

## 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,

$$z = (\bar{S}_1 - \bar{S}_2) / \sqrt{\text{var}(\bar{S}_1) + \text{var}(\bar{S}_2) - 2 \text{cov}(\bar{S}_1, \bar{S}_2)}$$

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, ...,  $\infty$  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 ( $\phi$ ), population size ( $P$ ), births ( $B$ ), 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  $\phi_i$  could be modeled (assumed) to be a single constant  $\phi$ . 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  $i$  ( $s_i$ ) equaling 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 restricted 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:

Release time	Recaptured history									
	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 <sup>1/</sup>	Published simulation studies	Well developed theory	Assumptions realistic <sup>2/</sup>	Generally a "useful" method <sup>3/</sup>
<b>This Handbook</b>																	
$M_0$	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	X
$M_3$	X		X		X	X	X	X	X	X	X	X	X	X	X	X	X
$H_{01}$	X	X	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
$H_2$	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	
$H_4$	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	
$H_6$	X	X	X		X	X	X	X	X	X	X	X	X		X	X	X
$H_7$	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
<b>Standard Methods Previously Used in Banding</b>																	
<b>Data Analysis<sup>9/</sup></b>																	
Composite dynamic life table (Hickey 1952)			X <sup>5/</sup>	X				X									
Composite time-specific life table (Hickey 1952)		X	X					X									
Dynamic life table (Hickey 1952)			X <sup>5/</sup>														
Time-specific life table (Hickey 1952)		X						X									
Haldane (1953, 1955) (for finite $l=k$ )	X		X		X			X	X	X	X						
Relative recovery rate (Ricker 1958)	X	X	X <sup>6/</sup>	X				X	X	X <sup>7/</sup>	X					X	
Williams (1963) (see Geis & Taber 1963)		X	X					X									
<b>Other Selected Methods</b>																	
Chapman-Robson (1960) <sup>11/</sup>	X		X <sup>8/</sup>		X	X	X	X	- <sup>10/</sup>	X	X	X	X	X	X	X	X
Cormack (1964)	X	X	X			X	X	X	- <sup>10/</sup>	X	X	X	X	X	X	X	X
Fisher-Ford (1947)	X		X		X	X	X	X	- <sup>10/</sup>	X	X	X	X	X	X	X	X
Fordham (1970)	X	X	X			X	X	X	- <sup>10/</sup>	X	X	X	X	X	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	X	X	X	X	X
Jolly (1965)	X	X	X			X	X	X	- <sup>10/</sup>	X	X	X	X	X	X	X	X
Manly-Parr (1968)	X	X	X	X		X	X	X	- <sup>10/</sup>	X	X	X	X	X	X	X	X
Roberts (1971)	X		X			X	X	X	- <sup>10/</sup>	X	X	X	X	X	X	X	X
Seber (1971)	X		X	X				X	X	X	X	X	X		X		

<sup>1/</sup> General, usually published, computer codes in FORTRAN.  
<sup>2/</sup> Relative to bird banding experiments, primarily.

(Footnotes continued next page)

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:

Release time $i$	First recapture after 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:

$i$	Known releases at time $i$	First recaptured at time $j$ given release 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  $j$  ( $= 2, \dots, k$ ) 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.

←

<sup>3/</sup> Our judgment based on a number of considerations.  
<sup>4/</sup> Individual parameters are not estimable under  $H_1$ .  
<sup>5/</sup> Age-specific if, and only if, additional, rather unlikely, assumptions are specified.  
<sup>6/</sup> Original development was age-independent: more recently used on data for both young and adults.  
<sup>7/</sup> Estimators of variance given if  $k = \ell = 2$ .  
<sup>8/</sup> Only if  $\ell = \infty$ .  
<sup>9/</sup> Many of these methods are described throughout the literature on banding data analysis (e.g., Geis and Taber 1963, Geis 1972a, 1972b).  
<sup>10/</sup> Not relevant in live recapture experiments.  
<sup>11/</sup> 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	$N_1$	$N_1\phi_1p_2$	$N_1\phi_1q_2\phi_2p_3$	$N_1\phi_1q_2\phi_2q_3\phi_3p_4$	...	$N_1\phi_1q_2\phi_2q_3\phi_3q_4\cdots\phi_kp_{k+1}$
2	$N_2$		$N_2\phi_2p_3$	$N_2\phi_2q_3\phi_3p_4$	...	$N_2\phi_2q_3\phi_3q_4\cdots\phi_kp_{k+1}$
3	$N_3$			$N_3\phi_3p_4$	...	$N_3\phi_3q_4\cdots\phi_kp_{k+1}$
.	.					
.	.					
.	.					
.	.					
k	$N_k$					$N_k\phi_kp_{k+1}$

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),
- $\phi_i$  = 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  $\phi_i, p_i$  are found as:

$$\hat{\phi}_i = \hat{f}_i + \hat{S}_i, \quad i = 1 \dots k - 1$$

$$\hat{p}_{i+1} = \hat{f}_i / (\hat{f}_i + \hat{S}_i), \quad i = 1 \dots k,$$

$\hat{f}_k = \widehat{\phi_k p_{k+1}}$  cannot be decomposed;  $\phi_k$  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:

$$\frac{E(m_j)}{N_i} = \begin{cases} \alpha_j, & j = i+1, \\ \beta_i \cdots \beta_{j-2} \alpha_{j-1}, & j > i+1. \end{cases}$$

Here,  $\alpha_j$  = 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,  $\beta_j$  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  $\lambda_i$  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  $\phi_i$ , 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  $\alpha_i = (1 - \phi_i)\lambda_i$ ; 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  $\phi_i$  and  $p_i$ . Then the Jolly-Seber estimate of population size ( $P_i$ ) is just

$$\hat{P}_i = \frac{n_i}{\hat{p}_i}, \quad i = 2, \dots, k \text{ (with } k+1 \text{ capture times),}$$

where  $n_i$  = total number of animals captured at time  $i$  ( $n_i = u_i + m_i$ ,  $m_i = m_{i1} + \dots + m_{i-1,i}$  = 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-re-sampling 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, ... , $\ell$
1	$N_1$	$R_{11}$	$R_{12}$	$R_{13}, \dots, R_{1\ell}$
2	$N_2$		$R_{22}$	$R_{23}, \dots, R_{2\ell}$

As in Chapter 2,  $C_1 = R_{11}$ , 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

$$\mathbf{E}\left(\frac{R_2}{N_2}\right) = \mathbf{E}\left(\frac{R_{12} + R_{13} + \dots + R_{1\ell}}{N_2^*}\right) = \mathbf{E}\left(\frac{R_1 - C_1}{N_2^*}\right).$$

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,

$$E(N_{i+1}^* | N_i^*, N_i) = (N_i^* + N_i) S_i.$$

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

$$E\left(\frac{R_i}{N_i}\right).$$

Of the  $N_i^*$  survivors of previously banded birds, there are  $T_i - R_i$  total recoveries. Thus, by the assumptions of Model 1

$$E\left(\frac{R_i}{N_i}\right) = E\left(\frac{T_i - R_i}{N_i^*} \mid N_i^*\right)$$

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  $T_i - C_i = T_{i+1} - R_{i+1}$ , 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 ( $\hat{S}_i, \hat{f}_i$  or  $\hat{S}, \hat{f}$ ), 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  $\bar{S}$  on estimates of average annual recovery rate  $\hat{f}$  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  $\bar{S}$  and  $\hat{f}$ . In Example 2.1, the estimated sampling correlation of  $\bar{S}$  and  $\hat{f}$  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  $\bar{S}$ ,  $\hat{f}$  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  $\bar{S}$  and  $\hat{f}$  tells us nothing about how true survival and recovery rates are related; it only reflects the extent to which the estimators  $\bar{S}$  and  $\hat{f}$  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  $k = l = 12$  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  $S_i = S$  and  $f_i = f$ . The pairs of estimators  $\bar{S}$  and  $\hat{f}$  and all 20 independent repetitions are shown below:

$\hat{f}$	$\bar{S}$
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  $\bar{S}$  and  $\hat{f}$  computed from these 20 pairs is  $-0.712$ . A plot of these points reveals a very evident linear relationship of  $\bar{S}$  on  $\hat{f}$ ; 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,  $\bar{S}$  on  $\hat{f}$  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  $\bar{S}_i$  and  $\hat{f}_i$  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  $t_1, t_2, \dots, t_k$ , and let  $d_i = t_{i+1} - t_i, i = 1, \dots, k-1$  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

$$H_0: S_i = S^{d_i}, i = 1, \dots, k-1.$$

The appropriate structure of Model 2 to represent a time-constant survival rate is now

$$E(R_{ij}) = \begin{cases} N_i f_i & , i = j \\ N_i S^{(d_i)} S^{(d_{i+1})} \dots S^{(d_{j-1})} f_j & , i < j \end{cases}$$

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  $d_1, \dots, d_{k-1}$ . 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

$$\tilde{S}_G = \left( \prod_{i=1}^{k-1} \tilde{S}_i \right)^{1/d}, \tag{8.5.1}$$

where

$$d = d_1 + \dots + d_{k-1}.$$

This estimator  $\tilde{S}_G$  then estimates average survival per time unit during the total time period  $t_1$  to  $t_k$  (which is of length  $d$ ).

#### 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  $d_i = d = 1$ , then (8.5.1) gives  $\tilde{S}_G$  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

$$\bar{S}_A = \frac{\sum_{i=1}^{k-1} \tilde{S}_i}{k-1}, \quad (8.5.2)$$

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

$$\bar{S} = \frac{\sum_{i=1}^{k-1} S_i}{k-1}$$

Another way of approaching this question is to ask if we know the true annual survival rates  $S_1, \dots, S_{k-1}$ , 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  $E(S_i)$  which is estimated by  $\bar{S}$ . Given this is the "average" we want to estimate, then the best estimator is  $\bar{S}_A$  as given in formula (8.5.2.) Moreover, for Model 1 (and others given here) we know

$$E(\tilde{S}_i) = S_i(1 - (1 - \rho_{i+1})^{N_i+1}),$$

which means for practical purposes  $\tilde{S}_i$  is unbiased. It follows  $\bar{S}_A$  is an unbiased estimator of  $\bar{S}$ . 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  $E(\tilde{S}_G) < \bar{S}$  will be true.

Also, because  $\bar{S}_A$  is a linear function of the  $\tilde{S}_i$ , the formula for the sampling variance of  $\bar{S}_A$  is known exactly in terms of variances and covariances of the  $\tilde{S}_i$ . Because  $\tilde{S}_G$  is a highly nonlinear function of the  $\tilde{S}_i$ , the formula for  $\text{VAR}(\tilde{S}_G)$  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  $\tilde{S}_1, \dots, \tilde{S}_{k-1}$  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  $\text{var}(\tilde{S}_i)$  and  $\text{cov}(\tilde{S}_i, \tilde{S}_j)$ ,  $i \neq j$ , be the estimated sampling variances and covariances of these estimates. For any set of constants  $c_1, \dots, c_{k-1}$ , the null hypothesis

$$H_0: c_1 S_1 + c_2 S_2 + \dots + c_{k-1} S_{k-1} = 0$$

can be tested with the test statistic

$$z = \frac{c_1 \tilde{S}_1 + c_2 \tilde{S}_2 + \dots + c_{k-1} \tilde{S}_{k-1}}{\sqrt{\sum_{i=1}^{k-1} (c_i)^2 \text{var}(\tilde{S}_i) + 2 \sum_{i=1}^{k-2} \sum_{j=i+1}^{k-1} c_i c_j \text{cov}(\tilde{S}_i, \tilde{S}_j)}}$$

The alternative hypothesis may be two-sided

$$H_a: \sum_{i=1}^{k-1} c_i S_i \neq 0,$$

or it may be one-sided, for example

$$H_a: \sum_{i=1}^{k-1} c_i S_i > 0.$$

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

$$H_0: \frac{S_1+S_2}{2} = \frac{S_3+S_4}{2}$$

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  $c_3=c_4=-1$ . Given this choice of constants the test statistic value is

$$z = \bar{\delta} / \sqrt{\text{var}(\bar{\delta})}$$

where  $\bar{\delta} = (\bar{S}_1 + \bar{S}_2) - (\bar{S}_3 + \bar{S}_4)$ , and  $\text{var}(\bar{\delta}) = \text{var}(\bar{S}_1) + \text{var}(\bar{S}_2) + \text{var}(\bar{S}_3) + \text{var}(\bar{S}_4) + 2\text{cov}(\bar{S}_1, \bar{S}_2) - 2\text{cov}(\bar{S}_2, \bar{S}_3) + 2\text{cov}(\bar{S}_3, \bar{S}_4)$ .

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

$$H_a: \frac{S_1+S_2}{2} > \frac{S_3+S_4}{2}$$

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

$i = \text{year}$	$\bar{S}_i$	$\text{se}(\bar{S}_i)$
1964	68.17	6.95
1965	64.65	6.26
1968	74.25	6.13

From Example 2.6 the sampling covariance of  $\bar{S}_{64}$  and  $\bar{S}_{65}$  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  $\bar{S}_i$  and  $\text{se}(\bar{S}_i)$  they must be multiplied by 10,000.) The sampling covariances of  $\bar{S}_{64}$  and  $\bar{S}_{68}$ , and  $\bar{S}_{65}$  and  $\bar{S}_{68}$  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

$$z = \frac{\bar{S}_{68} - \bar{S}_{64}}{\sqrt{\text{var}(\bar{S}_{68}) + \text{var}(\bar{S}_{64})}}$$

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

$$z = \frac{\frac{\bar{S}_{65} + \bar{S}_{68}}{2} - \bar{S}_{64}}{\sqrt{(1/2)^2 \text{var}(\bar{S}_{65}) + (1/2)^2 \text{var}(\bar{S}_{68}) + \text{var}(\bar{S}_{64}) - 2(1/2) \text{cov}(\bar{S}_{64}, \bar{S}_{65})}}$$

$$= \frac{\frac{64.65 + 74.25}{2} - 68.17}{\sqrt{(.25)(6.26)^2 + (.25)(6.13)^2 + (6.95)^2 + (29.48)}} = \frac{1.28}{9.85} = 0.13.$$

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.