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
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
2 postseason	M_2			Q_{23}	Q_{24}	Q_2
3 preseason	N_3			T_3 R_{33}	R_{34}	R_3
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
1 postseason	350		30	19	12	6	67
2 preseason	500		48	24	24	11	107
2 postseason	400			27	22	12	61
3 preseason	500			41	30	20	91
3 postseason	800				70	33	103
4 preseason	400				44	21	65
4 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_1f_1	$N_1h_1n_1f_2$	$N_1h_1n_1h_2n_2f_3$	$N_1h_1n_1h_2n_2h_3n_3f_4$
	postseason	M_1		$M_1n_1f_2$	$M_1n_1h_2n_2f_3$	$M_1n_1h_2n_2h_3n_3f_4$
2	preseason	N_2		N_2f_2	$N_2h_2n_2f_3$	$N_2h_2n_2h_3n_3f_4$
	postseason	M_2			$M_2n_2f_3$	$M_2n_2h_3n_3f_4$
3	preseason	N_3			N_3f_3	$N_3h_3n_3f_4$
	postseason	M_3				$M_3n_3f_4$
4	preseason	N_4				N_4f_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 var$(\hat{f}_i)$):

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

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

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

$$\text{var}(\hat{n}_i) = (\hat{n}_i)^2 \left[\frac{1}{Q_i} - \frac{1}{M_i} + \frac{1}{R_{i+1}} - \frac{1}{N_{i+1}}\right] \qquad , i = 1, \ldots, k-1\,,$$

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

Estimates of the corresponding standard errors are given by, for example, se$(\hat{f}_i) = \sqrt{\text{var}(\hat{f}_i)}$.

For example, using the data of Table 7.2,

$$\text{var}(\hat{f}_2) = (\hat{f}_2)^2 \left[\frac{1}{R_2} - \frac{1}{N_2} + \frac{1}{C_2} - \frac{1}{T_2}\right] = (0.0933)^2 \left[\frac{1}{107} - \frac{1}{500} + \frac{1}{109} - \frac{1}{250}\right] = 0.00010899$$

$$\text{se}(\hat{f}_2) = \sqrt{0.00010899} = 0.0104\,,$$

$$1.96 \times \text{se}(\hat{f}_2) = 0.0204\,,$$

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 $\hat{f}_i$, $\tilde{h}_i$, 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

$$\hat{S}_i = \hat{h}_i \hat{n}_i = \frac{R_i}{N_i} \frac{T_i - C_i}{T_i} \frac{N_{i+1}}{R_{i+1}} \quad , i = 1, \ldots, k-1 ,$$

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

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

$$\text{cov}(\tilde{h}_i, \tilde{n}_i) = -\tilde{h}_i \tilde{n}_i \left[\frac{1}{Q_i} - \frac{1}{M_i}\right] \quad , i = 1, \ldots, k-1 ,$$

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

$$\text{cov}(\tilde{h}_{i+1}, \tilde{n}_i) = -\tilde{h}_{i+1} \tilde{n}_i \left[\frac{1}{R_{i+1}} - \frac{1}{N_{i+1}}\right] \quad , i = 1, \ldots, k-2 .$$

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

$$\text{corr}(\hat{f}_i, \tilde{h}_i) = \frac{\text{cov}(\hat{f}_i, \tilde{h}_i)}{\text{se}(\hat{f}_i)\ \text{se}(\tilde{h}_i)} \quad , i = 1, \ldots, k-1 .$$

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 $(i=1,\ldots,k-1)$, as defined in Section 7.1, and

f_i^* = recovery rate in year i for banded adults released at the i^{th} preseason banding, $i=1,\ldots,k$,

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, $i=2,\ldots,k$.

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 and time of banding		Number banded	Year of recovery 1	2	3	4
1	preseason	N_1	$N_1f_1^*$	$N_1h_1n_1f_2$	$N_1h_1n_1h_2n_2f_3$	$N_1h_1n_1h_2n_2h_3n_3f_4$
	postseason	M_1		$M_1n_1f_2$	$M_1n_1h_2n_2f_3$	$M_1n_1h_2n_2h_3n_3f_4$
2	preseason	N_2		$N_2f_2^*$	$N_2h_2n_2f_3$	$N_2h_2n_2h_3n_3f_4$
	postseason	M_2			$M_2n_2f_3$	$M_2n_2h_3n_3f_4$
3	preseason	N_3			$N_3f_3^*$	$N_3h_3n_3f_4$
	postseason	M_3				$M_3n_3f_4$
4	preseason	N_4				$N_4f_4^*$

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

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

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

$$\text{var}(\tilde{h}_i) = (\tilde{h}_i)^2 \left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i} + \frac{1}{Q_i} - \frac{1}{M_i} \right] \qquad , i=1,\ldots,k-1,$$

$$\text{var}(\tilde{n}_i) = (\tilde{n}_i)^2 \left[\frac{1}{Q_i} - \frac{1}{M_i} + \frac{1}{R_{i+1} - R_{i+1,i+1}} - \frac{1}{N_{i+1}} + \frac{1}{T_{i+1} - R_{i+1} - C_{i+1} + R_{i+1,i+1}} - \frac{1}{T_{i+1} - R_{i+1}} \right] \qquad , i=1,\ldots,k-2.$$

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,

$$\text{var}(\hat{f}_3) = (\hat{f}_3)^2 \left[\frac{1}{R_3 - R_{33}} - \frac{1}{N_3} + \frac{1}{T_3 - R_3 - C_3 + R_{33}} + \frac{1}{C_3 - T_{33}}\right] = (0.0757)^2 \left[\frac{1}{50} - \frac{1}{500} + \frac{1}{115} + \frac{1}{87}\right] = 0.00021885\,,$$

and

$$\text{se}(\hat{f}_3) = \sqrt{0.00021885} = 0.0148\,.$$

To compute the 95% confidence interval for f_3 first compute $1.96 \times \text{se}(\hat{f}_3) = 0.0290$, 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

$$\text{cov}(\hat{f}_i^*, \hat{f}_i) = -\hat{f}_i^* \hat{f}_i / N_i \qquad , i = 2, \ldots, k-1\,,$$

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

$$\text{cov}(\hat{f}_{i+1}^*, \tilde{n}_i) = \hat{f}_{i+1}^* \tilde{n}_i / N_{i+1} \qquad , i = 1, \ldots, k-2\,,$$

$$\text{cov}(\hat{f}_i, \tilde{h}_i) = \hat{f}_i \tilde{h}_i \left[\frac{1}{R_i - R_{ii}} - \frac{1}{N_i}\right] \qquad , i = 2, \ldots, k-1\,,$$

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

$$\text{cov}(\tilde{h}_i,\tilde{n}_i) = -\tilde{h}_i\,\tilde{n}_i\left[\frac{1}{Q_i}-\frac{1}{M_i}\right] \qquad ,i=1,\ldots,k-2,$$

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

The corresponding correlations are then estimated in the usual way, for example,

$$\text{cov}(\hat{f}_3,\tilde{h}_3) = \hat{f}_3\tilde{h}_3\left[\frac{1}{R_3-R_{33}}-\frac{1}{N_3}\right] = (0.0757)\times(0.7702)\left[\frac{1}{50}-\frac{1}{500}\right] = 0.001049,$$

$$\text{corr}(\hat{f}_3,\tilde{h}_3) = \frac{\text{cov}(\hat{f}_3,\tilde{h}_3)}{\text{se}(\hat{f}_3)\,\text{se}(\tilde{h}_3)} = \frac{0.001049}{(0.0148)\times(0.1253)} = 0.566.$$

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 $\text{cov}(\hat{f}_i^*,\hat{f}_i)$	Correlation $\text{corr}(\hat{f}_i^*,\hat{f}_i)$	Covariance $\text{cov}(\hat{f}_i^*,\tilde{h}_i)$	Correlation $\text{corr}(\hat{f}_i^*,\tilde{h}_i)$
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
	$\text{cov}(\hat{f}_{i+1}^*,\tilde{n}_i)$	$\text{corr}(\hat{f}_{i+1}^*,\tilde{n}_i)$	$\text{cov}(\hat{f}_i,\tilde{h}_i)$	$\text{corr}(\hat{f}_i,\tilde{h}_i)$
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
	$\text{cov}(\hat{f}_{i+1},\tilde{n}_i)$	$\text{corr}(\hat{f}_{i+1},\tilde{n}_i)$	$\text{cov}(\tilde{h}_i,\tilde{n}_i)$	$\text{corr}(\tilde{h}_i,\tilde{n}_i)$
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
	$\text{cov}(\tilde{h}_{i+1},\tilde{n}_i)$	$\text{corr}(\tilde{h}_{i+1},\tilde{n}_i)$		
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 / 48 | 59	0.12
$i=3$	87 | 115 / 41 | 50	0.10
$i=4$	176 | 92 / 44 | 21	0.10

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).