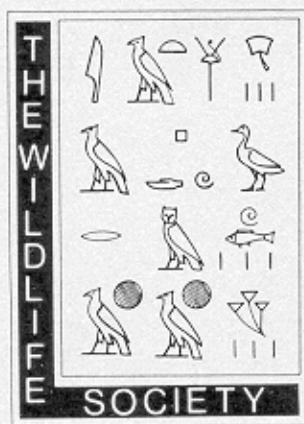


WILDLIFE MONOGRAPHS

A Publication of The Wildlife Society



STATISTICAL INFERENCE FROM
CAPTURE DATA ON CLOSED
ANIMAL POPULATIONS

by

DAVID L. OTIS, KENNETH P. BURNHAM,
GARY C. WHITE, AND DAVID R. ANDERSON

OCTOBER 1978

No. 62

WILDLIFE MONOGRAPHS

Louis A. Krumholz, *Editor*

Office of Academic Affairs, University of Louisville,
Louisville, Kentucky 40208

Consulting Editors for this Issue: A. Neil Arnason, The University of
Manitoba, Winnipeg, Manitoba, Canada
R3T 2N2

George A. F. Seber, The University of
Auckland, Private Bag, Auckland,
New Zealand

The Wildlife Society believes that increased awareness and appreciation of wildlife values is an important objective. Society publications are one means of doing this.

Wildlife Monographs was begun in 1957 to provide for longer papers than those normally accepted for *The Journal of Wildlife Management*. There is no set schedule for publication.

Each *Monograph* is sponsored financially by organizations or institutions interested in publication of the information contained therein. Usually, the sponsor is the organization that conducted the original research, but others interested in disseminating the information may assist in defraying *Monograph* costs. The sponsors pay for printing and distribution of each *Monograph*, and *The Wildlife Society* provides skilled editors to assist *Monograph* authors, and assures wide distribution through its worldwide mailing list to a select group of knowledgeable wildlife scientists, libraries, and others, and to members and subscribers who receive *The Journal of Wildlife Management*.

There is a perpetual need for additional funds to sponsor publication of worthwhile manuscripts in *Wildlife Monographs*. Any contribution will be accepted with gratitude by The Wildlife Society. Memorial funds collected to honor and perpetuate the names of deceased members of the profession probably could be put to no better use.

Individual issues of *Wildlife Monographs* will be brought out as suitable manuscripts are accepted and processed, and when financing has been arranged.

© The Wildlife Society, Inc. 1978

**The research reported here was done with
financial support through
COAL PROGRAM, OFFICE OF BIOLOGICAL SERVICES,
U.S. FISH AND WILDLIFE SERVICE;
with funds made available through
FEDERAL INTERAGENCY ENERGY/ENVIRONMENTAL
RESEARCH AND DEVELOPMENT PROGRAM,
U.S. ENVIRONMENTAL PROTECTION AGENCY**

**Publication costs were defrayed by
OIL SHALE PROGRAM, OFFICE OF BIOLOGICAL SERVICES,
U.S. FISH AND WILDLIFE SERVICE**

WILDLIFE MONOGRAPHS

Louis A. Krumholz, *Editor*

**Office of Academic Affairs, University of Louisville,
Louisville, Kentucky 40208**

***Consulting Editors for this Issue:* A. Neil Arnason, The University of
Manitoba, Winnipeg, Manitoba, Canada
R3T 2N2**

**George A. F. Seber, The University of
Aukland, Private Bag, Aukland,
New Zealand**

The Wildlife Society believes that increased awareness and appreciation of wildlife values is an important objective. Society publications are one means of doing this.

Wildlife Monographs was begun in 1957 to provide for longer papers than those normally accepted for *The Journal of Wildlife Management*. There is no set schedule for publication.

Each *Monograph* is sponsored financially by organizations or institutions interested in publication of the information contained therein. Usually, the sponsor is the organization that conducted the original research, but others interested in disseminating the information may assist in defraying *Monograph* costs. The sponsors pay for printing and distribution of each *Monograph*, and *The Wildlife Society* provides skilled editors to assist *Monograph* authors, and assures wide distribution through its worldwide mailing list to a select group of knowledgeable wildlife scientists, libraries, and others, and to members and subscribers who receive *The Journal of Wildlife Management*.

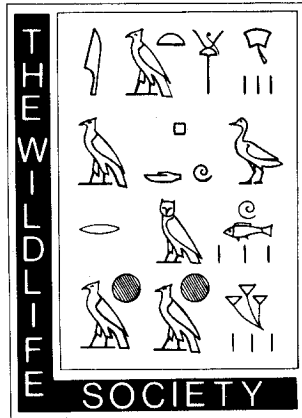
There is a perpetual need for additional funds to sponsor publication of worthwhile manuscripts in *Wildlife Monographs*. Any contribution will be accepted with gratitude by The Wildlife Society. Memorial funds collected to honor and perpetuate the names of deceased members of the profession probably could be put to no better use.

Individual issues of *Wildlife Monographs* will be brought out as suitable manuscripts are accepted and processed, and when financing has been arranged.

© The Wildlife Society, Inc. 1978

WILDLIFE MONOGRAPHS

A Publication of The Wildlife Society



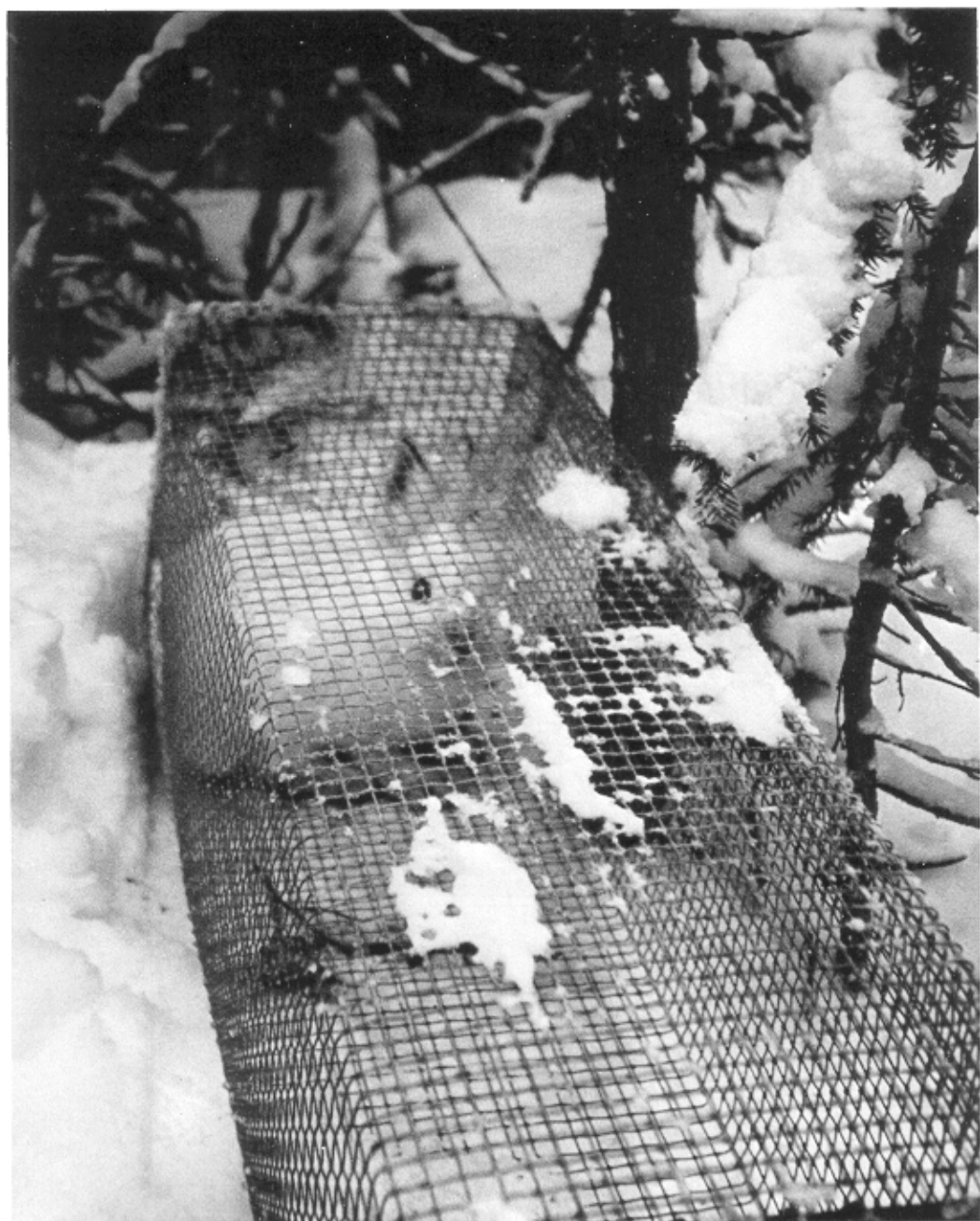
STATISTICAL INFERENCE FROM
CAPTURE DATA ON CLOSED
ANIMAL POPULATIONS

by

DAVID L. OTIS, KENNETH P. BURNHAM,
GARY C. WHITE, AND DAVID R. ANDERSON

OCTOBER 1978

No. 62



FRONTISPIECE. Capture-recapture studies are frequently conducted on small mammal populations such as snowshoe hares *Lepus americanus*. (Photograph by Leta Burnham.)

STATISTICAL INFERENCE FROM CAPTURE DATA ON CLOSED ANIMAL POPULATIONS

David L. Otis,¹ Kenneth P. Burnham,² Gary C. White,³
and David R. Anderson

Utah Cooperative Wildlife Research Unit, Utah State University, Logan, Utah 84322

CONTENTS

INTRODUCTION	7	<i>Example</i>	35
<i>Objectives</i>	8	<i>Example</i>	36
<i>Assumptions</i>	9	<i>Discussion</i>	37
<i>Unequal Capture Probabilities</i>	11	MODEL M _{tb} : CAPTURE PROBABILITIES VARY	
<i>Perspectives</i>	12	BY TIME AND BEHAVIORAL RESPONSE	
<i>Comments on the Use of this Monograph</i>	12	TO CAPTURE	37
ACKNOWLEDGMENTS	13	<i>Structure and Use of the Model</i>	37
FUNDAMENTAL CONCEPTS	14	<i>Discussion</i>	38
<i>Data and Parameters</i>	14	MODEL M _{th} : CAPTURE PROBABILITIES VARY	
<i>Statistics and Notation</i>	15	BY TIME AND INDIVIDUAL ANIMAL	38
<i>Parameter Estimation</i>	16	<i>Structure and Use of the Model</i>	38
<i>Interval Estimation</i>	17	<i>Simulation Results</i>	39
HISTORICAL OVERVIEW	17	<i>Discussion</i>	39
MODEL M _c : CAPTURE PROBABILITIES ARE		MODEL M _{bh} : CAPTURE PROBABILITIES VARY	
CONSTANT	21	BY INDIVIDUAL ANIMAL AND BY BE-	
<i>Structure and Use of the Model</i>	21	HAVIORAL RESPONSE TO CAPTURE	40
<i>Simulation Results</i>	22	<i>Structure and Use of the Model</i>	40
<i>Example</i>	23	<i>Simulation Results</i>	41
<i>Example</i>	24	<i>Example</i>	42
<i>Discussion</i>	24	<i>Discussion</i>	42
MODEL M _t : CAPTURE PROBABILITIES VARY		MODEL M _{tbh} : CAPTURE PROBABILITIES VARY	
WITH TIME	24	BY BEHAVIORAL RESPONSE TO CAPTURE,	
<i>Structure and Use of the Model</i>	24	TIME, AND INDIVIDUAL ANIMAL	43
<i>Simulation Results</i>	25	<i>Discussion</i>	43
<i>Example</i>	28	REMOVAL MODELS	44
<i>Discussion</i>	28	<i>Introduction</i>	44
MODEL M _b : CAPTURE PROBABILITIES VARY		<i>Structure and Use of the Generalized Re-</i>	
BY BEHAVIORAL RESPONSE TO CAPTURE		<i>moval Model</i>	44
<i>Structure and Use of the Model</i>	28	<i>Simulation Results</i>	46
<i>Simulation Results</i>	30	<i>Example</i>	46
<i>Example</i>	31	<i>Example</i>	48
<i>Example</i>	32	<i>Example</i>	49
<i>Discussion</i>	32	<i>Discussion</i>	49
MODEL M _h : CAPTURE PROBABILITIES VARY		TESTS OF MODEL ASSUMPTIONS	50
BY INDIVIDUAL ANIMAL	33	<i>Philosophy of the Approach</i>	50
<i>Structure and Use of the Model</i>	33	<i>Summary of Models and Estimators</i>	51
<i>Simulation Results</i>	34	<i>Specific Tests to Perform</i>	53

¹ Present address: U.S. Fish and Wildlife Service, Denver Wildlife Research Center, Denver, Colorado 80225.

² Present address: U.S. Fish and Wildlife Service, Office of Biological Services, Fort Collins, Colorado 80521.

³ Present address: Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87545.

<i>On the Need for an Objective Selection Procedure</i>	56	APPENDIX B	105
<i>An Objective Model Selection Procedure</i>	57	<i>Estimation in Model M_o</i>	105
<i>Estimation in Alternative Models</i>	61	APPENDIX C	106
<i>Additional Examples of Model Selection</i>	62	<i>Estimation in Model M_t</i>	106
<i>A Test for Closure</i>	66	APPENDIX D	107
DENSITY ESTIMATION	67	<i>Estimation in Model M_b</i>	107
<i>Introduction</i>	67	APPENDIX E	108
<i>Problem Formulation</i>	68	<i>Estimation in Model M_h</i>	108
<i>Statistical Treatment</i>	69	APPENDIX F	110
<i>Discussion</i>	72	<i>Discussion of Model M_{tb}</i>	110
STUDY DESIGN	74	APPENDIX G	111
<i>Livetrapping Versus Removal Methods</i>	75	<i>Discussion of Model M_{th}</i>	111
<i>Closure</i>	75	APPENDIX H	112
<i>Eliminating Variation due to Time, Behavior, and Heterogeneity</i>	75	<i>Estimation in Model M_{bh}</i>	112
<i>Sample Size</i>	77	APPENDIX I	113
<i>Recording Data</i>	80	<i>Discussion of Model M_{bth}</i>	113
<i>Data Anomalies</i>	80	APPENDIX J	114
COMPREHENSIVE EXAMPLES	81	<i>Estimation in Removal Models</i>	114
<i>A Taxicab Example</i>	81	APPENDIX K	115
<i>A Penned Rabbit Study</i>	84	<i>Tests of Model Assumptions</i>	115
<i>An Example of Trap Response</i>	87	APPENDIX L	121
<i>An Example Where No Model Fits</i>	93	<i>Density Estimation Based on Subgrids</i>	121
COMPREHENSIVE COMPUTER ALGORITHM	96	APPENDIX M	123
SUMMARY	96	<i>General Simulation Methods</i>	123
LITERATURE CITED	97	APPENDIX N	123
APPENDIX A	102	<i>Simulation Results</i>	123
<i>Notes on Estimation</i>	102	APPENDIX O	133
		<i>Interval Estimation</i>	133

INTRODUCTION

The estimation of animal abundance is an important problem in both the theoretical and applied biological sciences. Serious work to develop estimation methods began during the 1950s, with a few attempts before that time. The literature on estimation methods has increased tremendously during the past 25 years (Cormack 1968, Seber 1973).

However, in large part, the problem remains unsolved. Past efforts toward comprehensive and systematic estimation of density (D) or population size (N) have been inadequate, in general. While more than 200 papers have been published on the subject, one is generally left without a unified approach to the estimation of abundance of an animal population.

This situation is unfortunate because a number of pressing research problems require such information. In addition, a wide array of environmental assessment studies and biological inventory programs require the estimation of animal abundance. These needs have been further emphasized by the requirement for the preparation of Environmental Impact Statements imposed by the National Environmental Protection Act in 1970.

This publication treats inference procedures for certain types of capture data on closed animal populations. This includes multiple capture-recapture studies (variously called capture-mark-recapture, mark-recapture, or tag-recapture studies) involving livetrapping techniques and removal studies involving kill traps or at least temporary removal of captured individuals during the study. Animals do not necessarily need to be physically trapped; visual sightings of marked animals and electrofishing studies also produce data suitable for the methods described in this monograph.

To provide a frame of reference for what follows, we give an example of a capture-recapture experiment to estimate population size of small animals using live traps. The general field experiment is similar for all capture-recapture

studies (a removal study is, of course, slightly different). A typical field experiment is the following: a number of traps are positioned in the area to be studied, say 144 traps in a 12×12 grid, 7 m apart. At the beginning of the study ($j = 1$) a sample size of n_1 is taken from the population, the animals are marked or tagged for future identification, and then returned to the population, usually at the same point where they were trapped. After allowing time for the marked and unmarked animals to mix, a second sample ($j = 2$, often the following day) of n_2 animals is then taken. The second sample normally contains both marked and unmarked animals. The unmarked animals are marked and all captured animals are released back into the population. This procedure continues for t periods where $t \geq 2$. The animals should be marked in such a way that the capture-recapture history of each animal caught during the study is known. In practice, toes are often clipped to uniquely identify individual animals (Taber and Cowan 1969) or serially numbered tags are sometimes used on larger animals.

Such capture studies are classified by 2 schemes that are directly related to what class of models are appropriate and what parameters can be estimated. The first classification addresses the subject of closure. Closure usually means the size of the population is constant over the period of investigation, i.e., no recruitment (birth or immigration) or losses (death or emigration). This is a strong assumption and, of course, never completely true in a natural biological population. For greater generality, we define closure to mean there are no unknown changes to the initial population. In practice, this means known losses (trap death, or deliberate removals) do not violate our definition of closure. If the study is properly designed, closure can be met at least approximately. Open or nonclosed populations explicitly allow for one or more types of recruitment or losses to operate during the course of the experiment (Jolly 1965, Seber 1965, Robson 1969, Pollock 1975).

Only closed populations will be considered in this monograph.

The second classification depends on the type of data collected with 2 possibilities occurring (Pollock 1974, unpublished doctoral dissertation, Cornell University, Ithaca, New York):

- (1) only information on the recovery of marked animals is available for each sampling occasion, $j, j = 1, 2, \dots, t$.
- (2) information on both marked and unmarked animals is available for each sampling occasion, $j, j = 1, 2, \dots, t$.

In case (1), population size (N) is not identifiable, however, other parameters can be estimated (Brownie et al. 1978). In case (2), N can be estimated using a wide variety of approaches depending upon what we wish to assume. Only case (2) will be dealt with here.

Objectives

The objectives of this publication are twofold:

- (1) to give a thorough treatment of the estimation of population size given multiple capture occasions ($t > 2$) assuming
 - a. population closure,
 - b. there may exist 3 major types of variation in capture probabilities;
- (2) to extend and make available a procedure for estimating density (number of animals per unit area) from grid trapping studies.

This monograph is specifically oriented to the commonly done grid trapping and removal studies where closure can reasonably be assumed. Specifically, we do not treat the case of 2 livetrapping occasions ($t = 2$). This subject (i.e., the Petersen or Lincoln estimators and variations thereof) is adequately covered in the literature (see Seber 1973). In fact, to use the methods presented here for analysis of grid trapping data we suggest the study have 5 or more trapping occasions.

There are some types of study designs

and analysis methods we do not cover here. We do not treat sequential sampling studies (e.g., Samuel 1968), stratified populations (e.g., Darroch 1961, Arnason 1973), Bayesian schemes (e.g., Gaskell and George 1972), or change in ratio estimation (e.g., Paulik and Robson 1969). The subject of stratifying the data after the fact on such variables as sex, age, or species is not discussed primarily because there rarely are enough data for such a stratification. The contingency table approach to estimation from multiple capture studies is a promising new development (see Fienberg 1972), but currently it is relatively unexplored or developed; we do not discuss it. Finally, we do not treat studies or analysis methods for which the goal is to compute only an index to abundance (e.g., captures per 100 trap nights); standard statistical techniques are adequate for those types of studies.

Although our objective is to present comprehensive methods of analysis, the scientist must realize that no amount of sophisticated statistical analysis can compensate for poor study design or field technique (such as high trap losses). The experimenter can do far more to ensure valid estimates by having a properly planned and conducted study than he can by sophisticated analysis after the experiment. We have therefore included a section on statistical aspects of study design. That section includes comments on how to deal with anomalies such as trap losses.

This publication is intended for use by biologists. Such a goal is difficult to attain due to the generally technical and mathematically complex nature of the subject matter. We have developed a comprehensive computer program to compute estimates and test statistics for the various methods covered in subsequent sections (program CAPTURE). Biologists who wish to analyze data are urged to use the computer program rather than to try to compute the various estimates and test statistics by hand. Also, most of the mathematical and statistical details are contained in appendixes to this monograph.

We hope this publication and the associated computer program will be useful within the framework of the assumptions considered.

We undertook the theory development and the writing of this report for a variety of reasons. Several important advances have been made but are available only as unpublished dissertations (Burnham 1972, unpublished doctoral dissertation, Oregon State University, Corvallis, Oregon; Pollock). New methods have emphasized nonparametric approaches that are robust to the failure of certain assumptions. Further, the use of a sequence of statistical models seems appropriate. It is unreasonable to expect a single method to perform well on studies of various species in different habitats, or the same species at different times. Pollock (unpublished dissertation) treated 4 models, each based on specific assumptions, and suggested a statistical testing sequence. That general strategy, followed in this publication, allows models (assumptions) that are inadequate to be rejected for a particular data set. A method inappropriate for field mice *Peromyscus* spp. may work well for voles *Microtus* spp.

There exists a large body of standard statistical theory that is directly relevant and applicable to the estimation problem in capture-recapture and removal studies. Biologists need not, however, learn the theory to be able to use the results of these advanced methods. The methods employed here are often beyond the formal training of most biologists, although they should be able to make proper use of the results. We stress that we have examined the estimation and inference problems in a rigorous statistical framework as opposed to various ad hoc procedures.

Another objective of this monograph is to bring to the biologists' and statisticians' attention the computer program written to implement the complex analyses described here. Without the aid of a computer to do the calculations, development of sophisticated analyses is just an academic exercise. Our philosophy in this matter has been summed up by

Overton and Davis (1969:404): "Computers will soon prove of very great value in the routine processing of census and survey data. When they become generally available, it will be desirable to advance to even more realistic and complex solutions to the problems; there will be no premium on simplicity, so long as the users understand the principles and are able to comprehend the constraints and limitations of the models on which the computer solutions are based."

Assumptions

Every estimation method is based on a set of assumptions. The general assumptions for the capture-recapture methods we present here are listed and discussed below. The assumptions for the removal experiment are given in the section on removal studies. Four assumptions are necessary for the most restrictive experimental situations:

- (1) the population is closed,
- (2) animals do not lose their marks during the experiment,
- (3) all marks are correctly noted and recorded at each trapping occasion j , and
- (4) each animal has a constant and equal probability of capture on each trapping occasion. This also implies that capture and marking do not affect the catchability of the animal.

Before discussing the above, we must emphasize that the focal point of our work has been to relax Assumption 4. That assumption is not met in most capture-recapture studies, and a large percentage of past efforts have been directed at relaxing it. Assumptions 1-3 must be made for all models considered here. We briefly discuss the first 3 and then elaborate on the last in the following section.

(1) *Population closure*.—This assumption arises because population estimation models were initially conceptualized as extensions of urn models (Feller 1950). Such models are basically intended to provide a "snapshot" of the population size at a given point in space and time.

In that context, open and closed models become essentially noncompeting, since open models are more frequently used for purposes of monitoring populations over a longer period of time and obtaining information concerning such properties as survival and recruitment rates. If estimates of population size at a given time are also desired, however, competition between the 2 types of models does arise. In general, open models require more data than closed models due to the fact that assumptions are more rigorous and more parameters are involved. Therefore, feasibility often prohibits the use of very general stochastic models for estimating population size of open populations (Jolly 1965; Seber 1965; Robson 1969; Arnason 1972a, 1972b, 1973; Pollock 1975). If, for example, a 10-day experiment is considered, 17 basic parameters would have to be estimated using Jolly's (1965) model. Hence, data from many population estimation experiments are inadequate for obtaining estimates with acceptable precision and small bias using models for open populations. Moreover, unlike the models treated here, none of those open population models allows for unequal capture probabilities of individual animals. Let it be clear, we believe that well-developed, general models for capture data from open populations are essential in some studies. However, we also believe that for many populations of interest, the closure assumption can be met approximately and the models discussed in this monograph will be useful. For example, closure might be assumed for an 8-day study of cottontails *Sylvilagus* spp. during a nonbreeding period in a well-defined (sampled) area.

A number of tests for closure have been derived (Robson and Flick 1965, Robson and Regier 1968, Pollock et al. 1974), but they generally have little chance of rejecting closure unless the sample is large and there is a marked departure from closure. In addition, closure tests are often confounded with behavioral response to capture, e.g., an animal that becomes un-

catchable, or nearly so, is indistinguishable from one that dies or emigrates. Pollock (1972, unpublished master's thesis, Cornell University, Ithaca, New York) discussed a test for mortality in some detail. The tests for recruitment are more difficult. Thus, the biologist is forced to consider carefully the design of such studies in an effort to assure that the closure assumption is met. Finally, we note that the tests for closure implicitly assume equal capture probabilities; therefore, such tests can reject closure when in fact closure is true but equal capture probability is false. This greatly lessens the value and power of such tests. We believe closure will have to be assessed largely from a biological basis rather than from any definitive statistical tests.

The closure assumption can be relaxed in some cases. Seber (1973:70-71) showed that natural mortality will not bias some estimators if it acts equally on marked and unmarked segments of the population. In such cases, the population estimate then relates to the size of the population at the beginning of the study. However, if recruitment and mortality occur during the experiment, the estimate of N will be too high, on the average, for both initial and final population size (Robson and Regier 1968).

(2) *Permanency of marks.*—Loss of marks (tags) violates the closure assumption and will result in an overestimate of N . If the study is of short duration (to help assure the closure assumption), it seems that loss of marks will generally be a minor problem. Some exceptions, such as radioactive isotopes with a very short half-life, undoubtedly occur (cf. Seber 1973:93-100).

(3) *Reporting and recording marks (tags).*—This assumption can be easily assured by working carefully. Field reports and keypunched cards should be edited and verified. Often, a pilot study may be beneficial to train personnel and identify any problems with the marking method.

Unequal Capture Probabilities

The fourth assumption is particularly important and, for this reason, we focus on it here. It is now widely recognized that this assumption is commonly not met (e.g., Young *et al.* 1952, Geis 1955a, Huber 1962, Swinebroad 1964). Edwards and Eberhardt (1967), Nixon *et al.* (1967), and Carothers (1973a) provided clear evidence that accurate population estimation usually will require models that provide for unequal probabilities of capture. The effects of unequal capture probabilities on estimates derived from models that assume equal catchabilities have been studied by computer simulation by Burnham and Overton (1969), Manly (1970), Gilbert (1973), and Carothers (1973b). Estimators studied were generally found to be significantly biased when this assumption was violated.

This monograph presents a number of models and estimators developed to relax the critical assumption of equal catchability. We have drawn heavily from the work of Pollock (unpublished dissertation, pers. comm.) and Burnham (unpublished dissertation). Following Pollock (unpublished dissertation), we consider a sequence of models each allowing for different combinations of up to 3 types of unequal capture probabilities:

- (1) capture probabilities vary with time or trapping occasion—Model M_t ,
- (2) capture probabilities vary due to behavioral responses—Model M_b ,
- (3) capture probabilities vary by individual animal—Model M_h (h = heterogeneity among animals).

The *assumptions* regarding unequal capture probabilities are to be explicitly embodied in probability models that describe capture studies.

We agree with Carothers (1973b:146) that equal catchability is an unattainable ideal in natural populations (cf. Seber 1973:81–84). We discuss the 3 simplest ways to relax this assumption.

Model M_t allows capture probabilities to vary by time (e.g., each trapping oc-

casion). This situation may be common even though the number of traps might be fixed during the course of the study. For example, a cold rainy day during the study might reduce activity of the animals and reduce the probability of capture. Also, if different capture methods are used on each occasion, this model could be appropriate.

Model M_b allows capture probabilities to vary by behavioral response or “capture history,” and deals with situations in which animals become trap happy or trap shy. Carothers (1973a) referred to this as a contagion of catchability. This implies that an animal’s behavior tends to be altered after its initial capture (e.g., perhaps the animal was frightened or hurt during initial capture and marking and thereafter it will not likely enter another trap).

Model M_h allows capture probabilities to vary by individual animal. This situation has been modeled only with great difficulty and requires that additional distributional assumptions be made. Individual heterogeneity of capture may arise in many ways. Perhaps accessibility to traps (as influenced by individual home ranges), social dominance, or differences in age or sex can cause such an unequal probability structure. This is an important type of variation and has been rigorously treated by Burnham (unpublished dissertation), whose nonparametric approach is presented in a later section.

In addition to these 3 simple models, we consider all possible combinations of the 3 types of unequal capture probabilities (i.e., Models M_{tb} , M_{th} , M_{bh} , and M_{tbh}). We also treat the “null” case in which capture probability is constant with respect to all factors (Model M_0). Model M_0 corresponds to the 4 assumptions listed earlier. For simplicity, we denote estimators of population size for a specific model using the same subscript notation. For example \hat{N}_0 denotes the estimator derived from Model M_0 ; \hat{N}_t denotes the estimator derived from Model M_t ; \hat{N}_{bh} denotes the estimator derived from Model M_{bh} , and so on.

Perspectives

We wish to emphasize that a specific set of assumptions is the basis for a specific model. The assumptions and model then represent a *tentative* hypothesis when analyzing the results of a particular capture experiment conducted to estimate population size or density. Cormack (1968:456) stated, "In all cases every iota of information, both biological and statistical, must be gathered to check and countercheck the unavoidable assumptions." Statistical testing within and between models (assumptions) is emphasized here. In spite of this, more work in this direction is clearly indicated. Our approach is to derive models for an array of types of unequal probabilities of capture. We conducted statistical tests to enable selection of an appropriate model for the analysis of a particular data set (cf. Pollock unpublished dissertation). Some models are very sensitive to small departures from the underlying assumptions; therefore, testing between models and investigating the robustness of each estimator are essential.

The importance of such testing is reflected in the fact that use of an inadequate model will often lead to a highly biased estimate of population size. This is perhaps to be expected, if not obvious. More subtle is that estimates of the sampling variance (a measure of precision) are quite dependent on the correct model. Bias of the estimator may be small, but the estimate of variance may be very poor, even with large samples. This can cause, for instance, associated confidence intervals to have very poor properties. The importance of assumptions and their testing cannot be overemphasized. Paulik (1963) noted that an approximately correct estimate with low precision is always better than a highly precise incorrect estimate. Tests of assumptions concerning equal capture probabilities are especially important because estimators based on given sets of assumptions are usually not robust to departures from those assumptions (Seber 1970, Gilbert 1973).

We believe rigorous probability models explicitly incorporating various tentative assumptions represent the best approach toward estimating population size N , or density D . The tentative nature of the assumptions and the general uncertainty about biological processes make testing a key concern. As Seber (1973) pointed out, statistical models should be used with caution, due to lack of control over natural populations. All models depend on the validity of various underlying assumptions—that are often difficult to evaluate rigorously.

Finally, we believe that theory and application must be integrated. Either in the absence of the other will stifle progress. For this reason we have tried to integrate the statistical theory with the biological application. We have, however, tried to separate the more complex subjects and include them as a series of technical appendixes. We urge biologists to try to consider and understand the appendixes, and we ask statisticians to continue to be concerned with the biological complications and realities before attempting additional theory development. Through an integrated team approach we can expect further progress on this series of estimation problems.

Comments on the Use of This Monograph

We cover several topics here, and present mathematical as well as applied results. Topics covered include data analysis of short-term livetrapping and constant effort removal studies, design of such live trapping studies, and simulation results on inference procedures. Numerous examples are also given. A variety of uses of this monograph are anticipated by: (1) biologists who must analyze actual data, (2) biologists (and statisticians) faced with designing capture studies, (3) persons interested in performance of estimators presented here, (4) statisticians interested in developing more advanced models, and (5) educators who seek to teach courses on the subject of population size estimation.

Biologists who have data from closed population livetrapping studies will have to read quite a bit of this monograph before they can understand the methods. They do not need to read the appendixes. They would have to understand all sections through TESTS OF MODEL ASSUMPTIONS, except REMOVAL MODELS. We believe this can be done by anyone having had a solid course in college level algebra and beginning statistics. In order to understand the essence of what we present, the reader does not have to follow all the mathematical descriptions of models nor discussions of model properties. We have included numerous examples. In particular, the reader should benefit greatly from the section on COMPREHENSIVE EXAMPLES.

If the reader intends to do, say, data analysis according to these methods, it is virtually necessary to use program CAPTURE (see COMPREHENSIVE COMPUTER ALGORITHM). This program is available and there is a user's manual for it. Persons with many data sets to be analyzed should get the program. Conversely, we do not recommend trying to implement this computer program if one has only a few (or one) data sets to analyze. In this latter case, it is better to have the data run for you. The authors are willing to assist in running such data provided the user arranges his own keypunching of the data in the necessary format (we can supply this format).

If one's goal is to analyze some removal data, the relevant sections are those on Models M_b , M_{bh} , and the removal models (plus the introductory sections). Again, the authors would try to help users analyze removal data; within reasonable limits we may be able to run the data analyses or assist in setting them up.

Many readers will sometimes be faced with designing a capture study. The section on STUDY DESIGN covers some fundamental design aspects of livetrapping studies for closed populations. If your goal is to design a study, read that section at a minimum; to get full advantage of this monograph in terms of design, you

will need to read most sections, exceptions being HISTORICAL OVERVIEW, REMOVAL MODELS, and material following the STUDY DESIGN section.

If you are interested in obtaining insights into the performance of various estimators, you should put special effort into studying the numerous simulation results presented here. This would require reading almost all the text and careful study of Appendix N.

Persons interested in doing further research along the lines of the models and approach of this monograph will have to carefully study almost everything here, especially the appendixes.

Finally, this monograph and program CAPTURE have value for teaching and learning about population size estimation. The simulation feature of CAPTURE can be especially valuable in teaching the concepts of sampling variation and properties of estimators. Persons interested in performing such simulation of the methods presented here (either for design of studies, evaluation of estimators, or teaching purposes) will need to implement the program for their own use.

ACKNOWLEDGMENTS

Dr. K. H. Pollock, University of Reading, provided several ideas and criticisms through correspondence; in addition, we have drawn on the results of his doctoral program. Drs. M. H. Smith and J. B. Gentry, Savannah River Ecology Laboratory, contributed to this work through discussion and provided samples of their research to be used for program testing. Data to be used as examples were provided by A. D. Carothers, H. N. Coulombe, C. T. Cushwa, W. R. Edwards, S. W. Hoffman, E. C. Larsen, R. F. Raleigh, and V. H. Reid.

We appreciate extensive comments made by A. N. Arnason, A. D. Carothers, W. R. Clark, S. W. Hoffman, and R. P. Davison on an earlier draft of the manuscript. The comments by A. N. Arnason and G. A. F. Seber on the final version of

the manuscript are also much appreciated.

Dr. R. G. Streeter, U.S. Fish and Wildlife Service, provided encouragement during the investigations.

This work was performed under Contract 14-16-0008-1224 of the Coal Program, Office of Biological Services, U.S. Fish and Wildlife Service, to the Utah Cooperative Wildlife Research Unit. Funds for this work were made available to the Fish and Wildlife Service as part of the Federal Interagency Energy/Environment Research and Development Program, Office of Research and Development, U.S. Environmental Protection Agency (IAG-EPA-D5-E385). Publication costs were paid by the Oil Shale Program, Office of Biological Services, U.S. Fish and Wildlife Service, contract 14-16-0008-1197. Final debugging of the computer program, and preparation of examples were performed under the auspices of the U.S. Energy Research and Development Administration.

FUNDAMENTAL CONCEPTS

This section presents notation and discusses the statistical techniques used in this monograph. The subject matter is basically technical in nature, but we have tried to keep the presentation simple and refer the reader to appendixes for more complex details. We believe it is important for users of the methods described in this publication to understand the material presented in this section.

Data and Parameters

All the models discussed here assume population closure (except for known removals). Therefore, the parameter we wish to estimate is population size N which is constant. Moreover, because the same individual animals compose the population on each trapping occasion, $j = 1, 2, \dots, t$, we can conceive of the individuals as being numbered $i = 1, 2, \dots, N$.

The basic capture data are conveniently expressed in matrix form as

$$[X_{ij}] = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1t} \\ X_{21} & X_{22} & \dots & X_{2t} \\ \vdots & \vdots & \dots & \vdots \\ X_{N1} & X_{N2} & \dots & X_{Nt} \end{bmatrix}$$

where

$$X_{ij} = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ animal is caught on the } j^{\text{th}} \text{ occasion} \\ 0 & \text{otherwise.} \end{cases}$$

The X matrix is a simple way to record the capture or noncapture of each animal in the population on each trapping occasion. Row i gives the trapping results for individual i , while column j gives the results of the j^{th} trapping occasion. Note that the matrix X may not be observed in its entirety because some animals may never be captured; therefore, those rows of X are all zeros.

A series of specific models for capture data can be derived if we define the following general structural model:

$$p_{ij} = \text{the capture probability of the } i^{\text{th}} \text{ individual in the population on the } j^{\text{th}} \text{ trapping occasion, where } i = 1, 2, \dots, N, \text{ and } j = 1, 2, \dots, t.$$

For example, if we assume the restrictions $p_{ij} = p$ for all i and j we get Model M_0 , the simplest possible model. All other models we introduce may be thought of as generalizations of Model M_0 . Hence, in the following sections, models are developed based upon capture probabilities being time specific, behaviorally related, or differing among individual animals. Therefore, capture probabilities are the crucial element of the series of models we discuss.

The above structure and assumptions suffice to specify the marginal distribution of each individual X_{ij} (i.e., they are Bernoulli random variables); however, they do not specify the joint distribution of all X_{ij} . Therefore, we have assumed joint independence of the variables in order to have a completely specified general model structure. Specifically, we assume that given the correct model (i.e., the correct specification of capture probabilities p_{ij}), then the elements of X_{ij} are

mutually independent random variables. This assumption is not testable unless one first knows what the correct model is (which we never will for real data). It is our opinion this is not a restrictive assumption and it need not be a source of concern.

Statistics and Notation

Probability models from which estimators of population size N may be developed are discussed in following sections. A few simple statistics are needed for these models. They are defined and discussed below.

n_j = the number of animals captured in the j^{th} sample, $j = 1, 2, \dots, t$,

$$= \sum_{i=1}^N X_{ij},$$

$n.$ = the total number of captures during the study $= \sum_{j=1}^t n_j$,

u_j = the number of new (unmarked) animals captured in the j^{th} sample, $j = 1, 2, \dots, t$,

f_j = the captures frequencies = the number of individuals captured exactly j times in t days of trapping, $j = 1, 2, \dots, t$. f_0 will be used for the number of individuals never captured (obviously, f_0 is not observable).

M_{t+1} = the number of distinct individuals caught during the experiment (recall that t is fixed for a given experiment),

$$= \sum_{j=1}^t f_j = \sum_{j=1}^t u_j,$$

M_j = the number of marked animals in the population at the time of the j^{th} sample, $j = 2, 3, \dots, t$. (Note that $M_1 \equiv 0$),

$M.$ = sum of the M_j [does not include

$$M_{t+1}] = \sum_{j=1}^t M_j,$$

m_j = the number of marked animals captured in the j^{th} sample, $j =$

$2, \dots, t$. Note that $u_j = n_j - m_j$ and that $m_1 = 0$,

$$m. = \text{sum of the } m_j = \sum_{j=1}^t m_j.$$

The statistics u_j , f_j , M_j , and m_j may also be computed directly from the X matrix. However, the computation is not as straightforward as that of n_j and is not given here. We denote X_ω as the number of animals with a specific capture history ω . For example X_{10011} represents those individuals caught on trapping occasions 1, 4, and 5. The set of all the possible capture histories will be symbolized as $\{X_\omega\}$. In general, with t capture occasions there are 2^t possible capture histories.

Two other terms used frequently in this monograph are:

Robustness (of an estimator).—A robust estimator is one that is not sensitive to the breakdown of a particular assumption. A specific measure of robustness is difficult to define. Therefore, a somewhat subjective determination regarding the robustness of an estimator is made relative to the general performance of the estimator. Performance is evaluated with respect to the essential criteria of bias, precision, and confidence interval coverage. For example, the estimator developed under Model M_1 performs very poorly with respect to all criteria if individual heterogeneity to capture is present in the population. We say that this estimator is not robust to a particular assumption. In contrast, the jackknife estimator for Model M_h appears to be fairly robust for a number of specific assumptions.

Bias (of an estimator).—Bias is the difference between the expected value of an estimator and the true parameter being estimated, e.g., $B = E(\hat{N}) - N$. Percent relative bias, $100\{[E(\hat{N}) - N]/N\}$ is denoted as RB. Overton and Davis (1969) gave a good discussion of these and other related terms.

A final note concerns the differences between parameters (true values) and estimates. We are concerned chiefly with

making estimates of the parameters N and D , population size and density, respectively. We denote our sample estimators of these parameters as \hat{N} and \hat{D} . Biologists are referred to Kendall and Buckland (1971) for definitions of standard statistical terms.

Parameter Estimation

The data from capture-recapture or removal studies are *samples*. This imposes the need for a probabilistic treatment of the data to derive correct estimation and inference procedures. The models we consider here are termed stochastic models. Unlike the models for open populations, the only stochastic component for models under population closure relates to the sampling process: i.e., the capture probabilities. Model formulation in this context begins with a set of explicit assumptions. A probability model for the sampling distribution of the X matrix (the basic data) is derived to quantitatively express the assumptions. A probability function is a form of mathematical representation of the observed data under a specific set of assumptions. It provides a basis for quantitatively and explicitly incorporating the specific assumptions about capture probabilities and for developing the point and interval estimators by rigorous statistical estimation techniques.

Most parameter estimators in this publication were derived using the method of maximum likelihood (ML). Several models and their corresponding estimators were taken from existing literature (e.g., Zippin 1956, Darroch 1958), often with some modification. Other models and estimators were derived during the course of this study.

Estimators found by the ML method are optimal, at least for large samples. (For a discussion of optimality, refer to Appendix A). This is a generally accepted tenet of statistical estimation theory (Mood et al. 1974). In general, ML estimators of unknown parameters (e.g., N) are found by application of results from simple cal-

culus, using the likelihood function derived from the probability model. In some cases, the estimator may take a simple, easy to use form. For example, the ML estimator of N for Model M_t for 2 sample occasions ($t = 2$) is the Petersen estimator

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

where n_1 , n_2 , and m_2 have already been defined.

However, in capture-recapture models we rarely find that the exact ML estimators exist as a simple formula as above. To illustrate this, consider the model developed by Darroch (1958) when 4 sampling occasions are considered ($t = 4$) and the capture probabilities are assumed to vary only by time (i.e., p_1 , p_2 , p_3 , and p_4). The approximate ML estimator of N for this model (see Darroch 1958) is the unique solution of the equation

$$\left(1 - \frac{M_5}{N}\right) = \left(1 - \frac{n_1}{N}\right)\left(1 - \frac{n_2}{N}\right) \cdot \left(1 - \frac{n_3}{N}\right)\left(1 - \frac{n_4}{N}\right).$$

In general, for Model M_t the ML estimator is the solution of the equation

$$\left(1 - \frac{M_{t+1}}{N}\right) = \prod_{j=1}^t \left(1 - \frac{n_j}{N}\right).$$

For t greater than 2, this equation cannot be solved algebraically for N . In other words, it is not possible to arrange the symbols algebraically in such a way that only N appears on one side of the equation and all other terms appear on the other side. The equation can be solved, but only on a case by case basis using a numerical procedure. We say the equation does not have a simple, "closed form" solution. Complex probability models often do not have simple estimators and tests of assumptions; nonetheless, complex models appear necessary to describe many capture-recapture studies adequately.

Our work has shown that several of the

approximations of \hat{N} suggested for this model in the past are fairly poor. Furthermore, some of the iterative solutions given (e.g., Darroch 1958) produce only approximate ML estimates. We have obtained exact ML estimators for all the models in this publication (except Models M_h , M_{th} , M_{tb} , and M_{bth}) by employing numerical procedures on a digital computer (in fact no estimators can be derived for the latter 3 models). We have found the maximum of the likelihood function in such a way as to obtain exact integer valued ML estimators of N . The disadvantage here, of course, is that we cannot show simple closed form estimators. This subject is discussed further in Appendix A.

We find that the estimator of N under each model involves only simple statistics computed from the X matrix. Individual captures are not employed—only various sums (linear combinations) derived from the X matrix. Those sums are statistics such as n_j , n , u_j , and M_{t+1} . For any model we consider, there exists a set of simple statistics, called minimal sufficient statistics (MSS). Estimators should be based on only MSS. The use of the ML method results in estimators that are always functions of the MSS. This is a desired property because it can be shown that the MSS contains all the information available from the experiment for estimating the parameter(s) of interest (in our case N). An estimator based on statistics other than MSS is not using all available information and is, therefore, not optimal. Some statistical tests of assumptions will depend on information other than MSS.

The number of parameters that can be identified (estimated) is less than or equal to the number of elements in the MSS (regardless of the estimation method used). The subject of "identifiability" of parameters is important in the material that follows and, therefore, we place some emphasis on MSS. For example, under Model M_0 , the MSS is n , and M_{t+1} , whereas under Model M_h , the MSS is f_j , $j = 1, 2, \dots, t$. We make frequent use of the MSS in the following sections.

Interval Estimation

One of the several advantages of the probability model/ML approach is that estimates of sampling variance and covariances can be computed as part of the ML method. These measures of precision are essential in making inferences from the sample results of the experiment. The variance and covariance estimators are derived from "large sample" theory and usually are of unknown value as measures of precision in "smaller" samples. We have performed a large number of Monte Carlo simulation experiments (Appendixes M and N) to examine the small sample properties of such variance estimators and the confidence intervals that depend on them. Interval estimation is an old subject in the statistical literature, and we refer the interested reader to the text by Mood et al. (1974) for details. Seber (1973) also gave numerous examples.

Typically, the ML estimator of N is not normally distributed unless large samples are taken. Because confidence intervals commonly used depend upon an assumption of normality, we explored alternative interval estimation techniques (Appendix O). The alternative procedures were not totally satisfactory and we will use the standard procedure as follows to construct an approximate 95 percent confidence interval on N :

$$\hat{N} \pm 1.96\sqrt{\text{Var}(\hat{N})}$$

This procedure has its limitations but, all things considered, appears to be best at present.

HISTORICAL OVERVIEW

Although the basic concept of obtaining information about an animal population by marking some of its members may be traced as far back as the 17th century (Chapman 1948) and to Petersen's (1896) expression of the fundamental principle, one may argue that the practical beginnings of the literature concerning the marking method can be associated with

Lincoln's (1930) use of band returns to estimate the size of the North American waterfowl population. In the nearly 50 years since Lincoln's (1930) initial work, a voluminous literature has resulted from efforts directed toward deriving and refining techniques based on the capture-recapture method. In the past decade, 2 notable attempts to summarize the existing literature have been made. The first is by Cormack (1968) who provided a survey of mathematical models proposed for use in capture-recapture experiments. The second is an extensive text by Seber (1973) that attempted to bring together all the proposed techniques for estimating population abundance and related parameters. Included as a subset of those techniques are those concerned with the capture-recapture method. In the presentation of those techniques, a substantial amount of mathematical detail is provided, as are numerical examples. Furthermore, assumptions that must be met to ensure validity of a particular technique are presented; methods for testing the validity of some of the assumptions are given.

Because of the existence and quality of the cited works, no attempt is made here to present specific methods associated with the theory of capture-recapture and related experiments. Rather, we present a review that follows the chronological development of conceptual approaches in the literature.

The initial state of the art is well characterized by Lincoln's (1930:2) statement of the solution to his population estimation problem: "Given a fairly accurate statement showing the number of wild ducks killed in North America in any one season, then the total number of ducks present on the continent for that season may be estimated by a percentage computation, based upon the relation that the total number of banded ducks killed during their first season as band carriers bears to the total number banded." As one might expect, no mention is made of the statistical properties such as an estimator might possess or of the underlying

assumptions, such as random sampling, that influence the validity of the method. Such considerations were at least hinted at, however, in Schnabel's (1938) paper that extended the method to the situation in which members of the population were marked and released back into the population on more than 1 occasion. Mention is made of the percentage relative bias of the estimators in an experiment in which the population size is known, and the reader is cautioned that "none of the solutions can be expected to provide more than an estimate of the general order of magnitude of the total population" (Schnabel 1938:352). Presumably, some caution is generated by the fact that "assumptions of random sampling and constant population are only rough approximations to the actual situation" (Schnabel 1938:352). More consideration was given to the uses of the capture-recapture technique in a sequence of papers by Jackson (1933, 1937, 1939, 1940), who was concerned not only with estimating population size but also with birth-immigration and death-emigration parameters. Contained in the consideration of those parameters is the concept that the population is not "closed," i.e., population size is not constant throughout the sampling period. Those methods were then being applied mainly to fish and insect populations and not to terrestrial wildlife populations, although some exceptions did exist (e.g., Green and Evans 1940). Scepticism as to the worth of the method with respect to wildlife populations was expressed by Dice (1941:402), who stated that "the application of the proportional method of calculating mammalian populations may often require as much effort as the complete trapping or counting of the whole sample population." Nonetheless, effort continued in development of the theory. Schumacher and Eschmeyer (1943) provided an alternative solution to that of Schnabel (1938) by the use of regression techniques. Evidently, their work was spurred by the desire to develop an estimator that would be more robust to de-

partures from the underlying assumptions of the Schnabel method. Moreover, unlike Schnabel, they provided an estimator for the standard error of the estimate. Similar regression techniques were also investigated by Hayne (1949a) and DeLury (1958).

A significant change in both the quality and quantity of work in the field of capture-recapture theory occurred with the appearance of several important papers of the early 1950s. Those papers signaled the beginning of a more rigorous mathematical treatment of the theory in terms of both estimation and testing of assumptions. Bailey (1951), for example, proposed a binomial model for the single mark-release situation. He used the ML theory to develop an estimator of the precision of the population size estimator, the latter estimator being the same as Lincoln's (1930). Chapman (1952) considered a hypergeometric model for the multiple capture-recapture experiment and derived an approximate expression for the resulting ML estimator. He also gave a test for determining whether the probability of capture is independent of tagging. Alternative sampling schemes were proposed by such authors as Chapman (1952, 1954) and Goodman (1953). Such schemes were designed to avoid undesirable statistical properties associated with the direct sampling method that considers the total number of animals caught on each occasion as a fixed parameter. For instance, Chapman (1952) pointed out that an estimator of population size obtained via inverse sampling (i.e., considering the number of *marked* animals caught on each occasion as fixed) is unbiased, whereas the estimator associated with direct sampling is biased. In addition, removal data, similar to the type of data used in marking experiments, was used in alternative methods proposed by Moran (1951) and DeLury (1951). Also at that time, progress was made in the theory of estimation in open populations through a sequence of papers by Leslie and Chitty (1951), Leslie (1952), and Leslie et al. (1953). Those authors used ML

theory for estimating such parameters as death rate and population size, and devoted much effort to the examination of assumptions.

The appearance of such mathematical treatments generated most of the important immediately succeeding work on the development of the theory. As an example, one can consider the work of Zippin (1956), who provided a more complete statistical treatment of the removal method first suggested by Moran (1951). An important example is the work of Darroch (1958), who was responsible for the correct derivation of the probability model for the multiple capture-recapture experiment first treated by Schnabel (1938). Moreover, Darroch presented expressions for the asymptotic bias and approximate variance for his approximate ML estimator and a method for constructing confidence intervals. Darroch's (1958, 1959) work on the closed model, the birth only and death only models stands as a cornerstone in the development of the theory.

Given the methods available, it was now possible for researchers to direct effort toward the development of statistical tests of assumptions underlying the methods of estimation. One of the most generally invoked assumptions of proposed estimation techniques was (and still is) that all animals in the population, regardless of capture history and other individual characteristics, are equally at risk to capture on each trapping occasion. Leslie (1958) devised a test directed toward that hypothesis, which was later extended by Carothers (1971). Cormack (1966) made the important point that failure of the above assumption may be caused either by each animal in the population possessing an "innate catchability" which varies among individuals over the population, or by an individual's probability of capture being affected by its capture history, or both. Cormack (1966) provided a test for the former assuming the latter is false. Seber (1962, 1965) and Robson and Youngs (1971) considered the problem of testing whether

marking an animal affects its probability of capture on subsequent trapping occasions, and Manly (1971) provided a method for estimating the effect of marking on survival of the animal. During that period, Seber (1965) and Jolly (1965) independently developed what is now known as the Jolly-Seber method of estimating open population parameters from multiple capture-recapture experiments. That model, aspects of which were later generalized by Robson (1969) and Pollock (1975), allows the population to be experiencing death, recruitment, immigration, and permanent emigration. Arnason and Baniuk (1977) provided a comprehensive computer algorithm to compute estimates for various models for open populations. Existence of such open population models points out the need for tests for closure of the population under study. Unfortunately, good tests of that assumption are still not available.

The importance of developing and using valid tests of model assumptions was further emphasized by results appearing simultaneously in the literature concerning the operating characteristics of existing estimation techniques. Edwards and Eberhardt's (1967) study on a confined rabbit population of known size revealed large biases in both the Schnabel (1938) and Schumacher-Eschmeyer (1943) methods of estimation. The authors conjectured that those biases were due to "individual animals having different or changing probabilities of capture." A simulation study by Braaten (1969) indicated serious bias in the estimators derived from DeLury's (1947) catch-effort model if the assumption of "constant catchability" is violated. A similar lack of robustness to unequal capture probabilities among animals was exhibited by estimators examined in a computer simulation study by Burnham and Overton (1969), who generated "populations" using the family of beta distributions. More recently, Carothers (1973b) sampled a population of known size, the members of which were the taxicabs of the city of Edinburgh. The assumption of popula-

tion closure was reasonable, and non-homogeneous individual capture probabilities were caused by the sampling schemes used. Various "Schnabel type" estimators, that assume equal capture probabilities, were reported as having substantial bias. In addition, 2 regression type estimators proposed by Tanaka and Kanamori (1967) and Marten (1970), each of which assumed a certain form of unequal capture probabilities, failed to reduce significantly the magnitude of the bias of the "Schnabel" estimators.

The appearance in 1965 of the Jolly-Seber method of estimating parameters of open populations did not preclude the development of additional estimation techniques in the literature, in spite of the fact that Cormack (1968:487) believed the method to be "an extremely powerful general formulation" of the capture-recapture experiment. Although the method is general in the sense that it allows for such processes as recruitment and mortality, it is restricted by the assumption that all animals have the same probability of capture on a given trapping occasion. In many experimental situations, the assumptions of population closure and unequal capture probabilities constitute a more realistic set of assumptions than the set required for the Jolly-Seber model. Hence, parameter estimators derived from models based on different sets of assumptions than the Jolly-Seber models continued to be developed by researchers such as Tanaka and Kanamori (1967), Eberhardt (1969a), and Marten (1970). The jackknife technique for bias reduction proposed by Quenouille (1949, 1956) was used by Burnham (unpublished dissertation) to derive an estimator for the situation in which each member of the population has an "innate" probability of capture that varies among individuals. That effort represents a unique attempt to develop a robust estimator of population size that is non-parametric, i.e., one that does not need to assume how capture probabilities are distributed over the population. Such non-parametric approaches are appealing

because they are robust to specific assumptions regarding the experiment and tend not to suffer from breakdown of specific assumptions used to parameterize the model. Pollock (unpublished dissertation) also considered estimation of population size under the assumption of heterogeneity of capture probabilities, but with the added complication that an animal's probability of capture may be altered by its capture history. However, no specific estimation procedure for that model had been proposed in the literature until the appearance of the generalized removal method described in this monograph.

This overview would not be complete without making some observations concerning methods of density estimation in capture-recapture experiments. The notion that the effective area of trapping is greater than the actual area of the trapping grid (i.e., the so-called edge effect) has long been recognized. Dice (1938, 1941) corrected for the effect by adding to the grid area a strip of one-half the home range of the animal, and that remains the most common practice at present. Other authors (Stickel 1954, Mohr and Stumpf 1966, Smith et al. 1975) have used recapture radii to correct for edge effect bias. Assessment lines have also been used to estimate density (Kaufman et al. 1971, Smith et al. 1971). More recently, Burnham and Cushwa (pers. comm.) have formalized MacLulich's (1951) technique for estimating density that involves using concentric trap grids to allow simultaneous estimation of density and edge width.

An underlying theme of this historical overview is that any capture-recapture experiment requires that the researcher make specific assumptions concerning the many factors that affect the results of the experiment. The assumptions that are chosen determine which statistical estimation procedures should produce the best results available from the data. Many estimation procedures have been proposed because several different assumptions can often be made for a given factor.

Unfortunately, it has been shown that misinformation results if, for a given experiment, assumptions are not valid or statistical estimators are not appropriate or both. Thus, it should be obvious that a rigorous approach to parameter estimation in capture-recapture experiments will include a statistical testing algorithm that allows the data to aid in selection of the "best" set of assumptions for the experiment. Although some tests of specific assumptions have been introduced, unified approaches to the problem have not, for the most part, received attention in the literature (an exception is the work of Pollock, unpublished dissertation). The concept of a unified approach is the basis for the development of this monograph. We believe an approach based on such a concept is a step in the direction of improved analyses of data from capture-recapture experiments. Furthermore, we hope that future research will be directed to that same objective.

MODEL M_0 : CAPTURE PROBABILITIES ARE CONSTANT

Structure and Use of the Model

Assumptions and Parameters

The simplest of all models under consideration results from the assumption that all members of the population are equally at risk to capture on every trapping occasion. Moreover, the occasions themselves do not affect capture probabilities. We thus have a model in which there is no heterogeneity of capture probability, no behavioral response to capture, and no variation in the experimental situation over time. This model is designated Model M_0 , and involves only 2 parameters: N , the population size, and p , the probability that an animal is captured on any given trapping occasion.

Statistical Treatment

The probability distribution of the set of possible capture histories $\{X_{\omega}\}$ is given by (cf. Darroch 1958):

$$P\{\{X_{\omega}\}\} = \frac{N!}{\left[\prod_{\omega} X_{\omega}! \right] (N - M_{t+1})! \cdot p^n \cdot (1 - p)^{N-n}}$$

where $n = \sum_{j=1}^t n_j$ = total number of captures in the experiment, and

M_{t+1} = number of different animals captured in the experiment.

An algorithm for producing ML estimators of N and p is derived in Appendix B. (When $t = 2$, a closed form ML estimator of N exists and is given by $\hat{N}_0 = (n_1 + n_2)^2 / 4m_2$, where m_2 is the number of recaptures in the second sample.) These estimators are necessarily functions of the minimal sufficient statistic $\{n., M_{t+1}\}$. Thus, all the information relevant for estimation purposes is contained in the number of different animals captured and the total number of animals captured during the course of the experiment. Appendix B also gives an estimator for the asymptotic variance of \hat{N}_0 that we used in the construction of confidence intervals for N .

Simulation Results

Bias

A computer was used to simulate experiments from populations satisfying the assumptions of Model M_0 . By varying the population parameters N and p , some insight into the small sample bias of \hat{N} was obtained. Results indicate that the bias of \hat{N} is negligible for values of p at least as large as 0.10 and $t \geq 5$. For smaller probabilities of capture, however, positive relative biases of 15–20 percent are realized. For example, from Appendix N, Table N.1.b, for a population of size $N = 400$, one simulation consisting of 200 replications with $p = 0.10$ and $t = 5$ produced an average value of \hat{N}_0 of 406.0 (Trial 2), while another, based on 500 replications, produced an average value of \hat{N}_0 of 456.9 with $p = 0.05$ (Trial 3). Complete results of the simulation of \hat{N}_0

for Model M_0 are given in Table N.1.b of Appendix N.

Confidence Intervals

Achieved confidence coefficients of the confidence interval procedure simulated were consistently at or above the 0.90 level and hence were close to the claimed 0.95 coverage. However, the width of an average interval is so large for small values of p that not much information concerning true population size is provided. For instance, for $N = 400$, $t = 5$, and $p = 0.05$ (Trial 3) expected width $[= 2 \cdot 1.96 \cdot \text{Ave} \sqrt{\widehat{\text{Var}}(\hat{N})}]$ is 628.2; and for $N = 400$, $t = 5$, $p = 0.10$ (Trial 2) this value is 217.7. However, with $p = 0.30$ (Trial 1) expected width drops to an average of 46.9, indicating that the model provides useful information concerning N when p is reasonably large. The number of replications for these 3 examples were 500, 200, and 200, respectively. One should keep in mind that extremely wide confidence intervals tend to reveal poor experimental conditions, i.e., low values of p , and thus can be of use in providing the experimenter with information concerning the success or failure of the experiment. See Table N.1.b of Appendix N for further details of the simulation results.

Robustness

Because Model M_0 is built from the assumption that no factors that affect capture probabilities are present in the experiment, it is not surprising that simulation results reveal that the estimator derived from this model is not robust to any type of variability in the capture probabilities. In particular, if capture probabilities vary by animal \hat{N}_0 exhibits significant negative bias. This property has been documented in the literature (Robson and Regier 1964, Gilbert 1973, Carothers 1973b). Common sense and some reflection on the nature of the experiment should tell us not only that behavioral response will cause bias in the

estimator, but will also indicate the direction of that bias. That is, animals becoming trap shy will cause \hat{N}_0 to overestimate N , and vice versa when animals become trap addicted. These assertions are supported by the simulation results given in Tables N.3.b and N.4.b of Appendix N. Results also reveal that the estimator is somewhat robust to changes in capture probabilities over time. However, Seber (1973) recommended, on the basis of Darroch's (1959) work, that the estimator associated with Model M_0 not be used even if the capture probabilities are suspected of not varying with time. This is good advice if large numbers of animals are being caught but such a rule could result in some loss of efficiency for small sample sizes.

Example

A capture-recapture experiment that satisfies the conditions of Model M_0 can be analogous to an urn experiment, a sampling experiment conducted in order to estimate the number of marbles in a cylinder when all marbles are the same size. We may visualize a cylinder containing N white marbles (individuals), each of which has an equal probability (p) of being picked from the cylinder on any given occasion. On each of t occasions, the following sampling scheme is carried out. A "sampling cylinder," with a diameter that is 100 percent of the diameter of the cylinder containing the marbles, is inserted into the container and a random sample of marbles removed. The numbers of white (individuals not previously "captured") and black ("recaptures") marbles in the sample are recorded. All white marbles are painted black and returned to the container along with the black marbles, and all the marbles are randomly mixed. The number of white and black marbles in the sample is recorded. Using the data from these t samples, the estimation procedure associated with Model M_0 provides the appropriate estimator of N , the number of marbles in the cylinder. Notice that all

NUMBER OF TRAPPING OCCASIONS HAS	5
NUMBER OF ANIMALS CAPTURED, $M_{(t+1)}$, HAS	98
TOTAL NUMBER OF CAPTURES, n , HAS	238
ESTIMATED PROBABILITY OF CAPTURE, p -HAT = .4672	

POPULATION ESTIMATE IS	102 WITH STANDARD ERROR	2.3956
APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL		97 TO 107

FIG. 1. Example of population estimation with constant probability of capture under Model M_0 with simulated data based on $N = 100$, $t = 5$, and $p = 0.5$.

the assumptions of Model M_0 are fulfilled for this cylinder model. That is, the population is closed because marbles may not enter or leave the container, and every individual has the same probability of capture on every trapping occasion because (1) all marbles are the same size and thus are not "heterogeneous," (2) white and black marbles have the same capture probability and thus there is no "behavioral response to capture," and (3) the same "sampling cylinder" is used in the same manner on all t occasions and thus there is no "time variation."

The fact that an analogy can be drawn between a capture-recapture experiment modeled by Model M_0 and the simple urn experiment illustrates the point that it is not reasonable to expect that many capture-recapture studies can be adequately represented by Model M_0 . Therefore, to present an example of the estimation procedure of Model M_0 , we simulated capture-recapture sampling for 5 occasions on a population of 100 individuals, each of which had a 0.5 probability of capture. As Fig. 1 shows, the value of the minimal sufficient statistic $\{n, M_{t+1}\}$ is $\{238, 98\}$. These values, and the value of t , are used to produce the population estimate of 102. Because $N = 100$, this estimate is only 2 percent greater than the true value of N . Note also that the lower limit of the large sample 95 percent confidence interval extends below the number of different marbles seen. This undesirable operating characteristic is revealed throughout the results of this study, and is discussed in Appendix O. When this happens, the

NUMBER OF TRAPPING OCCASIONS HAS	5
NUMBER OF ANIMALS CAPTURED, $H(T+1)$, HAS	106
TOTAL NUMBER OF CAPTURES, N , HAS	149
ESTIMATED PROBABILITY OF CAPTURE, $P\text{-HAT}$	= .1726
POPULATION ESTIMATE IS	173 WITH STANDARD ERROR 17.9418
APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL	137 TO 209

FIG. 2. Example of population estimation with constant probability of capture under Model M_0 with meadow vole data from E. Larsen (pers. comm.).

lower limit should be increased to the number of distinct individuals seen. Finally, we mention that one would not expect the estimator \hat{N}_0 to be robust to departures from the assumptions of the urn experiment. For instance, if black (marked) marbles were larger than white (unmarked) ones, and thus had a higher probability of selection, we could expect \hat{N}_0 to exhibit significant negative bias.

Example

E. Larsen (pers. comm.) used live-trapping to estimate the population size of meadow voles *Microtus ochrogaster* on a grid near the Flint Hills of Kansas in June 1974. A 10×10 grid of live traps, spaced 40 feet (12.2 m) apart, was laid out in a tall-grass prairie that had been unburned and ungrazed for 3 years. On the first 2 nights of trapping, traps were placed on top of the deep, dense litter that uniformly covered the substrate, and as a result almost no animals were captured. On the third night, holes were dug in the litter and the traps were placed in the holes. That trapping occasion yielded only 12 animals captured, perhaps due to the adverse effect that disturbance of the environment may have had on the animals. On the last 5 nights of trapping, however, relatively large numbers of animals were captured consistently. Thus, we have chosen to analyze the data from only those occasions. When applied to those data, the discrimination procedure described in the section entitled TESTS OF MODEL ASSUMPTIONS chose Model M_0 as the appropriate model for the data.

The estimation procedure associated with this model produced the results presented in Fig. 2. Notice that the calculated 95 percent confidence interval is relatively narrow, probably due to the fact that the estimate of capture probability p is nearly 0.2.

Discussion

Model M_0 represents what might be called the "best" of all possible experimental situations considered here in that a minimum number of "nuisance" parameters is involved (one) if one is concerned only with estimation of population size N . This lack of nuisance parameters results of course from the restrictive assumptions on which the model is based. We believe that those assumptions are in most cases unrealistic, and, therefore, the estimator based on the model is, in general, of limited use. The case against the model is strengthened by the fact that its associated estimator \hat{N}_0 appears extremely nonrobust to variation in capture probabilities caused by behavioral response or heterogeneity. Moreover, it appears true in general that little is gained by using Model M_0 instead of Model M_t when only time specific changes in probabilities are present. Therefore, the greatest utility of Model M_0 lies in providing a "null" model useful in testing for sources of variation, and in providing a basic model that can be generalized in a number of different ways. Such generalizations are the subject of concern in the following 7 sections.

MODEL M_t : CAPTURE PROBABILITIES VARY WITH TIME

Structure and Use of the Model

Assumptions and Parameters

The set of assumptions used as a basis for Model M_t is the same set associated with the classical multiple capture-recapture experiment. It is assumed that all members of the population are equally at

risk to capture on the j^{th} trapping occasion. Thus, all animals have the same probability of capture on any particular trapping occasion, but that probability can change from one occasion to the next. The set of parameters involved in this model contains N , the population size, and the p_j , $j = 1, \dots, t$, where p_j is the probability of capture on the j^{th} occasion.

Statistical Treatment

Model M_t has received more statistical attention than any other in the literature (see Cormack 1968). Schnabel (1938) first used the above set of assumptions to develop a model from which the well-known Schnabel estimator was derived. Her model, however, assumed that the values of the M_j , the number of marked animals in the population at time j , are known a priori, for $j = 1, \dots, t$. It remained for Darroch (1958) to derive the correct model for the situation. Using his results, we may write the probability distribution of the set of possible capture histories $\{X_\omega\}$ as:

$$P[\{X_\omega\}] = \frac{N!}{\left[\prod_{\omega} X_\omega! \right] (N - M_{t+1})!} \cdot \prod_{j=1}^t p_j^{n_j} (1 - p_j)^{N - n_j}$$

where

- n_j = number of animals caught on the j^{th} occasion, and
- M_{t+1} = number of different animals captured in the experiment.

When $t = 2$, a closed form expression for the maximum likelihood estimator of N exists and is given by $\hat{N}_t = n_1 n_2 / m_2$, where m_2 is the number of recaptures in the second sample. This is the familiar Lincoln Index. Darroch (1958) derived an expression that may be solved iteratively to give an estimator of population size for $t > 2$. One is led to believe that this estimator produces estimates within unity of the true ML estimate of N , but this is not in fact the case. Details of the

algorithm necessary to produce the true ML estimate for a given set of data are given in Appendix C. The estimate will depend only on the value of the minimal sufficient statistic for the model, namely $\{n_1, n_2, \dots, n_t, M_{t+1}\}$. Note that the model involves $t + 1$ parameters and that the dimension of the MSS is also $t + 1$. This assures identifiability of all the parameters of the model. Darroch (1958) also developed an estimator of the asymptotic variance of the ML estimator that can be used in the construction of a confidence interval for N (see Appendix C).

Simulation Results

Bias

If the experimental situation is well represented by Model M_t , it is important to know what biases may be expected from the estimator of N discussed in the preceding section. Computer simulations of experiments on populations satisfying the assumptions of Model M_t produce the same general conclusions concerning the bias of \hat{N}_t as those produced in Model M_0 . That is, if the probabilities of capture p_j are, on the average, close to 0.1 or larger, the bias of \hat{N}_t is not significant. Again, however, if the p_j 's become smaller than 0.1, significant bias results. Some examples are given in Table 1 for experiments conducted for 5 and 7 trapping occasions (additional results are presented in Table N.2.b of Appendix N). In Table 1, as in all tables in this publication $\text{Ave}[\cdot]$ represents the average value of the quantity in brackets over all simulated replications. R represents the number of replications performed and RB represents percent relative bias.

Confidence Intervals

Confidence intervals were constructed from simulated experiments to compare achieved confidence coefficients to the stated value of 0.95. Those achieved levels depend on a number of factors of which the most important are the accuracy of the variance estimator of \hat{N}_t , the

TABLE 1.—RESULTS OF COMPUTER SIMULATION STUDIES OF THE BIAS OF \hat{N}_t (ALSO SEE APPENDIX TABLE N.2.b)

Percent relative bias RB	Average of estimates Ave[N]	True population N	Model M_t Probabilities of capture $p_j, j = 1, 2, \dots, t$	Number of reps R	Number of occasions t	Trial
22.0	1,015.6	800	0.01, 0.01, 0.02, 0.03, 0.03	200	5	11
10.7	442.7	400	0.03, 0.04, 0.05, 0.06, 0.07	200	5	5
2.2	408.6	400	0.10, 0.10, 0.10, 0.10, 0.01	200	5	3
-0.7	198.7	200	0.30, 0.40, 0.10, 0.40, 0.30	100	5	11
-0.4	398.6	400	0.50, 0.20, 0.10, 0.10, 0.10, 0.10, 0.10	100	7	7
-0.2	399.4	400	0.20, 0.40, 0.30, 0.10, 0.20, 0.30, 0.20	100	7	9

degree to which \hat{N}_t is normally distributed, and the accuracy of \hat{N}_t itself. Results indicated that the achieved confidence level was in most cases at least 0.90. Those results are encouraging because in many of the experimental situations investigated the assumptions necessary for the strict validity of the confidence interval used are not met (cf. Appendix O for a discussion of the assumptions). However, the amount and the kind of information contained in such an interval varies from one situation to the next depending on the values of the capture probabilities $p_j, j = 1, 2, \dots, t$. If those probabilities are on the average as large as 0.20 or 0.30, confidence interval coverage is good and interval width is small. Thus, very useful and reliable information concerning N can be obtained under such experimental conditions. As the capture probabilities p_j decrease, however, so does the information about population size provided by the confidence interval. That is, interval width tends to be so large as to not provide any useful information with respect to N. This result agrees with Chapman's (1951) argument, for the case $t = 2$, that variance

estimates tend to be prohibitively large when probabilities are small. In cases where capture probabilities are extremely low, lower confidence interval limits are negative, thus providing no information about N. As pointed out in the discussion of Model M_0 , however, such experimental results do inform the researcher that his data have been able to tell him essentially nothing about population size. In that respect, therefore, the calculated confidence interval is providing relevant information, not about population size but about the failure of the experiment itself.

To illustrate these points, Table 2 has been constructed, giving selected simulation results taken from Table N.2.b of Appendix N regarding confidence interval coverage (the proportion of replications R in which the confidence interval contained N), and average confidence interval width (Ave [C.I. width]).

Robustness

Because the maximum likelihood estimator (or approximations thereof) of population size N has been so frequently

TABLE 2.—SIMULATED CONFIDENCE INTERVAL WIDTHS AND COVERAGE USING \hat{N}_t (ALSO SEE APPENDIX TABLE N.2.b)

Average [C.I. width]	Coverage	Population size N	Model M_t Probabilities of capture					Number of reps R	Number of occasions t	Trial
			p_1	p_2	p_3	p_4	p_5			
1,348.3	0.89	400	0.01	0.01	0.02	0.03	0.03	1,000	5	
140.2	0.94	100	0.05	0.05	0.10	0.15	0.15	1,000	5	12
31.3	0.88	200	0.40	0.30	0.10	0.30	0.40	100	5	10
15.1	0.96	400	0.45	0.45	0.50	0.55	0.55	200	5	1

used in practice over the last 40 years, questions of robustness of those estimators with respect to departures from the assumptions of Model M_t are particularly relevant. Simulation results indicate that \hat{N}_t under Model M_t is nonrobust to failure of the assumption that all animals, regardless of capture history, are equally catchable on the j^{th} trapping occasion. As previously emphasized, departures from that assumption may be due to capture probabilities varying with the animal, or the animal's probability of capture being altered after first capture, or both. Although both departures may create significant bias in \hat{N}_t , the nature of the bias differs. That is, departure from equal catchability caused by heterogeneity results in \hat{N}_t being negatively biased, with the magnitude of the bias proportional to the amount of heterogeneity present in the population. This observation corresponds with the results of the study by Edwards and Eberhardt (1967) on a rabbit population of known size. Bias of \hat{N}_t that is a manifestation of animals exhibiting a behavioral response to capture, however, may be either positive or negative, according to whether the animals become trap shy or trap addicted, respectively. This is the same result noted for the estimator of N associated with Model M_o . The magnitude of the bias depends on the degree to which the animal's behavior is changed. Table 3 illustrates the bias of \hat{N}_t when it is used in simulated experiments conducted on populations satisfying the assumptions of Model M_b , in which animals exhibit be-

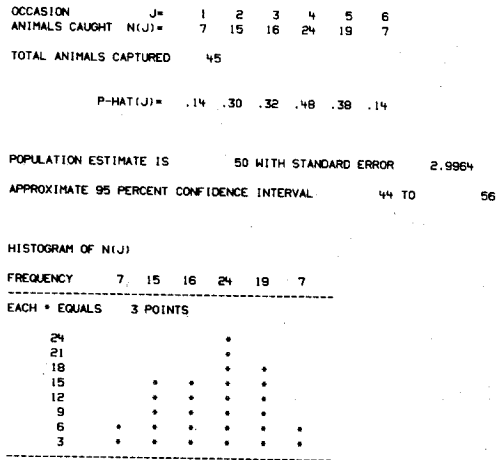


FIG. 3. Example of population estimation with time specific changes in probability of capture under Model M_t with least chipmunk data from V. Reid (pers. comm.).

havioral response to first capture, and Model M_b , in which heterogeneity of capture probabilities occurs. In Table 3, 2 examples of Model M_b are considered: Trial 8, where animals that have not been captured previously have a probability of capture of 0.2 on every trapping occasion and a probability of capture 0.05 after first capture, and Trial 9, where the probability of capture for previously captured animals changes from 0.2 to 0.5. Three examples of Model M_h are considered; for example Trial 1, in which 200 animals have a 0.05 probability of capture, 100 animals have a 0.15 probability of capture, and 100 animals have a 0.50 probability of capture. When the estimator \hat{N}_t was computed from the data generated

TABLE 3.—COMPUTER SIMULATION RESULTS ILLUSTRATING THE NONROBUSTNESS OF THE DARROCH ESTIMATOR \hat{N}_t USING DATA GENERATED UNDER OTHER MODELS. DATA MODELS ARE DESCRIBED IN TABLES N.3.a AND N.4.a OF APPENDIX N (ALSO SEE APPENDIX TABLES N.3.b AND N.4.b)

Percent relative bias RB	Ave[\hat{N}_t]	Population size N	Number of reps R	Number of occasions t	Data model
199.3	299.3	100	100	5	M_b , Trial 8 ¹
-28.3	71.8	100	100	5	M_b , Trial 9
-43.0	228.0	400	100	5	M_h , Trial 1
-13.2	173.6	200	100	10	M_h , Trial 6
-12.1	87.9	100	100	5	M_h , Trial 12

¹ For example, data generated under Model M_b were used to estimate N using estimator \hat{N}_t .

under this trial of Model M_b , an average percent bias of -43.0 resulted (Table 3).

Example

In the summer of 1975, V. Reid (pers. comm.) laid out a 9×11 livetrapping grid with traps spaced 50 feet (15.2 m) apart at a Colorado location in a bottom area dominated by sagebrush and snowberry and peripherally by gambel oak, serviceberry, and juniper. Least chipmunk *Eutamias minimus* were trapped for 6 consecutive days ($t = 6$). The discrimination procedure described in the TESTS OF MODEL ASSUMPTIONS section chose Model M_t as the most appropriate model for the data. Although that choice is subject to some suspicion (the goodness of fit test of Model M_t could not be performed because of insufficient data), the data were analyzed using the estimation procedure associated with Model M_t for purposes of illustration. Results (Fig. 3) indicate that estimates of the p_j are large enough for one to expect valid and useful confidence intervals, and such an interval does appear to result. The point estimate of 50 animals also seems to be in line with the observed data, i.e., the fact that u_6 was 0 (no new animals caught on day 6) leads us to believe that the value of N_t should not be much larger than the number of different animals captured. A histogram of the n_j values is included to help the researcher to visually examine the data.

Discussion

Given the simulation results of this section, one might conjecture that frequent use of the "Schnabel method," i.e., Model M_b , in practice has been unfortunate unless tests of the assumptions of Model M_t have indicated the model may be appropriate. Moreover, if Model M_t appears to be an adequate representation of the experimental situation, relatively large values (at least 0.2 on the average) of the parameters p_j usually are necessary to produce useful information with re-

spect to population size. As Cormack (1968) noted, "the higher the proportion of the population marked the more statistically precise will be the estimate of population size." If the capture probabilities are small, the variance estimate of \hat{N}_t , and hence the confidence interval for N_t , tends to be quite large, telling the experimenter that the estimate of population size N is unreliable. Because of the apparent positive correlation between average capture probability and the amount of useful information about N contained in the experiment, it is wise for the experimenter to calculate the maximum likelihood estimates of the p_j , $j = 1, 2, \dots, t$, from the data at hand. (Formulas for these ML estimators are given in Appendix C.) Such calculations should provide more feel for both the quantity and quality of information contained in the data.

MODEL M_b : CAPTURE PROBABILITIES VARY BY BEHAVIORAL RESPONSE TO CAPTURE

Structure and Use of the Model

Assumptions and Parameters

This model deals with the failure of the assumption that initial capture does not affect the probability of capture on subsequent occasions. That is, the model allows an animal to exhibit a behavioral response to capture and become either "trap addicted" or "trap shy." Overton and Davis (1969) pointed out that "it is well known that so-called trap-happy animals are often encountered." Moreover, the existence of trap response has been well documented (Geis 1955a, 1955 unpublished doctoral dissertation, Michigan State University, East Lansing, Michigan; Tanaka 1956, 1963; Flyger 1959; Bailey 1968; Pucek 1969). Formally, Model M_b assumes that on any given trapping occasion, all unmarked animals have one probability of capture, and all marked animals have another probability of capture. It is assumed that there is no

difference between trapping occasions, i.e., that capture probabilities do not vary with time. An implication of those assumptions is that all members of the population have the same probability of capture at the beginning of the experiment. Note also that the assumption is made that an animal's capture probability is altered only once, after first capture. Although one might think it more realistic to allow the capture probability to be changed more than once (e.g., after both first and second capture) this more general assumption turns out to have no effect on the estimation of population size N . Therefore, for simplicity of presentation the assumption is made that all marked animals, regardless of the number of times they might have been captured, have the same capture probability.

The assumptions of Model M_b result in 3 model parameters: N , population size; p , the probability of capture of an unmarked animal on any trapping occasion; and c , the probability that an animal is captured on any trapping occasion subsequent to the occasion on which it was first captured.

Statistical Treatment

The probability distribution of the set of possible capture histories $\{X_\omega\}$ from a multiple capture-recapture experiment on a population satisfying the assumptions of Model M_b may be written as (Pollock unpublished dissertation):

$$P\{\{X_\omega\}\} = \frac{N!}{\prod_\omega X_\omega!(N - M_{t+1})!} \cdot p^{M_{t+1}}(1 - p)^{N - M_{t+1} - M} \cdot c^m(1 - c)^{M - m},$$

where $M = \sum_{j=1}^t M_j =$ sum (over j) of the

number of marked animals in the population at the time of the j^{th} trapping occasion,

$m = \sum_{j=1}^t m_j =$ total number of marked animals, caught, and

$M_{t+1} =$ number of different animals caught during the entire experiment.

A MSS for this probability distribution is given by $\{M_{t+1}, m., M.\}$. Notice that this is a 3-dimensional vector of statistics, and because the model involves 3 parameters, each of those parameters is identifiable. What is most noteworthy concerning the above distribution is the fact that the estimation of c is independent of the estimation of the parameters N and p . That is, in Model M_b , once an animal has been captured, subsequent recaptures of the animal provide no information with respect to the estimation of N and p . The recapture information of the experiment is used only in the estimation of the nuisance parameter c , the probability of recapture. The reader is referred to Appendix D for details of the above argument and for the derivation of the maximum likelihood estimators of N and p . Those estimators are essentially equivalent to those given by Zippin (1956, 1958), who provided a basic statistical analysis of the removal experiment first proposed by Moran (1951). In Zippin's discussion, however, animals usually are physically removed by killtrapping or electrofishing, whereas in Model M_b , animals are "removed" from the population by being marked; in both, estimators of N and p depend only upon first captures.

Since estimation of population size under the conditions of Model M_b is equivalent to estimation in Zippin's removal model, the "failure criterion" associated with the removal method applies. Seber and Whale (1970) showed that valid ML estimators for N and p are obtained from the data when the criterion

$$\sum_{j=1}^t (t + 1 - 2j)(n_j - m_j) > 0$$

is satisfied. If that condition, which tests whether the population is being sufficiently "depleted" by the "removal" of new animals, is satisfied, a confidence interval for N may be constructed using

TABLE 4.—RESULTS OF COMPUTER SIMULATION STUDIES OF THE BIAS OF \hat{N}_b (ALSO SEE APPENDIX N, TABLE N.3.b)

Percent relative bias RB	Average of estimates Ave[\hat{N}_b]	True population N	Probability of capture p	Number of reps R	Number of occasions t	Trial
-18.0	82.0	100	0.10	161	5	10
15.4	461.5	400	0.10	196	5	4
9.8	109.8	100	0.20	199	5	8
-0.7	198.6	200	0.25	100	7	6
-1.1	197.9	200	0.30	200	5	5

the asymptotic variance estimator of \hat{N}_b , given by Zippin (Appendix D).

Simulation Results

Bias

Because behavioral response to capture is common in practice, an estimator of population size that exhibits relatively small biases in populations that satisfy the assumptions of Model M_b would prove useful in practice. Simulation results of experiments conducted on such populations indicate that \hat{N}_b , the ML estimator of N, is a good estimator if the probability of capture for an unmarked animal (p) is at least 0.2. Frequency of "failure" of the experiment is substantial (20%) with p values ≤ 0.1 , and when the experiment does succeed in these cases, \hat{N}_b tends to be significantly biased. A few examples taken from Table N.3.b of Appendix N of the simulated bias of \hat{N}_b are given in Table 4 which illustrates that biases on the order of 15–20 percent occur when p = 0.10, but that such bias gradually decreases as probability of first capture increases to reasonable levels greater than 0.20. Although the first 2 simulations included in Table 4 have the same probability of first capture, the case for N = 100 is negatively biased, whereas the case for N = 400 is positively biased. These seemingly contradictory results are caused by the estimator failing more often for the N = 100 case. When the estimator is close to failing (i.e., the failure criterion is close to zero), the estimates tend to be biased high. Because actual failure of the method occurs much more

frequently in cases where N is small, a relatively large number of potentially large estimates of N are "lost." This phenomenon therefore effects a significant reduction in the value of Ave[\hat{N}].

Confidence Intervals

Confidence intervals were simulated for various populations satisfying the assumptions of Model M_b (Table 5). A high correlation between the values of p and adequate performance of confidence intervals is indicated by Table 5. Evidence indicates that high confidence levels and small (hence informative) interval widths can be expected from experiments in which the probability of first capture is at least 0.30. On the other hand, in an experiment with N = 100 and p = 0.10 lower limits of confidence intervals were negative in more than 40 percent of those simulated experiments that succeeded. Such results illustrate the point that, in general, confidence intervals constructed from experiments with an insufficiently large value of p serve only to inform the researcher that his data cannot provide any real information with respect to population size.

Robustness

If one examines the failure criterion for estimation in Model M_b , it becomes clear that the number of new animals captured (removed) should decrease for each successive trapping occasion. Simulation results seem to indicate that steep declines in new captures over time produce good estimates of and informative confidence

TABLE 5.—SIMULATED CONFIDENCE INTERVAL WIDTHS AND COVERAGE USING \hat{N}_b (ALSO SEE APPENDIX N, TABLE N.3.b)

Average [C.I. width]	Coverage	Population size N	Probability of capture p	Number of reps R	Number of occasions t	Trial
253.9	0.71	100	0.10	161	5	10
145.0	0.95	200	0.20	98	5	
51.5	0.90	200	0.30	100	5	5
78.0	0.92	400	0.30	100	5	1

intervals for the parameter N . (This relates directly to the previous observation that larger values of the parameter p produce estimators with better properties.) One can see that if capture probabilities vary from one trapping occasion to the next, the vector of "removals" represented by (u_1, u_2, \dots, u_t) may be perturbed in such a way as to prevent the desired monotone decrease in the removals over time. This results in \hat{N}_b exhibiting a large bias and confidence intervals for N having large expected width. These observations support the conjecture that estimation of N based on Model M_b will be sensitive to significant changes in capture probabilities over time. Moreover, if capture probability varies among individuals, independently of the animal's capture history, \hat{N}_b tends to underestimate N . The magnitude of the bias depends directly on the number of animals in the population that are essentially untrappable, i.e., those that have small (<0.1) "innate" probabilities of capture. This non-robustness to heterogeneity of capture probability is of the same nature as that exhibited by \hat{N}_o and \hat{N}_t . These remarks are based on results obtained from simulation experiments on populations that satisfy the assumptions of Model M_h (see following section for a complete description of this model). A few examples follow concerning the bias of \hat{N}_b in experiments on Model M_h populations. In a population of size 100, with 40 animals that have a 0.05 probability of capture, 40 animals with a 0.10 probability, and 20 animals with a 0.30 probability (Trial 8), an experiment with $t = 10$ trapping occasions was simulated resulting in an average \hat{N}_b estimate of 78 for N . In a pop-

ulation of size 400, with 200 animals that have a 0.05 probability of capture, 100 with a 0.15 probability, and 100 with a 0.50 probability (Trial 1), an experiment with $t = 5$ trapping occasions produced an average estimate of 258 animals.

Example

E. Larsen (pers. comm.) reported the results of a livetrapping experiment conducted in the summer of 1976 as part of a study on community succession. A total of 55 live traps, spaced 15 m apart and covering 0.81 ha, were set out in a field in which vegetation was dominated by a mixture of sagebrush *Artemisia tridentata* and rabbitbrush *Chrysothamnus* spp. Trapping was conducted for 10 consecutive nights. Data collected on the deer mouse *Peromyscus maniculatus* from those 10 occasions (Fig. 4) were analyzed by the discrimination procedure described in the section entitled TESTS OF MODEL ASSUMPTIONS and it was determined that Model M_b would be an appropriate model for the data (the significance level of a goodness of fit test of Model M_b was approximately 0.47). Therefore, the estimation procedure associated with Model M_b was used to produce point and interval estimates for N . Two aspects of those estimates (Fig. 4) are worth noting: animals tend to become trap happy since recapture probability (\hat{c}) is more than twice the value of initial capture probability (\hat{p}) and the small value of $\hat{p} = 0.09$ is the major reason why the 95 percent confidence interval for N is wide and the lower limit of the computed interval extends far below the number of animals actually seen (as mentioned be-

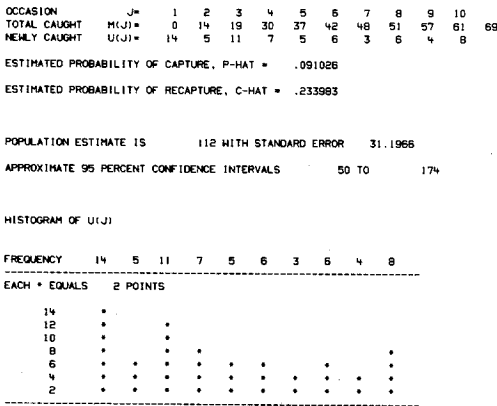


FIG. 4. Example of population estimation with constant probability removal estimator under Model M_b with deer mouse data from E. Larsen (pers. comm.).

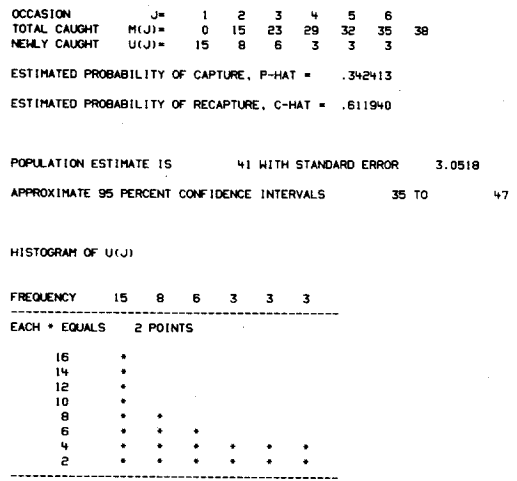


FIG. 5. Example of population estimation with constant probability removal estimator under Model M_b with deer mouse data from V. Reid (pers. comm.).

fore, the lower limit should be taken as 69, not 50, in this situation).

Example

V. Reid (pers. comm.) reported the results of livetrapping deer mice *Peromyscus maniculatus* in a drainage bottom of sagebrush, gambel oak, and serviceberry with pinyon pine and juniper on the uplands. The area, in Rio Blanco County, Colorado, was trapped for 6 consecutive nights in the summer of 1975. Traps were arranged in a 9 x 11 grid and spaced 50 feet (15.2 m) apart. Analysis of the resulting data by the model selection procedure indicated that Model M_b would be appropriate for use in estimating N (e.g., the significance level of a goodness of fit test of Model M_b was approximately 0.43). The data and the results of using \hat{N}_b to produce point and interval estimates for N are given in Fig. 5. Note that the mice tend to become trap happy, as in the previous example. In that study, however, the estimate of first capture probability is $\hat{p} = 0.34$, a fact that largely accounts for the narrow width of the 95 percent confidence interval. Again, the lower limit of the confidence interval could be taken as 38, the number of different animals captured.

Discussion

Simulation results concerning the performance of \hat{N}_b in populations that satisfy the assumptions of Model M_b seem encouraging in that the estimator and its associated confidence intervals provide useful information in the presence of behavioral response to first capture, if the probability of first capture is sufficiently large (>0.1). (Hence, calculation of the ML estimate of p should assist in assessing the amount of information concerning N that is contained in the data). However, the estimator of N appears nonrobust to other factors that may have an effect on an animal's probability of capture. Moreover, the estimation procedure associated with Model M_b is a special case of a more general estimation procedure which involves fewer assumptions than those associated with Model M_b . This procedure is described in the section on removal models. For these reasons, the estimator of population size N associated with Model M_b is useful only in those instances where all unmarked animals have the same capture probability on all trapping occasions.

MODEL M_h : CAPTURE PROBABILITIES
VARY BY INDIVIDUAL ANIMAL

Structure and Use of the Model

Assumptions and Parameters

The assertion that each member of the population has its own probability of capture independent of all other members of the population is the basis of Model M_h . The assumptions are made that there is no difference between trapping occasions and no behavioral response to capture, but that there is heterogeneity among the capture probabilities of individuals. Cormack (1968) stated that a test for the assumption of "heterogeneity of individuals" is impossible unless an independent experiment is conducted on a population of known size that is "representative" of the population of interest. Because of that apparent difficulty (and others), models similar to Model M_h have been largely ignored, although the assumptions behind the model fit more experimental situations than do most of those commonly used. As Eberhardt (1969a) pointed out, "various sets of data indicate . . . that the equal-probability-of-capture assumption is not fulfilled."

Conceptually, Model M_h involves $N + 1$ parameters: the population size N and the set of capture probabilities $\{p_i\}$, $i = 1, 2, \dots, N$, where p_i is the probability of capture of the i^{th} animal on any trapping occasion. This more general formulation of Model M_h does not allow estimation of population size N because of the presence of too large a number of nuisance parameters. Therefore, for most of the development in this chapter, it is more useful to think of $\{p_i\}$ as a random sample of size N from some probability distribution $F(p)$ defined on the interval $[0, 1]$.

Unfortunately, there is no completely satisfactory estimator when Model M_h is true. Therefore, one should design the study to minimize heterogeneity. In fact, it is worth repeating that it is very important to design any capture study carefully with the goal of having as few factors as possible affecting capture probabilities (see STUDY DESIGN).

Statistical Treatment

Recall that the random variable X_{ij} takes on the value 1 when the i^{th} animal is caught on the j^{th} trapping occasion and is zero otherwise. If it is assumed that $\{p_i\}$ results from a random sample from $F(p)$ then the probability distribution for the set of variables $\{X_{ij}\}$ reveals that a sufficient statistic for Model M_h is given by the capture frequencies $\{f_1, f_2, \dots, f_j\}$, where $f_j =$ the number of animals caught exactly j times in the experiment. Hence, all the information for estimating N is contained in the frequency of capture statistics. Estimators based on frequency of capture statistics have been proposed (Craig 1953, Tanton 1965, Eberhardt 1969a), but are the result of essentially ad hoc approaches. If the particular family of distributions of which $F(p)$ is a member is specified, a maximum likelihood estimator for N may be derived. For example, Burnham (unpublished dissertation) assumed that $\{p_i\}$ results from a random sample from a 2-parameter beta distribution and investigated the properties of the resulting ML estimator of N . Theoretical and simulation results indicated that this estimator has quite unsatisfactory operating characteristics, and this inspired Burnham to develop a nonparametric estimation procedure that would not require specification of $F(p)$. Such an estimator was developed using an extension of the jackknife method of bias reduction first proposed by Quenouille (1949, 1956). The resulting form of the estimator \hat{N}_h can be written as

$$\hat{N}_h = \sum_{j=1}^t a_j f_j.$$

Each set of constants a_1, a_2, \dots, a_t , generated by using a different "order" of the jackknife procedure, corresponds to a different "jackknife" estimator. Burnham (unpublished dissertation) suggested a sequence of statistical tests designed to choose the best one of the estimators for any given data set. For a more detailed presentation of that estimation procedure, consult Appendix E which also pre-

TABLE 6.—RESULTS OF COMPUTER SIMULATION STUDIES OF THE BIAS OF \hat{N}_h (ALSO SEE APPENDIX, TABLE N.4.b)

Percent relative bias RB	Average of estimates Ave(N_h)	True population N	Probabilities of capture $p_i, i = 1, 2, \dots, N$	Number of reps R	Number of occasions t	Trial
-17.2	331.1	400	$p_1 = 0.05, i = 1,200; p_1 = 0.15, i = 201,300;$ $p_1 = 0.50, i = 301,400.$	200	5	1
0.2	100.2	100	$p_1 = 0.05, i = 1,40; p_1 = 0.10, i = 41,80;$ $p_1 = 0.30, i = 81,100.$	200	10	8
15.0	460.1	400	$p_1 = 0.10, i = 1,100; p_1 = 0.20, i = 101,200;$ $p_1 = 0.25, i = 201,300; p_1 = 0.30, i = 301,400.$	200	5	3
4.3	417.1	400	$p_1 = 0.01, i = 1,50; p_1 = 0.15, i = 51,200;$ $p_1 = 0.25, i = 201,300; p_1 = 0.30, i = 301,400.$	100	5	4
3.7	207.0	200	$p_1 = 0.05, i = 1,50; p_1 = 0.15, i = 51,150;$ $p_1 = 0.25, i = 151,200.$	100	10	6
11.0	443.9	400	$p_1 = 0.20, i = 1,100; p_1 = 0.30, i = 101,200;$ $p_1 = 0.40, i = 201,300; p_1 = 0.50, i = 301,400.$	100	5	5
13.4	226.8	200	$p_1 = 0.15, i = 1,50; p_1 = 0.20, i = 51,100;$ $p_1 = 0.25, i = 101,150; p_1 = 0.30, i = 151,200.$	100	7	7

sents a formula for a variance estimator of \hat{N}_h .

Simulation Results

Bias

Cormack (1968:497) noted that when animals exhibit heterogeneous capture probabilities "any available estimate of population size [is] markedly biased." We performed simulation experiments with Model M_h populations to determine whether, in a given population, "marked bias" is associated with \hat{N}_h . Table 6 lists some examples of such results. The complete details of the simulation are given in Tables N.4.a and N.4.b of Appendix N. Results of those simulations and of those conducted by Burnham (unpublished dissertation) seem to indicate 2 major points concerning the bias of the jackknife estimator: (1) general statements concerning the magnitude and direction of the bias are not easily made since the estimator exhibits both small and large and positive and negative bias, depending on the values of N and the set $\{p_i\}$. One may conjecture, however, that if the number of trapping occasions is sufficiently large (say greater than 5) and if a negligible number of animals is for all practical purposes untrappable, then the bias of \hat{N}_h will be within a tolerable

range; and (2) even though significant bias may sometimes be present in the jackknife estimator, the magnitude of such bias will, in general, be substantially less than the bias that would be associated with any other estimator thus far discussed. Evidence of this assertion can be seen in Tables N.1.b, N.2.b, and N.3.b of Appendix N that present the results of using estimators other than \hat{N}_h on data simulated from Model M_h .

Confidence Intervals

Construction of confidence intervals from simulated experimental data generated from populations of Model M_h show a large variance in the achieved confidence coefficients (Table 7). Coverage is poor and ranges from an estimated 0 percent (Trials 2, 11) to 87 percent (Trial 6) (Appendix N, Table N.4.b). Because results indicate that the distribution of the jackknife estimator is approximately normal, poor coverage is caused either by significant negative bias of the variance estimator, a significant bias of the estimator \hat{N}_h , or both. Unless an experimenter has data from an adequate number of trapping occasions and feels that very few, if any, members of the population under study are untrappable, the confidence interval constructed by the

TABLE 7.—SIMULATED CONFIDENCE INTERVAL WIDTHS AND COVERAGE USING THE JACKKNIFE ESTIMATOR \hat{N}_h (ALSO SEE APPENDIX N, TABLE N.4.b)

Average [C.I. width]	Coverage	Population size N	Probabilities of capture $p_i, i = 1, 2, \dots, N$	Number of reps R	Number of occasions t	Trial
89.9	0.18	400	$p_1 = 0.05, i = 1,200; p_1 = 0.15, i = 201,300;$ $p_1 = 0.50, i = 301,400.$	200	5	1
104.9	0.40	400	$p_1 = 0.10, i = 1,100; p_1 = 0.20, i = 101,200;$ $p_1 = 0.25, i = 201,300; p_1 = 0.30, i = 301,400.$	200	5	3
45.4	0.87	200	$p_1 = 0.05, i = 1,50; p_1 = 0.15, i = 51,150;$ $p_1 = 0.25, i = 151,200.$	100	10	6
45.3	0.68	100	$p_1 = 0.05, i = 1,40; p_1 = 0.10, i = 41,80;$ $p_1 = 0.30, i = 81,100.$	200	5	8

methods associated with Model M_h should be considered unreliable. The reader may refer to Table N.4.b of Appendix N for complete results concerning simulated coverage of these confidence intervals.

Robustness

The jackknife estimator was constructed with the objective of having an estimator that is robust to heterogeneity of capture probabilities. The degree to which the objective is satisfied has been discussed previously. The question remains: is the jackknife estimator robust to other forms of departure from the assumption of equal catchability? Results indicate that of the estimators discussed in this monograph the jackknife is the most robust. This does not mean that it is a good estimator no matter what assumptions apply to the population under study. In general, it may be said that the

jackknife estimator will provide an adequate estimate of population size in an experiment in which many animals are caught a relatively large number of times. Some examples of the bias of the jackknife estimator when data are simulated from models other than M_h are given in Table 8. More information may be obtained from Tables N.2.b, N.3.b, and N.5.b of Appendix N. In the table, the information under the heading Data model refers to the population on which the simulated experiments were performed. For details of these populations, see Appendix N, Tables N.2.a, N.3.a, and N.5.a.

Example

Carothers (1973a) conducted a capture-recapture experiment on the "population" of taxicabs in Edinburgh, Scotland. The population, known to be of size 420 and assumed to be closed, was sam-

 TABLE 8.—COMPUTER SIMULATION RESULTS ILLUSTRATING THE ROBUSTNESS OF THE JACKKNIFE ESTIMATOR \hat{N}_h TO DATA GENERATED UNDER OTHER MODELS (ALSO SEE APPENDIX N, TABLES N.2.b, N.3.b AND N.5.b)

Percent relative bias RB	Ave[\hat{N}_h]	Population size N	Number of reps R	Number of occasions t	Data model
-12.5	349.2	400	200	5	M_t , Trial 3 ¹
- 1.3	98.7	100	1,000	5	M_t , Trial 12
-17.6	82.3	100	100	5	M_b , Trial 9
-13.1	347.7	400	200	5	M_b , Trial 4
0.7	402.9	400	100	5	M_{th} , Trial 1
19.7	478.7	400	100	5	M_{th} , Trial 2

¹ For example, data generated under Model M_t were used to estimate N using the estimator \hat{N}_h .

NUMBER OF TRAPPING OCCASIONS WAS 10
 NUMBER OF ANIMALS CAPTURED, N(I+1), WAS 283
 TOTAL NUMBER OF CAPTURES, N., WAS 500

FREQUENCIES OF CAPTURE, F(I)
 I = 1 2 3 4 5 6 7 8 9 10
 F(I) = 142 81 49 7 3 1 0 0 0 0

COMPUTED JACKKNIFE COEFFICIENTS

	N(1)	N(2)	N(3)	N(4)	N(5)
1	1.900	2.700	3.400	4.000	4.500
2	1.000	.269	-.878	-2.278	-3.722
3	1.000	1.000	1.476	2.535	4.042
4	1.000	1.000	1.000	.743	.077
5	1.000	1.000	1.000	1.000	1.103

THE RESULTS OF THE JACKKNIFE COMPUTATIONS

I	N(I)	SE(I)	.95 CONF. LIMITS	TEST OF N(I+1) VS. N(I)
0	283			25.877
1	410.8	15.58	380.3 441.3	4.225
2	466.8	25.20	417.4 516.2	1.807
3	495.0	36.42	423.7 566.4	1.735
4	516.9	49.98	418.9 614.9	0.000
5	540.4	65.29	412.4 668.4	

AVERAGE P-HAT = .1052

INTERPOLATED POPULATION ESTIMATE IS 471 WITH STANDARD ERROR 26.8266
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 418 TO 524

HISTOGRAM OF F(I)

FREQUENCY	142	81	49	7	3	1	0	0	0	0
EACH * EQUALS	15	15	15	15	15	15	15	15	15	15

NUMBER OF TRAPPING OCCASIONS WAS 6
 NUMBER OF ANIMALS CAPTURED, N(I+1), WAS 68
 TOTAL NUMBER OF CAPTURES, N., WAS 145

FREQUENCIES OF CAPTURE, F(I)
 I = 1 2 3 4 5 6
 F(I) = 25 22 13 5 1 2

COMPUTED JACKKNIFE COEFFICIENTS

	N(1)	N(2)	N(3)	N(4)	N(5)
1	1.833	2.500	3.000	3.333	3.500
2	1.000	.467	-.233	-.833	-1.167
3	1.000	1.000	1.225	1.542	1.750
4	1.000	1.000	1.000	.356	.914
5	1.000	1.000	1.000	1.000	1.001

THE RESULTS OF THE JACKKNIFE COMPUTATIONS

I	N(I)	SE(I)	.95 CONF. LIMITS	TEST OF N(I+1) VS. N(I)
0	68			CHI-SQUARE(1 D.F.)
1	88.8	6.18	76.7 100.9	1.410
2	93.8	9.40	75.4 112.2	.000
3	93.9	12.95	69.0 119.6	.078
4	92.8	15.45	62.5 123.1	.118
5	92.2	17.06	58.7 125.6	0.000

AVERAGE P-HAT = .2778

INTERPOLATED POPULATION ESTIMATE IS 87 WITH STANDARD ERROR 5.8220
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 75 TO 99

HISTOGRAM OF F(I)

FREQUENCY	25	22	13	5	1	2
EACH * EQUALS	3	3	3	3	3	3

FIG. 6. Example of population estimation with variable probability of capture by animal under Model M_h with the complete set of Scheme A taxicab data from Carothers (1973a).

FIG. 7. Example of population estimation with variable probability of capture by animal under Model M_h with snowshoe hare data from Burnham and Cushwa (pers. comm.).

pled for 10 consecutive days. The sampling was done by observing taxicabs that passed preselected points and recording "capture" or "recapture" according to whether or not the particular cab had been observed previously. In the author's sampling scheme "A," a different set of sampling points was selected each day, and the time of sampling (i.e., morning, afternoon, night) was also varied. We applied the model selection procedure (described in TESTS OF MODEL ASSUMPTIONS) to the data collected from this scheme. The procedure indicated that an appropriate model for the data would be Model M_h . (For example, a goodness of fit test of Model M_h has a significance level greater than 0.99). Thus, the estimation procedure associated with Model M_h was used to produce point and interval estimates for N (Fig. 6). Although \hat{N}_h overestimates the true value of N by approximately 12 percent, the confidence

interval constructed for N does in fact contain the true population value of 420. These results are as satisfactory as any produced by various estimation techniques used by Carothers.

Example

In 1972, Burnham and Cushwa (pers. comm.) laid out a livetrapping grid in a black spruce forest 30 miles (48.3 km) north of Fairbanks, Alaska. The basic grid was 10 x 10, with traps spaced 200 feet (61 m) apart. Trapping for snowshoe hares *Lepus americanus* was carried out for 9 consecutive days in early winter. Traps were not baited for the first 3 days, and therefore we have chosen to analyze the data from the last 6 days of trapping. The model selection procedure designed to identify a proper model for estimating N chose Model M_h as the most appropriate model for the data. A goodness of fit

test of Model M_h had a significance level of 0.06. The data, and the estimates produced by the estimation procedure of Model M_h are given in Fig. 7. The results of the entire 9 days of trapping show 74 different animals caught while the lower 95 percent confidence limit for N given in Fig. 7 is 75. Although this result is satisfying, we reemphasize that the true confidence coefficient of the confidence intervals associated with Model M_h is, for most populations, much less than the stated level of 0.95.

Discussion

Theoretically, 2 main complicating factors are associated with the existence of heterogeneity of individual capture probabilities. First, a parameter must be introduced into the model for every individual in the population, and this makes straightforward statistical estimation of population size impossible. Second, the assumption of heterogeneity can render some individuals nearly "invisible" with respect to any estimation procedure based on marking methods because such individuals have nearly zero catchability. The jackknife estimation procedure is the most robust procedure thus far proposed for dealing with problems associated with heterogeneity. Moreover, the estimator seems robust to other factors that may have an effect on capture probabilities. It should be pointed out, however, that in some cases the bias of the estimator is not negligible. In general, if the jackknife procedure is to be considered a good nonparametric and robust technique, the trapping must be performed on a large number of occasions, and the number of recaptures on each occasion must be substantial.

We remark that the observations contained in this section reflect Gilbert's (1973:524) opinion that "the presence of heterogeneity per se is not as important as the particular pattern of heterogeneity and whether the heterogeneity is near zero or one."

MODEL M_{tb} : CAPTURE PROBABILITIES VARY BY TIME AND BEHAVIORAL RESPONSE TO CAPTURE

Structure and Use of the Model

Assumptions and Parameters

If it is assumed that an animal's probability of capture changes after initial capture and that temporal changes also have an effect on capture probabilities, Model M_{tb} results. This model has been conceptualized in the literature (see Eberhardt et al. 1963, Seber 1973) but it has not received mathematical treatment with respect to estimation of population parameters. Although many variations of the basic model exist (see Appendix F), for purposes of this chapter we shall define the parameters of Model M_{tb} as:

- N = population size,
- p_j = the probability an unmarked animal is captured on the j^{th} trapping occasion, $j = 1, 2, \dots, t$,
- c_j = probability that a marked (previously captured) animal is captured on the j^{th} trapping occasion, $j = 2, 3, \dots, t$.

Note that again, as in Model M_b , an animal exhibits only 1 behavioral response to capture after its initial capture.

Statistical Treatment

The joint probability distribution of the set of possible capture histories $\{X_\omega\}$ can be written as

$$P\{\{X_\omega\}\} = \frac{N!}{\left[\prod_{\omega} X_\omega!\right] (N - M_{t+1})!} \prod_{j=1}^t p_j^{u_j} \cdot (1 - p_j)^{N - M_{j+1}} (c_j)^{m_j} (1 - c_j)^{M_j - m_j},$$

where

- u_j = number of unmarked animals caught on the j^{th} occasion, $j = 1, 2, \dots, t$,
- M_j = number of marked animals in the population the the time of the j^{th} sample, $j = 1, 2, \dots, t$, and
- m_j = number of marked animals captured on the j^{th} occasion, $j = 1, 2, \dots, t$.

A minimal sufficient statistic for this distribution is given by $\{u_1, u_2, \dots, u_t, m_2, m_3, \dots, m_t\}$ which has dimension $2t - 1$. Since the model involves $2t$ parameters, not all parameters can be estimated and maximum likelihood estimation of N proves to be impossible. The structure of $P[\{X_{\omega}\}]$ reveals that estimation of the parameters N, p_1, p_2, \dots, p_t depends only upon the vector of removals $\{u_1, u_2, \dots, u_t\}$. Thus, estimation of N in Model M_{th} would be equivalent to the estimation in Zippin's (1956) removal model generalized to allow the probability of removing an animal, p_j , to be different for every trapping occasion. This involves estimating $t + 1$ parameters with t statistics. Hence, in order to make N an identifiable parameter, one has to make the assumption that at least 2 of the p_j are equal. Because there appears to be no biological justification for making such an assumption, only ad hoc estimation procedures are available. For example, Tanaka (1951, 1952) has proposed a regression technique that involves plotting $y_j = m_j/n_j$ versus M_j on a logarithmic scale; however, Seber (1973) pointed out there are serious problems of interpretation involved in using this technique. It is true that graphical techniques such as this may possess some utility in that they encourage the researcher to examine the data carefully. However, regression methods are not contained in the class of estimation methods considered in this monograph. Therefore, Tanaka's (1951, 1952) method and those similar to it will not be considered here.

Discussion

In view of the previous discussion, we believe that Model M_{th} is not useful for estimation purposes. However, the model does have utility with respect to the problem of choosing a "best" estimation procedure given the data at hand. For details of how the model is used in such a testing procedure see the section on model selection.

MODEL M_{th} : CAPTURE PROBABILITIES VARY BY TIME AND INDIVIDUAL ANIMAL

Structure and Use of the Model

Assumptions and Parameters

If, on the j^{th} trapping occasion, the i^{th} animal has a capture probability of $p_{ij} = p_i p_j$ that is independent of its capture history (i.e., there is no behavioral response to capture), then Model M_{th} is the appropriate probability model for a capture-recapture experiment on such a population. Notice that the structure of p_{ij} implies that variation in capture probabilities due to time is independent of the variation caused by individual heterogeneity. In effect, this means that the factors responsible for time variation, e.g., environmental conditions, affect all members of the population similarly. For purposes of this section it shall be assumed that the $p_i, i = 1, 2, \dots, N$, are a random sample of size N from some probability distribution-function $F(p; \theta)$ that is parameterized by the vector θ and defined for p in the interval $[0, 1]$. Note that it is necessary that each of the parameters, $p_j, j = 1, 2, \dots, t$, be subject to the constraint that $0 \leq p_i p_j \leq 1$ for $i = 1, 2, \dots, N$. Conceptually, the form of the distribution function $F(p; \theta)$ and the parameters $\theta, p_1, p_2, \dots, p_t$ are unknown elements of the model, as is the parameter N .

Statistical Treatment

Estimation of population size under the assumptions of Model M_{th} has not been considered in the literature. If one is willing to completely specify the distribution $F(p; \theta)$, then maximum likelihood estimation of N becomes possible (cf. Appendix G). In general, such an assumption is not realistic. Furthermore, simulation results of Burnham (unpublished dissertation) indicate that an ML estimator resulting from such an assumption may have poor operating characteristics.

If no assumptions concerning $F(p; \theta)$ are made, ML estimation of N is not pos-

TABLE 9.—COMPUTER SIMULATION OF THE BIAS ASSOCIATED WITH THE ESTIMATORS \hat{N}_h AND \hat{N}_t FOR DATA GENERATED UNDER MODEL M_{th} (ALSO SEE APPENDIX N, TABLE N.5.b)

Percent relative bias RB	Ave[\hat{N}_h]	Percent relative bias RB	Ave[\hat{N}_t]	Population size N	Number of reps R	Number of occasions t	Trial ¹
0.7	402.9	-24.1	303.6	400	100	5	1
19.7	478.7	- 7.5	369.9	400	100	5	2
-11.5	353.9	-22.8	309.0	400	100	5	3
-11.8	352.8	-31.8	272.8	400	100	5	4
8.9	217.7	- 6.1	187.9	200	100	7	5
-10.9	178.2	-17.9	164.2	200	100	5	6

¹ The probability of capture for data generated under Model M_{th} is complex; therefore, the specific values are given in Appendix N, Table N.5.a.

sible. At present, we are not aware of any rigorous estimation technique that is appropriate for estimation in Model M_{th} .

Simulation Results

Bias

To gain some insight into the performance of estimators associated with models closely related to Model M_{th} , we simulated capture-recapture experiments on populations that satisfy the assumptions of Model M_{th} and calculated the estimators associated with Model M_t and Model M_h from the resulting data. Examples are given in Table 9. See Tables N.5.a and N.5.b of Appendix N for a description of the populations used and for results of other simulations. A choice between the 2 estimators with respect to performance is obviously not clearcut. The magnitude of the bias of \hat{N}_h seems in general to be less than that of \hat{N}_t . However, direction of bias of \hat{N}_h is not consistent, whereas the estimator \hat{N}_t appears to consistently exhibit negative bias. Thus, if the data seem to fit Model M_{th} , computing both estimates of N should give the experimenter some sense of the size of the population. However, it should be emphasized that the model does not seem particularly useful in providing a single satisfactory estimate of the value of N.

Confidence Intervals

Coverage of confidence intervals constructed both on the basis of the esti-

mators of Model M_h and Model M_t was poor (cf. Table N.5.b, Appendix N). The lack of coverage is due, in general, to the significant bias present in the estimates. Therefore, we recommend that confidence intervals constructed from data that apparently fit Model M_{th} be considered only as a *very crude indication* of the possible values of N and not as a valid statement of inference.

Robustness

A discussion of the robustness of the estimator associated with Model M_{th} is somewhat inappropriate because no single estimator is exclusively associated with the model. The robustness of \hat{N}_h and \hat{N}_t has been discussed previously. Those discussions indicated that the former estimator is relatively robust and that the latter is not. Such assertions seem to be supported by the results of this section concerning bias.

Discussion

Model M_{th} is one of the most difficult models dealt with in this monograph. That difficulty is reflected in the fact that no estimation procedure can be proposed that is specifically suited to all assumptions of the model. Estimators associated with 2 previous models have been investigated for possible use but neither seems completely adequate. Thus, at present, satisfactory techniques for constructing point and interval estimates of N are not

available. Hopefully, further research involving Model M_{th} will be forthcoming so that the model can be dealt with in a more adequate statistical manner. Until then, one must be satisfied with obtaining only very rough estimates of population size when Model M_{th} is the appropriate representation of the experiment.

MODEL M_{bh} : CAPTURE PROBABILITIES VARY BY INDIVIDUAL ANIMAL AND BY BEHAVIORAL RESPONSE TO CAPTURE

Structure and Use of the Model

Assumptions and Parameters

Model M_{bh} assumes that every member of the population has a specific pair of capture probabilities: p_i , the probability that the i^{th} animal is caught on any trapping occasion given that it has not been previously captured; and c_i , the probability that the i^{th} animal is recaptured given that it has been captured at least once previously. Thus, the model allows both behavioral response to first capture and individual heterogeneity of capture probabilities. An important and appealing characteristic of the model is that it allows the behavioral response to capture to vary with the animal, i.e., all members of the population do not exhibit an identical response to first capture.

The most general formulation of the model involves $2N + 1$ total parameters consisting of N (population size), and 2 capture probabilities for each of the N members of the population. Obviously, no real experiment will be able to provide enough information for the explicit estimation of this number of parameters. Further assumptions are necessary before estimation of population size is possible.

Statistical Treatment

Pollock (unpublished dissertation) introduced Model M_{bh} and assumed that the pairs (p_i, c_i) are a random sample of size N from a bivariate probability distri-

bution function $G(p, c; \theta)$ that is parameterized by the vector of unknown constants θ . If one assumes that the form of the distribution $G(p, c; \theta)$ is known and if the dimension of θ is no larger than $[t(t + 1)/2] - 1$, where t is the number of trapping occasions, then theoretically maximum likelihood estimation could be used to provide an estimator of N . However, a "nonparametric" approach is also possible. We discuss this approach in the context of the assumption that the bivariate density function $G(p, c; \theta)$ can be factored as $G_1(p; \theta_1) G_2(c; \theta_2)$. That is, an individual's initial capture probability and its probability of recapture are independent. With that assumption, Pollock (unpublished dissertation) showed that the probability distribution of the set of possible capture histories $\{X_\omega\}$ can be written as

$$P\{\{X_\omega\}\} = \frac{N!}{u_1!u_2! \cdots u_t!(N - M_{t+1})!} \cdot \pi_1^{u_1}\pi_2^{u_2} \cdots \pi_t^{u_t} \left(1 - \sum_{j=1}^t \pi_j\right)^{N - M_{t+1}} \cdot P^*\{\{X_\omega\} | u_1, u_2, \dots, u_t\}$$

where

u_j = number of unmarked animals caught at time j , $j = 1, 2, \dots, t$,

$$\pi_j = E[(1 - p)^{j-1}p] = \int_0^1 (1 - p)^{j-1}p \, dG_1$$

$j = 1, 2, \dots, t$, and

$P^*\{\{X_\omega\} | u_1, u_2, \dots, u_t\} =$
a conditional probability distribution that does not depend upon the parameter N or the distribution $G_1(p; \theta_1)$.

(Refer to Appendix H for further details). Thus, the "removals" $\{u_1, u_2, \dots, u_t\}$ are the relevant statistics for the purpose of estimating N and the parameters θ_1 in $G_1(p; \theta_1)$, and these statistics have a multinomial distribution with parameters N and $\pi_1, \pi_2, \dots, \pi_t$ (Seber 1973:316 gave this model for the special case $t = 3$). We

can transform the π_j by writing $\pi_j = (1 - \bar{p}_1)(1 - \bar{p}_2) \cdots (1 - \bar{p}_{j-1})\bar{p}_j$, $j = 1, \dots, t$, where \bar{p}_j is the average conditional probability of capture on the j^{th} occasion for those animals not previously captured. The distribution of the t -dimensional vector of removals now depends upon the $t + 1$ parameters N , $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_t$, and thus all the parameters of the model are not identifiable. This is similar to the situation that resulted in Model M_{tb} where it was concluded that ML estimation of N is not possible. If individual probabilities of first capture vary over the population, however, it is logical to assume that $\bar{p}_1 > \bar{p}_2 > \bar{p}_3 > \cdots > \bar{p}_t$, because the individuals with high first capture probability will tend to be removed first, the animals with slightly lower first capture probabilities removed second, and so on. Moreover, the assumption that $(\bar{p}_1 - \bar{p}_2) > (\bar{p}_2 - \bar{p}_3) > \cdots > (\bar{p}_{t-1} - \bar{p}_t)$, i.e., the bigger differences in the conditional probabilities of removal occur in the initial stages of the experiment, does not seem illogical. These assumptions are the basis for the development of an estimation procedure appropriate for Model M_{bh} . The estimator \hat{N}_{bh} derived under Model M_{bh} is obtained by sequentially testing (through goodness of fit tests) for differences among the \bar{p}_j . The process begins with testing whether or not all the \bar{p}_j are equal. If not, we let \bar{p}_1 be different and test whether or not $\bar{p}_2 = \bar{p}_3 = \cdots = \bar{p}_t$. This testing continues until it is concluded that the last $t - k + 1$ capture probabilities are not significantly different, whereas the first k capture probabilities do differ. In the case $k = 1$, we are saying the simple Model M_b (all \bar{p}_j equal) adequately fits the removals. For $k > 1$ we are concluding Model M_b does not fit the data, due to the presence of heterogeneity (or time variation in capture probabilities). The estimator \hat{N}_{bh} is the ML estimator under the selected model.

The above technique, called the generalized removal method, is further described in Appendix H. The appendix also points out that it is not necessary that the factorization $G(p, c; \theta) = G_1(p; \theta_1)$

$G_2(c; \theta_2)$ be possible in order to use the generalized removal method to estimate population size. It is important to realize that this method is a generalization of Zippin's (1956, 1968) removal method which assumes no variation in first capture probabilities. Therefore, the fact that the generalized removal method has utility in removal experiments as well as capture-recapture experiments is not surprising (cf. REMOVAL MODELS).

Finally, we mention that the method can "fail" if a mathematical criterion involving the removals u_1, u_2, \dots, u_t is not satisfied. This failure criterion, similar to the one involved in Zippin's removal method, ensures that a sufficient decline in the number of newly captured animals is being effected by successive trapping occasions. A formula expressing this criterion is given in Appendix H.

Simulation Results

Bias

Experiments were simulated on populations behaving according to the assumptions of Model M_{bh} . The estimator \hat{N}_{bh} was calculated to develop some idea of the bias involved. Results indicate that relative bias ranges from 3 to 38 percent for populations used in the simulations. That range is somewhat misleading however, because for all populations except one (in which half the population was essentially untrappable), the range of bias was 3 to 15 percent. Some examples are given in Table 10. The reader is referred to Tables N.6.a and N.6.b of Appendix N for further results and for descriptions of the populations. In general, it seems there is no serious bias in \hat{N}_{bh} if relatively few members of the population are essentially uncatchable (i.e., probability of first capture less than 0.05) and the number of trapping occasions is adequate. (Recall that since the estimation technique depends on removals only, probabilities of recapture have no effect on the performance of \hat{N}_{bh} .) Considering the complicated model structure and the assumptions required to produce the

TABLE 10.—COMPUTER SIMULATION OF THE BIAS ASSOCIATED WITH THE ESTIMATOR \hat{N}_{bh} FOR DATA GENERATED UNDER MODEL M_{bh} (ALSO SEE APPENDIX N, TABLE N.6.b)

Percent relative bias RB	Ave[\hat{N}_{bh}]	Population size N	Number of reps R	Number of occasions t	Trials ¹
-38.4	246.6	400	100	5	1
-14.8	340.8	400	100	5	2
- 5.9	94.1	100	100	5	7
-12.3	175.5	200	100	10	5
- 3.2	193.7	200	100	7	6
- 4.3	383.0	400	100	5	4

¹ The probability of capture for data generated under Model M_{bh} is complex; therefore, the specific values are given in Appendix N, Table N.6.a.

generalized removal estimator, these results are encouraging with respect to practical use of the estimator.

Confidence Intervals

Simulation results are not so encouraging, however, when it comes to placing a confidence interval on N . The variance estimator associated with \hat{N}_{bh} appears ill behaved and the distribution of \hat{N}_{bh} is nonnormal. Those factors resulted in simulated confidence coefficients averaging approximately 0.50. Further work is necessary to develop useful confidence intervals for N when Model M_{bh} seems appropriate because normal theory confidence intervals apparently are inappropriate for practical use.

Robustness

The generalized removal estimator can be expected to be robust to Models M_h and M_b , since both are special cases of Model M_{bh} . If capture probabilities exhibit significant variation over time, the utility of the method is undoubtedly decreased. That is particularly true if probabilities tend to increase with time, although the failure criterion should help to provide a safeguard against use of the method when it is inappropriate. In general, it is expected that the generalized removal method used in the context of a capture-recapture experiment will pro-

vide reasonable estimates of population size if the number of unmarked animals captured on each trapping occasion exhibits a definite decrease over time. Such a trend indicates that conditional probabilities of capture on the j^{th} occasion are reasonably large, which is the primary requirement for adequate estimation of population size.

Example

E. Larsen (pers. comm.) provided the results of livetrapping Great Basin pocket mice *Perognathus parvus* in a desert community, Curlew Valley, Utah. The area was trapped for 7 consecutive nights during June 1977. Traps were arranged in a 12×12 grid and spaced 15 m apart. Analysis of the resulting data by the model selection procedure (described in TESTS OF MODEL ASSUMPTIONS) indicated that either Model M_{tbh} or Model M_{bh} would be appropriate for use in estimating N . The data on first capture and the results of using \hat{N}_{bh} to produce point and interval estimates for N are given in Fig. 8. Note that the estimator rejects the model for $k = 1$, but accepts $k = 2$. Also of interest is that the estimation method fails for $k = 3$, but estimates for $k = 4$ and $k = 5$ are produced that appear reasonable. The estimated variance of \hat{N}_{bh} is quite large for $k = 2$. Possibly time variation is quite important and causes this imprecision. The probabilities of capture decline for $k = 2$, as they should under Model M_{bh} . A histogram of the u_j values is included to help the researcher visually examine his data.

Discussion

Model M_{bh} is one of the most realistic and useful models for a capture-recapture experiment considered in this monograph. Realism results from allowing every animal to possess its own probability of first capture and from not constraining the members of the population so that every individual's behavioral response to first capture is identical. The

OCCASION	J=	1	2	3	4	5	6	7	
TOTAL CAUGHT	M(J)=	0	23	32	35	40	46	52	55
NEWLY CAUGHT	U(J)=	23	9	3	5	6	6	3	

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, . . . , 7						
1	60.61	4.456394	9.965	.0762	.2802	.2802	.2802	.2802	.2802	.2802	.2802
2	76.73	24.66909	3.486	.4801	.2998	.1377	.1377	.1377	.1377	.1377	.1377
3	FAILURE CRITERION = -2 NO ESTIMATES FOR THIS STEP.										
4	67.44	18.18989	1.308	.5201	.3410	.2025	.0846	.2067	.2067	.2067	.2067
5	58.65	5.896785	.823	.3643	.3922	.2525	.1126	.2114	.3953	.3953	.3953

POPULATION ESTIMATE IS 77 WITH STANDARD ERROR 24.6691
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 28 TO 126

HISTOGRAM OF U(J)

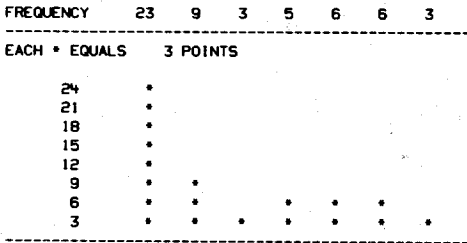


FIG. 8. Example of population estimation with variable probability removal estimator under Model M_{bh} with pocket mouse data from E. Larsen (pers. comm.).

model is useful because an estimation procedure based on the assumptions of the model is available for estimating population size. Although the operating characteristics of the estimator are not completely satisfactory, the estimator would seem superior to any other that might be used if, in fact the assumptions of Model M_{bh} are satisfied. Moreover, the method is nonparametric in the sense that one does not have to specify a probability distribution for the pairs (p_i, c_i) .

MODEL M_{tbh} : CAPTURE PROBABILITIES VARY BY BEHAVIORAL RESPONSE TO CAPTURE, TIME, AND INDIVIDUAL ANIMAL

Discussion

For purposes of this monograph, Model M_{tbh} is useful only conceptually. Because all 3 factors that may affect capture probabilities are assumed to be operating

in Model M_{tbh} , every other model considered here is of course a special case of this model. Thus it may be useful to think of Model M_{tbh} as the most realistic of all models that assume population closure. Unfortunately, this extreme realism precludes the model from having any practical use in the estimation of population size. If the experimenter is led to believe that all the assumptions of Model M_{tbh} are required for the capture-recapture experiment at hand, none of the models considered in this publication are capable of producing valid statistical information concerning the population parameters of interest. If such a failure occurs, the researcher is forced to reevaluate the experimental design to discern how the experiment might be conducted to obtain useful information. This reevaluation may cause the capture-recapture method to be discarded in favor of some other completely different technique, e.g., line transect methods. Whatever the conclu-

sions, such an approach is much preferred over one that ignores necessary assumptions associated with the appropriate capture-recapture model and proceeds to use some simpler but inappropriate model. This latter approach requires the researcher to give credence to statistically invalid information about the population, and therefore violates basic tenets of scientific research.

REMOVAL MODELS

Introduction

In a capture-recapture experiment, all members of the population captured on a given trapping occasion are marked and released back into the population. In a removal experiment, however, as the name implies, members of the population are permanently removed and are not reintroduced into the population after first capture. The removal may be accomplished by killtrapping, electrofishing, trawling, or merely livetrapping the animals and physically displacing them to another area.

Removal methods have traditionally been associated with catch-effort experiments, which have been used often in practice (Omand 1951, Ketchen 1953, Fischler 1965). Such experiments are based on the idea that the size of the population will be decreased gradually as will the catch per unit effort by applying a known amount of removal effort on a number of occasions. The basic assumption behind catch-effort techniques is that the number of animals removed from the population directly depends upon the amount of effort expended in removing them. By using the size of the successive decreases and the known efforts that effected those decreases, initial population size N can be estimated. The experiment usually is modeled by assuming that all animals have probability of capture $p_j = 1 - \exp(-ke_j)$ on the j^{th} trapping occasion, where k is often called the "Poisson catchability coefficient" and e_j represents the known effort expended on the j^{th} occasion. Under such conditions and the

condition that units of effort act independently, a number of estimation techniques have been proposed. Although maximum likelihood estimation of N is possible (cf. Seber 1973:297), the best known estimation techniques are the regression techniques proposed by Leslie and Davis (1939), DeLury (1947), and Ricker (1975). All those methods involve regressing catch per unit effort against some function of the "catchability coefficient" and the efforts expended, and then using least squares analysis to provide an estimator of N and its variance.

Such catch-effort techniques will not be considered in our approach to removal experimentation for a number of reasons. First, and most important, we restrict our consideration of removal studies to those in which effort is deliberately kept constant on all occasions. In such studies, the concept of effort has no utility. Second, because the general approach in this monograph has been to void parametric assumptions concerning capture probabilities, the assumption that $p_j = 1 - \exp(-ke_j)$ is inappropriate for our purposes. Third, it is necessary for purposes of least squares analysis to make certain assumptions that involve the variance structure of the observations. Because to each postulated variance structure there corresponds a formula by which the estimate of N is calculated, such procedures are to some extent arbitrary in the absence of valid tests for determining proper variance structure. Finally, Braaten (1969) investigated the robustness of the DeLury (1958) estimator to departures from some of the assumptions and found that the estimator may exhibit significant negative bias.

Structure and Use of the Generalized Removal Model

If it is assumed that sampling effort is constant over trapping occasions, and that all animals in the population have the same probability of removal, then the removal model and estimation method considered by Moran (1951) results. That