2S_2(1 - S_3)$	$N_2S_2S_3(1 - S_4)$	$N_2S_2S_3S_4$

tery failure. This simple example assumes no animals are lost due to radio failure or emigration from the study area.

Brownie et al. (1978:9–12) used the cell probabilities of models such as that in Table 1 (Boyd 1956, Seber 1970) to construct a likelihood function and derive ML estimators of the unknown parameters (in the example S_1, S_2, S_3, and S_4). The computer programs they described are the current "state-of-the-art" for estimating survival rates from band-recovery data. Several limitations occur with their approach. First, the parameter estimates are not constrained in a bounded region; thus, estimates of survival rates greater than unity occasionally occur, particularly if samples are small. Second, if a sparse data matrix happens to occur (i.e., certain critical cells have zeros), some estimators are undefined, even though the estimated survival for that cell could be unity (cf. North and Cormack 1981).

The restrictions of the programs provided by Brownie et al. (1978) can be alleviated by optimizing the likelihood function numerically, with the parameter space restricted to the range of admissible parameter values. Several other advantages are incurred by using numerical procedures. Sets of parameters can be constrained to a common value. Thus, in the simple example, S_1 and S_2 can be constrained to a common value, and likewise S_3 and S_4. Thus, the parameter space can be reduced from 4 to 2 and, if this model is still appropriate for the data, more precise estimates of the parameters are achieved.

An additional extension of the models presented in Brownie et al. (1978) is that models can be developed that allow the same survival parameters for 2 sexes at 1 age-class, but different survival rates at a later age-class. Consider as an example the estimation of survival of a cervid species. The survival of male and female young may be identical to develop a parsimonious model. However, when the age-class is reached where sex differences become obvious (such as antlers), the survival rates could differ. Thus, models to estimate the survival rates of cervid species should allow common survival rates for the young animals of both sexes, but different survival rates for the sexes at later age-classes. Tests of the hypothesis of identical survival of the young should be made before this hypothesis is supported, however.

The drawback to the numerical optimization approach is formulating the algebraic expression of the likelihood function. A computer program described in a later section is available to construct the likelihood function from the algebraic expressions for the expected cell probabilities, perform optimization (with constraints), and construct likelihood-ratio tests between models. However, the user must thoroughly understand the formulation of the expected cell probabilities to use this program.

A 2nd shortcoming of the numerically derived estimates is that the small sample bias of the estimators currently cannot be corrected as is done with analytical estimates (cf. Brownie et al. 1978:16). However, Monte Carlo simulation results (see Discussion) suggest that constraining the range of the estimate to its admissible range may partially offset the bias of the estimators.

A 3rd drawback of the numerically derived estimates is that the program may occasionally fail to converge on the global maximum of the likelihood function. Such failures of the method may be due to small data sets relative to the number of parameters being estimated, low numerical precision of the computer relative to the number of parameters being estimated, and finally lack of parameter identifiability. In the 1st 2 circumstances, failure to converge theoretically should not occur, but does due to the physical limitations of the machine. However, in the 3rd case, estimates should not be obtained. Due to the limit of numerical precision of digital computers, the program will attempt to find a solution, even though none exists. Generally such solutions are obviously wrong, because some parameters were not changed from their initial values and/or the variance–covariance matrix is not positive definite.

HYPOTHESIS TESTING VIA NUMERICAL PROCEDURES

Numerical maximization of the likelihood function produces ML estimates of the unknown parameters. Generally, the researcher also wants to test hypotheses from the data. Likelihood-ratio tests provide the means of testing elaborate hypotheses about complex data sets.

The appropriateness of estimating common parameter values can be easily determined with numerical methods because likelihood-ratio tests can be constructed from the values of the likelihood functions evaluated at the ML estimates. Thus, if the null hypothesis of H_0: $S_1 = S_2$ and $S_3 = S_4$ is to be tested, the likelihood function is first maximized with all 4 parameters individually estimated and then maximized again with $S_1 = S_2$ and $S_3 = S_4$. Then a likelihood-ratio test with the 2 likelihood values is constructed. Because the computer programs described by Brownie et al. (1978) use specific models, all hypothesis tests are for specific situations. The numerical optimization procedure allows a more general approach.

Likelihood-ratio tests are specific tests between 2 particular models where the null hypothesis is H_0: Model A is equivalent to Model B when Model B is a special case (nested) of Model A. The alternative hypothesis is H_a: Model A is not equivalent to Model B. Likelihood-ratio tests are constructed from the values of the likelihood functions for each of the models to be compared. By likelihood-ratio theory, the quantity

$$\chi^2 = -2 \ln \frac{L(\theta_0 | X)}{L(\theta | X)}$$

is known to be asymptotically chi-squared distributed with degrees of freedom equal to the number of parameters removed from the parameter space for the reduced model (model in the numerator with parameter vector θ_0) (Mood et al. 1974:440). Thus, to test the hypothesis that 2 survival rates are equivalent, the likelihood value for the general model with S_i and S_{i+1} both estimated is obtained. Then S_i is set equivalent to S_{i+1} (i.e., the parameter space of the reduced model is made smaller by 1 parameter) and the likelihood for this reduced model is obtained. The resulting chi-square statistic is distributed with 1 degree of freedom. Additional parameters can be equated to obtain more general hypotheses (i.e., $S_i = S_{i+1} = S_{i+2} \ldots$, or $S_i = S_j$, $S_{i+1} = S_{j+1}, \ldots$) to obtain a chi-square statistic with more than 1 degree of freedom.

Chi-square goodness-of-fit tests can also be constructed for any particular model by summing the (observed − expected)^2/expected values for each of the cells of the model (Mitra 1958). Generally, cells with expected values less than 2.0 are pooled with neighboring cells (Roscoe and Byars 1971) to obtain a chi-square statistic.

PROGRAM SURVIV

The practical application of the numerical optimization approach requires that a user-oriented computer program is available. That is, the chore of constructing the matrix of expected cell probabilities is difficult in itself and additional complications due to complex input for a computer program make the task nearly insurmountable. Program SURVIV has been developed to handle the numerical optimization task with straightforward and simple input to SURVIV. The advantages to numerical estimation would be offset by requiring a complex coding of model specifications and observed data. The simplicity of this input is demonstrated by a simple example for biotelemetry data.

Program SURVIV uses the procedures MODEL, ESTIMATE, and TEST to perform the numerical estimation calculations. PROC MODEL constructs the likelihood function from algebraic expressions describing the cell probabilities. The generality and ease of model specification are shown by the input in Fig. 1 for the simple example in Table 1. The PROC MODEL statement sets various options and alerts the program that the observed and expected cell probabilities are to follow. The COHORT card sets the number of animals for the 1st multinomial and is followed by the 3 multinomial cells with the observed value separated from the expected value by a colon. The entry of the expected cell probability is the feature of Program SURVIV that makes the numerical approach described feasible. The parameters to be estimated are denoted by the S(I) notation in the algebraic expression. This algebraic expression must be FORTRAN compatible, because the program manipulates this code to construct a FORTRAN subroutine to evaluate the likelihood function.

PROC MODEL continues to read COHORT, cell probability, and LABELS cards until the PROC ESTIMATE card is encountered. PROC ESTIMATE is called to make the actual parameter estimates. The INITIAL statement (Fig. 1) signifies that initial estimates of some parameters are provided (the default is S(I) = 0.5). The 1st time ESTIMATE is called in Fig. 1, only boundary constraints are provided (i.e., the admissible range of the parameter values are set). This call to PROC ESTIMATE evaluates the model with 4 individual values of S(1), S(2), S(3), and S(4).

The 2nd call (Fig. 1) to PROC ESTIMATE evaluates the model with S(1) = S(2) and S(3) = S(4) (i.e., the reduced model where 2 sets of parameters are assumed equal). The CONSTRAINTS statement is used to specify these equalities, and the degrees of freedom of the model are automatically reduced by 2. At this time, the constraints must be constants,

```
PROC TITLE SMALL BIOTELEMETRY EXAMPLE FOR JWM ARTICLE;
PROC MODEL NPAR=4 /* SIMPLE BIOTELEMETRY EXAMPLE */;
   COHORT = 40 /* NUMBER OF ANIMALS COLLARED IN YEAR 1*/;
    12:(1.-S(1));
     8:S(1)*(1.-S(2));
    13:S(1)*S(2)*(1.-S(3));
     7:S(1)*S(2)*S(3);
   COHORT = 50 /* NUMBER OF ANIMALS COLLARED IN YEAR 2*/;
    13:(1.-S(2));
    17:S(2)*(1.-S(3));
     7:S(2)*S(3)*(1.-S(4));
    13:S(2)*S(3)*S(4);
   LABELS;
     S(1)=SURVIVAL RATE IN YEAR 1;
     S(2)=SURVIVAL RATE IN YEAR 2;
     S(3)=SURVIVAL RATE IN YEAR 3;
     S(4)=SURVIVAL RATE IN YEAR 4;
PROC ESTIMATE NAME=GENERAL /* ALL PARAMETERS INDIVIDUALLY ESTIMATED */ ;
   INITIAL;   S(1)=0.6;   S(2)=0.7;   S(3)=0.5;   S(4)=0.6;
PROC ESTIMATE NAME=CONSTRAIN /* SETS OF 2 PARAMETERS CONSTRAINED */ ;
   INITIAL;   S(1)=0.6;   S(2)=0.7;   S(3)=0.5;   S(4)=0.6;
   CONSTRAINTS;   S(1)=S(2);   S(3)=S(4);
PROC TEST /* GENERATE LIKELIHOOD RATIO TEST OF ABOVE 2 MODELS */ ;
PROC STOP /* SIGNAL END OF ANALYSIS. */ ;
```

Fig. 1. Input to Program SURVIV to analyze the data and model in Table 1. Statements are separated by semicolons and observed recoveries are separated from the algebraic expression for the expected cell probabilities by colons.

hence they cannot be a function of the parameters.

PROC TEST performs a likelihood-ratio test between all pairs of models called in PROC ESTIMATE, given that the degrees of freedom available for each model are not equal. In this simple example, only 1 test is performed because only 1 pair of models is available: the general model with 4 parameters and the reduced model with 2 parameters. The null hypothesis tested is H_0: $S_1 = S_2$ and $S_3 = S_4$ vs. the alternative hypothesis H_a of at least 1 of the equalities not true.

Other procedures in Program SURVIV useful for performing survival-rate estimation experiments are SIMULATE and SAMPLE SIZE. PROC SIMULATE performs Monte Carlo simulation of a model entered via PROC MODEL. Use of PROC SIMULATE allows the researcher to determine the power of hypothesis tests and to determine expected confidence-interval length of interval estimates based on numbers of marked animals and estimated survival rates. PROC SAMPLE SIZE performs sample-size estimation for banding experiments, as described by Brownie et al. (1978:186–193). Both of these procedures are designed to aid the researcher in performing an optimal analysis to estimate survival rates. PROC SIMULATE is also useful in studying the operating characteristics of a model and the associated estimates.

NUMERICAL ALGORITHMS IN PROGRAM SURVIV

Program SURVIV uses a quasi-Newton method (Fletcher 1972) to minimize the likelihood function constructed from the cell probability expressions. Neither 1st nor 2nd partial derivatives of the likelihood function with respect to the parameters is required.

Parameters are constrained to the interval (S_L, S_B) with the transformation (Box 1966)

$$S_i = (S_U - S_L)(0.5 \sin[P_i] + 0.5) + S_L,$$

where P_i varies in the interval $(-\infty, +\infty)$, but S_i is always $S_L \leq S_i \leq S_U$.

EFFECTS OF PARAMETER BOUNDARIES ON ESTIMATES

A Monte Carlo simulation using PROC SIMULATE was performed to assess whether constraining the parameter estimates to their admissible range caused the remaining estimates to exhibit poor properties. A banding-analysis model (Model 1 of Brownie et al. 1978:15–20 or Seber 1970) was chosen for the simulation because both bias-adjusted and exact maximum-likelihood estimators (Brownie et al. 1978:16) were known. Banding was assumed to occur for $k = 10$ years, with $N_i = 75$ birds banded each year (a small sample size ensuring some estimates >1). Recoveries were followed for $l = 10$ years, thus 9 survival rates and 10 recovery rates can be estimated. Recovery rates were all set to 0.075, but the 9 survival rates were varied by year as 0.3, 0.4, 0.5, 0.6, 0.7, 0.6, 0.5, 0.4, and 0.3, respectively. Estimation was replicated for 1,000 simulations. The 9 estimates of survival rates were constrained to the interval [0, 1].

Simulation results varied (Table 2). At least 1 of the bias-adjusted survival estimates was >1 for 37.6% of the simulations. The numerical estimates are biased (as expected), but not as much as the unadjusted analytical estimates. Thus, the trade-off between the numerical estimates and the bias-adjusted estimates is bias vs. up to 12% inadmissible values for $S_5 = 0.7$. This trade-off would favor the numerical estimates as the true parameter proceeds toward unity because the numerical procedure shows less bias. It would favor the bias-adjusted analytical estimates as the parameter value proceeds toward 0.5 because there is less chance of obtaining inadmissible values as S decreases and the bias increases for the numerical estimator. However, as S approaches zero, small observed values will lead to greater probability of failure of the analytical estimator (i.e., zero denominator). The numerical procedure handles the zero cells by constraining 1 or more estimates to 0 or 1.

EXAMPLE USING BIOTELEMETRY DATA

Elk were radiocollared by Los Alamos National Laboratory personnel from 1978 to 1980 with radio collars having a 3-year battery life. Animals were initially collared only in the eastern Jemez Mountains of New Mexico, but in 1980 a 2nd herd was studied in the western Jemez Mountains. The age and sex of the animals collared (by location and year) and the collar returns varied (Table 3).

Program SURVIV can be used to test the hypothesis of interest: was the survival of western Jemez elk different from that of eastern Jemez elk? This hypothesis involves testing the equality of the 1980 survival rates for the western and eastern animals. That is, the survival parameters for western Jemez elk are set equal to the corresponding parameters for the eastern Jemez elk in the CONSTRAINTS section of PROC ESTIMATE. The results of this model are tested against the general model where all parameters are estimated individually.

Some of the survival rates are not estimable for this data set (e.g., adult male survival for 1978) because no animals were captured. Program SURVIV can handle such data because either the parameter can be left out of the cell probability statements, or the parameter can be fixed to a particular value and not estimated. The latter choice is preferred because, as the parameters are constrained equally in the reduced models, estimates can be generated.

The sequence of 5 models was tested

Table 2. Simulation of Model 1 of Brownie et al. (1978:15–20) for 1,000 replications using 3 different estimators of survival rates.

Parameter	True parameter value	Percentage of estimates inadmissible[a]	Mean of estimates	Standard error of estimates	95% CI	Percent bias[b]
Estimates From Numerical Procedure						
1	0.3	0.0	0.32	0.0064	0.31–0.33	6.5
2	0.4	0.0	0.43	0.0063	0.42–0.44	8.0
3	0.5	0.0	0.52	0.0063	0.51–0.54	4.6
4	0.6	0.0	0.62	0.0066	0.60–0.63	2.8
5	0.7	0.0	0.72	0.0068	0.70–0.73	2.6
6	0.6	0.0	0.64	0.0070	0.62–0.65	6.2
7	0.5	0.0	0.54	0.0075	0.52–0.55	7.6
8	0.4	0.0	0.45	0.0077	0.44–0.47	13.7
9	0.3	0.0	0.35	0.0079	0.33–0.36	15.2
Bias-adjusted Analytical Estimates						
1	0.3	1.0	0.30	0.0060	0.28–0.31	−1.2
2	0.4	0.8	0.41	0.0062	0.39–0.42	1.4
3	0.5	1.8	0.50	0.0065	0.48–0.51	−0.9
4	0.6	5.9	0.59	0.0076	0.58–0.60	−1.7
5	0.7	12.7	0.70	0.0095	0.68–0.72	0.5
6	0.6	9.2	0.61	0.0105	0.59–0.63	1.8
7	0.5	4.9	0.50	0.0087	0.49–0.51	−0.9
8	0.4	3.2	0.41	0.0084	0.40–0.43	3.3
9	0.3	1.9	0.30	0.0075	0.28–0.31	−0.7
Unadjusted Analytical Estimates						
1	0.3	1.0	0.33	0.0070	0.31–0.34	8.7
2	0.4	1.7	0.44	0.0069	0.42–0.45	9.3
3	0.5	2.9	0.53	0.0073	0.52–0.54	5.9
4	0.6	8.4	0.63	0.0085	0.61–0.65	5.1
5	0.7	19.6	0.77	0.0113	0.74–0.79	9.5
6	0.6	13.4	0.69	0.0154	0.66–0.72	14.6
7	0.5	8.5	0.57	0.0117	0.54–0.59	13.2
8	0.4	6.1	0.49	0.0122	0.47–0.51	22.4
9	0.3	5.0	0.38	0.0118	0.36–0.40	26.1

[a] $S_i > 1.0$.
[b] Percent bias $= 100(\hat{\theta} - \theta)/\theta$.

from these data. The most general model allows individual survival rates for each age- and sex class, year, and geographic area, making a total of 24 parameters. Only 20 of these parameters are estimable from the data available (Table 3). The 1st reduced model was that survival was the same in geographic areas, and the null hypothesis was not rejected ($P = 0.25$). Then, the hypothesis of survival equal between years was tested ($P = 0.51$), and again was not rejected. The next hypothesis was that survival was the same for age- and sex classes within the categories of hunted and nonhunted. That is, male calves and all female age-classes are pooled as 1 survival rate, and adult and yearling males as 1 survival rate, giving only 2 survival parameters to estimate. This highly reduced model was not significantly different from the previous model ($P = 0.17$). The last model of all survival rates constant was rejected by the goodness-of-fit test ($P \leq 0.01$) as well as by the likelihood-ratio test against the previous model ($P \leq 0.01$).

These results must be interpreted within the framework of statistical hypothesis testing, i.e., we "fail to reject" the null

Table 3. Return of radio collars from an elk biotelemetry study in the Jemez Mountains, New Mexico.

Year	Location	Sex	Age (yr)	Year returned				Battery failure or still alive
				N	1978	1979	1980	
1978	East Jemez	M	2½+	0	0	0	0	0
1979	East Jemez	M	2½+	1		0	1	0
1980	East Jemez	M	2½+	1			0	1
1980	West Jemez	M	2½+	1			1	0
1978	East Jemez	M	1½	1	1	0	0	0
1979	East Jemez	M	1½	1		1	0	0
1980	East Jemez	M	1½	2			1	1
1980	West Jemez	M	1½	5			0	5
1978	East Jemez	M	½	5	1	1	3	0
1979	East Jemez	M	½	2		0	1	1
1980	East Jemez	M	½	1			0	1
1980	West Jemez	M	½	0			0	0
1978	East Jemez	F	2½+	8	0	0	1	7
1979	East Jemez	F	2½+	3		2	0	1
1980	East Jemez	F	2½+	7			1	6
1980	West Jemez	F	2½+	5			0	5
1978	East Jemez	F	1½	0	0	0	0	0
1979	East Jemez	F	1½	0		0	0	0
1980	East Jemez	F	1½	4			0	4
1980	West Jemez	F	1½	0			0	0
1978	East Jemez	F	½	2	0	0	0	2
1979	East Jemez	F	½	4		0	0	4
1980	East Jemez	F	½	2			0	2
1980	West Jemez	F	½	2			0	2

hypothesis, but cannot accept it as true. The probability of a Type II error is high due to the small number of radio-collared animals relative to the number of years and age- and sex classes. Thus, the simple model with only 2 survival rates is probably not the true model for the elk population, but sufficient data are lacking to justify a more complex model. Therefore, this model is accepted as the model consistent with the observations taken. The estimated survival rates and their associated standard errors are 0.54 ± 0.10 and 0.93 ± 0.03 for the hunted and nonhunted segments of the radio-collared elk population, respectively.

However, a more extensive data set may indicate that survival rates do differ between age- and sex classes other than hunted and nonhunted. Certainly the small sample sizes in this example are not adequate for good estimates of survival, as indicated by the large estimated standard error of the 2 estimates.

Previous attempts to estimate survival rates in elk using tag recoveries (Kimball and Wolfe 1974, 1979) assumed that survival was constant from year to year. The advantage of the methods presented in this paper and those of Brownie et al. (1978) is that the estimates are less likely to be biased due to an incorrect model.

DISCUSSION

Instantaneous Survival Rates

Program SURVIV can also be used to estimate instantaneous survival rates. The finite survival rate (probability of survival for the time period between bandings) can be replaced with the appropriate algebraic expression containing the instantaneous survival rate. Rather than an estimate and standard error of a finite rate, the estimate and standard error of the instantaneous rate would be obtained. Boundary constraints on the instantaneous rate would differ from the default interval of [0, 1] appropriate for a finite rate.

Radio Failure

The failure of biotelemetry transmitters attached to experimental animals is assumed to occur only at the predicted time of battery depletion. Although failure with modern, commercially available equipment is generally uncommon, failures do occur. How radio failures are handled in the analysis depends on the cause of failure. If the failure is truly a random event (i.e., failure was not caused by some factor attributable to the animal's survival characteristics), then the animal can be discarded from the study and the number of radio-marked animals reduced by 1. However, often the failure of a radio is due to the animal's mortality, such as when it is illegally killed and the radio destroyed. In this case, the time of radio failure should be taken as the time of the animal's death. Obviously the biologist is not always sure why the radio failed. A small percentage of radio failures will not greatly change the results. However, significant numbers of radio failures lead to subjective decisions and will invalidate the results. The only solution to this problem is to use high-quality equipment so that random failures occur infrequently and time of death can be taken as the time of radio failure. Reliability studies of radio transmitters might also be performed by double-radio-tagging animals. Radio failure is not the same problem as that of band loss (Nelson et al. 1980), because radio failure is known to occur, whereas the loss of a band is rarely identified. The failure of a radio can be placed in a specified time interval, and the problem is 1 of excluding the animal from the experiment or assuming that the radio failure time is also the animal's time of death.

Radio-marked animals that move great distances from the population being studied present an additional complication. If they are known to have moved off the area, their survival rates may not be representative of the population and hence should be discarded. If the animal is not located again (i.e., the fate of the animal is not known), the animal can be classified as a radio failure.

Heterogeneous Survival Rates

Pollock and Raveling (1982) discussed the problem of heterogeneous survival rates of the members of the population. They found that the band-recovery models of Brownie et al. (1978) provide unbiased estimates of the average annual survival rate and the average recovery rate if both year-to-year recovery and survival rates are independent. They conjecture that such is unlikely because an individual with a high survival rate 1 year will probably continue to have a high survival rate the next year.

Estimates of survival rates from biotelemetry data using the techniques described would also be the average rates for heterogeneous populations, with the degree of bias depending on the correlation of survival rates among years. The ability of the biologist to separate the population based on age, sex, or any other identifiable characteristics of the animals helps in examining the causes of heterogeneity. The example of survival-rate estimation for elk in New Mexico illustrates the procedure for testing for differences in survival rates based on the geographic area where the animal was radiocollared. Tests of differences in survival rates of radio-marked animals could also be constructed based on the geographic area of recovery or past areas used. For example, Pollock and Raveling (1982) mentioned the case of waterfowl banded at the same location using different migration corridors. Radio-marked birds could be followed during migration to determine a category for separating the marked animals, and a test of the equality of survival rates for migration corridors could be conducted. This approach will not apply to waterfowl that are only banded because the model requires both survival and recovery rates to be estimated. Recovery rates for radio-marked birds are known except if radio failures become significant.

Current ecological theory suggests that individuals in a population must have heterogeneous survival rates over their lifetime because selection of the fittest implies that some animals are more fit than others. Presumably, fitness may be associated with survival rates. However, the variability of survival rates could not be too large or the proportion of the population with low survival rates would already be extinct. Thus, survival rates are undoubtedly heterogeneous, but not drastically so. The results of Pollock and Raveling (1982) show that the variance of the distribution of population survival rates must be fairly large to cause substantial bias in survival estimates. I suspect that the innate variability of a population's annual survival-rate distribution is not large enough to bias survival estimates significantly.

Sparse Data

One of the advantages of Program SURVIV over current methods available to handle the estimation of survival rates from band-recovery data is its ability to handle small sample sizes and sparse recovery matrices. However, this ability may become a hindrance to biologists analyzing such data because incorrect conclusions may often be reached if the user is unaware of the dangers of the statistical procedures. Small samples usually lead to hypothesis tests that lack power, and thus Type II errors are made (i.e., the null hypothesis is accepted when it should have been rejected). When the null hypothesis is accepted, usually nothing has been learned. Statistical tests only reject hypotheses, they do not prove them true. Thus, accepting a null hypothesis may mean the hypothesis is actually correct, or it may mean that a Type II error has been made. The only correct conclusion that can be made is that further research is required. Use of Program SURVIV on sparse data sets may lead to increased understanding of the biological situation if the principles of statistical hypothesis testing are well understood. However, if the program is used without this understanding, greater confusion of the biological picture will result.

Plant or Bird-nest Survivorship

The problem of estimating survival rates of sessile organisms such as plants or bird nestlings has been considered by Mayfield (1961, 1975), Johnson (1979), Hensler and Nichols (1981), and Bart and Robson (1982). The survival-rate estimation problem considered by Bart and Robson (1982) can be handled by SURVIV because maximum-likelihood estimates can be obtained from the cell probability expressions for the fates of the organisms under observation. Generally, for nest studies, the object is to estimate the daily survival rate (DSR) from which the survivorship during any interval of length l can be estimated as $(\mathrm{DSR})^l$.

Bart and Robson (1982) used mourning dove (*Zenaida macroura*) nest-success data from the North American Nest Record Card Program (NRCP). They estimated the daily nest-survival rate for doves from the center and periphery of the doves' North American range, and tested whether the daily nest-survival rate was constant for the 2 areas.

Nests are initially located then checked at a later time to determine if the nest is still active (i.e., has not been destroyed by predators or abandoned). The time between the initial location of the nest and the follow-up check will vary. Thus, the cell probability expression for nests that are observed L days apart in Program SURVIV would be S(1)**L, where S(1) is the daily survival rate. Suppose that 30 nests are located on day 3, and 15 are checked on day 6 with 13 still surviving, and the remaining 15 are checked on day 8 with 12 still surviving. Then the cell probability and cohort cards to estimate the daily survival rate [S(1)] for Program SURVIV would be

```
COHORT = 15;
  13: S(1)**3;
COHORT = 15;
  12: S(1)**5;
```

Program SURVIV would automatically supply the additional cell probability for each cohort by subtracting the above expressions from unity.

Reanalysis of the Bart and Robson (1982) data with Program SURVIV gives exactly the same estimates of daily nest survival for the 2 areas because they derive the maximum-likelihood estimates. However, the variances of the estimators are derived with slightly different assumptions about the number of nests located (cf. Bart and Robson 1982:1080–1081) and an ML estimator of the variance was not obtained. The variances calculated with Program SURVIV assume that the number of nests located is fixed (i.e., a predetermined number of nests is located). Thus, the standard errors of the estimates and the probability level of the test of differences between the center and peripheral areas from Program SURVIV do not agree with those of Bart and Robson (1982).

Bart and Robson (1982) also suggested extensions to the basic model they presented, such as a visitor-impact factor and changes in the daily survival rate through time. Such extensions are relatively easy to incorporate into models analyzed with program SURVIV. For example, the visitor-impact model they suggested does not fit the dove data significantly better than the standard model.

PROGRAM AVAILABILITY

Program SURVIV has been written in ANSI FORTRAN 77 (ANSI 1978) and is portable between machines having a FORTRAN 77 compiler. Conversion to any machine with a FORTRAN 77 compiler should not be difficult. However, conversion to a non-FORTRAN 77 compiler may be difficult because of the CHARACTER constructs used in Program SURVIV. A magnetic tape of the source code, example input, and a user's manual for the program are available upon request from the author.

LITERATURE CITED

ANSI. 1978. American national standard programming language FORTRAN. ANSI X3.9-1978. Am. Natl. Stand. Inst., New York, N.Y. (Not sequentially paginated.)

BART, J., AND D. S. ROBSON. 1982. Estimating survivorship when the subjects are visited periodically. Ecology 63:1078–1090.

BOX, M. J. 1966. A comparison of several current optimization methods, and the use of transformations in constrained problems. Comput. J. 9: 67–77.

BOYD, H. 1956. Statistics of the British population of the pink-footed goose. J. Anim. Ecol. 25:253–273.

BROWNIE, C., D. R. ANDERSON, K. P. BURNHAM, AND D. S. ROBSON. 1978. Statistical inference from band recovery data: a handbook. U.S. Dep. Inter., Fish and Wildl. Serv. Resour. Publ. 131. 212pp.

CORMACK, R. M. 1970. Statistical appendix. Pages 24–27 *in* R. A. Fordham. Mortality and population change of Dominican gulls in Wellington, New Zealand. J. Anim. Ecol. 39:13–27.

FLETCHER, R. 1972. FORTRAN subroutines for minimization by quasi-Newton methods. Atomic Energy Res. Establ. Rep. 7125. Harwell, U.K. 34pp.

HENSLER, G. L., AND J. D. NICHOLS. 1981. The Mayfield method of estimating nesting success: a model, estimators and simulation results. Wilson Bull. 93:42–53.

JOHNSON, D. H. 1974. Estimating survival rates from banding of adult and juvenile birds. J. Wildl. Manage. 38:290–297.

———. 1979. Estimating nest success: the Mayfield method and an alternative. Auk 96:651–661.

JOLLY, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52:225–247.

KIMBALL, J. F., JR., AND M. L. WOLFE. 1974. Population analysis of a northern Utah elk herd. J. Wildl. Manage. 38:161–174.

———, AND ———. 1979. Continuing studies of the demographics of a northern Utah elk population. Pages 20–28 *in* M. S. Boyce and L. D. Hayden-Wing, eds. North American elk: ecology, behavior, and management. Univ. Wyoming, Laramie.

MAYFIELD, H. 1961. Nesting success calculated from exposure. Wilson Bull. 73:255–261.

———. 1975. Suggestions for calculating nest success. Wilson Bull. 87:456–466.

MITRA, S. K. 1958. On the limiting power function of the frequency chi-square test. Ann. Math. Stat. 29:1221–1233.

MOOD, A. M., F. A. GRAYBILL, AND D. C. BOES. 1974. Introduction to the theory of statistics. 3rd ed. McGraw-Hill Book Co., New York, N.Y. 564pp.

MURTON, R. K. 1966. A statistical evaluation of the effect of wood-pigeon shooting as evidenced by the recoveries of ringed birds. The Stat. 16:183–202.

NELSON, L. J., D. R. ANDERSON, AND K. P. BURNHAM. 1980. The effect of band loss on estimates of annual survival. J. Field Ornithol. 51:30–38.

NORTH, P. M., AND R. M. CORMACK. 1981. On Seber's method for estimating age-specific bird survival rates from ringing recoveries. Biometrics 37:103–112.

POLLOCK, K. H., AND D. G. RAVELING. 1982. Assumption of modern band-recovery models, with emphasis on heterogeneous survival rates. J. Wildl. Manage. 46:88–98.

ROBSON, D. S. 1963. Maximum likelihood estimation of a sequence of annual survival rates from a capture-recapture series. Pages 330–335 *in* North Atlantic fish marking symp. Int. Comm. Northwest Atlantic Fish. Spec. Publ. 4.

ROSCOE, J. T., AND J. A. BYARS. 1971. An investigation of the restraints with respect to sample size commonly imposed on the use of the chi-square statistic. J. Am. Stat. Assoc. 66:755–759.

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

———. 1971. Estimating age-specific survival rates from bird bands when the reporting rate is constant. Biometrika 58:491–497.

———. 1972. Estimating survival rates from birdband returns. J. Wildl. Manage. 36:405–413.

YOUNGS, W. D., AND D. S. ROBSON. 1975. Estimating survival rate from tag returns: model tests and sample size determination. J. Fish. Res. Board Can. 32:2365–2371.

Received 5 April 1982.
Accepted 11 November 1982.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson.
Statistical Inference From Band Recovery Data — A Handbook. Second Edition. U.S. Fish Wildl. Serv., Resour. Pub. 156.
1985.

This Handbook presents 14 models for the analysis of bird banding or fish tagging studies. Maximum likelihood estimators of survival and recovery rates are given with their estimators of sampling variances and covariances. Goodness of fit tests are presented as well as log-likelihood ratio tests of hypotheses. Two computer software packages are illustrated in the Handbook and available.

Keywords: Banding, Estimation, Maximum likelihood, Models, Population dynamics, Survival, Tagging.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson.
Statistical Inference From Band Recovery Data — A Handbook. Second Edition. U.S. Fish Wildl. Serv., Resour. Pub. 156.
1985.

This Handbook presents 14 models for the analysis of bird banding or fish tagging studies. Maximum likelihood estimators of survival and recovery rates are given with their estimators of sampling variances and covariances. Goodness of fit tests are presented as well as log-likelihood ratio tests of hypotheses. Two computer software packages are illustrated in the Handbook and available.

Keywords: Banding, Estimation, Maximum likelihood, Models, Population dynamics, Survival, Tagging.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson.
Statistical Inference From Band Recovery Data — A Handbook. Second Edition. U.S. Fish Wildl. Serv., Resour. Pub. 156.
1985.

This Handbook presents 14 models for the analysis of bird banding or fish tagging studies. Maximum likelihood estimators of survival and recovery rates are given with their estimators of sampling variances and covariances. Goodness of fit tests are presented as well as log-likelihood ratio tests of hypotheses. Two computer software packages are illustrated in the Handbook and available.

Keywords: Banding, Estimation, Maximum likelihood, Models, Population dynamics, Survival, Tagging.

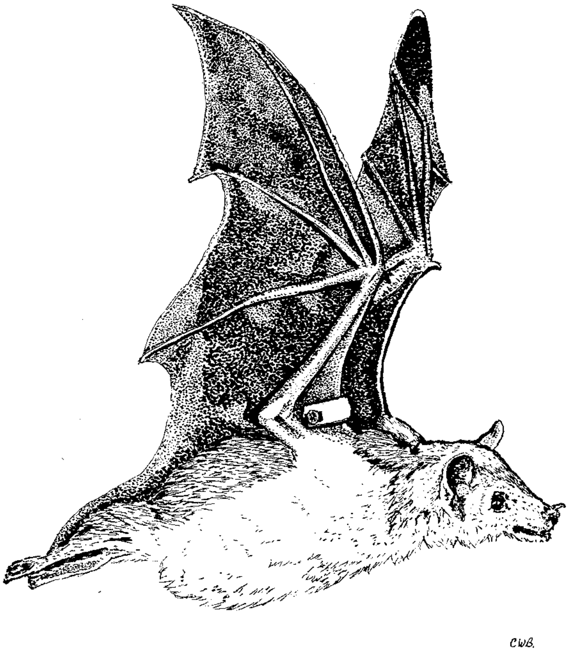
Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson.
Statistical Inference From Band Recovery Data — A Handbook. Second Edition. U.S. Fish Wildl. Serv., Resour. Pub. 156.
1985.

This Handbook presents 14 models for the analysis of bird banding or fish tagging studies. Maximum likelihood estimators of survival and recovery rates are given with their estimators of sampling variances and covariances. Goodness of fit tests are presented as well as log-likelihood ratio tests of hypotheses. Two computer software packages are illustrated in the Handbook and available.

Keywords: Banding, Estimation, Maximum likelihood, Models, Population dynamics, Survival, Tagging.

A list of current *Resource Publications* follows.

150. Microscopic Anatomy of Salmonids: An Atlas, by William T. Yasutake and Joseph H. Wales. 1983.
151. Avian Use of Forest Habitats in the Pembina Hills of Northeastern North Dakota, by Craig A. Faanes and Jonathan M. Andrew. 1983. 24 pp.
152. National Pesticide Monitoring Program: Organochlorine Residues in Freshwater Fish, by Christopher J. Schmitt, Michael A Ribick, J. Larry Ludke and Thomas M. May. 1983. 62 pp.
153. Handbook of Toxicity of Pesticides to Wildlife, by Rick H. Hudson, Richard K. Tucker and M.A. Haegele. 1984. 97 pp.
154. Nonconsumptive Use of Wildlife in the United States, by William W. Shaw and William R. Mangun. 1984. 20 pp.
155. Ecology and Management of the Bullfrog, by R. Bruce Bury and Jill A. Whelan. 1984. 23 pp.



March 3, 1849

UNITED STATES
DEPARTMENT OF THE INTERIOR
FISH AND WILDLIFE SERVICE
OFFICE OF INFORMATION TRANSFER
EDITORIAL SECTION
AYLESWORTH HALL, CSU
FORT COLLINS, COLORADO 80523