CHAPTER 3
CAPTURE-RECAPTURE MODELS

In the typical capture-recapture study, a main objective is to estimate population size N. Neither the true value of N nor the correct assumptions to make about capture probabilities are known. The scientific problem is, first, to formulate a model, or a series of models, and to select the most appropriate model based on the actual data, then, given the model, to compute the most efficient estimate of N and the reliability of that estimate.

Common practice has been to compute a few summary statistics from the entire X matrix of captures, then to compute an estimate of N based on these summary statistics, using one of dozens of published estimators, and to stop there without giving real justification for the selection of the estimator used. The assumptions are not tested and the sampling variance of $\hat{N}$ is not estimated. This is not an objective, scientific procedure. Because there are numerous published estimators, different persons can get quite different estimates with the same data.

Table 3.1 presents common summary statistics for a 10-occasion, simulated capture-recapture study. From just these summary statistics, we computed nine estimates of population size: 175, 183, 187, 197, 200, 202, 234, 245, and 260. Each of these numbers derives from a different published estimator of

TABLE 3.1. Some common summary statistics from a simulated 10-occasion (10-day) capture-recapture study. The number of captures each day is n_j. The number of unmarked animals caught on day j is u_j. The number of marked animals in the population just before the j^{th} capture occasion is M_j. At the end of the study, the number of animals captured exactly j times is f_j. Based on just these summary statistics, more than a dozen different estimators of population size can be computed. We computed, from this one set of data, some of the more common ones plus the estimators we are recommending; the range of values for $\hat{N}$ was 175 to 260.

Capture Occasion j	Animals Caught n_j	Newly Caught u_j	Total Caught M_j	Capture Frequencies f_j
1	38	38	0	26
2	45	34	38	43
3	57	27	72	39
4	56	23	99	28
5	65	17	122	26
6	72	11	139	13
7	59	15	150	4
8	62	5	165	1
9	64	6	170	0
10	67	4	176	0

population size for capture-recapture data for closed populations. Still other estimates could be computed from the data in Table 3.1 or from alternative summaries of the basic data. It is quite possible that a real study producing these data, if published by different people, could have an estimate ranging from 175 to 260, with no measure of the estimate's validity or precision. Yet, in this study, there are ample data to allow an objective assessment of the assumptions underlying any reasonable estimator and, thereby, to choose an appropriate estimator and to give its reliability.

Given a set of capture-recapture data, several questions need to be answered. What are the plausible sources of variation in capture probabilities? Hence, what is a plausible model? What is a good estimator of N? Is there a good estimator, based on the data in hand? Given an estimator, what is its reliability? Without such a rigorous framework for the analysis of capture-recapture data, estimates of N are not defensible. We present here the methodology for such an approach (cf. *Nichols et al. 1981; Pollock 1981b*).

This chapter provides the basic information on the eight models underlying our suggested methods of estimating population size from capture-recapture data under the assumption of population closure. Central to this chapter are the assumptions of the models, the tests of the assumptions, the estimators based on the models, and confidence interval construction. Tests of assumptions are either tests between models or goodness of fit tests applicable to individual models. Based on these tests there is an objective rule for the selection of the "best" model to describe any given set of capture data. Five of the eight models have a corresponding "good" estimator of N.

In this chapter, as elsewhere in this book, the emphasis is primarily on concepts; mathematical details are given in *Otis et al. (1978)*. Each example in this chapter is based on simulated data that fit one of the eight models exactly. To prepare the examples, 10 simulated data sets were generated for each model. The replications represented by these 10 cases for each model are used to illustrate the naturally occurring sampling variations to be expected in estimates of population size. The reader should fully comprehend the material in Chapter 3 before going to Chapter 6, which presents examples using real data.

The typical literature on capture-recapture methods with closed models concentrates on estimating population size N. However, there are three critical considerations in constructing capture-type models.

- What does population size N mean? Because no capture model has anything analogous to the sides of the urn in ball and urn models, consideration often must be given to converting N to a density of animals per unit area, N/A, where A represents the size of the area being used by the population. Thus, one must ask, to what area does N relate? We discuss this problem (for closed models) in Chapter 5.
- Should the model be demographically open or closed? Some comments on this problem are in Chapter 8.
- How can the parameters of the model vary over the three factors time, behavior, and heterogeneity? We are dealing only with closed models in this chapter (and in most of this primer); consequently, the only parameters of the capture model are the capture probabilities and the unknown population size N.

Modeling Capture Probabilities

Development of the early capture models was motivated by thoughts of ball and urn studies. Imagine an urn filled with 100 small white balls. One reaches in and removes, say, 30 ($= n_1$) balls, marks them by coloring them black, and returns them all to the urn. Thus, there are $n_1 = M_2$ marked balls in the urn when the second sample is taken. (In our notation the number of marked balls, or animals, just before the j^{th} capture sample is M_j. Given 100% survival of marked animals, M_j is the total number of animals marked and released before the j^{th} sample is taken.) The 100 balls in the urn are mixed well and a second random sample of size, say 36 ($= n_2$), is drawn. Some of these, say 10 ($= m_2$), will be black (previously marked) and the rest (26 in this example) will be white (unmarked). We let $u_2 = n_2 - m_2 = 36 - 10$ represent the unmarked balls in the sample. The basic assumption is that on the average, that is, in terms of statistical expectations, the ratio of marked balls to total balls in the population will be the same as the ratio of marked balls in the sample—$(n_1/N) = (m_2/n_2) \equiv M_2/N$.

In this example, therefore, the ratios to be set equal (and solved for $\hat{N}$) are $30/\hat{N}$ and $10/36$:

$$\frac{30}{\hat{N}} = \frac{10}{36}$$

or

$$\hat{N} = \frac{30 \times 36}{10}$$

$$= 108 \ .$$

In terms of the symbols, n_1 (animals caught, hence marked, in the first sample), n_2 (total animals caught in the second sample), and m_2 (marked animals caught in the second sample),

$$\hat{N} = \frac{n_1 n_2}{m_2} \ .$$

In ecology this equation is known as the Petersen-Lincoln estimator (see *Seber 1973:60*).

One can continue to draw samples, recording on each occasion the numbers of marked, m_j, and the numbers of unmarked, u_j, balls. Each time, white balls are colored black before all are returned to the urn. This conceptual "model" of capture studies has dominated the ecology literature for 30 years. Yet, it is illogical to apply such a ball and urn model to biological populations because capture probabilities vary in real populations and because there is not always an analogy in biological populations to the sides of the urn. This lack of analogy is what creates difficulties in interpreting what N means.

The process of capturing living organisms is not analogous to the process of stirring up balls in an urn and drawing a random sample. One cannot mix the population after each capture occasion; moreover, animals will not mix themselves randomly and the capture process itself is potentially very complex. Capture probabilities can vary over time, because of weather or the amount of effort expended on any occasion to capture animals. Individual capture probabilities can vary because of innate factors (heterogeneity), such as the age and sex of the animal, its social status, the number of traps in its home range, or its inquisitiveness. Finally, animals often exhibit a behavioral response to capture; hence the capture probability of an individual can easily change after first capture. Ball and urn models never have allowed for heterogeneity and only rarely (and recently) have allowed for limited degrees of behavioral variations in capture probabilities, but models for estimating the abundance of populations of living organisms must allow for these sorts of variations.

The most general conceptual model of capture probabilities allows each individual to have a unique capture probability on every capture occasion. Symbolically, the set of capture probabilities is (p_{ij}), where i ranges from 1 to N individuals (not all of which will be caught during a study) and j ranges from 1 to t occasions. This model has far more capture parameters (the p_{ij}) than there will be data, although it may be the only truly realistic model. To derive simple models (that is, models having few enough parameters that they can be estimated), we must make simplifying assumptions about the capture probabilities. The models presented below are described in terms of their assumptions about capture probabilities.

It is important to understand the interpretation of the capture probabilities as we use them here. Conceptually, on each trapping occasion of the study, every individual has an unknown probability of capture, symbolized as p_{ij}. For the estimation of population size, it does not matter in which trap an animal is caught, because these capture probabilities do not apply to given traps. Capture probability means the probability of an animal's being caught in any trap. The capture process in capture-recapture and removal sampling basically catches individual animals (not groups of animals in the sense of taking a handful of balls from an urn) on a series of separate occasions; the capture probabilities of our models reflect this aspect of the process.

Model M_0, Constant Capture Probabilities

If every animal has the same capture probability p on every capture occasion, we have Model M_0, the simplest model (*Otis et al. 1978:21-24*). It allows for no sources of variation in its capture probabilities. Model M_0 is valuable primarily as a necessary starting point for testing assumptions about capture probabilities. We also use it to introduce many ideas about analysis, such as confidence intervals, that will apply to all estimators. In terms of a restriction on the most general possible model, Model M_0 is equivalent to the assumption that $p_{ij} \equiv p$ for every animal at risk of capture ($i = 1, \ldots, N$) on every trapping occasion ($j = 1, \ldots, t$).

Simulated data to illustrate Model M_0 were generated with 6 capture occasions, a true population size of 50, and a capture probability of $p = 0.3$. The complete X matrix for the first simulation run of this capture model is presented in Table 3.2. We deliberately ordered the rows in the table as if the data were

TABLE 3.2. The complete X matrix from the first of 10 simulations of Model M_0 with $N = 50$, $p = 0.3$, and $t = 6$ occasions. All capture-recapture summary statistics can be computed from this data representation by various counting methods. For example, the sum (count) of the 1's in column 3 is $n_3 = 15$, the number of captures on day 3. The number of new animals captured on day 3 ($u_3 = 8$) is found by counting the number of 1's in column 3 for which no previous captures are recorded on days 1 or 2. Capture frequencies also can be obtained by counting; for example, 17 rows have only a single 1 in them, thus $f_1 = 17$. Three rows (animals) have four 1's each, thus $f_4 = 3$.

	Occasion							Occasion					
Animal	1	2	3	4	5	6	Animal	1	2	3	4	5	6
1	1	1	1	1	0	0	24	0	0	1	0	1	0
2	1	0	0	0	0	0	25	0	0	1	0	0	0
3	1	0	1	0	0	1	26	0	0	1	0	0	1
4	1	0	0	0	0	1	27	0	0	1	0	0	0
5	1	0	0	0	0	0	28	0	0	1	1	0	0
6	1	1	0	0	0	0	29	0	0	1	0	1	0
7	1	1	0	0	0	0	30	0	0	1	0	0	1
8	1	0	1	0	1	1	31	0	0	1	0	0	1
9	1	0	0	0	1	0	32	0	0	0	1	0	0
10	1	1	1	0	0	0	33	0	0	0	1	0	0
11	1	0	0	0	0	0	34	0	0	0	1	0	0
12	1	0	0	0	0	0	35	0	0	0	1	0	1
13	1	0	0	1	0	0	36	0	0	0	1	0	0
14	1	0	0	1	1	0	37	0	0	0	1	0	1
15	1	0	1	0	0	0	38	0	0	0	1	1	0
16	1	0	1	0	0	0	39	0	0	0	1	1	1
17	0	1	0	0	0	1	40	0	0	0	1	0	0
18	0	1	0	0	0	1	41	0	0	0	0	1	0
19	0	1	0	0	1	0	42	0	0	0	0	1	0
20	0	1	0	0	0	0	43	0	0	0	0	1	1
21	0	1	1	1	0	1	44	0	0	0	0	1	1
22	0	1	0	0	1	1	45	0	0	0	0	0	1
23	0	1	0	0	1	0	46	0	0	0	0	0	1
							47	0	0	0	0	0	1

from a real study. Specifically, the first 16 rows show the capture histories of the 16 ($= n_1$) animals first captured on day 1, and the next 7 rows show the 7 ($= u_2$) new (previously unmarked) animals caught on day 2. This pattern continues for subsequent days. All basic summary statistics can be determined from this type of table by counting in various ways.

Several basic summary statistics from Model M_0 are illustrated in Fig. 3.1. Program CAPTURE always produces this very condensed summary and prints it with the various test results discussed in the section on testing assumptions. From Fig. 3.1 we see that the numbers caught on each occasion (n_1 to n_6) are 16, 11, 15, 14, 14, and 18. Under Model M_0 we expect relatively little variation and no trends in the numbers of captures from day to day; these results illustrate the expected constancy of data under Model M_0.

For estimation of population size N under Model M_0, the entire X matrix of capture-recapture data can be reduced to two summary statistics: M_{t+1} = the total number of different individuals captured during the entire study (the number of rows in the X matrix), and n. = the total number of all captures (the sum of all the 1's in the X matrix). For the example simulation data in Table 3.2 and Fig. 3.1, these summary statistics are $M_7 = 47$ and n. = 88.

There is no simple (closed-form) formula for the maximum likelihood (ML) estimator of N under Model M_0 when there are more than $t = 2$ capture occasions. There is a simple estimator for two capture occasions; however, in this situation it is better to use an alternative estimator valid under the more general assumptions of Model M_t, discussed in the next section.

The exact ML estimate of N computed by program CAPTURE using numerical methods *(Otis et al. 1978:105)* is shown in Fig. 3.2. From Fig. 3.2, we have $\hat{N} = 55$, with an estimated standard error of 4.157. Initially, the approximate 95% confidence interval is computed as $55 \pm 1.96(4.157)$, then the upper limit is rounded up to the nearest integer and the lower limit is rounded down to the nearest integer. All confidence intervals that we recommend or that program CAPTURE computes are computed in this way. In this example, the lower limit of 46 is below the number of individuals actually seen ($M_7 = 47$). This occurrence is not uncommon; but it requires that the lower limit be moved up to 47 in this case (or

```
OCCASION             J=      1    2    3    4    5    6
ANIMALS CAUGHT    N(J)=     16   11   15   14   14   18
TOTAL CAUGHT      M(J)=      0   16   23   31   40   44   47
NEWLY CAUGHT      U(J)=     16    7    8    9    4    3
FREQUENCIES       F(J)=     17   22    5    3    0    0
```

Fig. 3.1. A print-out of basic summary statistics for the first of 10 simulations of Model M_0 with $N = 50$, $p = 0.3$, and $t = 6$ occasions. (Table 3.2 shows the complete X matrix.) The statistics include n_j = the number of animals caught on occasion j and M_j = the total number of marked animals in the population just before the j^{th} capture occasion. Also shown are $M_7 = 47$, which is the total number of different individuals seen in this "study," u_j = the number of unmarked animals caught on occasion j, and finally, the capture frequencies f_j = the number of animals caught exactly j times during the study (for $j = 1, ..., t$, because f_0 is not known). Note the lack of any large variations in the numbers caught on each day (n_j) and the fairly consistent decrease in the numbers of unmarked animals caught on each occasion. These are characteristics of the data expected from Model M_0.

```
NUMBER OF TRAPPI'3 OCCASIONS WAS              6
NUMBER OF ANIMALS CAPTURED, M(T+1), WAS      47
TOTAL NUMBER OF CAPTURES, N., WAS            88

ESTIMATED PROBABILITY OF CAPTURE, P-HAT =  0.2654


POPULATION ESTIMATE IS          55 WITH STANDARD ERROR      4.1570

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL         46 TO          64
```

Fig. 3.2. Print-out of the results of estimating N and p from the simulation data in Table 3.2 and Fig. 3.1. The ML estimate of N is 55, with an estimated standard error of 4.2. The approximate 95% confidence interval computed for N from $\hat{N}$ and $\hat{se}(\hat{N})$ is 46 to 64. However, because $M_7 = 47$ (47 individuals were caught), we must replace the lower limit of 46 by 47 and report this interval as 47 to 64. The ML estimate of capture probability is $\hat{p} = 0.2654$. In the true, underlying Model M_0, $p = 0.3$ and $N = 50$.

up to M_{t+1}, in general). In *Otis et al. (1978:133-135)* we investigated alternative methods of constructing confidence intervals, partly to avoid this problem of inadmissible lower limits. After considering a variety of approaches, we concluded that the most practical approach is to use $\hat{N} \pm 1.96 \times \hat{se}(\hat{N})$ for approximate 95% confidence intervals and then, if the lower limit is below M_{t+1}, to replace that lower limit by M_{t+1}. We deliberately did not implement this replacement feature in CAPTURE so that the user could see whether or how much the computed lower limit falls below M_{t+1}. A lower limit far below M_{t+1} indicates very poor experimental results. In this example, the actual, unrounded lower limit is 46.85 = 55 − 1.96(4.157).

The results from this one simulation study are acceptable. If it were a real study, we would have a very good estimate of N. The true value of p is 0.3; its estimate $\hat{p}$ is 0.27. Also, the true value of N is 50; its estimate $\hat{N}$ is 55. In real studies we do not know N or p; therefore, the judgment of whether the results are reliable rests on clues from the data and their analyses, and on the consistency of external information about the population with the results of the analyses.

The estimated (average) capture probability $\hat{\bar{p}}$ gives some insight into the reliability of results as do, the standard error of the estimate of N, and the resultant confidence interval. With typical studies of populations in the size range from 50 to several hundred, results are not reliable unless the true average capture probability is at least 0.1 and preferably at least 0.2. In Model M_o, the estimate of p is the same as the estimated average capture probability. A $\hat{p}$ of 0.27 and the small standard error on $\hat{N}$ of 4.2 indicate the results are trustworthy.

There is no fail-safe way to be sure that the true average capture probability $\bar{p}$ for a study exceeds 0.10, because we will have only an estimate of this value. We have not presented confidence intervals about p itself in Model M_o (or about $\bar{p}$ in other models). Such intervals would be reliable only if the model were true. Our reason for trying to judge whether $\bar{p}$ is at least 0.10 is to provide a basis for deciding whether we can reasonably trust the model selection procedure and the resultant estimate of N. If the data are poor, the estimate $\hat{\bar{p}}$ can be quite a bit greater than 0.10, yet the true $\bar{p}$ can be less than 0.10. Thus, the problem of judging the data's reliability from the data themselves is one of circularity. If the study is reliable (that is, if the capture probabilities are high enough and the true population size is sufficiently large), one probably will conclude this from the data. But if the data are poor, especially because of very low capture probabilities, this fact cannot always be determined from analysis of the data.

The standard error of $\hat{N}$ depends upon the value of N. Therefore, to judge the relative precision of results one can look at the coefficient of variation of $\hat{N}$, $cv(\hat{N})$.

$$cv(\hat{N}) = \frac{\hat{se}(\hat{N})}{\hat{N}} .$$

In the example of Fig. 3.2 we have

$$cv(\hat{N}) = \frac{4.157}{55}$$

$$= 0.0756 \text{ or } 7.56\% .$$

This value reflects good precision for the estimated population size. In our opinion, reliable scientific studies require a coefficient of variation of $\hat{N}$ of no more than 20% and investigators should try for $cv(\hat{N}) \leq 0.1$ (10%). Less exacting management studies, including long-term monitoring studies, may be acceptable with a coefficient of variation of 20 to 50%. Studies producing $cv(\hat{N})$ values above 0.5 (50%) can indicate only order-of-magnitude changes in population abundance, for example, densities changing from 1 to 10 or vice versa.

We have illustrated Model M_o with the first simulation results of 10 repetitions, using N = 50, p = 0.3, and t = 6. Summary results of all 10 simulations for this "study," given in Table 3.3, illustrate natural sampling variability. For example, M_7 varies from 41 to 47, n. varies from 81 to 99, and $\hat{N}$ varies from 45

TABLE 3.3. Summary results from all 10 simulations of Model M_0 with $N = 50$, $p = 0.3$, and $t = 6$. The variations observed here in quantities like M_7, n., and $\hat{N}$ are entirely the result of the stochastic nature of the "capture" process. In particular, $\hat{N}$ ranges from 45 to 55 and only one value is exactly 50.

Repetition	M_7	n.	$\hat{N}$	$\hat{se}(\hat{N})$	$\hat{p}$
1	47	88	55	4.2	0.27
2	47	89	55	4.0	0.27
3	43	81	50	3.9	0.27
4	44	91	49	2.9	0.31
5	45	99	49	2.5	0.34
6	41	87	45	2.6	0.32
7	41	83	46	3.0	0.30
8	42	82	48	3.4	0.28
9	43	89	48	2.9	0.31
10	43	95	46	2.4	0.34
Averages	43.6	89.0	49.1	3.2	0.30
Sampling standard deviation	2.2	5.7	3.5		

to 55. These levels of variation are very small compared to those often seen with other models or with smaller $\bar{p}$ values. If the data from Model M_0 were real, with $p = 0.3$, $t = 6$, and $N = 50$, the estimator of N nearly always would be very good.

Model M_t, Variation by Time

The most common model for capture-recapture studies allows capture probabilities to vary only by time. Thus for t capture occasions, there are t possible capture probability parameters, $p_1, \ldots, p_t$ where p_j = the probability that any individual animal will be captured on occasion j. This same capture probability is assumed to apply to all N animals in the population on the j^{th} capture occasion. Hence, the past capture history of an animal is not allowed to influence its current capture probability. In particular, unmarked (not previously caught) animals are assumed to have the same probability of capture as marked (previously caught) animals. Either behavioral response or innately varying capture probabilities will invalidate this model.

The basic summary statistics from the first of 10 simulations of Model M_t, with $N = 150$, $t = 5$, and daily capture probabilities of $p_1 = 0.20$, $p_2 = 0.40$, $p_3 = 0.30$, $p_4 = 0.35$, and $p_5 = 0.25$, are presented in Fig. 3.3. On the average for such a study we would expect to catch $Np_j = E(n_j)$ animals on the j^{th} day, and we can readily compute the following for this example.

$E(n_1) = 150.0 \times 0.20 = 30$
$E(n_2) = 150.0 \times 0.40 = 60$
$E(n_3) = 150.0 \times 0.30 = 45$
$E(n_4) = 150.0 \times 0.35 = 52.5$
$E(n_5) = 150.0 \times 0.25 = 37.5$

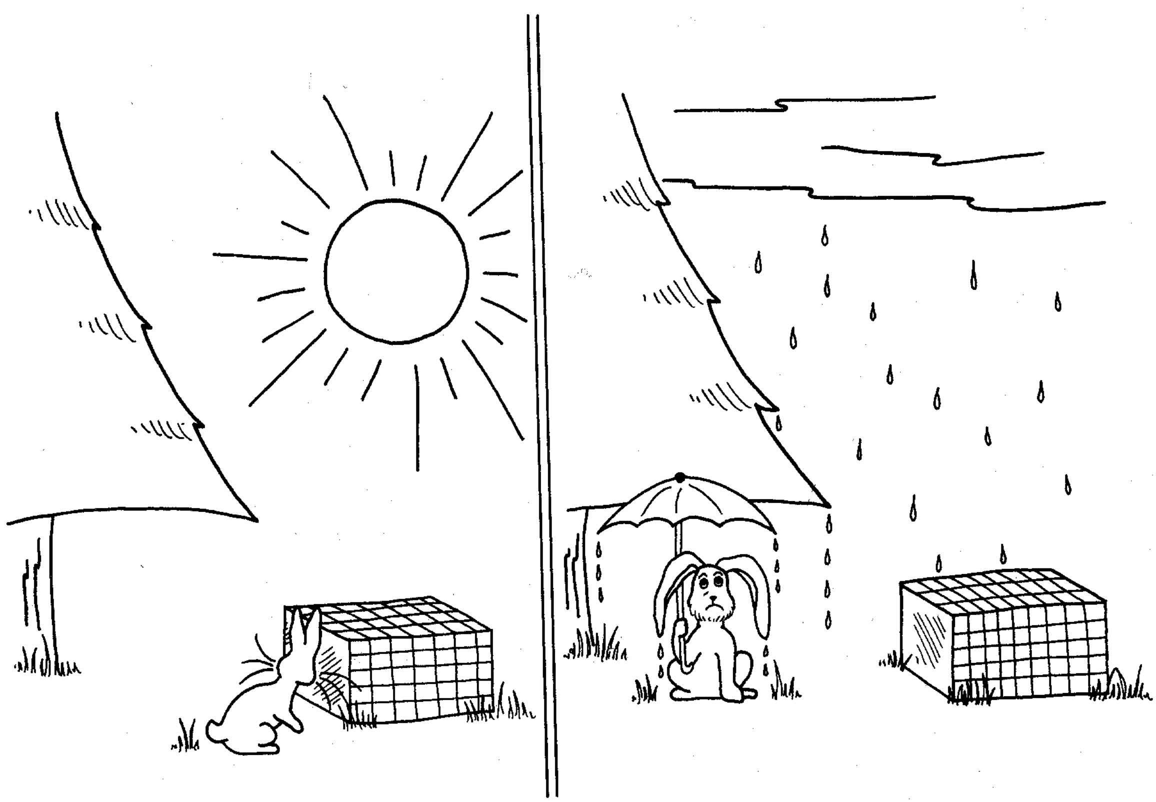
The observed daily captures n_j shown in Fig. 3.3 reflect these expected numbers.

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	28	53	50	60	37	
TOTAL CAUGHT	M(J)=	0	28	70	101	121	127
NEWLY CAUGHT	U(J)=	28	42	31	20	6	
FREQUENCIES	F(J)=	52	52	20	3	0	

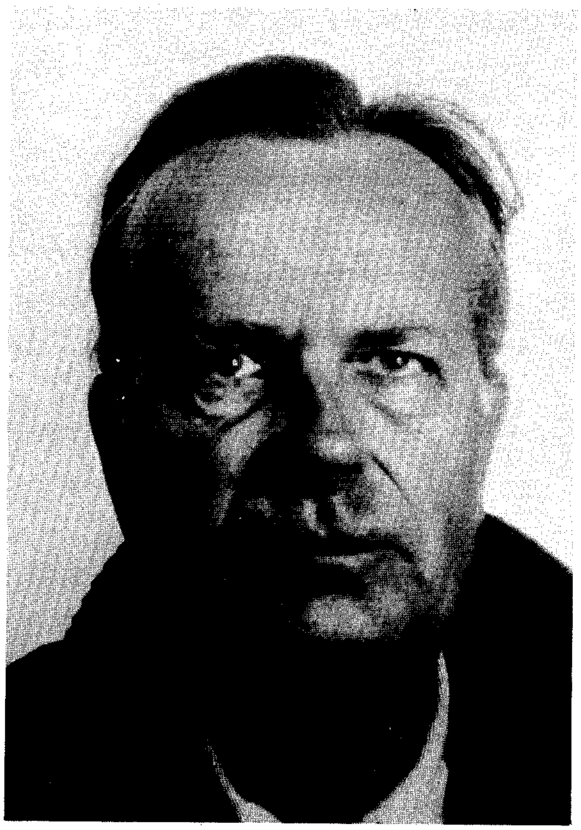
Fig. 3.3. Basic summary statistics for the first of 10 simulations of Model M_t, with $N = 150$, $t = 5$ occasions, and daily capture probabilities $p_1 = 0.20$, $p_2 = 0.40$, $p_3 = 0.30$, $p_4 = 0.35$, and $p_5 = 0.25$. The average capture probability is $\bar{p} = 0.30$. Observe that the variation in the daily numbers caught is substantial, and the $u_2 = 42$ even though $u_1 = 28$; that is, substantially more animals were caught for the first time on the day 2 than on day 1. This data pattern suggests time effects in the capture probabilities.

Under Model M_t, we expect the significant variation in the numbers of captures on each occasion shown in Fig. 3.3. Under Model M_0, we would expect a consistent, somewhat smooth decrease in the numbers of captures of previously uncaught animals (the u_j), but we can expect no such decrease under Model M_t. The first-capture data $u_1, \ldots, u_t$ (newly caught) are likely to be erratic. In Fig. 3.3, $u_1 = 28$, $u_2 = 42$, $u_3 = 31$, $u_4 = 20$, and $u_5 = 6$. The large increase of new animals on day 2 strongly suggests that Model M_0 would not be the correct model. Rather, there is some difference between the population capture probabilities on capture occasions 1 and 2. If we can assume closure to be true, this data pattern suggests an increased capture probability on day 2 compared with day 1.

There is ample evidence in the literature that varying environmental conditions affect capture probabilities: *Paloheimo (1963)* found that water temperature affects the catchability of lobsters; *Gentry et al. (1966)* and *Getz (1961)* found that weather affects the catchability of small mammals when live traps are used; and *Bailey (1969)* found that weather also affects the capture probabilities of rabbits. Varying effort over time also causes time variation in capture probabilities, as when the number of operational traps or the number of times traps are checked each day varies during the study.



Capture probabilities may vary over time because of varying weather conditions.



Norman Bailey's research on open- and closed-population models arose 30 years ago in response to problems discussed with Sir Ronald Fisher. While working at Cambridge University Medical School, Bailey developed the so-called "triple catch" method, which was used widely for many years. Bailey's primary interest has been medical-statistical problems, and his work on capture-recapture was an aside, although the contribution to biologists has been quite substantial.

Bailey took B.A. and M.A. degrees from Cambridge and a Doctor of Science from Oxford in 1959. He is the author of 7 books and more than 90 research papers in medical statistics and biomathematics. (Recent photograph.)

Norman T. J. Bailey

Model M_t is very much a ball and urn model, assuming as it does that marked and unmarked animals have the same capture probability on any given capture occasion. This concept was clearly the basis for one of the earliest papers dealing with this model (*Schnabel 1938*); in fact, in that paper Schnabel tested her estimator with a physical simulation using beans in an "urn."

The literature on Model M_t since Schnabel's pioneering work is extensive. The first exact treatment was given by *Darroch (1958)*, although he presented only a close approximation to the ML estimator. We recommend use of the ML estimator and believe that the only value in all the approximate and *ad hoc* estimators presented for this model is their ease of computation and their simplicity for teaching purposes. An introduction to the extensive literature on Model M_t is presented in *Seber (1973:130-164)*.

All relevant information for estimation of population size under Model M_t is contained in the summary statistics $n_1, \ldots, n_t$ (the number of animals captured on each day) and M_{t+1}(the total number of different individuals captured). The exact ML estimator of N (*Otis et al. 1978:106-107*) does not exist in closed form; however, for $t = 2$ the usual Petersen-Lincoln estimate closely approximates the ML estimator. We advise against the use of only two capture occasions because assumptions cannot be tested. However, this approach may be reasonable at times, especially when different capture methods are used on the two occasions. For $t = 2$ we recommend *Chapman's (1951)* modification of the Petersen-Lincoln estimator.

$$\hat{N} = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1 ,$$

where m_2 = the number of marked animals recaptured. We will not deal further with the case of $t = 2$; it is well covered in the literature, for example, in *Seber (1973:59-70)*.

Douglas G. Chapman

Douglas Chapman has contributed to the theory and application of capture-recapture and related sampling problems over the past 30 years. He took an undergraduate degree in mathematics and economics at the University of Saskatchewan in Canada. He entered the University of California at Berkeley to do graduate work in mathematics, but came under the influence of Professor Jerzey Neyman and switched to statistics.

Chapman's career has been a rich mix of statistical theory and application, at first closely associated with fishery problems and later with marine mammal problems. Consultation concerning fishery problems began in 1946 with the International Pacific Salmon Commission. After moving to the University of Washington, he became involved with the fur seal research group of the (then) Bureau of Commercial Fisheries. Chapman has long been active in the Center for Quantitative Science at the University of Washington and was Dean of the College of Fisheries there until recently.

He has long been interested in the general problems of population dynamics rather than in the narrower issues of parameter estimation. This leaning stems from his work on whales through the International Whaling Commission. (Photograph taken in the mid-1960s.)

Many closed-form estimators in the literature are based on Model M_t. The best known is the Schnabel estimator (*Schnabel 1938; Seber 1973:139*). The Schnabel estimator is easy to compute and often is a good approximation to the ML estimator. However, the comprehensive analysis of any multiple recapture data requires complex testing of assumptions and subsequent selection of a model. For these analyses a computer routine is essential. Given such a routine it is better to compute the exact ML estimator of N under Model M_t than to bother with approximations, such as *Darroch (1958)* and *Schnabel (1938)*.

The ML estimate of N under Model M_t applied to the example data in Fig. 3.3 is given in Fig. 3.4. The value of $\hat{N} = 151$ is closer to the true N of 150 than we have a right to expect. The standard error of this estimate is 6.975. The approximate 95% confidence interval on N is 137 to 165 [computed as $151 \pm (1.96 \times 6.975)$]. The lower limit of this confidence interval is 137, whereas 127 animals were caught. Thus, in this example the lower limit does not fall below M_{t+1}.

The estimated daily capture probabilities $\hat{p}_1 = 0.19$, $\hat{p}_2 = 0.35$, $\hat{p}_3 = 0.33$, $\hat{p}_4 = 0.40$, and $\hat{p}_5 = 0.25$ also are given in Fig. 3.4. These values should be compared with the true daily capture probabilities. The rough, but useful, histogram (frequency plot) of the daily numbers of captures, also shown in Fig. 3.4, provides a visual display of the n_j.

Table 3.4 presents the results of 10 repetitions of simulating Model M_t with the parameters used in Figs. 3.3 and 3.4. The variation observed among these 10 repetitions is entirely the result of the stochastic nature of the data (catching or not catching animals on each occasion). We see that $\hat{N}$ varies from 138 to 153. Also compare the average of the daily captures, n_j, over these 10 repetitions with the expected values of n_j given above. For example, $E(n_1) = 30$, and $\bar{n}_1 = 29.3$.

```
OCCASION              J=    1    2    3    4    5
ANIMALS CAUGHT   N(J)=     28   53   50   60   37

TOTAL ANIMALS CAPTURED     127

           P-HAT(J)=  0.19 0.35 0.33 0.40 0.25


POPULATION ESTIMATE IS           151 WITH STANDARD ERROR        6.9754

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL          137 TO           165


HISTOGRAM OF N(J)


FREQUENCY     28   53   50   60   37
----------------------------------------
EACH * EQUALS     6 POINTS

     60                        *
     54             *          *
     48             *    *     *
     42             *    *     *
     36             *    *     *    *
     30        *    *    *     *    *
     24        *    *    *     *    *
     18        *    *    *     *    *
     12        *    *    *     *    *
      6        *    *    *     *    *
----------------------------------------
```

Fig. 3.4. Print-out of Model M_t results of estimating N and capture probabilities from the simulation data in Fig. 3.3. These data were generated under Model M_t, with $N = 150$, $p_1 = 0.20$, $p_2 = 0.40$, $p_3 = 0.30$, $p_4 = 0.35$, and $p_5 = 0.25$. Compare the estimated daily capture probabilities $\hat{p}_j$ with the true ones. The lower limit of the 95% confidence interval does not fall below $M_6 = 127$ in this example. The coefficient of variation of $\hat{N}$ is $cv(\hat{N}) = 6.98/151 = 0.046$, or 4.6%, which is quite good.

TABLE 3.4. Summary results from all 10 simulation repetitions of Model M_t with $N = 150$, $t = 5$, and daily capture probabilities, $p_1 = 0.20$, $p_2 = 0.40$, $p_3 = 0.30$, $p_4 = 0.35$, and $p_5 = 0.25$. The average capture probability is $\bar{p} = 0.30$. The daily numbers of captures and the total number of different individuals captured (M_6) illustrate the variability in the capture data as well as the variability in the estimates of N. Notice that $\hat{N}$ ranges from 138 to 153.

	Daily Captures							
Repetition	n_1	n_2	n_3	n_4	n_5	M_6	$\hat{N}$	$\hat{se}(\hat{N})$
1	28	53	50	60	37	127	151	7.0
2	26	63	50	49	42	125	146	6.3
3	33	51	44	55	27	122	150	8.0
4	31	54	41	48	41	120	143	7.0
5	33	62	42	52	43	129	153	7.1
6	34	53	42	56	38	119	138	5.9
7	27	58	37	62	37	123	145	6.7
8	19	65	43	51	36	118	138	6.2
9	33	61	40	53	31	125	152	7.6
10	29	60	46	53	42	124	144	6.2
Averages	29.3	58.0	43.5	53.7	37.4	123.2	146.0	6.8

Model M_b, Behavioral Response

Animals frequently exhibit a behavioral response to capture, especially to first capture. This means that, after first capture, their capture probability on subsequent capture occasions changes, often greatly. The biological literature on this phenomenon is extensive; for example, see *Tanaka (1956), Crowcroft* and *Jeffers (1961), Getz (1961), Hunter* and *Wisby (1964), Bailey (1969),* and *Beukema* and *de Vos (1974).* To deal with this situation *Otis et al. (1978:28-32)* used Model M_b, wherein the probability of recapture, c, is allowed to be arbitrarily different from the probability of first capture, p. If the recapture probability is lower than the first-capture probability ($c < p$), the animals are exhibiting trap avoidance (they have become trap shy), whereas if the capture probability increases ($c > p$) after first capture, animals are showing trap fascination (they have become trap happy). We give simulation examples of both situations.

We again emphasize the interpretation of these capture probabilities: they apply to individual animals for each separate trapping occasion. For example, let the probability of first capture be $p = 0.5$, and consider a single animal. On the first trapping occasion, that animal has a 50% chance of being caught. If the animal is not caught, then $p = 0.5$ on the second occasion, and it again has a 50% chance of being caught. If it is not caught on occasions 1 and 2, it continues to have a 50% chance of being caught on the third occasion. Once the animal is caught, however, its behavior changes with respect to the traps: it tends to either avoid traps or return to them. Assume that an animal is caught on occasion 3, likes the bait, and becomes trap happy, with recapture probability $c = 0.80$. On capture occasion 4, this animal has an 80% chance of being caught. Whether or not it is caught on occasion 4, it continues to have an 80% chance of being on caught each subsequent, separate capture occasion.

Model M_b does not incorporate any relation between the probability of first capture and the probability of recapture. The recapture data therefore contain no information about the unknown population size N. [This is a key point; because proving it would require presenting the full-blown mathematics of Model M_b, we refer the reader to *Otis et al. (1978:107-108).*] As a consequence, the estimate of N for Model M_b is based entirely on the first-capture information. Because recaptures are not used in the estimation of population size for Model M_b, the data analysis methods are the same as for removal data, as detailed in Chapter 4.

We let u_1 = the number of animals captured on day 1; these are, of course, all first captures. On day 2, $N - u_1$ animals remain uncaught. Let u_2 be the number of animals caught for the first time on day 2. These u_2 animals come entirely from the $N - u_1$ animals not caught on day 1. In general, we let u_j represent the number of animals captured for the first time on day j. Estimation of N and of p, the first-capture probability, is based entirely on the first-capture data, $u_1, u_2, \ldots, u_t$.

The X matrix from a simulation example of Model M_b using $N = 100$, $p = 0.25$, $c = 0.55$, and $t = 7$ is presented in Table 3.5. This is a trap-happy example because after first capture, individual capture probabilities increase to 0.55. Summary data computed from this X matrix (Fig. 3.5) are $u_1 = 19$, $u_2 = 17$, $u_3 = 24$, $u_4 = 11$, $u_5 = 8$, $u_6 = 4$, and $u_7 = 8$.

The statistics for n_j and u_j in Fig. 3.5 generally reflect the pattern expected under trap happiness. There is a general decrease in the numbers of first captures. Conversely, there is a general increase in the numbers of daily captures n_j. We expect, on the average, a decrease in first captures (the u_j), because the number of uncaught animals continues to decline. In essence, we are removing animals from the (uncaught) population by marking them. In this sense, the numbers of first captures constitute data from a removal study, and their analysis exactly follows methods for the analysis of removal data, to be studied in Chapter 4.

The increase in the n_j, the numbers of daily captures, in Fig. 3.5 is due to the increasing number of animals in the population with the higher capture probability of 0.55. We see $n_1 = 19$, $n_2 = 28$, $n_3 = 44$, $n_4 = 37$, $n_5 = 44$, $n_6 = 49$, and $n_7 = 53$. By day 3, $u_1 + u_2 = 36$ of the 100 animals in the population have a (re)capture probability of 0.55,

TABLE 3.5. The complete X matrix from the first of 10 simulations of Model M_b with $N = 100$, $p = 0.25$, $c = 0.55$, and $t = 7$ occasions. All capture-recapture summary statistics in Fig. 3.5 can be computed from this data representation by various counting methods.

	Occasion								Occasion						
Animal	1	2	3	4	5	6	7	Animal	1	2	3	4	5	6	7
1	1	1	1	1	0	0	1	47	0	0	1	1	1	0	0
2	1	0	1	1	1	1	1	48	0	0	1	0	1	1	0
3	1	1	1	0	0	1	1	49	0	0	1	1	1	0	0
4	1	1	0	0	0	1	1	50	0	0	1	1	0	1	1
5	1	0	1	1	1	1	0	51	0	0	1	1	0	1	0
6	1	1	0	0	0	1	0	52	0	0	1	0	0	1	1
7	1	0	0	0	1	0	0	53	0	0	1	0	0	1	0
8	1	0	1	0	0	1	0	54	0	0	1	1	1	0	1
9	1	1	1	1	0	1	1	55	0	0	1	0	0	1	0
10	1	1	0	0	0	1	1	56	0	0	1	1	1	1	0
11	1	1	0	1	1	1	0	57	0	0	1	1	1	1	1
12	1	0	1	0	0	0	0	58	0	0	1	0	0	0	0
13	1	1	1	0	1	0	0	59	0	0	1	0	0	1	1
14	1	0	0	1	0	0	1	60	0	0	1	0	1	1	0
15	1	1	0	1	1	1	0	61	0	0	0	1	0	0	1
16	1	1	1	1	0	0	1	62	0	0	0	1	1	1	0
17	1	0	0	0	0	0	0	63	0	0	0	1	1	1	0
18	1	1	0	0	0	1	1	64	0	0	0	1	0	1	0
19	1	0	1	0	1	1	0	65	0	0	0	1	1	0	0
20	0	1	1	1	0	0	1	66	0	0	0	1	1	1	1
21	0	1	0	0	1	0	1	67	0	0	0	1	0	0	0
22	0	1	1	0	1	1	1	68	0	0	0	1	0	1	1
23	0	1	0	1	1	0	0	69	0	0	0	1	0	0	1
24	0	1	1	0	1	0	0	70	0	0	0	1	0	1	1
25	0	1	0	0	0	1	1	71	0	0	0	1	0	0	0
26	0	1	0	0	1	0	1	72	0	0	0	0	1	1	0
27	0	1	1	1	1	1	0	73	0	0	0	0	1	0	0
28	0	1	0	0	1	0	1	74	0	0	0	0	1	1	1
29	0	1	0	0	0	1	1	75	0	0	0	0	1	0	1
30	0	1	1	0	0	1	1	76	0	0	0	0	1	1	1
31	0	1	1	1	1	0	1	77	0	0	0	0	1	0	0
32	0	1	1	0	0	0	1	78	0	0	0	0	1	0	1
33	0	1	1	1	1	1	1	79	0	0	0	0	1	1	1
34	0	1	1	0	1	1	1	80	0	0	0	0	0	1	1
35	0	1	0	1	1	1	1	81	0	0	0	0	0	1	0
36	0	1	1	1	1	1	1	82	0	0	0	0	0	1	1
37	0	0	1	1	1	1	0	83	0	0	0	0	0	1	1
38	0	0	1	0	1	1	0	84	0	0	0	0	0	0	1
39	0	0	1	0	0	0	0	85	0	0	0	0	0	0	1
40	0	0	1	0	1	1	0	86	0	0	0	0	0	0	1
41	0	0	1	0	1	1	0	87	0	0	0	0	0	0	1
42	0	0	1	1	1	0	1	88	0	0	0	0	0	0	1
43	0	0	1	0	0	0	0	89	0	0	0	0	0	0	1
44	0	0	1	1	0	0	1	90	0	0	0	0	0	0	1
45	0	0	1	0	1	0	1	91	0	0	0	0	0	0	1
46	0	0	1	1	0	0	1								

OCCASION	J=	1	2	3	4	5	6	7	
ANIMALS CAUGHT	N(J)=	19	28	44	37	44	49	53	
TOTAL CAUGHT	M(J)=	0	19	36	60	71	79	83	91
NEWLY CAUGHT	U(J)=	19	17	24	11	8	4	8	
FREQUENCIES	F(J)=	17	14	31	13	12	4	0	

Fig. 3.5. Print-out of basic summary statistics for the first of 10 simulations of Model M_b, with $N = 100$, $p = 0.25$, $c = 0.55$, and $t = 7$; Table 3.5 shows the complete X matrix. The numbers of captures each day, n_j, increase from 19 on day 1 to 53 on day 7, because animals are being caught, and hence their capture probability is changing from 0.25 to 0.55. Notice that even at a first-capture probability of 0.25, 90% of the population has been caught at least once after 7 days. The numbers of new captures u_j generally decline.

while the remaining 64 animals, not yet caught at the start of day 3, still have a capture probability of 0.25. Thus, at the start of day 3 in this example the average capture probability in the population of 100 is

$$\bar{p} = \frac{64 \times 0.25 + 36 \times 0.55}{100} = 0.358 .$$

This probability is increased quite a bit from the 0.25 first-capture probability.

For any Model M_b, the expected average daily capture probabilities are given by the formula

$$E(\bar{p}_j) = [1 - (1 - p)^{j-1}](c - p) + p, \qquad j = 1,2, \ldots, t,$$

and the expected number caught on day j is $E(n_j) = NE(\bar{p}_j)$. For the model underlying Fig. 3.5 ($N = 100$, $p = 0.25$, and $c = 0.55$) we have

j	$E(\bar{p}_j)$	$E(n_j)$
1	0.250	25.0
2	0.325	32.5
3	0.381	38.1
4	0.423	42.3
5	0.455	45.5
6	0.479	47.9
7	0.497	49.7

Note that the formula on p. 57 of *Otis et al. (1978)* is wrong; the formula given above for $E(\bar{p}_j)$ is correct.

Recall that the expected values of n_j and $\bar{p}_j$ increase or decrease over time in Model M_b. This change in $\bar{p}_j$ has no relation to the change in Model M_t, where capture probabilities vary over time for external reasons, not because of behavioral response to capture. Yet when we look at capture-recapture data, especially the n_j, we find it difficult to distinguish between the two causes of variation. Choosing between the models requires tests of assumptions, discussed later in this chapter. Behavioral response can "look" like time variation, and this similarity causes difficulties in the proper analysis of capture-recapture data.

We recommend ML estimates of N, p, and c. The estimator of recapture probability is simple, but it is not of primary interest. The ML estimates of N and p do not exist in closed form. Program CAPTURE can compute $\hat{N}$ and $\hat{p}$ from the "removal" data $u_1, \ldots, u_t$. (See Chapter 4 for discussion of estimation based on removal data.)

The estimates of parameters based on the simulated removal data in Fig. 3.5 are presented in Fig. 3.6. We see that a total of 91 (= M_8) of 100 animals were caught at least once.

The ML estimate of N is 114, with an estimated standard error of 12.9. This gives a coefficient of variation on $\hat{N}$ of 11.3%. The approximate 95% confidence interval on N is computed to be 88 to 140. However, because 91 animals were actually seen, the lower bound of 88 must be replaced by 91 when these results are reported.

The estimated first-capture probability from the data in Fig. 3.6 is $\hat{p} = 0.20$. The estimated recapture probability is 0.53. Recall that the true parameters of this simulation were N = 100, p = 0.25, and c = 0.55.

The basic results of all 10 simulations of this trap-happy capture-recapture model are given in Table 3.6. The results are in close agreement with the known parameters.

A simulation example of trap-avoidance response to first capture will further illustrate the behavioral response model (M_b). The basic summary statistics from the first of 10 simulations of Model M_b, with N = 100, p = 0.40, c = 0.20, and t = 7 are shown in Fig. 3.7. In trap-shy behavioral response, the daily capture probabilities decrease over time. This decrease, however, is not due to time variation in capture probabilities in the sense of Model M_t, but rather to the animals' becoming less catchable (trap shy) after first capture. Using the formula $E(\bar{p}_j) = [1 - (1 - p)^{j-1}](c - p) + p$, we compute the expected daily capture probabilities and daily captures as follows.

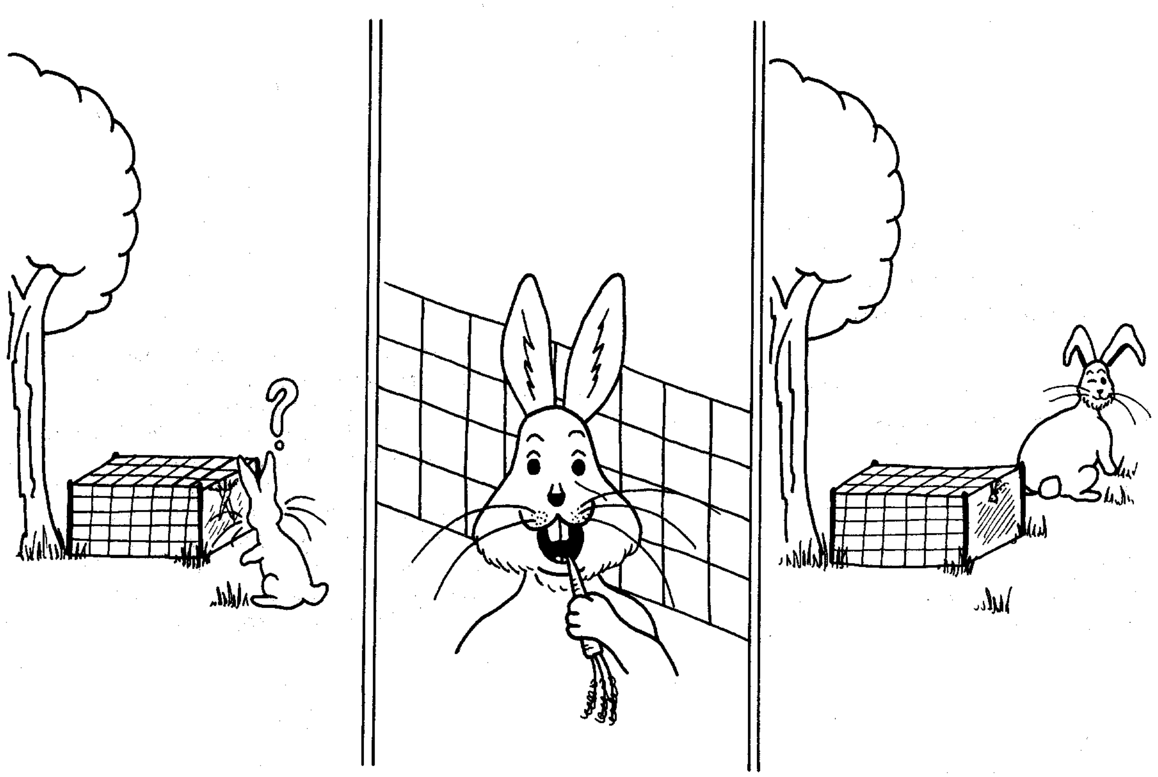
j	$E(\bar{p}_j)$	$E(n_j)$
1	0.400	40.0
2	0.320	32.0
3	0.272	27.2
4	0.243	24.3
5	0.226	22.6
6	0.216	21.6
7	0.209	20.9

```
OCCASION          J=     1   2   3   4   5   6   7
TOTAL CAUGHT   M(J)=     0  19  36  60  71  79  83  91
NEWLY CAUGHT   U(J)=    19  17  24  11   8   4   8

ESTIMATED PROBABILITY OF CAPTURE, P-HAT =   0.201181

ESTIMATED PROBABILITY OF RECAPTURE, C-HAT = 0.525862


POPULATION ESTIMATE IS        114 WITH STANDARD ERROR     12.9195

APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS        88 TO        140


HISTOGRAM OF U(J)


FREQUENCY     19    17    24    11     8     4     8
----------------------------------------------------
EACH * EQUALS     3 POINTS

     24                 *
     21                 *
     18     *     *     *
     15     *     *     *
     12     *     *     *     *
      9     *     *     *     *     *           *
      6     *     *     *     *     *           *
      3     *     *     *     *     *     *     *
----------------------------------------------------
```

Fig. 3.6. Results of estimating N, p, and c under Model M_b from the data in Fig. 3.5. The ML estimate of N is 114, with an estimated standard error of 12.9. The approximate 95% confidence interval is computed to be 88 to 140; however, the lower limit must be replaced by 91 when the interval is reported because 91 animals were seen. The ML estimates of p (first-capture probability) and c (recapture probability) are 0.20 and 0.53, respectively. The true underlying model is M_b with N = 100, p = 0.25, and c = 0.55.

TABLE 3.6. Summary results for all 10 simulations of Model M_b (trap-happy case) with $N = 100$, $p = 0.25$, $c = 0.55$, and $t = 7$. The variation in the numbers of first captures by occasion (u_i), total individuals captured (M_8), and estimates of parameters is entirely the result of the stochastic nature of the capture processes. Notice that $\hat{N}$ varies from 83 to 119, but there is no evidence of bias, as the average of all 10 estimates is 102.4. The average estimates of p and c are also very close to the true values of these parameters.

	Numbers First Captured on Occasion j											
Replication	u_1	u_2	u_3	u_4	u_5	u_6	u_7	M_8	$\hat{N}$	$\hat{se}(\hat{N})$	$\hat{p}$	$\hat{c}$
1	19	17	24	11	8	4	8	91	114	12.9	0.20	0.53
2	30	24	8	7	9	4	7	89	97	5.2	0.29	0.53
3	19	24	14	7	8	2	4	78	85	4.7	0.29	0.51
4	26	13	16	12	6	7	9	89	112	12.9	0.20	0.52
5	23	20	13	8	8	6	6	84	98	8.0	0.24	0.53
6	22	16	17	17	9	3	3	89	105	9.2	0.23	0.55
7	28	17	12	7	7	4	3	78	83	3.5	0.33	0.55
8	23	20	12	14	6	5	5	85	97	7.0	0.26	0.54
9	20	14	16	12	7	11	5	85	114	17.2	0.18	0.53
10	26	14	11	14	12	6	8	91	119	15.4	0.19	0.58
Averages	23.6	17.9	14.3	10.9	8.0	5.2	5.8	85.9	102.4	9.6	0.24	0.54
Standard deviations								4.8	12.5	4.8	0.05	0.02



Trap-happy behavioral changes in capture probabilities may result from a favorable first-capture experience.

OCCASION	J=	1	2	3	4	5	6	7	
ANIMALS CAUGHT	N(J)=	36	34	25	23	18	25	24	
TOTAL CAUGHT	M(J)=	0	36	62	76	87	92	93	94
NEWLY CAUGHT	U(J)=	36	26	14	11	5	1	1	
FREQUENCIES	F(J)=	29	43	18	4	0	0	0	

Fig. 3.7. Basic summary statistics for the first of 10 simulations of Model M_b with N = 100, p = 0.4, c = 0.2, and t = 7. The daily numbers caught decrease because of trap shyness in individuals after first capture. Because the probability of first capture is substantial (0.4), the daily "removals" (by marking) decrease substantially over the seven capture occasions. This decrease suggests that not many animals are left uncaught.

[These values of $E(\bar{p}_j)$ are correct for the situation considered on p. 57 of *Otis et al. (1978).*] A consistent decrease in daily captures over time may be à clue that the population is exhibiting trap-shy behavioral response.

The estimates of N, p, and c under Model M_b for the simulation data in Fig. 3.7 are shown in Fig. 3.8. As in the example for trap happiness, the estimates of N and p are based entirely on the first-capture data (u_1 to u_7 here), which may be considered as removal data. (Animals are "removed" from the unmarked population by marking them.) In this example, 94 of 100 animals were captured. Also, the numbers of first captures decrease markedly, from 36 ($= n_1$) on day 1 to 1 ($= n_7$) on day 7. From such data we can expect precise estimates of the parameters. Indeed, from Fig. 3.8 we see that $\hat{N}$ is 96, with an estimated standard error of 1.8. The computed 95% confidence interval on N is 92 to 100, but 94 animals were caught so we would report the results as $\hat{N} = 96$, with a confidence interval of 94 to 100. The estimated capture probabilities are $\hat{p} = 0.422$ and $\hat{c} = 0.204$, compared with true values of p = 0.4 and c = 0.2.

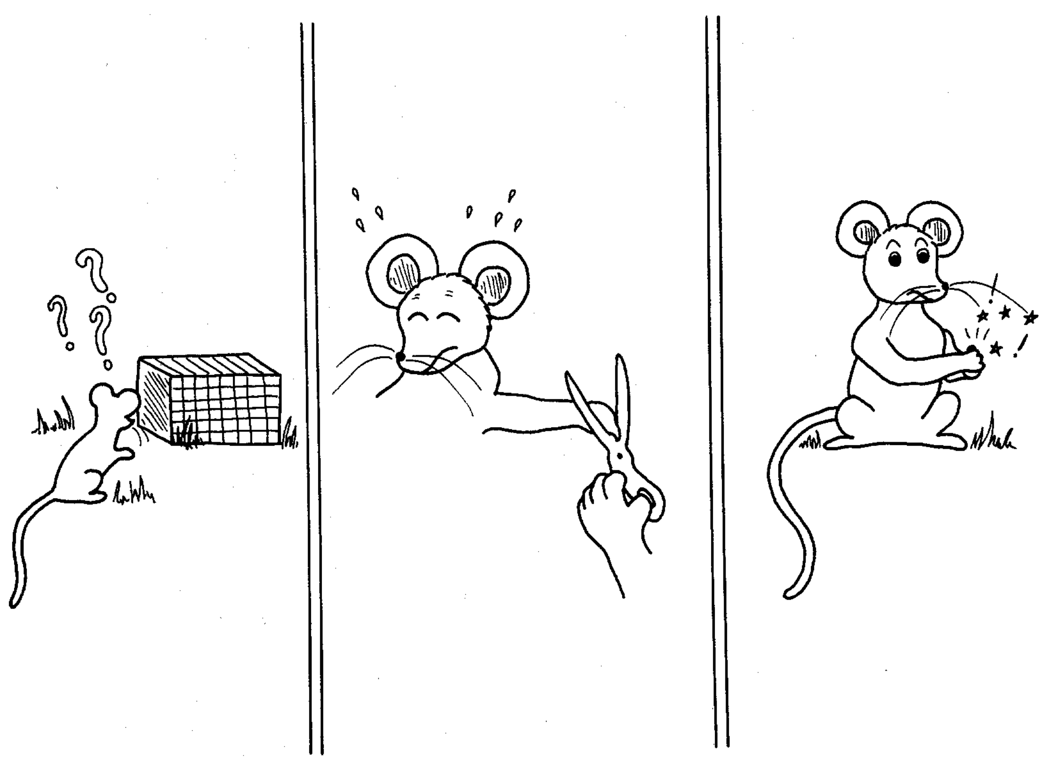
The basic results of all 10 simulations of the Model M_b trap-shy example are given in Table 3.7. Note that the sampling variance of $\hat{N}$ in this example is smaller than in the previous example, where p was 0.2 (Fig. 3.6 and Table 3.6). Because the estimate of N for Model M_b uses only the first-capture data, the difference in recapture probabilities is irrelevant in comparing the estimates of N in the two examples; in both, N = 100 and t = 7. But in the first example (Figs. 3.5 and 3.6, and Tables 3.5 and 3.6), p = 0.2, whereas in the second (Figs. 3.7 and 3.8 and Table 3.7), p = 0.4. From Table 3.7 we see that the average standard error of $\hat{N}$ is 2.2. This value represents a fourfold increase in precision of the estimate $\hat{N}$, achieved by increasing p from 0.2 to 0.4. The average coefficient of variation with p = 0.4 is about 2%, which is excellent.

```
OCCASION           J=      1    2    3    4    5    6    7
TOTAL CAUGHT    M(J)=      0   36   62   76   87   92   93   94
NEWLY CAUGHT    U(J)=     36   26   14   11    5    1    1

ESTIMATED PROBABILITY OF CAPTURE, P-HAT =     0.421699

ESTIMATED PROBABILITY OF RECAPTURE, C-HAT = 0.204036


POPULATION ESTIMATE IS          96 WITH STANDARD ERROR       1.7898

APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS         92 TO         100


HISTOGRAM OF U(J)


FREQUENCY     36   26   14   11    5    1    1
------------------------------------------------
EACH * EQUALS     4 POINTS

    36        *
    32        *
    28        *    *
    24        *    *
    20        *    *
    16        *    *    *
    12        *    *    *    *
     8        *    *    *    *
     4        *    *    *    *    *
------------------------------------------------
```

Fig. 3.8. Results of estimating N, p, and c under Model M_b from data in Fig. 3.7. The ML estimate of N is 96, with an estimated standard error of 1.8. The confidence interval should be taken as 94 to 100. The true parameter values underlying these Model M_b simulated data are N = 100, p = 0.4, and c = 0.2.

TABLE 3.7. Summary results for all 10 simulations of Model M_b (trap-shy case) with $N = 100$, $p = 0.4$, $c = 0.020$, and $t = 7$. The variation in the results across the 10 repetitions is entirely the result of the stochastic nature of the capture process. Notice that $\hat{N}$ varies only from 96 to 104. Compare these results with those of Table 3.6, where $N = 100$, $t = 7$, but $p = 0.2$. Clearly, values of $p = 0.4$ and $t = 7$ lead to very good estimates of N when Model M_b is true.

	Numbers First Captured on Occasion j											
Replication	u_1	u_2	u_3	u_4	u_5	u_6	u_7	M_8	$\hat{N}$	$\hat{se}(\hat{N})$	$\hat{p}$	$\hat{c}$
1	36	26	14	11	5	1	1	94	96	1.8	0.42	0.20
2	32	32	17	7	6	1	2	97	99	2.1	0.40	0.21
3	46	26	9	5	5	4	4	99	101	1.8	0.42	0.18
4	38	30	13	9	6	1	1	98	99	1.6	0.44	0.18
5	42	21	15	7	2	7	4	98	101	2.6	0.38	0.23
6	39	21	17	11	5	2	2	97	99	2.2	0.40	0.21
7	39	23	17	7	2	6	2	96	98	2.1	0.40	0.19
8	37	19	15	13	5	5	1	95	99	2.8	0.37	0.18
9	38	21	16	8	7	6	3	99	104	3.4	0.35	0.22
10	39	27	10	13	3	1	2	95	96	1.6	0.43	0.21
Averages	38.6	24.6	14.3	9.1	4.6	3.4	2.2	96.8	99.2	2.2	0.40	0.20
Standard deviations								1.8	2.4	0.6	0.03	0.02

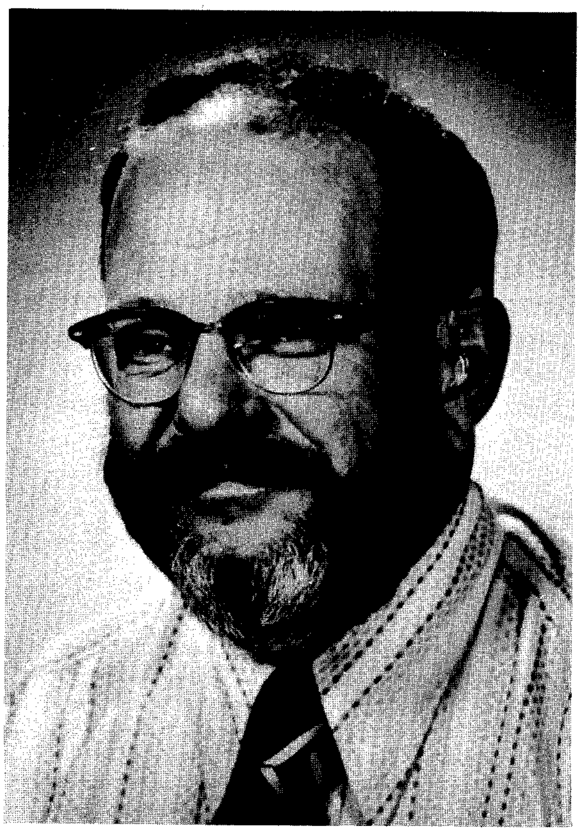


Trap-shy behavioral changes in capture probabilities may result from an unfavorable first-capture experience.

The behavioral response model for capture-recapture data is formulated to allow only one behavioral response (to first capture) and only that same response for all animals. However, the estimate of N based on the removal data $u_1, u_2, \ldots, u_t$ is robust to any extent of behavioral response. Given that one uses the Model M_b estimator, it does not matter that every animal has a different behavioral response to capture. In particular, some animals can become trap happy, some trap shy. This more general behavioral response would affect only the model selection procedure.

Model M_h, Heterogeneity

Capture probabilities often vary by animal, sometimes for obvious reasons (differences in species, sex, or age), but there also may be unrecognized sources of variation in capture probability by animal (social dominance, number and placement of traps in an animal's home range, or innate level of activity). Because these factors result in capture probabilities that vary among animals, we refer to this source of variation as heterogeneity. Numerous studies reported in the ecological literature clearly show heterogeneity and other sources of variation in capture probabilities for a wide range of species and many types of studies. Examples include *Young et al. (1952), Tanaka (1956), Crowcroft and Jeffers (1961), Huber (1962), Edwards and Eberhardt (1967), Bailey (1969), Gliwicz (1970), Carothers (1973a), Beukema* and *de Vos (1974), Jensen (1975),* and *Montgomery (1979)*. In studies where the true population size was known, the commonly used estimators were biased severely by heterogeneity of capture probabilities: the estimates were very much too low. (See, for example, *Edwards and Eberhardt 1967* and *Carothers 1973a.*) Computer simulation studies also have shown that heterogeneity can cause



L. Lee Eberhardt

Lee Eberhardt's interest in capture-recapture studies dates to the 1950s, when he was the biometrician with the Michigan Department of Conservation. An estimation method based on capture frequencies stemmed from his work with others on rabbits, which revealed the obvious inadequacies of existing methods. He was among the first to recognize that heterogeneity is a common violation of the equal-catchability assumption. More recently, he has studied sample size prediction in capture-recapture sampling.

Eberhardt took a B.S. degree in education from Minot State Teachers College and a Ph.D. degree in wildlife management from Michigan State University. He did postdoctoral work in statistics at the University of California at Berkeley under Jerzey Neyman, a founder of modern statistical theory. Since 1965, he has worked on a wide variety of quantitative ecological problems at Battelle Memorial Institute, including work with seals in Antarctica and New Zealand. (Recent photograph.)

substantial negative bias in the commonly used estimators. (See, for example, *Burnham and Overton 1969; Manly 1970; Gilbert 1973; Carothers 1973b* and *1979*; and *Otis et al. 1978*.) In spite of the evidence that heterogeneity exists and invalidates the usual estimates under Model M_t, only recently has Model M_h been formalized and has an estimator been derived for it, because it is a very difficult model. (See *Burnham 1972*; *Otis et al. 1978:33-37*; and *Burnham and Overton 1978* and *1979*.)

Model M_h assumes that each animal has a possibly unique individual capture probability p_i for the $i = 1, \ldots, N$ individuals. In terms of the most general capture probability structure p_{ij} for individual i on occasion j, this model assumes that $p_{ij} = p_i$ independent of capture occasion. Thus, neither time variation nor behavioral response is allowed in capture probabilities for this model.

In Model M_h, different individuals can have quite different capture probabilities . For example, animal *a* may have $p_a = 0.20$, and animal *b* may have $p_b = 0.60$. As always, these capture probabilities apply to each separate capture occasion; hence, on the first occasion, animal *a* has a 20% capture probability, but animal *b* has a 60% capture probability. By assumption, the catching of animal *a* will not influence whether animal *b* is caught. Of course, for this to be true there must be a sufficient number of traps to avoid having all traps fill up with animals. Also, if animal *b*, for example, is caught on day 1, it still will have a 60% capture probability on day 2, day 3, and so on.

Like Models M_t and M_b, Model M_h has only one source of variation in capture probabilities. Unlike those models and Model M_o, which have only a few parameters (for example, N and p in Model M_o), Model M_h can have as many as $N + 1$ parameters: N and $p_1, p_2, \ldots, p_N$. Estimating this many parameters from capture-recapture data is not possible. We must either reduce the number of capture probabilities in some way or find a way to estimate N without having to estimate all the capture probabilities. Both approaches have been explored *(Burnham 1972)*. We discuss here the only known method derived specifically to estimate N under Model M_h: the "jackknife" estimator.

The summary statistics from the first of 10 simulations of Model M_h with $N = 200$ and $t = 7$ are presented in Fig. 3.9. The specified capture probability structure has 10 values of p_i, each value holding for 20 of the 200 animals. The set of 10 capture probabilities is 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, and 0.55. For the example, 20 individuals have a daily capture probability of only 0.10, while a different 20 individuals have a daily capture probability of 0.55. The average of all the capture probabilities is $E(\bar{p}) = 0.325$. Thus, the expected number of captures on each occasion is $E(n_j) = NE(\bar{p}) = 65$, a value in close agreement with the results shown in Fig. 3.9.

Estimation of N under Model M_h is based on the capture frequency data. For the example in Fig. 3.9, these data are $f_1 = 50, f_2 = 46, f_3 = 35, f_4 = 24, f_5 = 14, f_6 = 5$, and $f_7 = 0$. Thus, 50 animals were caught only once, but 5 animals were caught on 6 of the 7 capture occasions. The jackknife estimator of N is computed as a linear combination of these capture frequencies: $\hat{N} = a_1f_1 + a_2f_2 + \ldots + a_tf_t$. The key to this estimator is the derivation of the coefficients a_i. In keeping with our emphasis on concepts, we will not delve into the mathematics behind the jackknife estimator of N. For those details see *Burnham* and *Overton (1978, 1979)* and *Otis et al. (1978:33-37, 108-109)*.

The estimation of N based on the simulation data of Fig. 3.9 is shown in Fig. 3.10. The capture frequencies are followed by a table of computed jackknife coefficients. Each column gives the first five coefficients (a_i) of a different estimator. (The values of coefficients 6, 7, etc., are all 1.) The program uses

OCCASION	J=	1	2	3	4	5	6	7	
ANIMALS CAUGHT	N(J)=	65	68	60	68	67	48	67	
TOTAL CAUGHT	M(J)=	0	65	107	133	149	162	167	174
NEWLY CAUGHT	U(J)=	65	42	26	16	13	5	7	
FREQUENCIES	F(J)=	50	46	35	24	14	5	0	

Fig. 3.9. Basic summary results from the first of 10 simulations of Model M_h, with N = 200, t = 7, and 10 different capture probabilities, each of which is applied to 20 different animals; the p_i values are 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, and 0.55. Notice that there is no apparent time variation in the numbers caught on each occasion, n_j, and there is a distinct decrease in the numbers of first captures. Finally, nothing in the capture frequencies f_j distinguishes these results (visually) from those of Model M_o.

```
NUMBER OF TRAPPING OCCASIONS WAS                  7
NUMBER OF ANIMALS CAPTURED, M(T+1), WAS         174
TOTAL NUMBER OF CAPTURES, N., WAS               443

FREQUENCIES OF CAPTURE,F(I)
   I=  1  2  3  4  5  6  7
F(I)= 50 46 35 24 14  5  0

            COMPUTED JACKKNIFE COEFFICIENTS

            N(1)        N(2)        N(3)        N(4)        N(5)
      1   1.857       2.571       3.143       3.571       3.857
      2   1.000       0.405       -.452      -1.310      -1.976
      3   1.000       1.000       1.305       1.833       2.357
      4   1.000       1.000       1.000       0.904       0.749
      5   1.000       1.000       1.000       1.000       1.013

            THE RESULTS OF THE JACKKNIFE COMPUTATIONS

    I      N(I)     SE(I)    0.95 CONF. LIMITS        TEST OF N(I+1) VS. N(I)
    0      174                                            CHI-SQUARE(1 D.F.)
    1      216.9     8.92      199.4      234.3              1.667
    2      225.2    13.82      198.1      252.3              0.001
    3      225.0    19.52      186.7      263.3              0.062
    4      223.2    25.49      173.2      273.1              0.072
    5      221.6    30.48      161.9      281.3              0.000

AVERAGE P-HAT = 0.2944

INTERPOLATED POPULATION ESTIMATE IS          215 WITH STANDARD ERROR       8.6844

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL        198 TO         233


HISTOGRAM OF F(I)

FREQUENCY       50    46    35    24    14     5     0
-------------------------------------------------------
EACH * EQUALS    5 POINTS

        50       *
        45       *     *
        40       *     *
        35       *     *     *
        30       *     *     *
        25       *     *     *     *
        20       *     *     *     *
        15       *     *     *     *     *
        10       *     *     *     *     *
         5       *     *     *     *     *     *
-------------------------------------------------------
```

Fig. 3.10. Results of estimating N under Model M_h from the simulated data of Fig. 3.9; $N = 200$ and $t = 7$. Ten different capture probabilities are spread evenly over $p = 0.10$ to 0.55, and each value is applied to 20 different animals. Program CAPTURE computes $\hat{N}$ and its standard error and prints out some intermediate results of the estimation process, which are not explained here. For this example, $\hat{N} = 215 \pm 8.7$, and the 95% confidence interval on N is therefore 198 to 233. The histogram of capture frequencies provides a visual display of these data.

the coefficients to compute a sequence of five estimators, then selects one estimator to simultaneously minimize bias and sampling variance. The second table of values in Fig. 3.10 gives the results of this procedure for estimator selection. Finally, CAPTURE gives the estimate of N and its estimated standard error. The user need not be concerned about the intermediate computations in the figure; they are explained in *Otis et al. (1978:108-109)* and in *Burnham and Overton (1979)*.

In the example of Fig. 3.10 we have $\hat{N} = 215$ and $\hat{se}(\hat{N}) = 8.68$, thus the coefficient of variation of $\hat{N}$ is estimated as $8.68/215 = 4.0\%$. The 95% confidence interval is 198 to 233, covering the true population size of 200. The average capture probability during the entire study is estimated as $\bar{p} = 0.294$; the true value is $E(\bar{p}) = 0.325$. The probability is high enough to make the results seem reliable.

The jackknife estimator is not an ML estimator. We cannot derive a useful ML estimator for Model M_h because of the many parameters. This model is mathematically very difficult to deal with, but it probably is very realistic for many studies. Consequently, having an estimator for it is important, even though the jackknife estimator does not perform well under some patterns of heterogeneity. Specifically, if many animals have very small capture probabilities (say, less than 0.05), the jackknife estimator will underestimate N, as will every other known estimator. There is no mathematical solution to this problem; if some animals are essentially uncatchable, no estimation method can estimate N properly.

The results of estimating N for all 10 simulations of the heterogeneity model underlying Figs. 3.9 and 3.10 are shown in Table 3.8. The 10 sets of capture frequencies again illustrate sampling variation. The average of the 10 values of $\hat{N}$ is 209.4. For this particular model, $\hat{N}$ probably has a slight positive bias. In general, $\hat{N}$ under the Model M_h is not free of bias, but is more robust (has smaller bias) than the previously discussed estimators for Models M_o, M_t, and M_b when these estimators are applied to data that really fit Model M_h.

Model M_{bh}, Behavioral Response and Heterogeneity

In real populations, capture probabilities may vary by animals, as in Model M_h, and there may also be behavioral response to first capture, as in Model M_b. The presence of both sources of variation in individual capture probabilities results in Model M_{bh}. Under this model, each animal is allowed to have its own probability of first capture, p_i, $i = 1, \ldots, N$. This part of Model M_{bh} is exactly like Model M_h. However, the animal also may have a behavioral response to first capture, which alters its subsequent (daily) capture probability. Thus, we let c_i be the probability of recapture for the i^{th} animal. There is no relation between p_i and c_i built into this model (although there could be).

To explain Model M_{bh}, consider three individuals with first-capture probabilities of 0.25, 0.5, and 0.65. There is a 25% chance of catching individual 1 on day 1. If individual 1 is not caught on day 1, there remains a 25% chance of catching it on day 2. Let individual 1 become trap shy if caught, say, $c_1 = 0.1$, but let animal 3 become trap happy if caught, say, $c_3 = 0.9$. Finally, assume that animal 2 has no behavioral response to capture, hence $c_2 \equiv p_2 = 0.5$. These types of capture and recapture probabilities are allowed under Model M_{bh}.

TABLE 3.8. Summary results for all 10 simulations of Model M_h with $N = 200$, $t = 7$. There are 10 different individual capture probabilities assumed for this population; 20 animals have $p = 0.10$, 20 have $p = 0.15$, and so on with sets of 20 animals each having a capture probability of $p = 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50$, or 0.55. The estimate of N is generally reasonable, with an average value over the 10 repetitions of 209.4.

	Capture Frequencies										
Replication	f_1	f_2	f_3	f_4	f_5	f_6	f_7	M_8	$\hat{N}$	$\hat{se}(\hat{N})$	$\hat{p}$
1	50	46	35	24	14	5	0	174	215	8.68	0.294
2	49	37	38	26	13	3	0	166	211	9.75	0.287
3	41	53	38	22	17	3	1	175	207	7.63	0.317
4	41	37	38	26	19	4	0	165	198	7.80	0.326
5	48	36	34	25	20	3	0	166	210	9.60	0.299
6	39	51	42	23	13	3	1	172	203	7.42	0.316
7	44	43	41	37	8	6	0	179	215	8.02	0.317
8	34	50	32	34	15	5	0	170	197	6.94	0.342
9	42	54	38	24	15	3	0	176	209	7.73	0.310
10	54	49	40	27	9	5	0	184	229	9.08	0.284
Averages	39.2	45.6	37.6	26.8	14.3	4.0	0.2	172.7	209.4	8.3	0.309
Standard deviations								6.2	9.3		0.018

The estimator of population size for use with Model M_{bh} is based entirely on the first-capture data, just as it is in Model M_b. First-capture data are $u_1, \ldots, u_t$, the numbers of animals caught for the first time on occasion 1 through t. Because recapture information does not enter the estimator, the nature of the behavioral response for each animal is irrelevant to the estimation of N. In principle, one is "removing" animals from the population by marking them and estimating population size as if this were a removal study. In this example, however, these first-capture data do not now fit Model M_b because of the presence of heterogeneity. Thus a more general estimation method is required for Model M_{bh} than for Model M_b. See comments on this method given below, in Chapter 4, and in *Otis et al. (1978:40-43, 112-113)*.

Summary data from the first of 10 repetitions simulating a Model M_{bh} study are presented in Fig. 3.11. This simulation uses $N = 200$ and $t = 8$ occasions. The heterogeneity structure assumed for first-capture probabilities is the same as that used in the previous section on Model M_h; namely, 20 animals with $p_i = 0.1$, 20 with $p_i = 0.15$, and so on up to 20 animals with $p_i = 0.55$. On the average, results for capture occasion 1 will be the same for this example as for the Model M_h example. After first capture, however, a trap-shy behavioral response is assumed to occur for all 200 animals. To simulate such a response, we generated an individual's recapture probability as $c_i = 0.6 \times p_i$ after its first capture. If an individual had a first-capture probability of 0.1, its recapture probability became 0.06 ($= 0.6 \times 0.1$); if its first-capture probability was 0.50, its recapture probability became 0.30.

Perhaps a useful way to visualize this example is to see it as 10 separate Model M_b studies that have been pooled. Specifically, each set of 20 animals, with their own common first-capture probability, satisfies the assumptions of a Model M_b study.

The concept that, on the average, the first-capture data (u_j) will decrease over the t capture occasions is illustrated by Fig. 3.11. The expected decrease in the numbers of animals caught for the first time on occasions 1, 2, 3, and so on, is the only certain feature of Model M_{bh} data. The characteristics of total daily captures (n_j) are not predictable because the recapture probabilities are not predictable. However, if all animals show a trap-shy response, some decrease in the n_j will be expected over time. Such a decrease is not very evident in the data of Fig. 3.11, especially after occasion 1. In general, there are no easily perceived clues in the summary data from a Model M_{bh} study to distinguish it from the results of Model M_b or several other models, such as some versions of Models M_t and M_{tb}. Making a judgment on the best underlying model for a capture study requires sophisticated data analyses to test model assumptions.

The results of applying the Model M_{bh} estimator to the first-capture data of Fig. 3.11 are shown in Fig. 3.12. The computed estimate is 192 with an estimated standard error of 5.0. The approximate 95% confidence interval covers the true value of $N = 200$, and the lower limit of the interval (182) is not less than the number of different individuals caught ($M_9 = 181$). In addition to the estimate of N, the print-out gives information on the estimation process and provides a histogram of the "removal" data, $u_1, \ldots, u_8$. To explain the main body of numbers in Fig. 3.12, we must discuss the ideas behind estimation of population size with Model M_{bh}.

We define the expected probability of first capture on occasion j as $\bar{p}_j = E(u_j)/[N - E(u_1) - \ldots - E(u_{j-1})]$, $j = 1, \ldots, t$. Under Model M_b, these probabilities are constant; that is, $\bar{p}_j \equiv p$. With such a model for the first-capture data (conveniently called removal data), estimators of p and N are

OCCASION	J=	1	2	3	4	5	6	7	8	
ANIMALS CAUGHT	N(J)=	64	47	54	39	40	50	46	39	
TOTAL CAUGHT	M(J)=	0	64	96	128	142	156	168	175	181
NEWLY CAUGHT	U(J)=	64	32	32	14	14	12	7	6	
FREQUENCIES	F(J)=	74	50	30	21	5	1	0	0	

Fig. 3.11. Basic summary statistics for the first of 10 simulations of Model M_{bh}, with $N = 200$ and $t = 8$. First-capture probabilities have the same structure as the simulation example of Figs. 3.9 and 3.10; namely, the 10 sets of 20 animals each have capture probabilities 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, and 0.55. However, after the first capture each animal becomes trap shy; its new capture probability becomes 0.6 times its previous capture probability. These data give no visual clue to the complex, underlying probability-of-capture model, although there is some basis for thinking that average daily capture probabilities are changing.

OCCASION	J=	1	2	3	4	5	6	7	8	
TOTAL CAUGHT	M(J)=	0	64	96	128	142	156	168	175	181
NEWLY CAUGHT	U(J)=	64	32	32	14	14	12	7	6	

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J),J=1,...., 8							
1	191.60	5.026287	5.518	0.4792	0.2998	0.2998	0.2998	0.2998	0.2998	0.2998	0.2998	0.2998
2	195.86	7.537645	3.456	0.6300	0.3268	0.2647	0.2647	0.2647	0.2647	0.2647	0.2647	0.2647
3	192.77	7.023416	2.876	0.5788	0.3320	0.2485	0.2915	0.2915	0.2915	0.2915	0.2915	0.2915
4	201.16	15.53062	0.994	0.8026	0.3182	0.2333	0.3043	0.2238	0.2238	0.2238	0.2238	0.2238
5	192.44	10.27116	0.434	0.8050	0.3326	0.2492	0.3318	0.2173	0.3029	0.3029	0.3029	0.3029
6	187.77	8.008416	0.466	0.4946	0.3408	0.2586	0.3487	0.2342	0.3059	0.3888	0.3888	0.3888

POPULATION ESTIMATE IS 192 WITH STANDARD ERROR 5.0263

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 182 TO 202

HISTOGRAM OF U(J)

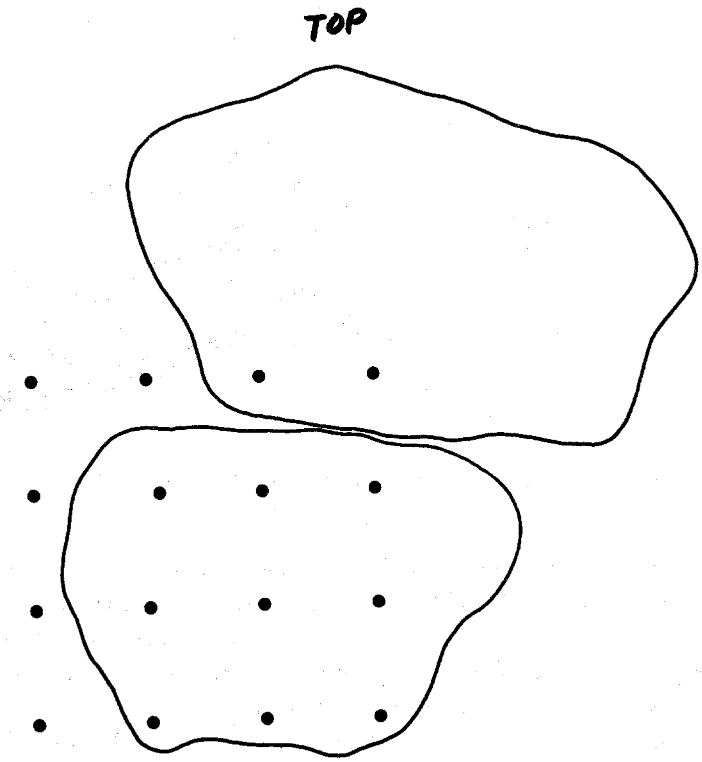
```
FREQUENCY     64    32    32    14    14    12     7     6
-----------------------------------------------------------
EACH * EQUALS   7 POINTS

     63        *
     56        *
     49        *
     42        *
     35        *     *     *
     28        *     *     *
     21        *     *     *
     14        *     *     *     *     *     *
      7        *     *     *     *     *     *     *     *
-----------------------------------------------------------
```

Fig. 3.12. Results of estimation of N from the simulation data of Fig. 3.11; the true population size is 200, and the true model is M_{bh}. The estimator under Model M_{bh} computed by CAPTURE is 192, with an estimated standard error of 5.03. The confidence interval covers N and does not overlap the total number of animals caught, 181.

possible without further assumptions. When there is heterogeneity of capture probabilities, all the $\bar{p}_j$ are different, but in a qualitatively predictable way: $\bar{p}_1$ is greater than $\bar{p}_2$, $\bar{p}_2$ is greater than $\bar{p}_3$, and so on. The expected capture probabilities decrease, $\bar{p}_1 > \bar{p}_2 > \ldots > \bar{p}_{t-1} > \bar{p}_t$, and there is a reason for this decrease. Individuals with greater first-capture probabilities tend to be caught earlier in the trapping than individuals with smaller first-capture probabilities. For example, consider the simulation study underlying Figs. 3.11 and 3.12. Of the 20 individuals having a first-capture probability of 0.50, half (or 10) of them would be caught, on the average, on day 1. But of the 20 individuals with capture probability 0.10, we would expect to catch only 2 on day 1. In only these 2 groups of 20 animals, on day 2 there would be an expected $10 + 18 = 28$ animals left, and their expected average capture probability would be $\bar{p}_2 = (0.5 \times 10 + 0.1 \times 18)/28 = 0.24$, down from 0.30 on day 1. Thus on day 2, on the average, only 10 individuals with a capture probability of 0.5 would be left uncaught, but 18 individuals with a capture probability of 0.1 still would be uncaught. It is this phenomenon that causes the first-capture probabilities $\bar{p}_j$ to decrease over time when heterogeneity is present.

For these same two groups of animals, the value of $\bar{p}_3$ is $(0.5 \times 5 + 0.1 \times 16.2)/21.2 = 0.19$, and similar computations yield $\bar{p}_4 = 0.158$ and $\bar{p}_5 = 0.135$. Although these expected first-capture probabilities are computed for only 2 groups of 20 animals in this population, they illustrate two points: (1) the expected probabilities of first capture, $\bar{p}_j$, decrease over time, and (2) this decrease is most rapid for the first few days. The second point is hard to see, but it is important. The differences $\bar{p}_j - \bar{p}_{j+1}$ get smaller as time (j) progresses; in a sense, the later values of $\bar{p}_j$ tend to stabilize. For instance, from the sample values above we have $\bar{p}_1 = 0.30$, $\bar{p}_2 = 0.24$, $\bar{p}_3 = 0.19$, $\bar{p}_4 = 0.158$, and $\bar{p}_5 = 0.135$.

If we are to estimate N from the data $u_1, \ldots, u_t$, we must reduce the number of parameters ($\bar{p}_1, \ldots,$ $\bar{p}_t$. Because of the characteristic pattern of $\bar{p}_j$ values decreasing toward a limit, we have devised the following scheme *(Otis et al. 1978:40-43)*. We fit a sequence of increasingly general models to the



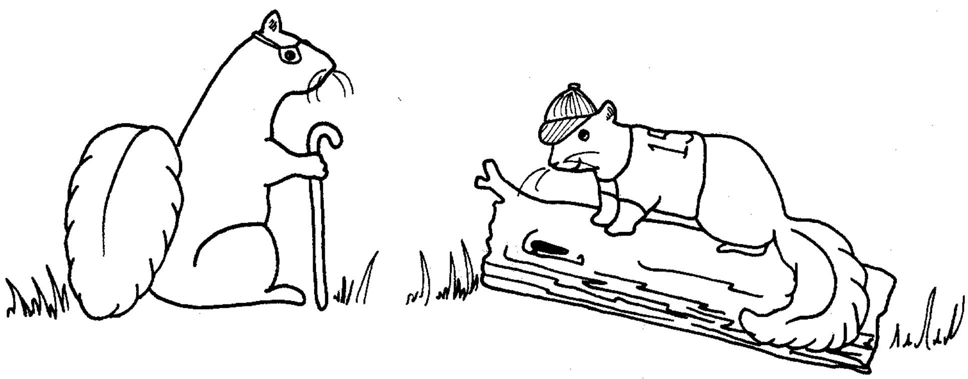
Capture probabilities can be affected by the number of traps in an individual's home range. The two traps in the range above versus the nine in the range below are a source of heterogeneity.

"removal" data $u_1, \ldots, u_t$ and stop when a suitable fit is found. For each model, N can be estimated. The first model is that of a constant first-capture probability: $\bar{p}_1 = \bar{p}_2 = \ldots = \bar{p}_t \equiv p$, the situation assumed for Model M_b (no heterogeneity). The parameters p and N are estimated by the ML method, and a goodness of fit test is computed for this model. In Model M_b there is only one capture probability parameter ($k = 1$). If heterogeneity is slight, this model may be a reasonable one to use. In Fig. 3.12, the row for $k = 1$ gives the results of this model: $\hat{N} = 191.6$, $\hat{se}(\hat{N}) = 5.03$, a chi-square goodness of fit test statistic of 5.518 with an associated observed significance level of 0.4792, and finally $\hat{\bar{p}}_1 = \hat{\bar{p}}_2 = \ldots = \hat{\bar{p}}_8 = 0.2998$.

The next model fit by CAPTURE allows two capture probabilities: $\bar{p}_2 = \bar{p}_3 = \ldots = \bar{p}_t$ are assumed all equal, but $\bar{p}_1$ is allowed to differ from these. Results of estimation for this model are shown on row $k = 2$ in Fig. 3.12. In particular, note that $\hat{\bar{p}}_1 = 0.3268$, whereas $\hat{\bar{p}}_2 = \hat{\bar{p}}_3 = \ldots = \hat{\bar{p}}_8 = 0.2647$. This model is also judged to fit the data by the chi-square goodness of fit test; the observed significance level is $P = 0.6300$.

In general, the k^{th} model of the sequence allows k different capture probabilities: $\bar{p}_1, \ldots, \bar{p}_{k-1}$ are allowed to differ, but $\bar{p}_k = \ldots = \bar{p}_t$ are forced to be equal. Thus, k is the number of different capture probability parameters in the model. Such models are fitted for $k = 1, 2, \ldots, t-2$. In Fig. 3.12, row $k = 6$, we see that the first 5 values of $\hat{\bar{p}}_j$ all differ (slightly) but the values $\hat{\bar{p}}_6 = \hat{\bar{p}}_7 = \hat{\bar{p}}_8 = 0.3888$ all are forced to be equal.

The heuristic intent in fitting a sequence of increasingly general models is to find the simplest one that gives an adequate fit to the data. Results from that model are used to estimate N. Often $k = 1$ is chosen unless heterogeneity is extreme. The criterion of "fit" implemented in CAPTURE is that $P \geq 0.20$ must hold where P is the observed significance level of the chi-square goodness of fit test; that is, the probability of a test statistic as large as, or larger than, the computed statistic (see Fig. 2.9). The large significant value (0.2) was chosen to minimize Type II errors (selecting too simple a model and thereby getting a biased but precise estimator).



Individual capture probabilities (heterogeneity) may depend on the age of the animals.

The results for all 10 simulations of the Model M_{bh} example used in this section are presented in Table 3.9. In general, $\hat{N}$ tends to be an underestimate in the presence of heterogeneity, even when the scheme of selecting from a series of models is used. Also, as replication 10 illustrates, a very poor result occasionally occurs when both heterogeneity and behavioral response are present. However, the estimated standard error of 76.7 clearly shows that this estimate is very unreliable.

Heterogeneity of capture probabilities makes estimation of N difficult. We saw this for Model M_h, and we observe it again here, for Model M_{bh}. Unbiased estimation of N cannot be expected when heterogeneity is present. But the estimation scheme used for Model M_{bh}, which is basically the generalized removal method discussed in Chapter 4, reduces the bias as compared with results from using the estimator for Model M_b.

TABLE 3.9. Summary results for all 10 simulations of Model M_{bh} with $N = 200$, $t = 8$. Initial capture probabilities have the same structure as that of the Model M_h example; namely sets of 20 animals each have capture probability 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, or 0.55. After an animal is first captured, it becomes trap-shy, and its recapture probability becomes 0.6 times its probability of first capture. For example, an animal with first-capture probability of 0.40 has a recapture probability of 0.24 = 0.40 x 0.60. The results for repetition 10 are no mistake; rather, they demonstrate that occasionally one gets very poor results from capture studies.

	Numbers First Captured on Occasion j										
Replication	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	M_9	$\hat{N}$	$\hat{se}(\hat{N})$
1	64	32	32	14	14	12	7	6	181	192	5.0
2	57	45	44	14	9	0	0	0	169	178	4.7
3	61	39	24	17	16	10	7	6	180	191	5.0
4	54	43	29	21	18	4	5	1	175	179	2.9
5	76	37	26	9	9	7	9	4	177	184	4.3
6	75	39	26	11	8	9	8	5	181	189	4.6
7	66	48	22	17	8	7	7	3	178	182	2.8
8	63	38	29	20	10	6	9	5	180	188	4.1
9	63	34	29	19	15	6	9	5	180	189	4.6
10	66	35	38	10	12	10	11	7	189	254	76.7
Averages	64.5	39.0	29.9	15.2	11.9	7.1	7.2	4.2	179.0	192.6	25.8
Standard deviations									5.1	22.1	

Model M_{th}, Time Effects and Heterogeneity

Of the three remaining models, none has an associated estimator. Rather, these models are necessary to complete the set of conceptual models for the three sources of variation we have recognized (time, behavior, and heterogeneity). They are needed for testing, and they certainly can arise as descriptions of real capture-recapture data. As with the five models previously described, we have done a simulation example of each model; most of the results are presented later in this chapter for purposes of comparing estimators over models.

If both time and heterogeneity affect daily capture probabilities, we have Model M_{th}. This model can be conceptualized by starting with the heterogeneity-only Model M_h and then by assuming that external factors, such as weather or unequal effort, cause an upward or downward shift in all individual capture probabilities on each capture occasion. Let $p_1, p_2, \ldots, p_N$ be individual capture probabilities, and assume that on the j^{th} capture occasion the actual capture probabilities (as in Model M_h) are $p_{ij} = p_i \times a_j$, where the parameter a_j represents a time effect. For example, let us again consider the heterogeneity structure used for the capture probabilities in the simulation of Models M_h and M_{bh} (sets of 20 animals each at 10 different capture probability levels, $p = 0.1$ to 0.55). Then let there be $t = 5$ capture occasions with $a_1 = 1.0$, $a_2 = 0.6$, $a_3 = 1.5$, $a_4 = 0.7$, and $a_5 = 0.9$. On day 1, the capture probabilities are the original ones given by the heterogeneity structure above. But on day 2, each animal's capture probability (p_{i2}) is p_{i1} multiplied by 0.6: $p_{i2} = p_{i1} \times 0.6$. Therefore, on day 2 the set of 10 basic capture probabilities becomes 0.06, 0.09, 0.12, 0.15, 0.18, 0.21, 0.24, 0.27, 0.3, and 0.33, with each set of 20 animals having one of these capture probabilities, as before. There is a considerable difference in average capture probability on days 1 and 2. In fact, in this example the average capture probability for day 1 is 0.325, but it drops to $0.195 = 0.6 \times 0.325$ for day 2. The change in capture probability is applied to all animals, regardless of whether they were caught on day 1. Changes in capture probabilities thus are not a result of behavioral response of animals, but rather are due to external factors, which we lump under the name of time effects.

In this example, which is the basis for our simulation example of Model M_{th}, $a_3 = 1.5$, so individual capture probabilities are greater on day 3 than they were on day 1. The capture probabilities of day 3 are based on the original heterogeneous capture probabilities of day 1 times 1.5, not on those of day 2 times 1.5. Thus on day 3, animals with the "base" capture probability of 0.40 have a capture probability of $0.6 = 0.40 \times 1.5$. The full set of capture probabilities in this example is given in Table 3.10. There are 5 occasions and 10 groups of 20 animals ($N = 200$); each group has a different capture probability.

Summary statistics for the first repetition of 10 simulations are presented in Fig. 3.13 for this example of Model M_{th}. These data show time variation in the average daily capture probabilities: $n_1 = 73$, $n_2 = 43$, $n_3 = 91$, $n_4 = 45$, and $n_5 = 57$. The n_j compare well with the expected values $E(n_j) = N \times \bar{p}_j = N \times 0.325 \times a_j$; for example, $E(n_3) = 200 \times 0.325 \times 1.5 = 98$. However, it is impossible to tell by looking at just these summary statistics that the data arise from a case of Model M_{th}.

Model M_{tb}, Time Effects and Behavioral Response

When both time and behavioral response affect capture probabilities, we have Model M_{tb}. We conceptualize a set of time-varying daily capture probabilities ($p_1, p_2, \ldots, p_t$) that apply to all animals not yet caught. Thus if an individual is not caught on day 1, its probability of capture on day 2 is p_2. If an animal is caught on day 1, however, it exhibits a behavioral response to this capture, and its subsequent daily capture probabilities alter. They become $c_2, c_3, \ldots, c_t$. The recapture probabilities also are allowed to vary by time, but $c_2 \neq p_2$, $c_3 \neq p_3$, and so on. Notice that if we assume no time variation in capture or recapture probabilities ($p_1 = p_2 = \ldots = p_t$ and $c_2 = c_3 = \ldots = c_t$), we have Model M_b (behavior only), or if we assume capture and recapture probabilities are the same ($c_i = p_i$, $i = 2, \ldots, t$), we have Model M_t.

Example data for Model M_{tb} were simulated with the following parameters. First note that $N = 150$ and $t = 5$ were used. The probabilities of first capture on days 1 through 5 were $p_1 = 0.3$, $p_2 = 0.2$, $p_3 =$

TABLE 3.10. The capture probability structure used as the basis of a simulation example of Model M_{th}, with $N = 200$ and $t = 5$. The population is composed of 10 groups of 20 animals each. Heterogeneity of capture probabilities extends over these 10 groups. Superimposed on the heterogeneity structure is a multiplicative time effect; $p_{ij} = p_i \times a_j$, where $a_1 = 1$, $a_2 = 0.6$, $a_3 = 1.5$, $a_4 = 0.7$, $a_5 = 0.9$, and the p_i have the same values as for the Model M_h example.

	Day				
Animal Group	1	2	3	4	5
1	0.10	0.06	0.15	0.07	0.09
2	0.15	0.09	0.23	0.11	0.14
3	0.20	0.12	0.30	0.14	0.18
4	0.25	0.15	0.38	0.18	0.23
5	0.30	0.18	0.45	0.21	0.27
6	0.35	0.21	0.53	0.25	0.32
7	0.40	0.24	0.60	0.28	0.36
8	0.45	0.27	0.68	0.32	0.41
9	0.50	0.30	0.75	0.35	0.45
10	0.55	0.33	0.83	0.39	0.50
Averages	0.325	0.20	0.49	0.23	0.30

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	73	43	91	45	57	
TOTAL CAUGHT	M(J)=	0	73	95	144	153	164
NEWLY CAUGHT	U(J)=	73	22	49	9	11	
FREQUENCIES	F(J)=	65	64	24	11	0	

Fig. 3.13. Basic summary statistics for the first of 10 simulations of Model M_{th}, with $N = 200$ and $t = 5$. The heterogeneity structure in the population is the same as that used in the simulation of Models M_h and M_{bh}. Time variation in capture probabilities is imposed on these individual capture probabilities. See Table 3.10 for the complete capture probability structure of this (simulation) model. Notice that time variation is evident in the data, both in the n_j and in the fact that $u_2 = 22$ while $u_3 = 49$. There is no way to look at summary statistics like these and tell that the model is M_{th} rather than M_t, M_{tb}, or M_{tbh}.

0.4, $p_4 = 0.35$, and $p_5 = 0.25$. Of course, on day 1 all animals have $p_1 = 0.3$ as their capture probability, but on day 2 only those animals not caught on day 1 have capture probability $p_2 = 0.2$. Here, the recapture probabilities are set at one-half the original capture probabilities, thus $c_2 = 0.1$, $c_3 = 0.2$, $c_4 = 0.175$, and $c_5 = 0.125$. This is a case of trap-shy response. We have set c_i equal to a constant multiple of p_i only for convenience when we simulate Model M_{tb} data. The basic model assumes no constant relation between p_i and c_i.

There are only five relevant different capture histories for this example, corresponding to the day on which the animals were first caught. Table 3.11 shows the applicable set of capture probabilities, and Fig. 3.14 gives the summary statistics from one simulation repetition of this example of Model M_{tb}. Because of the trap-shy behavioral response, recapture probabilities are less than first-capture probabilities. Under Model M_b, this relationship would cause a decline in the daily numbers caught (n_j), over time. However, such a decline is masked here by the time variation in capture probabilities.

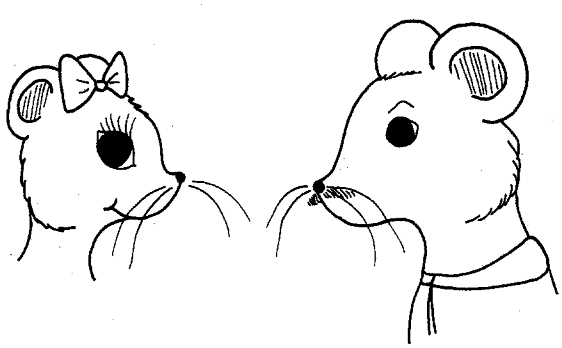
TABLE 3.11. A representation of the capture probabilities applicable to different capture histories of animals in the simulation example of Model M_{tb}. If an animal is captured on day 1, the first row of the table gives its capture probabilities. On days 2 through 5, the animal is subject to the lower recapture probabilities as a result of a trap-shy behavioral response. At the other extreme, animals not captured at all or not captured until day 5 are subject to the daily capture probabilities shown in row 5. Viewed another way, if we look at column 3, all individuals caught on either day 1 or 2 or on both days have capture probability 0.2 on day 3, but individuals not caught by day 3 have capture probability 0.4 on that day.

	Capture Probability On Each Capture Occasion				
Occasion When First Caught	1	2	3	4	5
1	0.3	0.1	0.2	0.175	0.125
2	0.3	0.2	0.2	0.175	0.125
3	0.3	0.2	0.4	0.175	0.125
4	0.3	0.2	0.4	0.350	0.125
5	0.3	0.2	0.4	0.350	0.250

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	50	17	40	37	15	
TOTAL CAUGHT	M(J)=	0	50	64	89	112	116
NEWLY CAUGHT	U(J)=	50	14	25	23	4	
FREQUENCIES	F(J)=	79	31	6	0	0	

Fig. 3.14. Basic summary statistics for the first of 10 simulations of Model M_{tb}, with $N = 150$ and $t = 5$. See text and Table 3.11 for the underlying capture probability structure. Merely looking at these data does not make clear what the true model is.

Because no relation is assumed between recaptures and first captures, only the first-capture data, u_1, u_2, . . ., u_t, are relevant for estimating population size N under Model M_{tb}. These removal data are the appropriate basis for estimating N whenever there is a behavioral response to first capture. For Model M_b, the data $u_1, u_2, \ldots, u_t$ depend on only two parameters, N and p; hence, N can be estimated if there are at least two capture occasions. For Model M_{bh}, there are t + 1 parameters (N, $\bar{p}_1, \ldots, \bar{p}_t$), but we



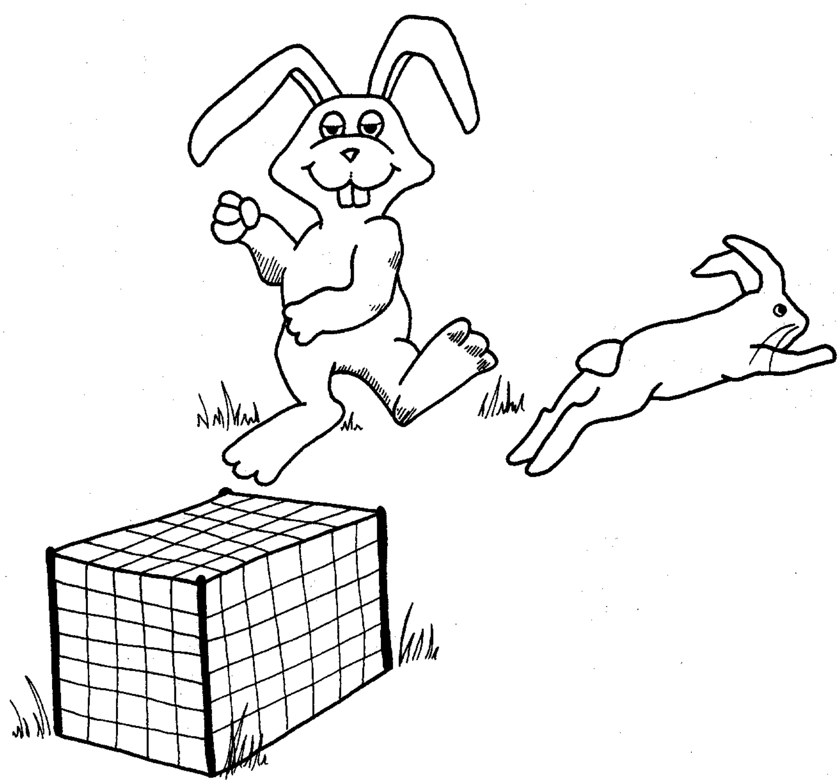
Heterogeneity of individual capture probabilities can be caused partly by differences in sex-specific capture rates.

know enough about the nature of the average first-capture probabilities $\bar{p}_j$ to devise an estimation scheme. For Model M_{tb}, there also are $t + 1$ parameters, for only t "bits" of data. However, now there are no logical relations among the (average) first-capture probabilities, $\bar{p}_1, \ldots, \bar{p}_t$, and therefore we cannot devise a reasonable method of estimating population size from Model M_b data.

Model M_{tbh}, Time Effects, Behavioral Response, and Heterogeneity

In Model M_{tbh}, the most general (closed) model of capture-recapture studies, all three factors (time, behavior, and heterogeneity) are operating. To formulate the factors as a mathematical model, we take the heterogeneity and time effects to be multiplicative—exactly as with Model M_{th}. A behavioral response change in basic capture probabilities is assumed after an animal's first capture. There is no unique way to formulate this model. However, we can illustrate the concept by reference to a simulation example of Model M_{tbh}, with $N = 200$ and $t = 6$.

We start on the first trapping day with only heterogeneity evident in the population, because it takes at least 2 days for behavior or time effects, or both, to become evident. The heterogeneity structure on capture probabilities assumed here is the same as in the examples of Model M_h, M_{bh}, and M_{th}: 10 groups of 20 animals each with capture probabilities 0.1 (1st group), 0.15 (2nd group), up to 0.55 (10th group). Thus the expected results on day 1 are the same for Model M_{tbh} as for Models M_h, M_{bh}, and M_{th}. This heterogeneity structure is modified during the course of six trapping occasions for time effects and behavioral response. For animals not previously captured, the probability of first capture is $p_i a_j$, for the i^{th} individual on the j^{th} day. The time effects here are $a_1 = 1$, $a_2 = 0.7$, $a_3 = 1.3$, $a_4 = 1.4$, $a_5 = 0.6$, and $a_6 = 1.2$. If no behavioral response were allowed, the capture probability structure, $p_{ij} = p_i a_j$ would be an instance of Model M_{th}.



Very active and socially dominant individuals may have high individual capture probabilities—a possible source of heterogeneity.

After the first capture, the recapture probability for an animal is equal to $c_i a_j$, where c_i is not the same as p_i. For simulation of this example we assume a fixed relation between initial and subsequent capture probability. Thus we set the recapture probability as $c_i a_j \equiv p_i a_j b$ and use $b = 1.3$ in the simulation. Consider the second trapping day. All animals not caught on the first day have capture probability $p_i \times 0.7$, where p_i is their initial (or "basic") capture probability. Any animal caught on the first day has a recapture probability on day 2 of $p_i \times 0.7 \times 1.3 = p_i \times 0.91$.

Consider what happens on day 4. The capture probability of an animal not previously caught is $p_i \times 1.4$. For example, if $p_i = 0.55$, that animal's capture probability on day 4 is $0.55 \times 1.4 = 0.77$. If the animal had been captured previously, its recapture probability is computed as $p_i \times 1.4 \times 1.3$. For $p_i = 0.55$, this computation gives $0.55 \times 1.4 \times 1.3 = 1.001$. Of course, a capture probability exceeding 1 is not meaningful. The full-blown mathematical versions of all the models we present here do not allow capture probabilities outside the range of 0 to 1. Program CAPTURE truncates back to 1 any capture probability computed as more than 1. Thus, on day 4 in this example, all previously captured animals with the basic capture probability of 0.55 will be caught.

Another way to view the example is to see each group of 20 animals as a case of Model M_{tb}. Because the capture probabilities differ between the groups, heterogeneity is also present, and the whole population becomes a case of Model M_{tbh}.

The summary statistics from the first of 10 repetitions of this simulation example of Model M_{tbh} are presented in Fig. 3.15. As with data from other models, we cannot identify the underlying model just by looking at these data. However, the increase in first-capture data on day 3 over day 2 and again on day 6 over day 5 ($u_1 = 63$, $u_2 = 23$, $u_3 = 44$, $u_4 = 21$, $u_5 = 4$, and $u_6 = 14$) tends to rule out Model M_b, and hence Model M_o. Also, the time variation evident from the u_i and n_i tends to rule out Model M_h. Further determination of the best fitting model would require tests of assumptions.

Summary of Models

Because the reader must have the eight models and the relations among them clearly in mind before proceeding further, we summarize them here briefly. The reader also should see *Otis et al. (1978:50-52).* Table 3.12 shows the models by symbol, the sources of variation in capture probabilities that enter into each model, and the estimator associated with the model. For convenience we have associated a name with each estimator; the names are used by program CAPTURE and appear in the CAPTURE output.

The eight models have distinct relations to each other. For example, the simplest model, M_o, is a special case of all other models. Figure 3.16 diagrams the relations between the models, some of which have been pointed out in the preceding discussions. In the figure, each arrow points from one model to another, which is a special case of the first model. Mathematically, the relations are true because we can assume that some parameters of the more general model are equal to each other to "produce" the simpler model. For example, in Model M_b if first-capture probability p is assumed equal to recapture probability c, we have Model M_o. Assumptions such as these about capture probabilities often can be tested.

OCCASION	J=	1	2	3	4	5	6	
ANIMALS CAUGHT	N(J)=	63	46	100	102	38	96	
TOTAL CAUGHT	M(J)=	0	63	86	127	148	152	166
NEWLY CAUGHT	U(J)=	63	23	41	21	4	14	
FREQUENCIES	F(J)=	36	45	42	24	17	2	

Fig. 3.15. Basic summary statistics for the first of 10 simulations of Model M_{tbh}, with $N = 200$ and $t = 6$. The capture probability structure for this simulation is discussed in the text. Briefly, there is heterogeneity (as for Model M_h), 10 groups of 20 animals each at "basic" capture probabilities 0.1 to 0.55. These capture probabilities are modified for time variation and for behavioral response. The probability of first capture on any day is $p_i \times a_j$ ($a_1 = 1$, $a_2 = 0.7$, $a_3 = 1.3$, $a_4 = 1.4$, $a_5 = 0.6$, $a_6 = 1.2$). The probability of recapture on any given day is $c_i \times a_j \times 1.3$, $j = 2, ..., 6$.

TABLE 3.12. The eight models summarized by symbol, sources of variation in capture probabilities, and the associated estimator, if any. Program CAPTURE uses these names for the estimators.

Model	Sources Of Variation In Capture Probabilities	Appropriate Estimator
M_o	none	null
M_t	time	Darroch
M_b	behavior	Zippin
M_h	heterogeneity	jackknife
M_{tb}	time, behavior	none
M_{th}	time, heterogeneity	none
M_{bh}	behavior, heterogeneity	generalized removal
M_{tbh}	time, behavior, heterogeneity	none

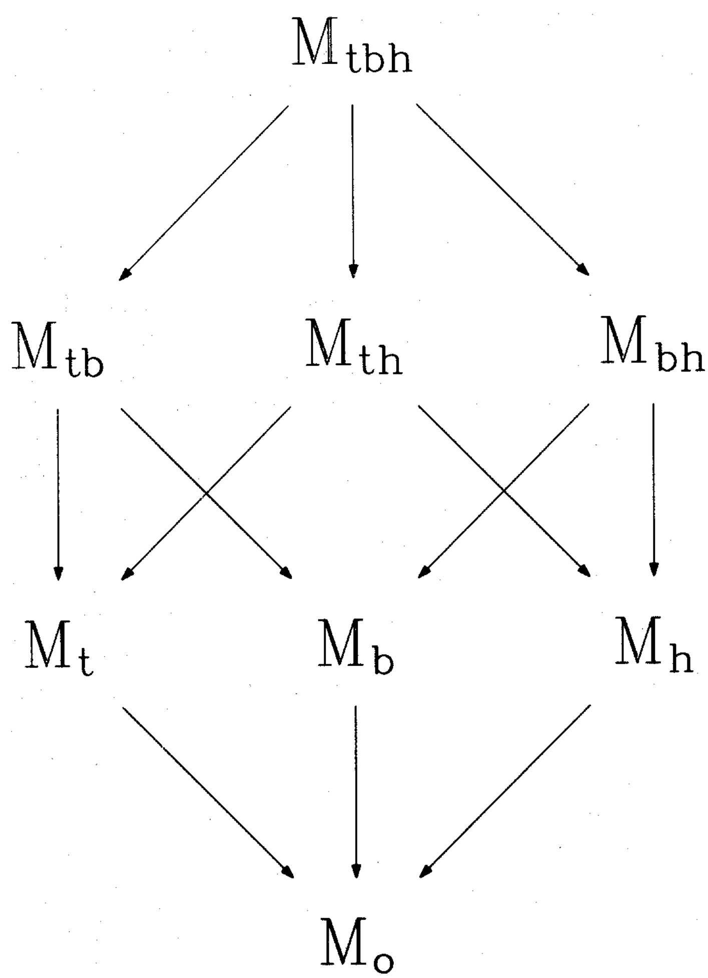


Fig. 3.16. Relations among the models. Each arrow points from one model to an immediate special case of that model. Tests of assumptions are based partly on these relations among the models.

Testing Model Assumptions and Model Selection

Overview. Our objective in developing these models and their associated estimators is to allow selection of the "best" estimator, of those we present, for any given set of capture-recapture data (assuming closure). We test the assumptions about capture probabilities by comparing the absolute and relative fits, to the data, of the various models, and then we select the simplest, best-fitting model. Unfortunately, three of the eight models do not have estimators. If one of these three models is selected as best, we must either forego estimation or continue to search for the simplest model that is relatively best-fitting and for which an estimator exists. Foregoing estimation is theoretically desirable, but usually unacceptable in practice. In this section we introduce the seven statistical tests of the assumptions on which model selection is based; they are summarized in Table 3.13.

Tests 1, 2, and 3 compare models to detect the presence or absence of heterogeneity, behavior, and time, respectively. Tests 4, 5, and 6 test the goodness of fit to the data of Models M_h, M_b, and M_t, respectively. Finally, test 7 compares Models M_h and M_{bh} to detect behavioral response in the presence of heterogeneity. There is an estimator for each of these models. The tests are illustrated in Fig. 3.17, which uses the simulation example of Model M_o described in Figs. 3.1 and 3.2, and Table 3.2. All seven test statistics have a chi-square distribution under the null hypothesis.

Model M_o was simulated with true $N = 50$ and $p = 0.3$. Figure 3.17 shows the summary statistics of the first simulation repetition, the seven tests of assumptions, and the model selection criteria, along with the suggested model and estimator. The model selection criteria represent an automated procedure implemented in program CAPTURE to suggest the appropriate model. It uses the results of all seven tests and generally is better than human judgment.

Test 1, Heterogeneity. If Model M_h provides a significantly better description of (a significantly better fit to) the data than Model M_o provides, we conclude that some form of heterogeneity is affecting

TABLE 3.13. Summary of the statistical tests of assumptions about capture probabilities, which form the basis of model selection. Test numbers are identical to those used by program CAPTURE.

Test Number	Test Purpose
1	Compares the relative fits of Models M_o and M_h to detemine whether there is evidence of heterogeneity in capture probabilities.
2	Compares the relative fits of Models M_o and M_b to determine whether there is evidence of behavioral effects on capture probabilities.
3	Compares the relative fits of Models M_o and M_t to determine whether there is evidence of time variation in capture probabilities.
4	Judges the goodness of fit of Model M_h; the test result is that Model M_h either fits or fails to fit the data.
5	Judges the goodness of fit of Model M_b; the test result is that Model M_b either fits or fails to fit the data.
6	Judges the goodness of fit of Model M_t; the test result is that Model M_t either fits or fails to fit the data.
7	Compares the relative fits of Models M_h and M_{bh} to determine whether there is evidence of behavioral response in the presence of heterogeneity.

```
OCCASION              J=     1    2    3    4    5    6
ANIMALS CAUGHT     N(J)=    16   11   15   14   14   18
TOTAL CAUGHT       M(J)=     0   16   23   31   40   44   47
NEWLY CAUGHT       U(J)=    16    7    8    9    4    3
FREQUENCIES        F(J)=    17   22    5    3    0    0


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =     3.524   DEGREES OF FREEDOM =   3   PROBABILITY OF LARGER VALUE =  0.31767


2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =     0.002   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.96170


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =     2.670   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.75075

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =     2.485   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.77878

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        -------------------------------------------------

              1                3.824     5     0.57509
              2                2.614     5     0.75929


5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =     6.433   DEGREES OF FREEDOM =   8   PROBABILITY OF LARGER VALUE =  0.59886

    5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =     3.183   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.52774

    5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =     3.250   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.51686

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        EXPECTED VALUES TOO SMALL.  TEST NOT PERFORMED.


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =     5.410   DEGREES OF FREEDOM =  11   PROBABILITY OF LARGER VALUE =  0.90970

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL      M(O)      M(H)      M(B)      M(BH)     M(T)      M(TH)     M(TB)     M(TBH)
CRITERIA   1.00      0.85      0.29      0.57      0.00      0.42      0.29      0.65


APPROPRIATE MODEL PROBABLY IS M(O)
SUGGESTED ESTIMATOR IS NULL.
```

Fig. 3.17. The seven tests of assumptions used in model selection applied to the simulated data from Model M_o with true N = 50 and p = 0.30. The underlying data are exactly the same as those used for Figs. 3.1 and 3.2 and given in Table 3.2.

capture probabilities. Program CAPTURE gives the chi-square test statistic used for test 1, its degrees of freedom (df), and the probability of a more extreme (larger) test statistic value if, in fact, the null hypothesis is true. Roughly stated, the null hypothesis for test 1 is that there is no heterogeneity of capture probabilities. Rigorously stated, the null hypothesis is that Model M_h does not provide a better description of the data than Model M_o provides. However, no matter which way the test result goes (in favor of M_o or of M_h), it provides no evidence that either model actually fits the data.

From Fig. 3.17, for test 1, the chi-square value is 3.524, with 3 df for these data. The probability P of a larger value is only 0.31767. The observed significance level is far from significant; hence we have no basis to reject the null hypothesis for test 1. The conclusion is that the simpler Model M_o is to be preferred over Model M_h for these data. Of course, we expected this result because Model M_o is the true model for this example.

Test 2, Behavioral Response. By testing the relative fit to the data of Model M_o versus Model M_b, we are testing whether behavioral variation in capture probabilities is likely. As with all of the tests, a chi-square test statistic is used. This test always has just 1 df because it is testing the assumption that first-capture probabilities p are equal to recapture probabilities c, given that either Model M_b or Model M_o is the true model.

From Fig. 3.17, for test 2, the chi-square value is 0.002 with an observed significance level P of 0.96170. That is, 96% of the time the test statistic value will be this large or larger if there is no behavioral response. If this test had rejected the null hypothesis (that Model M_o provides a better fit to the data than Model M_b), we would conclude that some form of behavioral response was affecting capture probabilities. In this example we conclude that Model M_o is to be preferred to Model M_b.

Test 3, Time Effects. By comparing the relative fit to the data of Model M_o versus Model M_t, we are testing for any time variation in capture probabilities. If $p_1, p_2, \ldots, p_t$ represent the average first-capture probabilities for the population on capture occasions 1, 2, . . ., t, then this test is testing the null hypothesis that $p_1 = p_2 = \ldots = p_t$, given that either Model M_t or Model M_o is the true model.

From Fig. 3.17, the chi-square value of test 3 is 2.670, with 5 df and an observed significance level of 0.75075. (Here, significance requires $P < 0.05$, or perhaps even $P < 0.01$.) Because this P value is far from significant, we conclude that, in a comparison of Models M_o and M_t, M_o fits the data just as well as M_t. As with tests 1 and 2, this is not an absolute test of the goodness of fit of Model M_o. To determine absolute goodness of fit, we use tests 4, 5, and 6.

Test 4, Goodness of Fit of Model M_h. The null hypothesis for test 4 is that Model M_h fits the data versus the alternative that Model M_h does not fit the data. From Fig. 3.17, the chi-square value of this test is 2.485 with 5 df. The P value of 0.77878 is not significant. We conclude that Model M_h adequately fits these data. Note that Model M_o is a special case of Model M_h (Fig. 3.16).

Test 4 can be applied to the capture data partitioned by frequency of capture. Although program CAPTURE gives these partitioned results, it is rarely necessary to look at them. The overall value of test 4 is all we need to examine. (Partitioned results in Fig. 3.17 are chi-square values of 3.824 and 2.614 for animals caught once and twice, respectively.)

Test 5, Goodness of Fit of Model M_b. The null hypothesis for test 5 is that Model M_b fits the data versus the alternative that Model M_b does not fit the data. This null hypothesis can be broken into two parts: (1) first-capture probabilities are constant over time, and (2) recapture probabilities are constant over time. The parts constitute tests 5a and 5b, respectively. The sum of the two chi-square test statistics gives the overall goodness of fit statistic for Model M_b.

From Fig. 3.17, the overall goodness of fit chi-square for Model M_b is 6.433 with $P = 0.59886$. The P value is not significant, as we would expect, given that Model M_o is the true model. Similarly, we judge first-capture probabilities and recapture probabilities to be adequately modeled as constant over time.

Test 6, Goodness of Fit of Model M_t. The null hypothesis for test 6 is that Model M_t fits the data. The alternative is that Model M_t does not fit the data. This test requires more data for computation than the other 6 tests require, and it sometimes cannot be computed for small numbers of captures. For the Model M_o simulation ($N = 50$), insufficient numbers of captures and recaptures were available to compute this test. However, Fig. 3.18 presents the same tests and the model selection criteria for the first simulation case of Model M_t. In this simulation $N = 150$, with $t = 5$ and an average capture probability of about 0.3. Under these conditions sufficient data were available. From Fig. 3.18, the chi-square goodness of fit statistic is 56.082 with 68 df; the observed significance level of 0.84862 is not significant. We conclude that Model M_t adequately fits these data.

Test 7, Behavioral Response Given Heterogeneity. Test 7, like tests 1, 2, and 3, compares the relative fits of two models, in this case, Models M_h and M_{bh}. The null hypothesis is that Model M_{bh} does not provide any better fit to the data than Model M_h provides. The alternative is that Model M_{bh} is a better fitting model for the data at hand than Model M_h. From Fig. 3.17, the chi-square value for test 7 is 5.410 with 11 df; the observed significance level of 0.90970 is not significant. We conclude that Model M_h provides an adequate model as opposed to Model M_{bh}. Again, the result is to be expected, given that the true model for these data is M_o.

Comment. From Fig. 3.17 we see that none of the seven null hypotheses were rejected in this example, where the true model is M_o. We expect this result when Model M_o is the correct model (a rare situation), or when the data are very poor (unfortunately a common situation). Conversely, any time that all seven tests are nonsignificant, the appropriate conclusion is that Model M_o best represents the data. We do not conclude that it is the true model, only that it is the best model to represent the given data.

The Model Selection Procedure

It is very difficult to evaluate the results of these seven tests subjectively (for example, on the basis of observed significance levels) and to decide on the most appropriate model. We need an objective, mathematical procedure (see *Otis et al. 1978:56-57).* Such a procedure has been developed and implemented in program CAPTURE: it is an application of multivariate discriminant function analysis (see *Otis et al. 1978:57-66).* The basic idea is that the pattern of observed significance levels, on the average, will be different for each of the eight models. If these patterns can be characterized, then a mathematical function can be constructed to classify an unknown set of data into the most likely pattern, and hence to judge which model would be most appropriate for those data. The details are very complex, because the "patterns" are really regions in a seven-dimensional space.

The model selection procedure involves computing a selection criterion (a number) for each model from the seven observed significance levels. Using a different mathematical function for each model, the computation gives seven "raw" selection criteria, whose absolute values are not important. The values are all nonnegative numbers. The most appropriate, or likely, model is the one corresponding to the maximum selection criterion. Therefore, CAPTURE finds this maximum value and divides it into all seven model selection criteria. As a result, one of the selection criteria (or more in the case of ties), is equal

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	28	53	50	60	37	
TOTAL CAUGHT	M(J)=	0	28	70	101	121	127
NEWLY CAUGHT	U(J)=	28	42	31	20	6	
FREQUENCIES	F(J)=	52	52	20	3	0	

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)

CHI-SQUARE VALUE = 1.460 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.69161

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)

CHI-SQUARE VALUE = 1.654 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.19842

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)

CHI-SQUARE VALUE = 23.479 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00010

4. GOODNESS OF FIT TEST OF MODEL M(H)
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)

CHI-SQUARE VALUE = 20.405 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00042

TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
(FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	10.885	4	0.02789
2	5.051	4	0.28208
3	11.000	4	0.02656

5. GOODNESS OF FIT TEST OF MODEL M(B)
NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)

CHI-SQUARE VALUE = 21.235 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00166

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

CHI-SQUARE VALUE = 14.827 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00197

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

CHI-SQUARE VALUE = 6.408 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.09336

6. GOODNESS OF FIT TEST OF MODEL M(T)
NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)

CHI-SQUARE VALUE = 56.082 DEGREES OF FREEDOM = 68 PROBABILITY OF LARGER VALUE = 0.84807

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)

CHI-SQUARE VALUE = 21.904 DEGREES OF FREEDOM = 10 PROBABILITY OF LARGER VALUE = 0.01567

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
CRITERIA	0.13	0.00	0.10	0.12	1.00	0.75	0.49	0.26

APPROPRIATE MODEL PROBABLY IS M(T)
SUGGESTED ESTIMATOR IS DARROCH.

Fig. 3.18. The seven tests of assumptions used in model selection applied to the simulated data from Model M_t with true N = 150. The underlying data are exactly the same as those used for Figs. 3.3 and 3.4. (See Table 3.4 and the discussion of the simulation of Model M_t for the capture probability parameters.)

to 1.00; that criterion corresponds to the most likely model. The other selection criteria lie between zero and 1.00. Roughly speaking, the higher the model selection criterion, the more likely that model is to be appropriate for the data. (However, these criteria, as we have constructed them, cannot be interpreted as probabilities in favor of the individual models.) These "normalized" model selection criteria are printed by CAPTURE just below the results of the seven tests of assumptions. The program also indicates the apparent appropriate model and the estimator based on that model.

In Fig. 3.17, the simulation case of Model M_o, the maximum selection criterion of 1.00 is for Model M_o, hence that is the indicated appropriate model for these data. Figures 3.18-3.25 give the results of the 7 tests and the model selection criteria for the first of 10 simulation runs of Models M_t through M_{tbh}. The simulation examples were used as the basis for illustrating the data and estimators (where one exists) for these models. The summary data are repeated as part of the new figures.

Consider Fig. 3.18, for which Model M_t is the true model. Test 1 does not reject; thus, in a choice between Models M_o and M_h, the simpler model is just as suitable as Model M_h. Because both models are false, there is in fact no reason to use the more complex Model M_h, so this failure to reject M_o is logical. Alternatively, we can interpret this result as showing no evidence of heterogeneity of capture probabilities. Similarly, test 2 provides no evidence of behavioral variation in capture probabilities. Test 3, however, strongly suggests the presence of time variation in the capture probabilities (chi-square = 23.479, 4 df, P = 0.0010).

Considering the three goodness of fit tests in Fig. 3.18, we see that neither Model M_h nor Model M_b fits these data (chi-square = 20.405, 4 df, P = 0.00042 and chi-square = 21.235, 6 df, P = 0.00166, respectively). From test 6, we do not reject the null hypothesis that Model M_t fits the data.

Finally, test 7 suggests some behavioral response (chi-square = 21.904, 10 df, P = 0.01560). Because Model M_t is the true model for these data, this is a type I error (rejection of a true null hypothesis). Any time multiple tests are made, we must expect to encounter some type I errors.

The pattern of observed significance levels in Fig. 3.18 strongly suggests that Model M_t is the best model for these data. This belief is corroborated by the results of the model selection procedure. From Fig. 3.18, the model criterion value is 1.00 for Model M_t, whereas the next highest value is only 0.75 for Model M_{th}, and the next is 0.49 for M_{tb}. None of the other criteria are greater than 0.26. Model M_t is clearly the most appropriate model for these data.

At this point, the reader should study Figs. 3.19-3.25 in detail, bearing in mind in each case what the true model is and observing the model selection criteria. In every case for these examples, the model selection procedure selected the correct model. Thus, the pattern of observed significance levels is fairly typical of what we can expect for each type of model, with reasonably good data.

For Fig. 3.19, the true model is M_b. Notice the suggestion of heterogeneity from test 1; it is appropriate in a sense. After the first capture day the population has two types of animals: those not captured and those previously captured. The two groups have different capture probabilities. Thus, a kind of heterogeneity is induced by behavioral response. It is not, of course, the kind we mean by the term "heterogeneity," but test 1 is sensitive to this kind of "heterogeneity" when, in fact, Model M_b is correct.

In Fig. 3.19, notice that test 3 clearly shows time variation in average daily capture probabilities. This inference is also correct, because the behavioral response to capture increases (Fig. 3.19) or decreases (Fig. 3.20) average daily capture probabilities. Because of effects like these, determination of the correct model just by casual examination of the seven test results is not easy and whether we could correctly judge Model M_b to be appropriate for these data is not clear. However, the model selection procedure clearly indicates Model M_b as the choice. The next closest model, M_{tb}, has a selection criterion of only 0.67; then M_{bh} has a criterion of 0.58. Notice that behavior enters both models.

For the trap-shy case in Fig. 3.20, Model M_b is selected as appropriate; it also fits the data in test 5. But in this example many of the remaining selection criteria are higher than in the Fig. 3.19 example (the trap-happy case). In part, this is because the trap-happy case generates more recaptures, hence more data. These two examples also illustrate that the selection criteria are only relative measures of the appropriate model.

```
OCCASION          J=      1   2   3   4   5   6   7
ANIMALS CAUGHT  N(J)=    19  28  44  37  44  49  53
TOTAL CAUGHT    M(J)=     0  19  36  60  71  79  83  91
NEWLY CAUGHT    U(J)=    19  17  24  11   8   4   8
FREQUENCIES     F(J)=    17  14  31  13  12   4   0


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =   13.877   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.01641


2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =   38.965   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.00000


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =   47.445   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.00000

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =   40.018   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.00000

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        -------------------------------------------------

               1              17.176     6     0.00866
               2              12.000     6     0.06197
               3              23.194     6     0.00073



5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =   12.468   DEGREES OF FREEDOM =  10   PROBABILITY OF LARGER VALUE =  0.25497

    5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =    9.269   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.09879

    5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =    3.198   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.66943

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        CHI-SQUARE VALUE =   23.572   DEGREES OF FREEDOM =  23   PROBABILITY OF LARGER VALUE =  0.42781


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =   74.257   DEGREES OF FREEDOM =  19   PROBABILITY OF LARGER VALUE =  0.00000

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL      M(O)      M(H)      M(B)      M(BH)     M(T)      M(TH)     M(TB)     M(TBH)
CRITERIA   0.22      0.10      1.00      0.58      0.00      0.31      0.67      0.31


APPROPRIATE MODEL PROBABLY IS M(B)
SUGGESTED ESTIMATOR IS ZIPPIN.
```

Fig. 3.19. The seven tests of assumptions used in model selection applied to the simulated data from Model M_b (the trap-happy case) with true N = 100. The underlying data are exactly the same as those used for Figs. 3.5 and 3.6. (See Table 3.6 and the discussion of the simulation of Model M_b, trap-happy case, for the capture probability parameters.)

```
OCCASION              J=     1    2    3    4    5    6    7
ANIMALS CAUGHT   N(J)=      36   34   25   23   18   25   24
TOTAL CAUGHT     M(J)=       0   36   62   76   87   92   93   94
NEWLY CAUGHT     U(J)=      36   26   14   11    5    1    1
FREQUENCIES      F(J)=      29   43   18    4    0    0    0


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =     5.419   DEGREES OF FREEDOM =   3   PROBABILITY OF LARGER VALUE =  0.14354


2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =    23.315   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.00000


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =    11.901   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.06422

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =    11.696   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.06911

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        -------------------------------------------------

              1               17.586      6     0.00735
              2                5.219      6     0.51610
              3                8.833      6     0.18317


5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =     9.164   DEGREES OF FREEDOM =  10   PROBABILITY OF LARGER VALUE =  0.51658

    5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =     3.324   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.65010

    5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =     5.840   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.32211

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        CHI-SQUARE VALUE =    53.857   DEGREES OF FREEDOM =  60   PROBABILITY OF LARGER VALUE =  0.69836


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =    22.630   DEGREES OF FREEDOM =  15   PROBABILITY OF LARGER VALUE =  0.09232

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL      M(O)     M(H)     M(B)     M(BH)     M(T)     M(TH)     M(TB)     M(TBH)
CRITERIA   0.59     0.52     1.00     0.94      0.00     0.47      0.71      0.72


APPROPRIATE MODEL PROBABLY IS M(B)
SUGGESTED ESTIMATOR IS ZIPPIN.
```

Fig. 3.20. The seven tests of assumptions used in model selection applied to the simulated data from Model M_b (the trap-shy case) with true $N = 100$. The underlying data are exactly the same as those used for Figs. 3.7 and 3.8. (See Table 3.7 and the discussion of the simulation of Model M_b for the capture probability parameters.)

```
OCCASION          J=      1    2    3    4    5    6    7
ANIMALS CAUGHT  N(J)=    65   68   60   68   67   48   67
TOTAL CAUGHT    M(J)=     0   65  107  133  149  162  167  174
NEWLY CAUGHT    U(J)=    65   42   26   16   13    5    7
FREQUENCIES     F(J)=    50   46   35   24   14    5    0


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =   24.084   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.00008


2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =    0.092   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.76154


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =    8.597   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.19756

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =    8.190   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.22448

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        -------------------------------------------------

              1                4.040     6     0.67126
              2                4.096     6     0.66373
              3                6.200     6     0.40116
              4                6.333     6     0.38690
              5                5.400     6     0.49362


5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =    9.636   DEGREES OF FREEDOM =  10   PROBABILITY OF LARGER VALUE =  0.47300

    5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =    2.273   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.81018

    5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =    7.363   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.19504

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        CHI-SQUARE VALUE =  167.216   DEGREES OF FREEDOM = 130   PROBABILITY OF LARGER VALUE =  0.01412


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =   14.741   DEGREES OF FREEDOM =  21   PROBABILITY OF LARGER VALUE =  0.83571

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL      M(O)     M(H)     M(B)     M(BH)    M(T)     M(TH)    M(TB)    M(TBH)
CRITERIA   0.74     1.00     0.22     0.52     0.00     0.43     0.26     0.56


APPROPRIATE MODEL PROBABLY IS M(H)
SUGGESTED ESTIMATOR IS JACKKNIFE.
```

Fig. 3.21. The seven tests of assumptions used in model selection applied to the simulated data from Model M_h, with true N = 200. The underlying data are exactly the same as those used for Figs. 3.9 and 3.10. (See Table 3.8 and the discussion of the simulation of Model M_h for the capture probability parameters.)

```
OCCASION           J=      1    2    3    4    5    6    7    8
ANIMALS CAUGHT  N(J)=     64   47   54   39   40   50   46   39
TOTAL CAUGHT    M(J)=      0   64   96  128  142  156  168  175  181
NEWLY CAUGHT    U(J)=     64   32   32   14   14   12    7    6
FREQUENCIES     F(J)=     74   50   30   21    5    1    0    0


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =   11.852   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.01849


2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =    8.073   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.00449


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =   13.570   DEGREES OF FREEDOM =   7   PROBABILITY OF LARGER VALUE =  0.05938

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =   14.735   DEGREES OF FREEDOM =   7   PROBABILITY OF LARGER VALUE =  0.03956

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        ------------------------------------------------

               1               9.459     7     0.22133
               2               9.520     7     0.21745
               3              17.609     7     0.01387
               4               5.667     7     0.57917


5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =    8.698   DEGREES OF FREEDOM =  12   PROBABILITY OF LARGER VALUE =  0.72846

   5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =    5.518   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.47924

   5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =    3.180   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.78595

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        CHI-SQUARE VALUE =  162.378   DEGREES OF FREEDOM = 125   PROBABILITY OF LARGER VALUE =  0.01251


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =   23.398   DEGREES OF FREEDOM =  21   PROBABILITY OF LARGER VALUE =  0.32314

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL       M(O)      M(H)      M(B)      M(BH)     M(T)      M(TH)     M(TB)     M(TBH)
CRITERIA    0.58      0.66      0.67      1.00      0.00      0.40      0.39      0.63


APPROPRIATE MODEL PROBABLY IS M(BH)
SUGGESTED ESTIMATOR IS GENERALIZED REMOVAL.
```

Fig. 3.22. The seven tests of assumptions used in model selection applied to the simulated data from Model M_{bh}, with true N = 200. The underlying data are exactly the same as those used for Figs. 3.11 and 3.12. (See Table 3.9 and the discussion of the simulation of Model M_{bh} for the capture probability parameters.)

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	73	43	91	45	57	
TOTAL CAUGHT	M(J)=	0	73	95	144	153	164
NEWLY CAUGHT	U(J)=	73	22	49	9	11	
FREQUENCIES	F(J)=	65	64	24	11	0	

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)

CHI-SQUARE VALUE = 3.960 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.26585

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)

CHI-SQUARE VALUE = 1.725 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.18903

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)

CHI-SQUARE VALUE = 39.771 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

4. GOODNESS OF FIT TEST OF MODEL M(H)
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)

CHI-SQUARE VALUE = 39.346 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
(FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	10.462	4	0.03333
2	25.063	4	0.00005
3	10.167	4	0.03771
4	9.455	4	0.05069

5. GOODNESS OF FIT TEST OF MODEL M(B)
NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)

CHI-SQUARE VALUE = 43.018 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00000

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

CHI-SQUARE VALUE = 32.652 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00000

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

CHI-SQUARE VALUE = 10.366 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.01570

6. GOODNESS OF FIT TEST OF MODEL M(T)
NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)

CHI-SQUARE VALUE = 117.857 DEGREES OF FREEDOM = 93 PROBABILITY OF LARGER VALUE = 0.03993

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)

CHI-SQUARE VALUE = 42.816 DEGREES OF FREEDOM = 10 PROBABILITY OF LARGER VALUE = 0.00001

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
CRITERIA	0.13	0.00	0.11	0.10	0.88	1.00	0.49	0.28

APPROPRIATE MODEL PROBABLY IS M(TH)
NO ESTIMATOR RESULTS FROM THIS MODEL.

Fig. 3.23. The seven tests of assumptions used in model selection applied to the simulated data from Model M_{th}, with true $N = 200$. The underlying data are exactly the same as those used for Fig. 3.13. (See Table 3.10 and the discussion of the simulation of Model M_{th} for the capture probability parameters.)

```
OCCASION          J=      1   2   3   4   5
ANIMALS CAUGHT  N(J)=    50  17  40  37  15
TOTAL CAUGHT    M(J)=     0  50  64  89 112 116
NEWLY CAUGHT    U(J)=    50  14  25  23   4
FREQUENCIES     F(J)=    79  31   6   0   0


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =    0.605   DEGREES OF FREEDOM =   2   PROBABILITY OF LARGER VALUE =  0.73897


2. TEST FOR BFHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTH"SIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =   12.829   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.00034


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =   39.776   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.00000

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =   34.454   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.00000

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        ------------------------------------------------

              1               25.494      4     0.00004
              2                9.591      4     0.04790


5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =   29.390   DEGREES OF FREEDOM =   6   PROBABILITY OF LARGER VALUE =  0.00005

    5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =   19.987   DEGREES OF FREEDOM =   3   PROBABILITY OF LARGER VALUE =  0.00017

    5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =    9.403   DEGREES OF FREEDOM =   3   PROBABILITY OF LARGER VALUE =  0.02439

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        CHI-SQUARE VALUE =   47.145   DEGREES OF FREEDOM =  49   PROBABILITY OF LARGER VALUE =  0.54860


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =   29.927   DEGREES OF FREEDOM =   8   PROBABILITY OF LARGER VALUE =  0.00022

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL      M(O)     M(H)     M(B)     M(BH)    M(T)     M(TH)    M(TB)    M(TBH)
CRITERIA   0.12     0.00     0.31     0.18     0.42     0.59     1.00     0.37


APPROPRIATE MODEL PROBABLY IS M(TB)
NO ESTIMATOR RESULTS FROM THIS MODEL.
```

Fig. 3.24. The seven tests of assumptions used in model selection applied to the simulated data from Model M_{tb}, with true $N = 150$. The underlying data are exactly the same as those used for Fig. 3.14. (See Table 3.11 and the discussion of the simulation of Model M_{tb} for the capture probability parameters.)

```
OCCASION              J=      1    2    3    4    5    6
ANIMALS CAUGHT  N(J)=    63   46  100  102   38   96
TOTAL CAUGHT    M(J)=     0   63   86  127  148  152  166
NEWLY CAUGHT    U(J)=    63   23   41   21    4   14
FREQUENCIES     F(J)=    36   45   42   24   17    2


1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(H)

        CHI-SQUARE VALUE =   14.763   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.00522


2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(B)

        CHI-SQUARE VALUE =   18.741   DEGREES OF FREEDOM =   1   PROBABILITY OF LARGER VALUE =  0.00002


3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.
   NULL HYPOTHESIS OF MODEL  M(O)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(T)

        CHI-SQUARE VALUE =  116.194   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.00000

4. GOODNESS OF FIT TEST OF MODEL  M(H)
   NULL HYPOTHESIS OF MODEL  M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(H)

        CHI-SQUARE VALUE =  104.054   DEGREES OF FREEDOM =   5   PROBABILITY OF LARGER VALUE =  0.00000

        TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE
        (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

        NUMBER OF CAPTURES  CHI-SQUARE  D.F.  PROBABILITY
        ------------------------------------------------

              1               16.333      5     0.00596
              2               24.500      5     0.00018
              3               50.794      5     0.00000
              4               18.125      5     0.00280
              5               18.647      5     0.00224


5. GOODNESS OF FIT TEST OF MODEL  M(B)
   NULL HYPOTHESIS OF MODEL  M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(B)

        CHI-SQUARE VALUE =   93.254   DEGREES OF FREEDOM =   8   PROBABILITY OF LARGER VALUE =  0.00000

    5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME

        CHI-SQUARE VALUE =   23.418   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.00010

    5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME

        CHI-SQUARE VALUE =   69.836   DEGREES OF FREEDOM =   4   PROBABILITY OF LARGER VALUE =  0.00000

6. GOODNESS OF FIT TEST OF MODEL  M(T)
   NULL HYPOTHESIS OF MODEL  M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL  M(T)

        CHI-SQUARE VALUE =  176.095   DEGREES OF FREEDOM = 124   PROBABILITY OF LARGER VALUE =  0.00114


7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.
   NULL HYPOTHESIS OF MODEL  M(H)  VS. ALTERNATE HYPOTHESIS OF MODEL  M(BH)

        CHI-SQUARE VALUE =   50.998   DEGREES OF FREEDOM =  15   PROBABILITY OF LARGER VALUE =  0.00001

MODEL SELECTION CRITERIA.  MODEL SELECTED HAS MAXIMUM VALUE.

MODEL      M(O)    M(H)    M(B)    M(BH)   M(T)    M(TH)   M(TB)   M(TBH)
CRITERIA   0.47    0.38    0.23    0.59    0.00    0.46    0.53    1.00


APPROPRIATE MODEL PROBABLY IS M(TBH)
NO ESTIMATOR RESULTS FROM THIS MODEL.
```

Fig. 3.25. The seven tests of assumptions used in model selection applied to the simulated data from Model M_{tbh}, with true N = 200. The underlying data are exactly the same as those used for Fig. 3.15. (See the discussion of the simulation of Model M_{tbh} for the capture probability parameters.)

Contrast the results in Fig. 3.25 (Model M_{tbh}) with those in Fig. 3.17 (Model M_o). In Fig. 3.25 each of the seven tests clearly rejects the null hypothesis. The only possible conclusion is that all three factors (time, behavior, and heterogeneity) must be present in these data. Thus Model M_{tbh} is the appropriate model. How, then, can we estimate N? There is no satisfactory solution. In the example of Fig. 3.25 we look for the model with the next highest selection criterion; here, it is M_{bh}. Because there is an estimator for M_{bh}, we can use that model as a basis for estimation. If, however, the selected generalized removal estimator corresponds to a model that does not fit, we can place little confidence in the estimator. (Goodness of fit tests are given along with the generalized removal estimator for each submodel of M_{bh} examined.) Even if the model does fit, we cannot be very confident that the estimator is unbiased, because the model initially selected was Model M_{tbh}.

Consider the example in Fig. 3.23; here the true model is M_{th}. To compute an estimator from these data, however, we would be led to use Model M_t; it has the next highest selection criterion (0.88). Having made that decision, we would look at the goodness of fit test results for Model M_t (test 6). The chi-square test value of 117.857 (93 df) has an observed significance level of 0.042; therefore, we would conclude that Model M_t does not fit these data, or at least we would be very suspicious of it as an adequate model. In fact, the estimate of N based on Model M_t cannot be considered very reliable for these data; it is liable to be biased and the estimated sampling variance will be too small.

A Comprehensive Look at the Simulation Examples

Throughout this chapter we have used simulation examples to illustrate the models, estimators, and model selection method. So far, only the first of 10 simulations has been presented for each model. In this section we present summary results on estimation and model selection for all 10 repetitions for each model. The summary results illustrate some common features of these methods. The general statements about the properties of the estimators are based on the theoretical and simulation results in *Otis et al. (1978)* and on practical experience with these methods.

Robustness of the Different Estimators. Each of the five estimators was derived under a different model (see Table 3.12). When the correct model is assumed for the capture-recapture data being analyzed, the given estimator performs well. That is, it has small bias, and estimated sampling variances are also relatively unbiased. When the wrong model is assumed, the computed estimator is generally biased, often badly so, and the estimate of its sampling variance is unreliable. In the worst case, the use of too simple a model, like Model M_o, can lead to a very biased estimate with a severe underestimate of its sampling variance.

Table 3.14 presents the average values of all five estimators, for all 10 repetitions of the 8 models. (Two cases of Model M_b were simulated.) The values in parentheses are the observed standard errors of these means. For Model M_o with the null estimator, for example, there were 10 independent values of $\hat{N}_o$. The sampling variance among these 10 values was computed. Dividing that estimated variance by 10 (the sample size) and taking its square root gave 1.1 as the standard error of the mean, 49.1, of these 10 values.

From Table 3.14, we can see that any given estimator does well when used with the true model. For example, when Model M_o was true ($N = 50$), the null estimator averaged 49.1 (1.1), but when Model M_{tbh} was true ($N = 200$), the null estimator averaged 176.2 (1.0), which is clearly biased. In contrast, all the estimators ought to give reasonable results when the true model is M_o. The first row of Table 3.14 shows that this is indeed the case. The reader should examine all of the table, bearing in mind that these results are not representative of all possible cases: high average capture probabilities (around 0.3) were used in the simulations, hence estimators are more reliable when used with the wrong model than would be the case with lower average capture probabilities, (around 0.2 or 0.15).

TABLE 3.14. Average value of each estimator, over the 10 simulations for each model. Numbers in parentheses are the standard errors of these averages, based on the 10 replications. The standard errors provide a basis for judging the degree of bias of the estimator. The capture probability parameters for each model have been described in the text for that model.

		Estimator (Model) Used				
True Model	True N	Null (M_o)	Darroch (M_t)	Zippin (M_b)	Jackknife (M_h)	Generalized Removal
M_o	50	49.1(1.1)	49.0(1.0)	48.5(1.1)	53.4(0.9)	49.7(1.8)
M_t	150	147.1(1.7)	146.0(1.7)	155.8(2.4)	168.1(3.0)	136.7(2.7)
M_b [a]	100	86.9(1.6)	86.8(1.6)	102.4(3.9)	94.5(2.4)	105.0(4.6)
M_b [b]	100	114.2(1.2)	113.3(1.1)	99.2(0.8)	129.1(3.4)	100.0(1.2)
M_h	200	180.4(2.3)	180.4(2.3)	182.2(2.5)	209.4(2.9)	182.3(2.5)
M_{bh}	200	201.3(2.4)	200.7(2.4)	186.8(2.1)	239.4(5.0)	192.6(7.0)
M_{th}	200	182.2(1.9)	178.8(1.9)	178.9(1.7)	216.8(3.2)	161.3(1.3)
M_{tb}	150	212.8(7.3)	208.1(7.1)	160.5(5.3)	252.3(5.3)	131.9(2.4)
M_{tbh}	200	176.2(1.0)	174.8(0.9)	190.3(2.0)	201.0(2.1)	181.9(4.4)

[a] Trap-happy case.
[b] Trap-shy case.

In terms of their robustness we rank the five estimators in the order $\hat{N}_h$ (jackknife, Model M_h); $\hat{N}_{bh}$ (generalized removal, Model M_{bh}); $\hat{N}_b$ (Zippin, Model M_b); $\hat{N}_t$ (Darroch, Model M_t); and $\hat{N}_o$ (null, Model M_o). Darroch's estimator is always valid when Model M_o is true; moreover, very little precision is lost by using $\hat{N}_t$ when Model M_o is indicated. Also, $\hat{N}_b$ is a special case of $\hat{N}_{bh}$. Thus, we can reduce the choices to three estimators: $\hat{N}_h$, $\hat{N}_{bh}$, and $\hat{N}_t$. Certainly, when the selection procedure suggests Model M_{tb}, M_{th}, or M_{tbh}, and estimation is necessary, only one of these three estimators should be considered for use. Of the three, $\hat{N}_h$ will generally be the best choice, although the particular set of data, circumstances, and model selection criteria may cause one to select either $\hat{N}_{bh}$ or $\hat{N}_t$.

Model Selection and Estimation. In practice, we attempt to select the correct model before estimating N. Thus, the real test of the methods we recommend first involves model selection, then estimation based on the most appropriate model and concurrent evaluation of whether that model fits the data. These results, for each simulated case, are shown in Table 3.15. For example, the model selection procedure correctly selected Model M_o in 8 of the 10 simulation cases. In the other two cases, Model M_h was used as the basis for estimation, even though M_{th} was chosen once. Results of point estimation, confidence interval coverage, and model goodness of fit were all very good when Model M_o was the true model. (A significance level of 5% was used to judge model fit.)

Results from Table 3.15 are also good for Models M_t, M_b, and M_h. The model selection procedure in these examples generally led to use of the correct model. For Model M_t, the correct model was selected only 7 (of 10) times, but when Model M_{th} was selected, the selection criteria clearly indicated that Model M_t (not M_h) should be used for estimation.

Results for Model M_{bh} are good, but not as good as for the four simple (one-factor) models (M_o, M_t, M_b, and M_h). In general, the results when Model M_{bh} is true tend to underestimate N; the presence of heterogeneity "causes" this tendency toward a negative bias.

For Models M_{th}, M_{tb}, and M_{tbh}, estimation is clearly inferior to estimation under the other models. For example, for Model M_{th}, the correct model was selected only 4 (of 10) times. This is not critical from an estimation viewpoint because there is no estimator for Model M_{th}. However, it is clearly misleading that 6

TABLE 3.15. Model selection and estimation results for the 10 simulations of each model. "Model selected" means the model recommended by the model selection procedure of program CAPTURE. "Coverage" shows whether the computed 95% confidence interval on N included the true value of N. When the selected model had no estimator, we examined the model selection results and chose the apparent "best" model that had an estimator. "Goodness of fit" shows whether the model used as a basis for estimation fit the data.

Model M_o, $N = 50$

Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_o	55	Yes	M_o	Yes
2	M_o	55	Yes	M_o	Yes
3	M_o	50	Yes	M_o	Yes
4	M_o	49	Yes	M_o	Yes
5	M_o	49	Yes	M_o	Yes
6	M_o	45	Yes	M_o	Yes
7	M_h	52	Yes	M_h	Yes
8	M_o	48	Yes	M_o	Yes
9	M_o	48	Yes	M_o	Yes
10	M_{th}	50	Yes	M_h	Yes
Mean		50.1			

TABLE 3.15. (cont)

Model M_t, $N = 150$

Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_t	151	Yes	M_t	Yes
2	M_{th}	146	Yes	M_t	Yes
3	M_t	150	Yes	M_t	Yes
4	M_o	144	Yes	M_o	Yes
5	M_t	153	Yes	M_t	Yes
6	M_{th}	138	Yes	M_t	Yes
7	M_t	145	Yes	M_t	Yes
8	M_t	138	Yes	M_t	Yes
9	M_t	152	Yes	M_t	Yes
10	M_t	144	Yes	M_t	Yes
Mean		146.1			

TABLE 3.15. (cont)

Model M_b, (trap-happy case), $N = 100$

Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_b	114	Yes	M_b	Yes
2	M_b	97	Yes	M_b	No
3	M_b	85	No	M_b	Yes
4	M_b	112	Yes	M_b	Yes
5	M_b	98	Yes	M_b	Yes
6	M_b	105	Yes	M_b	Yes
7	M_b	83	No	M_b	Yes
8	M_b	97	Yes	M_b	Yes
9	M_b	114	Yes	M_b	Yes
10	M_b	119	Yes	M_b	Yes
Mean		102.4			

TABLE 3.15. (cont)

Model M_b, (trap-shy case), $N = 100$

Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_b	99	Yes	M_b	Yes
2	M_b	101	Yes	M_b	Yes
3	M_b	99	Yes	M_b	Yes
4	M_{tb}	103	Yes	M_{bh}	Yes
5	M_b	96	Yes	M_b	Yes
6	M_b	99	Yes	M_b	Yes
7	M_b	98	Yes	M_b	Yes
8	M_b	96	Yes	M_b	Yes
9	M_b	99	Yes	M_b	Yes
10	M_{bh}	104	Yes	M_{bh}	Yes
Mean		99.4			

TABLE 3.15. (cont)

Model M_h, N = 200					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_h	215	Yes	M_h	Yes
2	M_h	211	Yes	M_h	Yes
3	M_h	207	Yes	M_h	Yes
4	M_h	198	Yes	M_h	Yes
5	M_{bh}	181	No	M_{bh}	Yes
6	M_h	203	Yes	M_h	Yes
7	M_h	215	Yes	M_h	Yes
8	M_h	197	Yes	M_h	Yes
9	M_h	209	Yes	M_h	Yes
10	M_h	229	No	M_h	Yes
Mean		206.5			

TABLE 3.15. (cont)

Model M_{bh}, (trap-shy case), N = 200					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_{bh}	192	Yes	M_{bh}	Yes
2	M_h	219	Yes	M_h	Yes
3	M_b	191	Yes	M_b	Yes
4	M_{bh}	179	No	M_{bh}	Yes
5	M_b	181	No	M_b	Yes
6	M_b	186	No	M_b	Yes
7	M_b	182	No	M_b	Yes
8	M_{bh}	188	No	M_{bh}	Yes
9	M_{bh}	189	No	M_{bh}	Yes
10	M_{bh}	254	Yes	M_{bh}	Yes
Mean		196.1			

TABLE 3.15. (cont)

Model M_{th}, N = 200

Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_t	177	No	M_t	Yes
2	M_t	179	No	M_t	Yes
3	M_t	179	No	M_t	Yes
4	M_{th}	188	Yes	M_t	No
5	M_{th}	176	No	M_t	Yes
6	M_{th}	175	No	M_t	Yes
7	M_t	170	No	M_t	Yes
8	M_t	176	No	M_t	Yes
9	M_{th}	190	Yes	M_t	Yes
10	M_t	178	No	M_t	Yes
Mean		178.8			

TABLE 3.15. (cont)

Model M_{tb}, N = 150

Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_t	207	No	M_t	Yes
2	M_t	208	No	M_t	Yes
3	M_{tb}	192	No	M_t	Yes
4	M_{tb}	221	No	M_t	Yes
5	M_{tb}	145	Yes	M_b	Yes
6	M_{th}	207	No	M_t	Yes
7	M_{tb}	149	Yes	M_b	No
8	M_{th}	182	Yes	M_t	Yes
9	M_{tb}	236	No	M_t	Yes
10	M_{th}	182	No	M_t	Yes
Mean		192.8			

TABLE 3.15. (cont)

Model M_{tbh}, N = 200					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	M_{tbh}	186	Yes	M_{bh}	No
2	M_{tbh}	174	No	M_{bh}	No
3	M_{tbh}	174	No	M_{bh}	No
4	M_{tbh}	179	No	M_{bh}	No
5	M_{tb}	175	No	M_{bh}	No
6	M_{tbh}	177	No	M_{bh}	No
7	M_{tbh}	169	No	M_{bh}	Yes
8	M_{tbh}	216	Yes	M_{bh}	No
9	M_{tbh}	194	Yes	M_{bh}	No
10	M_{tbh}	175	No	M_{bh}	Yes
Mean		181.8			

of 10 times Model M_t was selected and that Model M_t was not rejected by the goodness of fit test. This is just one example; the model selection procedure often will do better, even for Model M_{th}.

These results illustrate a general truth: the goodness of fit test for Model M_t has low power. That is, even when Model M_t is false, this goodness of fit test does not have a large probability of rejecting Model M_t. (This is a type II error: failure to reject a false null hypothesis). The low power of the Model M_t goodness of fit test must be kept in mind when the adequacy of Model M_t is judged, especially when the selection criteria suggest a model such as M_{th}, M_{tb}, or M_{tbh}. Then it will often be best, if an estimator must be computed, to take Model M_h or M_{bh} rather than M_t, unless M_t has a selection criterion very close to 1.

Testing for Closure

We have emphasized the need to make assumptions explicit and to test those assumptions. All of the models we have presented assume closure. Thus, it is natural and appropriate to want a statistical test of the closure assumption. Unfortunately, such a test is impossible. The problem is that true failure of closure cannot be distinguished from behavioral changes in capture probabilities or from certain patterns of time-varying capture probabilities. Thus, only when we assume that either Model M_h or Model M_o is the underlying model (the null hypothesis), can we test for closure. *Pollock et al. (1974)* present a test for closure assuming Model M_o is the null hypothesis. *Otis et al. (1978:66-87, 120-121)* present a test for closure assuming Model M_h is the null hypothesis. We emphasize that neither test is valid if closure is true but a different model holds, such as M_b or M_{tbh}. Program CAPTURE however, computes the closure test assuming Model M_h is the null hypothesis.

For all simulation examples in this chapter, the closure test in CAPTURE was performed. That test is one of the first items presented by CAPTURE. The results for the example data case of simulating Model M_o (wherein $p = 0.3$) are shown in Fig. 3.26. The closure test statistic has a standard normal distribution under the null hypothesis, which is Model M_h and closure. As expected under Model M_o, the closure test does not reject the null hypothesis: $z = -0.328$ and $P = 0.37145$. A partitioned version of this closure test is computed for subsets of the data defined by frequencies of capture. However, both the partitioned version and the overall closure test should be ignored because of serious problems in their interpretation.

The results of the closure test for the first simulation repetition of Model M_b, the trap-happy case ($p = 0.25$, $c = 0.55$), are given in Fig. 3.27. The test rejects the null hypothesis of population closure: $z =$

```
OVERALL TEST RESULTS --
Z-VALUE                                   -.328
PROBABILITY OF A SMALLER VALUE          0.37145


TEST OF CLOSURE BY FREQUENCY OF CAPTURE.
(FREQUENCIES LESS THAN 10 ARE NOT COMPUTED.)

NUMBER OF CAPTURES  Z-VALUE  PROBABILITY
----------------------------------------

        2            0.285      0.61214
```

Fig. 3.26. Results of the closure test for the first simulation case of Model M_o. Under the null hypothesis of either Model M_o or M_h (only) and closure, the computed z-value has a standard normal distribution. The program gives the observed significance level of the test, P = 0.37145 in this example.

```
OVERALL TEST RESULTS --
Z-VALUE                                  -3.010
PROBABILITY OF A SMALLER VALUE          0.00131

TEST OF CLOSURE BY FREQUENCY OF CAPTURE.
(FREQUENCIES LESS THAN 10 ARE NOT COMPUTED.)

NUMBER OF CAPTURES  Z-VALUE  PROBABILITY
----------------------------------------

        2           -1.494      0.06758
        3           -1.988      0.02341
        4           -1.246      0.10647
        5           -1.215      0.11216
```

Fig. 3.27. Results of the closure test for the first simulation case of Model M_b, trap-happy case. Closure is true, but the test cannot distinguish failure of closure from behavioral change in capture probabilities. The test rejects the null hypothesis (P = 0.00131) because of this behavioral response.

−3.01 and P = 0.00131. In fact the population is closed, but the test is "reacting" to the behavioral change in capture probabilities, which "looks" like recruitment.

In all, for examples in this chapter, we simulated 10 repetitions of 9 models (two cases of Model M_b). For Models M_o and M_h, the closure test did not reject the model even once at the 5% significance level. This result is not strange, given that only 20 tests were made and that the test is valid for these two models. But for the 70 remaining simulations, the closure test rejected the model 23 times at the 5% significance level; of course, closure was really true in all these cases. The number of rejections corresponds to a 33% rejection rate, when it should be 5%, and illustrates that this closure test is invalid whenever time or behavior affects capture probabilities. We emphasize that this problem is fundamental; no valid statistical test of closure can be constructed on the basis of only the capture-recapture data.

Summary

1. There are three critical considerations in constructing capture-recapture models: what population size N means (this relates to geographic closure), whether the model should be demographically closed or open, and how to model capture probabilities.

2. Ball and urn models have motivated most of the thinking about capture-recapture models. However, in real populations, capture probabilities vary and there is no analogy to the sides of the urn.

3. We have not dealt with the case of t = 2 because we cannot test any assumptions in this case, and it is covered adequately in the literature.

4. Models are based on the concept of capture probabilities: p_{ij} = the probability of capturing the i^{th} animal (in the population at risk of capture) on the j^{th} sampling occasion.

5. Three factors can affect capture probabilities: time effects, behavioral response to capture, and innate heterogeneity (that is, variations among individuals in capture probabilities). On the basis of the three factors, eight different basic models for closed-population capture-recapture studies are presented (see

Table 3.12 and Fig. 3.16). The reader should be able to name these models and describe their nature before proceeding to the next chapters.

6. The biological literature clearly shows that heterogeneity and behavioral effects on capture probabilities are common.

7. Five of the eight models have associated estimators; only the jackknife estimator, which was derived originally for Model M_h, exists as a simple algebraic formula. All other estimators require a numerical computer solution of complicated equations. The reader should be able to name the estimators and their associated models and should know which three models have no estimator.

8. When Model M_b is true, there will be time variation in average daily capture probabilities, and there will be a type of heterogeneity of capture probabilities after day 1. This will cause difficulty in selecting the correct model.

9. Seven tests of assumptions about capture probabilities are presented in Table 3.13. A mathematical model selection procedure based on the tests produces eight normalized selection criteria. At least one criterion will be equal to 1, thereby indicating the appropriate model for the data.

10. If the selected model has no estimator, then there is apparently no valid estimator for those data. One can, however, select the most appropriate remaining model that has an estimator and use it to estimate N. The result is likely to be a biased estimate.

11. The estimator for Model M_h (the jackknife, $\hat{N}_h$) is the most robust of the five estimators, followed by $\hat{N}_{bh}$, $\hat{N}_b$, $\hat{N}_t$, and $\hat{N}_o$. Also, in practice $\hat{N}_t$ always can be used, rather than $\hat{N}_o$.

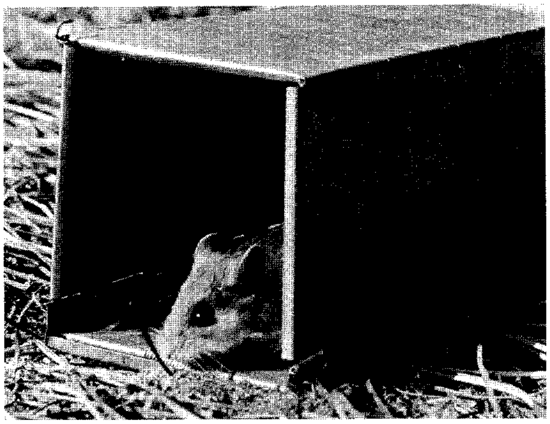
12. The full-blown procedure of doing all seven tests, and then selecting the model leads to a robust estimation procedure for good capture-recapture data on closed populations.

13. A valid test of closure cannot be devised because behavioral responses and time trends in capture probabilities cannot be distinguished from failure of closure.

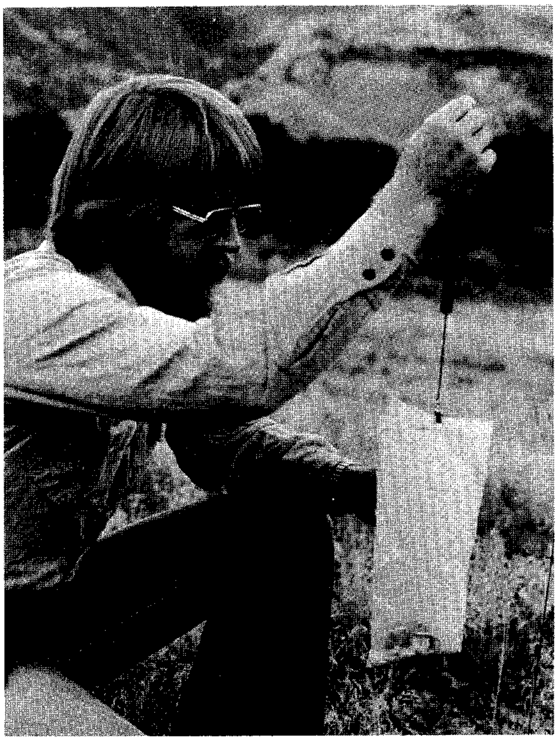
Questions and Exercises

1. Is equal probability of capture usually attainable in field studies, if enough traps are available and trapping is done for at least 4 days?
2. Are most capture-recapture and removal models used to estimate N sensitive (not robust) to failure of assumptions regarding capture probabilities?
3. If you get no recaptures in a live-trapping study, can you estimate N using a capture-recapture method?
4. Are capture-recapture methods useful for very small populations such as condors or whooping cranes?
5. Can N be estimated if the study is conducted for only one trapping occasion; that is, if $t = 1$?
6. Are testing and model selection possible if $t = 2$?
7. If you mark animals by trapping and "recapture" them by hunting ($t = 2$), what model is likely to apply? Is this a reasonable type of study?
8. The daily capture probability of an individual animal may be related to which of the following: home range size, social dominance, innate activeness of the animal, trap spacing, or number of days of trapping?

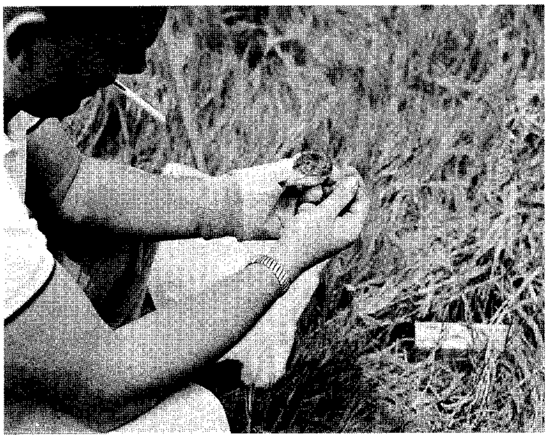
9. If you trap within the same fixed area with 100 traps on days 1, 2, and 3, then with 150 traps on days 4, 5, and 6, and finally with 200 traps on days 7, 8, and 9, can Model M_o fit the data? Can Model M_h, M_b, or M_t fit the data?
10. The capture probability p (say, in Model M_o) is which of the following?
    a. The probability that a trap will catch an animal,
    b. The probability that an individual animal will be caught on a given trapping occasion,
    c. The probability that an individual animal will be caught at least once during the study.
11. Will ML estimators of N be developed for Models M_{th}, M_{tb}, and M_{tbh}?
12. Is Model M_o robust in trap-shy populations?
13. If the model selection criteria are the same for Models M_o and M_h, why do we recommend selecting Model M_h as the basis for estimating N?
14. Is it just a matter of time until a general, completely valid test for closure is developed?



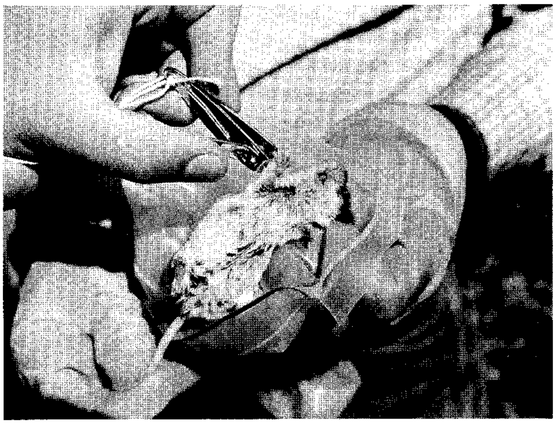
This deer mouse *(Peromyscus maniculatus)* is about to be captured in a Sherman live trap. (Photograph by Gary White.)



Plastic bags provide an easy method to handle small mammals, as long as they are not left in the bag too long. The animal is dumped from the trap directly into the bag. (Photograph by Gary White.)



House mice *(Mus musculus)* should be handled with gloves to protect the investigator and the animal. Improper handling of the captured animals can change recapture probabilities due to injuries or behavioral response. (Photograph courtesy of Harry Coulombe.)



Nail clippers are convenient for marking the toes of small mammals. Note how the animal has been grasped by the scruff of the neck through the plastic bag. (Photograph courtesy of David McInroy.)