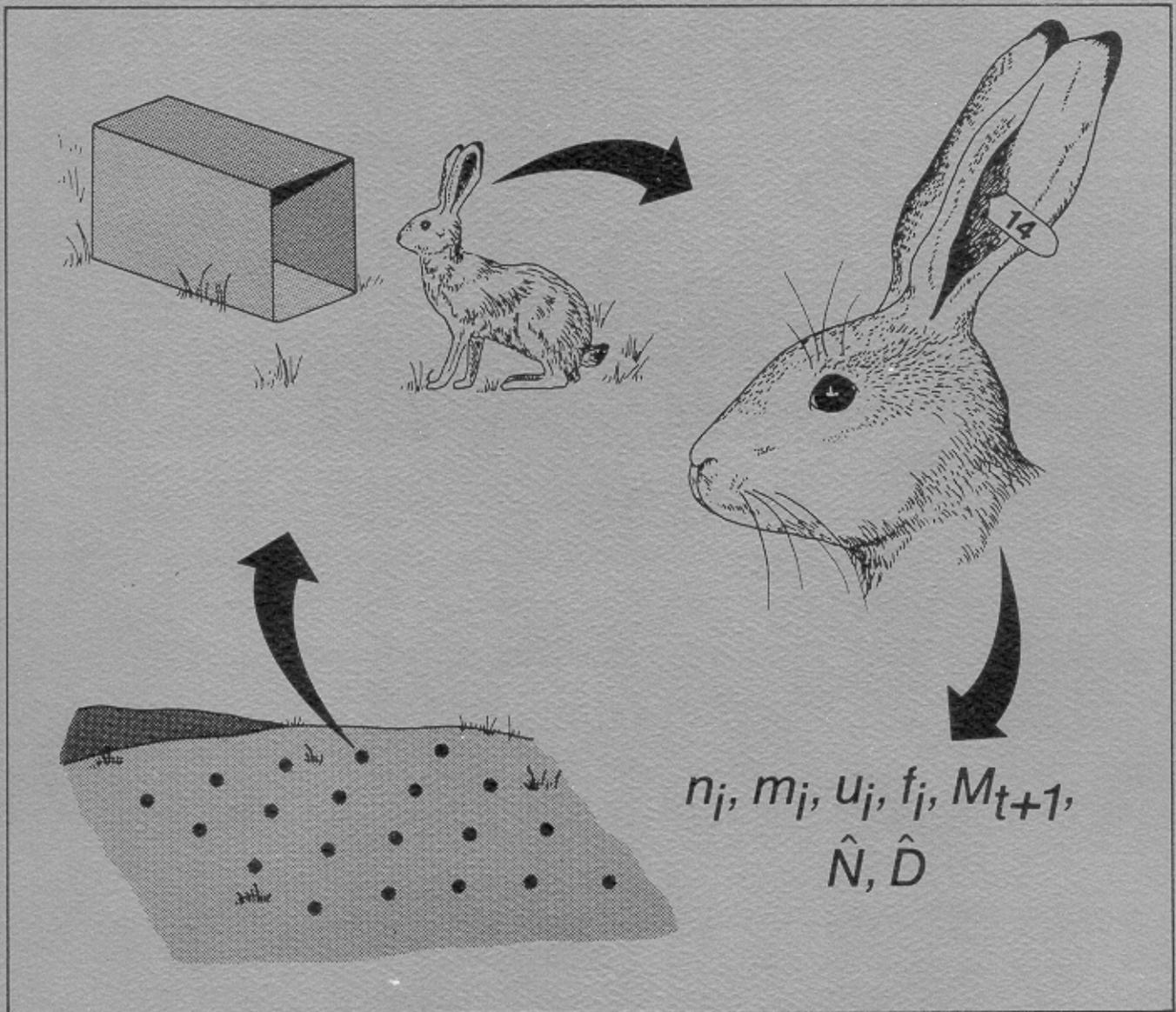


## Capture-Recapture and Removal Methods for Sampling Closed Populations



This work was supported by the US Department of Energy, Office of Health and Environmental Research, National Environmental Research Park.

Edited by Betty Leffler  
Photocomposition by Barbara J. Velarde

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

LA-8787-NERP

UC-11

Issued: August 1982

## **Capture-Recapture and Removal Methods for Sampling Closed Populations**

Gary C. White  
David R. Anderson\*  
Kenneth P. Burnham\*\*  
David L. Otis†

\*Utah Cooperative Wildlife Research Unit, Utah State University, Logan UT 84322.

\*\*US Fish and Wildlife Service, 2625 Redwing Road, Ft. Collins, CO 80526.

†US Fish and Wildlife Service, Denver Federal Center, Denver, CO 80225.

**Los Alamos** Los Alamos National Laboratory  
Los Alamos, New Mexico 87545

# ABSTRACT

The problem of estimating animal abundance is common in wildlife management and environmental impact assessment. Capture-recapture and removal methods are often used to estimate population size. *Statistical Inference From Capture Data On Closed Animal Populations*, a monograph by *Otis et al. (1978)*, provides us with a comprehensive synthesis of much of the wildlife and statistical literature on the methods, as well as some extensions of the general theory. In our primer, we focus on capture-recapture and removal methods for trapping studies in which a population is assumed to be closed and do not treat open-population models, such as the Jolly-Seber model, or catch-effort methods in any detail. The primer, written for students interested in population estimation, is intended for use with the more theoretical monograph.

## CONTENTS

PREFACE .....	ix
ACKNOWLEDGMENTS .....	xi
GLOSSARY .....	xii
<b>CHAPTER 1. INTRODUCTION .....</b>	<b>1</b>
Closure: An Important Assumption .....	3
Data .....	5
Parameters .....	6
Statistics .....	7
Fundamental Assumptions .....	7
Capture-Recapture Models .....	8
Removal Models .....	8
Program CAPTURE .....	10
Questions and Exercises .....	11
<b>CHAPTER 2. STATISTICAL CONCEPTS .....</b>	<b>14</b>
Theory, Reality, and Models .....	15
Estimation .....	16
Parameter .....	16
Estimator .....	16
Accuracy .....	18
Bias .....	18
Precision .....	19
Stochastic Processes and Models .....	20
Variation .....	21
Spatial and Temporal Variation .....	21
Stochastic Variation .....	22
Standard Errors and Sample Size .....	24
A Further Example of Variation .....	25
Properties of a Good Estimator .....	26
Estimation Methods .....	26
Random Sampling .....	27
Robustness of an Estimator .....	27
Closed-Form Solutions to the ML Estimator .....	27
Numerical Solutions to the ML Estimator .....	29
Likelihood Function .....	30
Basis for Rigorous Inference .....	32
Confidence Intervals .....	33
Tests of Hypotheses .....	34
Error Types and Distributions under the Null Hypothesis .....	35
Error Types and Distributions under the Alternative Hypothesis .....	37
Hypothesis Testing in Capture-Recapture and Removal Studies .....	37
Simulation Methods .....	39
Summary .....	41
Questions and Exercises .....	42

CHAPTER 3. CAPTURE-RECAPTURE MODELS .....	45
Modeling Capture Probabilities .....	46
Model $M_o$ , Constant Capture Probabilities .....	48
Model $M_t$ , Variation by Time .....	51
Model $M_b$ , Behavioral Response .....	56
Model $M_h$ , Heterogeneity .....	63
Model $M_{bh}$ , Behavioral Response and Heterogeneity .....	66
Model $M_{th}$ , Time Effects and Heterogeneity .....	71
Model $M_{tb}$ , Time Effects and Behavioral Response .....	71
Model $M_{tbb}$ , Time Effects, Behavioral Response, and Heterogeneity .....	74
Summary of Models .....	75
Testing Model Assumptions and Model Selection .....	77
Overview .....	77
Test 1, Heterogeneity .....	77
Test 2, Behavioral Response .....	79
Test 3, Time Effects .....	79
Test 4, Goodness of Fit of Model $M_h$ .....	79
Test 5, Goodness of Fit of Model $M_b$ .....	79
Test 6, Goodness of Fit of Model $M_t$ .....	80
Test 7, Behavioral Response Given Heterogeneity .....	80
Comment .....	80
The Model Selection Procedure .....	80
A Comprehensive Look at the Simulation Examples .....	90
Robustness of the Different Estimators .....	90
Model Selection and Estimation .....	91
Testing for Closure .....	96
Summary .....	97
Questions and Exercises .....	98
 CHAPTER 4. REMOVAL METHODS .....	 101
The Removal Model, Assuming Constant Capture Probability .....	103
The Model and Its Assumptions .....	103
Some Simulated Data .....	104
Estimation for the Case $t = 2$ .....	105
Estimation and Goodness of Fit Testing for Three or More Capture Occasions .....	106
Another Simulation Example .....	107
Failure of Removal Experiments .....	108
Generalized Removal Method .....	109
The Basic Idea .....	109
Examples of the Generalized Removal Model .....	111
An Example Using Fish Removal Data .....	113
Regression, or Catch Per Unit Effort Methods .....	114
Summary .....	117
Questions and Exercises .....	118
 CHAPTER 5. DENSITY ESTIMATION .....	 120
Theory .....	120
Example .....	122
Design Requirements .....	124
Summary .....	130
Questions and Exercises .....	131

CHAPTER 6. EXAMPLES .....	132
Example 1. Interpreting the Data .....	132
Example 2. Trap Happy or Trap Shy—No Difference? .....	134
Example 3. Closure? .....	135
Example 4. Separating the Sexes .....	135
Example 5. Time Is of the Essence .....	142
Example 6. How Many Are Not Enough? .....	149
Example 7. Density Estimation .....	151
 CHAPTER 7. STUDY DESIGN .....	 161
Livetrapping Versus Removal Methods .....	161
Closure .....	162
Eliminating Variation Due to Time, Behavior, and Heterogeneity .....	163
Sample Size .....	164
Unique Identification of Each Captured Animal .....	168
Recording Data .....	170
Data Anomalies .....	171
Simulation Example .....	173
Check List .....	178
Questions and Exercises .....	178
 CHAPTER 8. OPEN MODELS .....	 180
Basic Concepts .....	180
State of the Art .....	181
Discussion .....	184
Summary .....	186
Questions and Exercises .....	186
 CHAPTER 9. THE FUTURE .....	 188
Biological Developments .....	188
Statistical Developments .....	189
Closed Models .....	190
Open Models .....	192
Computer Programs .....	193
 POSTSCRIPT .....	 195
 REFERENCES .....	 196
 APPENDIX A. FORTRAN 77 USER'S MANUAL FOR PROGRAM CAPTURE .....	 206
Introduction .....	206
Overview of Program Input .....	206
Reserved Files .....	210
Continuations .....	210
Comments .....	210
Specific Task Formats .....	210
(1) TITLE= .....	210
(2) TASK READ CAPTURES .....	210
(3) TASK CLOSURE TEST .....	215
(4) TASK MODEL SELECTION .....	215
(5) TASK POPULATION ESTIMATE .....	216

(6) TASK UNIFORM DENSITY TEST .....	217
(7) TASK DENSITY ESTIMATE .....	217
(8) TASK SIMULATE .....	218
Other Tasks .....	221
<b>APPENDIX B. ANSWERS TO QUESTIONS AND EXERCISES .....</b>	<b>222</b>
Chapter 1 .....	222
Chapter 2 .....	223
Chapter 3 .....	224
Chapter 4 .....	225
Chapter 5 .....	226
Chapter 7 .....	226
Chapter 8 .....	227
<b>APPENDIX C. GENERAL READING LIST .....</b>	<b>228</b>
General Reviews .....	228
Estimation Methods for Closed Populations .....	228
Estimation Methods for Open Populations .....	229
General Computer Program Packages .....	229
Related Reading .....	229
<b>INDEX .....</b>	<b>230</b>

# PREFACE

The problem of estimating animal abundance is common in wildlife management and environmental impact assessment. Capture-recapture and removal methods are often used to estimate population size. *Statistical Inference From Capture Data On Closed Animal Populations*, a monograph by *Otis et al. (1978)*, provides us with a comprehensive synthesis of much of the wildlife and statistical literature on the methods, as well as some extensions of the general theory. In our primer, we focus on capture-recapture and removal methods for trapping studies in which a population is assumed to be closed and do not treat open-population models, such as the Jolly-Seber model, or catch-effort methods in any detail. The primer, written for students interested in population estimation, is intended for use with the more theoretical monograph.

In the monograph, we attempted to produce a state-of-the-art document related to model building, rigorous statistical treatment, and exact maximum likelihood estimators of model parameters. We developed an algorithm, or computational method, to automate model selection and implemented the entire analytical procedure in a computer program called CAPTURE. CAPTURE contains many features and options, including new algorithms for density estimation and simulation experiments. The monograph was intended for biologists as well as applied biometricians. However, at recent workshops and seminars with biologists and students, we found that they frequently did not understand the monograph's key points fully and therefore could not use the methods effectively.

In this primer, we present the basic concepts and methods of sampling. Readers should read *Otis et al. (1978)* for derivations of the methods and tests and for other technical material that is not included here; we have cited specific pages and sections of the monograph to enable ready access to the relevant material. Our emphasis here is on concepts and practical information useful to biologists.

In designing sampling studies, biologists must be aware of what the assumptions are and must make proper transitions from model assumptions to field problems. The design of effective sampling studies requires some familiarity both with the random (or stochastic) nature of the sampling process and with such fundamentals as sampling variation, bias, precision, parameter identifiability, and the criteria for selection of an estimator. We therefore have included an extensively illustrated review of important statistical concepts; an understanding of these is fundamental to an understanding of the rest of the material presented here. We urge a careful study of Chapter 2. Throughout the primer, figures and their captions are used to emphasize key concepts. To further facilitate understanding of the main points, we present the more technical material in small type; figures that involve more technical aspects are denoted by an asterisk. On a first reading one can ignore the small type.

The primer is intended for classroom use by college seniors and graduate students. Suitable for biologists and ecologists, it does not require substantial quantitative training beyond a course in basic statistics. Details on the use of program CAPTURE are given by *White et al. (1978)*, and the uses of some program features are given here in Appendix A. A set of questions and exercises appears at the end of most chapters; answers are provided in Appendix B. Appropriate sections of *Seber (1973)* should be considered auxiliary to the present work, and the General Reading List (Appendix C) should be consulted for closely related material.

The theory and practice of capture-type studies have had a long history, and many people have contributed to our present understanding of the subject. To recognize these individuals and to stimulate the reader's interest in capture-type studies, we have selected for special attention the people whose contributions at the time were most significant. Although our selections involved some arbitrary judgment, we feel that readers will benefit from knowing something about each of these 21 people. We

attempted to obtain a photograph of each person at the time he or she was active in the subject of capture-type studies; for this reason, the quality of the photographs is not ideal in some cases. Some photographs were made from published half-tones or small passport pictures, and one was made from a mid 1960s South African driver's license.

In a sobering review of a sample of the biological literature *Hayne (1978)* suggests that biologists often have not thought deeply enough about the most fundamental question relating to their research—why it is being undertaken. Unless this basic question is thought through, the material presented in this primer can be of little value.

December 1980

G. C. White, D. R. Anderson,  
K. P. Burnham, and D. L. Otis

## POSTSCRIPT

One last capture-recapture example was prepared for this primer—to estimate the number of typographical errors in the text. Each of the four authors plus the editor proofread the manuscript, thus generating five “capture occasions.” An  $X$  matrix was constructed from the results. The values of  $n_j$  were 26, 47, 59, 60, and 79, and  $f_j$  values were 68, 43, 18, 7, and 7, and  $M_{t+1}$  equaled 143.

Each of the five occasions was independent; that is, none of the five reviewers worked together, so there cannot be behavioral response. However, some errors are more difficult to spot than others, so there is heterogeneity of “capture” (typographical error detection) probabilities. Also the reviewers spent different amounts of time proofreading, and one author’s spouse assisted in the process, so there reasonably should be variation of capture probabilities by occasion. Thus, the appropriate model is  $M_{th}$ , for which there is no estimator. We do not believe that Model  $M_b$  or Model  $M_{bh}$  is at all appropriate for this case because no behavioral response is involved, and because the estimators for these models are dependent on the ordering of the capture occasions. For these data, no logical ordering of capture occasions can be made. Note that the other three estimators are not dependent on the ordering of the capture occasions. Study of the simulation results in Table N.5.b of *Otis et al. (1978:129)* suggests the jackknife estimator associated with Model  $M_h$  is more appropriate for the analysis than Darroch’s estimator for Model  $M_t$ . The estimate of  $N_h$  was 217 with the 95% confidence interval (189, 246). Thus, subtracting the 143 errors located in the manuscript, there are still  $217 - 143 = 74$  typographical errors remaining, with the 95% confidence interval (46, 103). Good luck in finding them.

# INDEX

## A

accuracy, 18  
animal identification, 168-170  
animal births, 181  
animal deaths, 120, 162, 171, 173, 181.  
*See also* emigration.  
animal influx, 135, 181. *See also*  
immigration.  
area, effective (A), 120, 130, 164  
area relating to N, 3  
Arnason, A. Neil, 162  
assumption of closure. *See* closure.  
assumption of equal effort, 104  
assumption of finite size, 120  
assumption of constant density, 127, 130  
assumptions for capture-recapture studies, 3, 7  
assumptions for density estimation, 124, 127  
assumptions for removal studies, 7  
assumptions of removal models, 103

## B

Bailey, Norman T. J., 53  
ball and urn model, 3, 4, 7, 27, 46-47, 53  
Bayesian statistical methods, 190  
behavioral response, 8, 47, 56, 71, 74, 80.  
*See also* Model  $M_b$ , Model  $M_{bh}$ .  
clue to, 133  
control of, 163  
in open models, 192  
test for, 77, 82, 90  
benthos, stream bottom, 112  
bias, 18, 22, 23  
effect of recaptures on, 134  
in estimators, 26, 90, 175  
minimization of, 114  
in models, 19  
sources of, 120  
biotelemetry, 188  
births of animals, 181  
bootstrap method, 190  
boundary strip, 120, 127  
boundary strip width W, 120-124, 130, 164  
Burnham, Kenneth P., 9

## C

CAPTURE, 10-11, 33, 39, 41, 206-221  
computations for closure test, 96  
computation in removal studies, 105  
print-out, 49, 132  
density estimation by, 154  
capture occasions in removal studies 105,  
106, 118  
capture probabilities,  
constant, 48,  
magnitude of, 41  
unique, 47  
variation in, 47, 52  
capture probability p, 6, 47  
capture-recapture studies, 1  
captures, "sufficient numbers of," 165  
caret (^) notation, 6, 16  
Carothers, Andrew D., 10  
catchability, 181  
catchability structure, 132  
catchabilities, equal, 7  
catch-effort methods, 161  
catch per unit effort (CPUE) methods, 114-115  
census, 23, 24  
Chapman, Douglas G., 54  
closure, 3  
methods to ensure, 162  
in removal studies, 117  
tests for 36, 38, 96-97, 135, 162  
closure, demographic, 3-5, 46, 120, 186  
closure, geographic, 3-5, 101, 184, 186  
coefficient of variation, 24  
computer programs, 10, 192  
computer simulation, 39  
confidence interval, 33  
computation of, 49  
construction of, 33-34  
coverage on N, 190  
"contagion of catchability," 8.  
*See also* Model  $M_b$   
Cormack, Richard, 2  
cotton rat data, 134  
count, 100%, 101

## D

D and W, estimation of, 120, 130, 154  
 $\hat{D}$  and  $\hat{W}$ , correlation between, 124  
Darroch, John W., 29  
Darroch's estimator, 53, 54, 91, 216  
Darroch's model, 29, 30  
data, 5  
  anomalies in, 171  
  interpretation of, 132  
  pooling of, 168  
  recording of, 170  
  in removal studies, 8  
  real-world, 133  
data, fish removal, 113-114  
data, simulated, 218-221  
  for model  $M_b$ , 56-63  
  for Model  $M_{bh}$ , 66-70  
  for Model  $M_h$ , 63-66  
  for Model  $M_o$ , 48-49  
  for Model  $M_t$ , 51-56  
  for Model  $M_{tb}$ , 71-74  
  for Model  $M_{tbh}$ , 74-75  
  for Model  $M_{th}$ , 71  
  for removal studies, 104  
data stratification, 170  
  example of, 135, 139  
  in removal studies, 104, 113, 118  
  to remove heterogeneity, 163  
data, U-shaped, 133  
deaths of animals, 120, 181.  
  *See also* emigration.  
  effects on livetrapping studies, 162  
    171, 173  
density D, 120, 130  
density estimation, 120  
  examples of, 122, 151-154  
  future research on, 191-192  
  sample size in, 164-165  
density, naive estimate of, 120  
density, uniform, 151, 154  
design. *See* study design.

## E

Eberhardt, L. Lee, 63  
edge, effect, 120, 164  
effort, equal, 104, 163  
electrofishing, 101, 104  
  data stratification in, 118, 170  
emigration, 120, 181  
entry rate, 186  
error, measurement, 26

error, Type I, 35, 37, 162  
error, Type II, 35, 37, 69  
estimate, 16  
estimates, maximum likelihood, 30, 58  
estimation, 16  
  of D and W, 120, 130, 154  
estimation methods, 26, 115  
estimators, 16  
  for eight models, 75-76  
  robustness of, 27, 90-91  
estimator, closed form, 27, 29, 54  
estimator, maximum likelihood (ML), 27, 29  
  distribution of, 34  
  under generalized removal models, 111, 118  
  under Model  $M_b$ , 58-63  
  under Model  $M_{bh}$ , 66-70  
  under Model  $M_h$ , 64-66  
  under Model  $M_o$ , 49  
  under Model  $M_t$ , 53-54  
  for removal data, 105  
  solutions to, 29-30  
estimators,  
  Darroch, 53, 216  
  generalized removal, 90, 216-217  
  jackknife, 64-66, 91, 142, 190, 216  
  null, 49, 216  
  removal, 216-217  
  Petersen-Lincoln, 16, 18, 27, 29, 32, 47  
  Schnabel, 27, 53-54  
  Zippin, 91, 216  
estimator of population mean, 16  
experiments, "failure criterion" in, 108  
experiments, simulated, 173

## F

failure criterion in removal experiments, 108  
fantail darter data, 114  
field forms, 170  
field methods, effect on data analysis of, 161  
 $f_i$  notation, 7, 64  
Fisher, Sir Ronald A., 27, 28

## G

generalized removal estimator, 90, 216-217  
grids, 120-130, 151-154  
  design of, 120, 164  
  effective area of, 120  
  minimum area of, 166  
  nested, 121  
  subgrids in, 127, 151  
  shape of, 127, 164-165

size of, 3, 164  
ground squirrel data, 122, 151

## H

$H_A$ . *See* hypothesis, alternative.  
 $H_0$ . *See* hypothesis, null.  
hare data, 109, 113  
"hat" (^)  
heterogeneity, 8, 47, 63-66, 71, 74. *See also*  
  Model  $M_h$ , Model  $M_{bh}$ .  
  clues to, 132  
  due to unrecognized attributes, 163  
  elimination of, 163-164  
  in open-population models, 186, 192  
  in removal studies, 104, 109, 112, 118  
  test for, 77, 82, 90  
"heterogeneity" induced by behavioral  
  response, 82  
home range, 120, 165  
hypotheses, tests of 34-35, 79-80, 122-124  
hypothesis, alternative, 35  
hypothesis, null, 35, 79-80, 90, 139  
hypothesis, statistical, 34

## I

identification of animals, 168-170  
immigration, 120, 135, 162, 181  
indexes, use of, 32  
"index to abundance," 32  
inference, inductive, 14  
inferences, invalid, 182  
influx of animals, 181

## J

jackknife estimator, 64-66, 91, 142, 190, 216.  
  *See also* estimator under Model  $M_h$ .  
Jackson, C. H. H., 34  
Jolly, George M., 182  
Jolly-Seber model, 180-181, 184, 192

## L

La Place, P. S., 17  
Leslie, Patrick H., 40  
likelihood function, 27, 30-32  
Lincoln, Frederick C., 17  
Lincoln Index, 17. *See also*  
  Petersen-Lincoln estimator.

livetrapping methods, 161, 162, 164  
loss rate, 186

## M

Manly, Bryan F. J., 185  
marking systems, 168-170, 188-189  
marking techniques, 161  
matrix,  $X$ . *See*  $X$  matrix.  
maximum likelihood (ML) method, 18, 27,  
  31-32  
meadow vole population estimate, 3  
measurement error, 26  
methods, catch per unit effort (CPUE), 114-116  
microtaggants, 189  
migration, 162, 181  
 $m_j$  notation, 7, 47  
ML estimates, 30, 58  
ML estimators. *See* estimator, maximum  
  likelihood (ML)  
model, 15  
model, ball and urn, 3, 4, 7, 27, 46-47, 53  
model, Cormack's, 184  
model, constant capture probability, 48, 115.  
  *See also* Model  $M_0$ .  
model formulation, 26-27  
model, Jolly-Seber, 180, 181, 184  
model, log-linear, 191  
model, mathematical, 41  
Model  $M_b$ , 8, 27  
  estimator under, 58-63  
  examples of, 134, 135  
  likelihood function for, 31  
Model  $M_{bh}$ , 8, 27  
  estimator under, 66-70  
  example of, 154  
Model  $M_h$ , 8, 27  
  estimator under, 64-66  
  examples of, 132, 139  
Model  $M_0$ , 7, 8, 48-51  
  estimator under, 49  
  examples of, 132, 151  
  likelihood function for, 31  
Model  $M_R$ , 8, 10  
Model  $M_t$ , 8, 51-55  
  estimator under, 53-54  
  likelihood function for, 31  
Model  $M_{tb}$ , 8, 27, 71-74  
Model  $M_{tbh}$ , 8, 27, 74-75  
  example of, 133  
Model  $M_{th}$ , 8, 27, 71  
  example of, 133

model, Pollock's, 184  
 model, probability, 27  
 model, regression, 123  
 model, removal, 103, 113, 191  
 model, Robson's, 184  
 models, capture-recapture, 7, 45-75, 185  
 models, closed  
   difficulties in, 190  
   use of, 185  
 model selection, 41  
   criteria for, 77, 80  
   procedure for, 50, 63, 80-90,  
     133-134, 151, 170  
 model selection rule, 190-191  
 models, open, 180, 192  
   use of, 163, 185  
 models, realistic, 189-190  
 models, relationships among, 75-76  
 models, removal, 8, 103-105, 109-111  
 models, stochastic, 15, 41  
 models, tests of, 77  
 Monte Carlo simulations, 173  
 Moran, P. A. P., 103

## N

N. *See* population size.  
 notation  
   in computer output, 132  
   for estimates ( $\hat{\cdot}$ ), 6, 16  
    $f_j$ , 7, 64  
    $m_j$ , 7, 47  
    $n_j$ , 7, 47, 51  
   for summation, 7  
    $u_j$ , 7, 50  
 null estimator, 49, 216. *See also*  
   estimator under Model  $M_0$ .  
 null hypothesis, 35, 79-80, 90, 139

## O

open populations, 180  
 outflux, 181. *See also* emigration.

## P

p. *See* capture probability.  
 parameter, 6, 16  
 parameters, effect of increased number, 180-181  
 parameters of capture model, 46  
 partitioning. *See* stratification.  
 pattern of observed significance levels, 80, 82  
 patterns in real-world data, 133

pellet counts, 161  
 Peterson, Carl George Johannes, 17  
 Petersen-Lincoln estimator, 16, 18, 27, 29,  
   32, 47,  
   Chapman's modification of, 53  
 pocket mice data, 173  
 Pollock, Kenneth H., 9  
 POPAN-2, 192  
 population density. *See* density.  
 population dynamics, 182-184  
 population mean, 16  
 population size, N, 6  
   area relating to, 3, 46  
   estimation of, 1  
 population variance, 22  
 populations, open, 180  
 populations, simulated, 39-40  
 precision, 19  
   effect of recaptures on, 134  
   in removal studies, 106  
   sacrifice of, 114  
 probability, constant capture, 103  
 program CAPTURE, *See* CAPTURE

## R

random sampling, 27  
 recapture probability  $c$ , 6, 56  
 recaptures, number of, 165  
 recruitment, 120, 162, 181  
 regression (CPUE) methods, 114-116  
 removal estimator, 216-217  
 removal by marking, 10, 101, 161  
 removal methods, 1, 101  
   generalized, 109-114  
   Zippin, 27  
 removal model,  
   assumptions of, 112  
   variability in, 107-108  
 removal models. *See also* Model  $M_b$ , Model  
 $M_{bh}$ .  
   generalized, 111  
   simplest, 8, 103-108  
   Zippin's, 27, 91  
 removal, physical, 10  
 removal sampling, 101  
 removals, number of, 105-106  
 removal studies, 1  
 response, behavioral, 8, 47, 56, 66,  
   71, 74, 80  
   clue to, 133  
   control of, 163  
   in open models, 192  
   test for, 79, 82, 90

Robson, Douglas S., 111  
robustness, 3, 27, 90-91, 139

## S

sample size, 41, 164, 166  
  example of, 149, 151  
sampling concepts, 14  
sampling methods, 27  
sampling, random, 27  
sampling variation, 24-26, 104  
Schnabel, Zoe Emily, 28  
Schnabel estimator, 27, 53-54  
scientific method, 14, 16  
Seber, George A. F., 2  
significance levels, 80, 82  
simulation of experiments, 168, 173  
  simulation methods, 39  
  simulation, Monte Carlo, 173  
Smith, Michael, 121  
Statistical Analysis System (SAS), 193  
standard errors, 24  
statistics, 7, 14  
  mathematical, 34-35  
  statistics, summary, 45  
  under Model  $M_b$ , 50  
  under Model  $M_h$ , 64  
  under Model  $M_{bh}$ , 67  
  under Model  $M_o$ , 49  
  under Model  $M_t$ , 51-52  
stochastic processes, 20  
strip width, *See* boundary strip width  $W$ .  
studies, invalid inferences from, 182  
studies, length of, 163, 185-186  
study design, 161-178  
  effect of behavioral response on, 134-135  
  example of, 173  
  requirements for, 124, 161  
  tests of, 168  
surveys, 161  
survival, 6, 181, 192

## T

tags, electronic, 188  
test, Leslie's, 39  
test, simple alternative, 38  
tests,  
  for closure, 96-97, 135, 162  
  of data stratified by sex, 139, 142  
  of design, 168  
  for goodness of fit, 37, 69, 79, 80,  
  90, 96

  for homogeneity in stratified  
  data, 163  
  of hypotheses, 34-35, 79-80, 122-124  
  of model assumptions, 33, 77-80,  
  in removal studies, 106-107, 113  
tests, power of, 36, 96, 191  
tests, problems of, 38-39  
test statistics, 34-35, 37, 38, 39  
theory vs reality, 15  
time-associated variation, 8, 47, 71, 74,  
  142. *See also* Model  $M_t$ .  
  clue to, 133  
  control of, 163, 168, 175  
  example of, 142  
  in removal studies, 109, 118  
  test for, 79, 82, 90  
tracks, animal, 164  
trap-happy animals, 56, 82, 133-134  
trap location, 127, 135, 164, 170  
trap numbering, 170,  
trapping area, 186  
trapping effort, 104, 163  
trapping grids, *See grids*.  
trapping, kill, 101  
trapping occasions, 165, 168  
trapping, small-mammal, 104, 110, 164  
trapping, snap, 117  
trap-shy animals, 56, 59, 67, 72, 82, 133  
trap spacing, 164, 165  
Type I error, 35, 162. *See also* tests of  
  hypotheses.  
Type II error, 35, 37, 69. *See also* tests of  
  hypotheses.

## U

$u_j$  notation, 7, 50  
uniform unmarked minimum variance estimator  
(UMVUE), 190

## V

variation, 21-26  
variation in capture probabilities, 8, 47, 163,  
  188. *See also* behavioral response,  
  heterogeneity, time-associated variation.  
variation, coefficient, of, 24  
variation, sampling, 24-26, 104  
variation, sources of, 25, 63  
variation, spatial, 21, 25-26  
variation, stochastic, 22-23  
variation, temporal, 21, 26

## W

W. *See* boundary strip width W  
weather conditions, 163

## X

X matrix, 5, 6, 7, 27, 30, 31  
construction of, 168-169  
examples of, 6, 11, 12, 48, 57, 169  
for Model  $M_0$ , 48

reading in CAPTURE, 210-215  
from simulated data, 40

## Z

Zippin, Calvin, 28  
Zippin estimator, 91  
Zippin removal method, 27  
z-statistic, 38  
z-test, 139, 142

# ACKNOWLEDGMENTS

Funding for the computer time to do this work was provided by the National Environmental Research Park at Los Alamos National Laboratory through contract W-7405-ENG.36 from the US Department of Energy. We heartily thank K. V. Bostick, assistant coordinator, and E. M. Wewerka and W. R. Hansen, coordinators of the Research Park.

Appreciation is expressed to several programs within the US Fish and Wildlife Service for their support of this work: Office of Cooperative Units, Denver Wildlife Research Center, and the Office of Biological Services.

The Denver Wildlife Research Center provided funds for typing the final manuscript. Many people were involved in typing various drafts: we wish to acknowledge E. Montoya, G. Mower, A. Krambule, M. Sievein, and C. Snelling.

Special thanks go to K. Redman, who provided an outstanding editorial review of the final working draft of the manuscript. Her expertise allowed a significant improvement in the readability of the primer.

Thanks also to Elizabeth White for the many hours she spent in proofing the final manuscript.

Four colleagues, R. Anthony, W. Clark, F. Guthery, and L. MacDonald, provided their comments on an early draft of the primer. Their help is appreciated and their suggestions allowed substantial improvement to the present work.

We thank J. Shoemaker for producing the cover drawing and the illustrations in Chapters 3 and 4.

People who provided data from their studies for our use as examples include K. A. Fagerstone, L. E. Lefebvre, R. Mahon, G. H. Matschke, and R. F. Raleigh. We appreciate their help in illustrating the various methods using real data.

Many people who had made major contributions to the theory and practice of capture-recapture and removal studies kindly sent photographs and short narratives for use in this primer. We appreciate their cooperation and, in many cases, their encouragement. In addition, the following aided us in locating addresses, photographs, or information concerning people who had made important contributions, but who are deceased or were difficult to locate: A. N. Arnason, J. G. Box, E. Bursell, J. M. Dunkley, E. R. Easton, C. Elton, G. W. Frame, J. M. B. Harley, W. Jepson, S. C. Kleene, J. L. Laake, H. F. Lamprey, P. M. Mwambu, T. R. E. Southwood, W. H. Stickel, P. Stigger, and H. B. Tordoff.

# APPENDIX A FORTRAN 77 USER'S MANUAL FOR PROGRAM CAPTURE

## INTRODUCTION

The computations necessary to calculate many of the capture-recapture population estimates we have described are lengthy and essentially impossible to perform without a computer. Therefore, we have written a FORTRAN computer program to complement this primer. Input to the program was written in a free-form and natural style to provide ease of use by the computer users. The major program input options are described here, and a complete user's manual is provided in *White et al. (1978)* for earlier versions of CAPTURE. The material in this appendix is applicable to ANSI FORTRAN 77 versions of CAPTURE dated 1980 or later, and except for the file-naming conventions, everything in *White et al. (1978)* applies equally to ANSI FORTRAN 77 versions of CAPTURE. In addition, several modifications have increased the capabilities of the program. The main reason for changing the program was to provide greater portability between machines.

## Overview of Program Input

The basic unit of the program is a TASK. TASK, a reserved word in the program input, specifies that a particular set of computations or data input is requested. The computations necessary to calculate a population estimate based on a particular model are assumed to be a TASK. Input of the  $X_{ij}$  matrix of captures is also a TASK. Various model estimators and hypothesis tests of the validity of the models make up the available TASKs. TASKs are summarized in Table A.1a; reserved words and phrases for the CAPTURE program are listed in Table A.1b.

As a simple example of input to program CAPTURE, consider the job stream shown in Table A.2. This CAPTURE run would analyze the  $X_{ij}$  matrix shown in Table 3.1. The TITLE card specifies a title to be printed on each page of output. TASK READ CAPTURES specifies 6 columns in the X matrix, and the data make up an  $X_{ij}$  matrix of zeros and ones. The DATA statement specifies an identifier for the particular data set, also printed on each page of output for this data set. The FORMAT card specifies the  $X_{ij}$  matrix format on the file CAPTDT. In this case, the first three columns of the data set are the animal identification, read with the A3 format. Then three columns are skipped. The six columns of the  $X_{ij}$  matrix are read with the (3X, F1.0) specification, repeated six times. TASK MODEL SELECTION causes the model selection procedure to select the model best fitting this data set, and the TASK POPULATION ESTIMATE card produces a population estimate for the APPROPRIATE model as determined by the model selection procedure.

Many TASKs require only one input card, for example, the following.

**TABLE A.1a.** TASK cards available in program CAPTURE, and the parameters and options available for each card. Optional parameter specifications are in brackets; mutually exclusive options are in braces, with the default value underscored if a default exists.

---

TASK READ CAPTURES {  
 XY REDUCED  
 XY COMPLETE  
 NON XY  
 X MATRIX } OCCASIONS= [FILE=] [CAPTURES=] [SUMMARY]

Optional additional input cards are  
 DATA='information on data'  
 FORMAT='format specified'  
 READ INPUT DATA

TASK CLOSURE TEST [OCCASIONS=]  
 TASK MODEL SELECTION [OCCASIONS=]  
 TASK UNIFORM DENSITY TEST [OCCASIONS=]

TASK POPULATION ESTIMATE [OCCASIONS=]  
 ALL  
 APPROPRIATE  
 NULL  
 JACKKNIFE  
 REMOVAL  
 DARROCH  
 ZIPPIN

TASK DENSITY ESTIMATE INTERVAL= CONVERSION=  
 ALL  
 APPROPRIATE  
 NULL  
 JACKKNIFE  
 REMOVAL  
 DARROCH  
 ZIPPIN

From two to eight additional input cards define grids:

X= Y= [OCCASIONS=]

· · ·  
 · · ·  
 · · ·

X= Y= [OCCASIONS=]

END OF GRID DEFINITIONS [DENSITY=] [STRIP=]

TASK SIMULATE [SEED=] [POPULATION=] [OCCASIONS=] [REPLICATIONS=] [PRINT] &

{  
 NULL  
 JACKKNIFE  
 REMOVAL } [MATRIX]  
 {  
 DARROCH  
 ZIPPIN

Up to four additional input cards define capture probability structure or provide identifying information:

HETEROGENEITY=

BEHAVIOR=

TIME=

DATA='identifying information about simulation.'

TITLE='a heading to be printed at the top of each page of output.'

---

**TABLE A.1b.** Program CAPTURE reserved words and phrases.

TITLE	HETEROGENEITY
TASK	BEHAVIOR
READ CAPTURES	TIME
CLOSURE TEST	PRINT
MODEL SELECTION	SEED
UNIFORM DENSITY TEST	POPULATION
POPULATION ESTIMATE	REPLICATIONS
DENSITY ESTIMATE	X
READ POPULATION	Y
READ DENSITY	CONVERSION
SIMULATE	INTERVAL
XY REDUCED	END OF GRID DEFINITIONS
XY COMPLETE	DENSITY
NON XY	STRIP
X MATRIX	ALL
SUMMARY	APPROPRIATE
OCCASIONS	NULL
FILE	JACKKNIFE
CAPTURES	REMOVAL
DATA	DARROCH
FORMAT	ZIPPIN
READ INPUT DATA	

### TASK MODEL SELECTION

A second example is a TASK card requiring only one input card, but on which additional key words may be specified to provide an option in the computations.

### TASK POPULATION ESTIMATE JACKKNIFE

This card specifies that a population estimate is desired, specifically the jackknife estimator appropriate for Model  $M_h$ . Other TASK cards require that parameters be specified on the card. For example,

### TASK READ CAPTURES OCCASIONS=10

indicates there were 10 trapping occasions for the data set to be read. The more complicated input requires additional cards after the TASK card. An example is TASK DENSITY ESTIMATE, which requires one card for each grid to specify the grid's dimensions and location.

TASKs may be performed in almost any order, although there is a logical order of determining which estimator is appropriate before estimating the population or density. The captures must be read in before any of the TASKs that require these data can be executed.

### INPUT AND ERRORS LISTING

is a listing at the beginning of the program, made of the input cards as they are read. Each input card is

**TABLE A.2.** The input cards needed for program CAPTURE to analyze the  $X_{ij}$  matrix in Table 3.1. The data (X matrix) is located on file CAPTDT, while CAPTURE reads instructions from the file CAPTIN.

*INSTRUCTIONS read from file CAPTIN*

TITLE='ANALYSIS OF X MATRIX TAKEN FROM TABLE A.2'  
 TASK READ CAPTURES OCCASIONS=6 X MATRIX  
 DATA='X MATRIX FROM TABLE 3.1'  
 FORMAT='(A3,3X,6(3X,F1.0))'  
 TASK MODEL SELECTION  
 TASK POPULATION ESTIMATE APPROPRIATE

*CAPTURES read from file CAPTDT*

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

listed with

**INPUT**\_\_\_\_\_

in front of the statement to separate it from the errors and warnings that also are printed. Warnings provide the default values of parameters not specified on the preceding card, and when an option is taken by default. So long as the default values are satisfactory, the optional parameters need not be set.

Errors usually are printed immediately after the input statement that caused an error to be detected. However, if an earlier statement caused the error, it may not be detected until the time of listing. The errors and warnings printed in the INPUT AND ERRORS LISTING generally concern only program input statements. Errors resulting from poor data, such as no recaptures, are printed in the output from the TASK.

If the program terminates properly (that is, when the last card has been read from the instructions), the following message is printed.

## SUCCESSFUL EXECUTION

Although this message indicates that the program terminated properly, it does not mean that all TASKs were executed. An error on a TASK card may have caused that TASK to be skipped.

**Reserved Files.** The program requires that instructions be read from the file CAPTIN. The default data input file is assumed to be CAPTDT. Output is printed on the file CAPTLP.

These file conventions apply to versions of CAPTURE written in ANSI FORTRAN 77 and not to the older FORTRAN IV versions. If the version of CAPTURE available at your computer center is dated before 1980, see *White et al. (1978)* for a description of the files needed and used by CAPTURE.

**Continuations.** Any card may be continued by putting an ampersand (&) as the last nonblank character on the previous card. Up to two continuations may be used, for a total of 240 characters of input.

**Comments.** The user may punch comments on any of the input cards in the space that remains after the necessary options and parameters have been set. Any words but the reserved words given in Table A.1b may be used. The reserved words, which specify information to the program, should not be used because they may be unintentionally read as instructions.

## Specific Task Formats

### (1) TITLE=

The TITLE= card is the same as a TASK card, but without the word TASK. It provides a title to be printed at the top of each page of output. Title changes are made by placing a TITLE card directly before a TASK card.

Title information is specified by single quotation marks as in the following example.

```
TITLE='PUT YOUR INFORMATION HERE'
```

Note that there are no embedded blanks between the key word TITLE, the equals sign, and the first single quote. Blanks may appear between the two quotes, as needed. However, no single quotes are allowed in the information because the next quote encountered after TITLE=' is taken as the end of the title.

### (2) TASK READ CAPTURES

This task reads the raw data (the  $X_{ij}$  matrix) required to select a model, estimate population size, and so on. The program assumes that the capture histories of the animals are coded on cards in one of the four methods discussed below. If density estimates are required, you must include the coordinates of each trap at which the animal was captured, a process that complicates the input slightly. The coordinates of the trap on the upper left corner of the grid should be (1,1). Coordinates of (0,0) are not permitted because zero values signify that the animal was not captured on this occasion. The two options for reading trap coordinates are XY COMPLETE and XY REDUCED. Option XY REDUCED is the default input format for the program and thus is easier to use than XY COMPLETE. The standard input of the XY REDUCED option is

*animal id, occasion i, x-coordinate, y-coordinate, occasion j, x-coordinate, y-coordinate, . . .*

where *occasion i* is the number of the trapping occasion for which the animal was caught, and *x-coordinate* and *y-coordinate* are the Cartesian coordinates of the trap in which the animal was caught. This input allows the user to specify information only when an animal is caught. If an animal is caught only once, the *occasion, x-coordinate, y-coordinate* repetition is given only once, whereas if an animal is caught three times, the repetition is given three times. The rest of the card is ignored after the first blank or zero set of coordinates and occasion number.

As stated earlier, the program assumes the upper left trap of the grid is numbered (1,1). Numbering systems where other corners are labeled (1,1) can be used and will give correct estimates of population and density. However, when the matrix of captures per trap station is printed, it will be transposed or reflected (or both). The corner trap cannot be numbered (0,0) because zero values indicate the animal was not captured.

The XY COMPLETE option assumes the complete capture history of each animal is being read. The information appears in the form

*animal id; x,y coordinates for occasion 1; x,y coordinates for occasion 2; . . . ; x, y coordinates for last occasion.*

With this option, x,y coordinates are entered on the card only when the animal is captured, with each card representing a separate animal. When an animal is not captured on a particular occasion, the columns are left blank. For an animal captured only once, most of the card will be blank. The number of pairs of x,y coordinates to be read is determined from the OCCASIONS= parameter, to be discussed later.

The third input option, NON XY, is used if the experiment is conducted without coordinates for the traps, or if trap coordinates are to be ignored. With this option, all but the density estimates can be computed. The general form of the input is

*animal id, 1st capture occasion, 2nd capture occasion, 3rd capture occasion. . .*

where capture occasion specifies the number of the trapping occasion on which the animal was captured. The number of trapping occasions is determined by the OCCASIONS= parameter, to be discussed later. The rest of the card is ignored when the first blank or zero occasion is encountered.

The X MATRIX option assumes that the complete  $X_{ij}$  matrix is being read, as described in *Otis et al. (1978)*. The general form is

*animal id, string of ones and zeros to signify capture history (1 = capture, 0 = no capture).*

Three parameters can be specified on the TASK card. OCCASIONS= specifies the number of trapping occasions. For example, if the population was trapped for 7 days and the traps were checked daily, the parameter would be set as follows.

OCCASIONS=7

Note that there can be no embedded blanks because the program is scanning for the end of specification, signified by the first blank. This limitation is true for all parameter specifications in the program. Remember that there can be no blanks between the key word, the equals sign, and value specification.

The other two parameters that can be specified for this TASK relate to the raw capture data input file. The raw data are assumed to be read from file CAPTDT, using the default format (A3, 12(3F2.0)), which means that only one set of population data can be read per run, unless a multifile data set is used. Hence, with the FILE='name' parameter, files other than file CAPTDT can be read, and multiple sets of data

can be analyzed in one run. The examples provided with the program assume that a multifile data set will be used with file CAPTDT. Note that if your version of CAPTURE is dated before 1980, a different file-handling system is used, and you must follow the instructions given in *White et al. (1978:11)*.

The CAPTURES= parameter specifies the number of captures per card. The default value is the number of occasions specified if the number is 12 or less, which is consistent with the default format of up to 12 captures per animal on one card for the XY REDUCED option. The CAPTURES= parameter is used only for the XY REDUCED and NON XY options because the OCCASIONS parameter specifies the number of fields to read for the other options.

In addition, summary information about the distance moved between captures can be obtained by specifying the word SUMMARY on the TASK READ CAPTURES cards, as follows.

#### TASK READ CAPTURES OCCASIONS=10 SUMMARY

The program will summarize the average and maximum distances that the animal moved between successive captures, and the average of the maximum distances moved for all animals by the frequency of capture. This information is used to check the reliability of the estimates of density or, as described in *Otis et al. (1978)*, an estimate of density may be obtained based on distance moved.

Three optional cards may be included after the TASK READ CAPTURES card. The first is the FORMAT= card, which defines the format with which to read the captures. The format is put in quotes as in the following example.

```
FORMAT='(A3, 36 (F2.0))'
```

This format corresponds to the default used in the program, although any of the ANSI FORTRAN 77 format conventions may be used. Format interpretation will depend on which input option is used. Animal identification must be in the A format, and the maximum number of columns allowed depends on the word size of the computer. For example, IBM, DEC-VAX, and Xerox computers allow 4 characters per word; Burroughs, Univac, and Honeywell computers allow 6; and CDC computers allow 10. The x,y coordinates and the occasion number must be read in F format for all input options. For the X MATRIX option, the zeros and ones must be read in F format.

The second optional card is the DATA= parameter, which specifies information about the captures read in addition to that given on the TITLE= card. For example, if a set of three grids is to be run, the TITLE= card specifies general information about the run, whereas the three DATA= cards specify information specific to the individual grids. The order in which FORMAT= and DATA= cards appear is not crucial—either may precede the other.

The third optional card (and always the last) is READ INPUT DATA. Often the user does not want to have a separate data file for the  $X_{ij}$  matrix, but would prefer to read the captures from the input file CAPTIN. If the card READ INPUT DATA is encountered as the last card in the TASK READ CAPTURES input, the  $X_{ij}$  matrix cards are assumed to follow. Data will be read from file CAPTIN until a TASK or TITLE= card is encountered. This method of inputting data does not require that FILE= be specified. Note, however, that a program abort will occur if the first nondata card is not a TASK or TITLE= card.

Because the above descriptions are abstract without examples, we will now give some specifics. First, consider an example of the XY REDUCED option with all the default values. The listing in Table A.3 represents the simplest form of the TASK READ CAPTURES statement. Although listing the entries for each occasion on which an animal was captured in chronological order is not mandatory, we suggest that you order them this way. Multiple cards with the same animal identification will not cause problems, but if there is a conflict, the second card will override the first.

Table A.4 gives a second example of the XY REDUCED option, in which all parameters are specified to illustrate input for which none of the default values apply.

**TABLE A.3.** Example of TASK READ CAPTURES with all the default values taken.

*INSTRUCTIONS read from file CAPTIN*

TITLE='EXAMPLE INPUT FOR TABLE A.3'  
TASK READ CAPTURES OCCASIONS=8

*CAPTURES read from file CAPTDI*

A01	1	5	2	3	7	2	4	8	1	6	6	2	7	7	2
A02	1	4	3	3	6	2	4	6	2	6	3	4	7	1	4
A03	1	8	3	7	8	2	8	10	2						
A04	2	9	2	3	9	2	6	9	2	7	8	3	8	9	3
A05	2	9	3												
A06	4	8	9												
A07	4	10	6	6	10	3									
A08	6	10	6	7	10	7	8	10	6						

**TABLE A.4.** Example of TASK READ CAPTURES with XY REDUCED option set, five captures per card, with input from file MATRIX. Animal identification appears in columns 73-76.

*INSTRUCTIONS read from file CAPTIN*

TASK READ CAPTURES XY REDUCED FILE='MATRIX' OCCASIONS=8 CAPTURES=5  
DATA='EXAMPLE INPUT FOR TABLE A.4'  
FORMAT='(72X,A4,T1,5(F2.0,2F3.0))'

*CAPTURES read from file MATRIX*

1	5	2	3	7	1	4	8	1	6	6	2	7	7	1	A001
1	4	3	3	6	2	4	6	2	6	3	4	7	1	1	A002
1	8	3	7	8	1	8	10	2							A003
2	9	2	3	9	2	6	9	2	7	8	3	8	9	3	A004
2	9	3													A005
4	8	9													A006
4	10	6	6	10	3										A007
6	10	6	7	10	7	8	10	6							A008

Table A.5 gives an example of the XY COMPLETE option. A nondefault format illustrates the use of two cards to record the coordinates for one animal. The default format cannot be used to read two cards (as in the example).

Table A.6 gives an example of the NON XY option input. Note that the animal identified as A01 was caught on occasions 1, 3, 4, and 6, and a later card specifies that it was also caught on occasion 7.

Table A.7 gives an example of X MATRIX option. The first four columns are the animal identification.

TASK READ CAPTURES produces a summary table of output on the INPUT AND ERRORS listing. The listing gives the number of trapping occasions, number of different animals captured, and the maximum x- and y-coordinates. These values will help the user to determine whether the input was coded correctly, because misspunched cards often cause irregular x,y coordinates.

**TABLE A.5.** Example of TASK READ CAPTURES with XY COMPLETE option set and multiple cards per record. The captures are read from CAPTIN with the READ INPUT DATA card. The TASK CLOSURE TEST stops the reading of CAPTURES.

*INSTRUCTIONS read from file CAPTIN*

TASK READ CAPTURES XY COMPLETE OCCASIONS=8

FORMAT='(A4,5(2F5.0)/4X,5(2F5.0))'

DATA='EXAMPLE INPUT FOR TABLE A.5'

READ INPUT DATA

A001	5	2			7	2	8	1
A001	6	2	7	2				
A002	4	3			6	2	6	2
A002	3	4	1	4				
A003	8	3						
A003			8	2	10	2		
A004			9	2	9	2		
A004	9	2	8	3	9	3		
A005			9	3				
A005								
A006							8	9
A006								
A007							10	6
A007	10	3						
A008								
A008	10	6	10	7	10	6		
TASK CLOSURE TEST								

**TABLE A.6.** Example of TASK READ CAPTURES with the NON XY option set. A nondefault file of MYDATA and the default format are used.

---

*INSTRUCTIONS read from file CAPTIN*

TASK READ CAPTURES NON XY OCCASIONS=8 FILE='MYDATA'  
DATA='EXAMPLE INPUT FOR TABLE A.6'

*CAPTURES read from file MYDATA*

A01 1 3 4 6  
A02 1 3 4 6 7  
A03 1 7 8  
A04 2 3 6 7 8  
A05 2  
A06 4  
A07 4 6  
A08 6 7 8  
A01 .7

### **(3) TASK CLOSURE TEST**

The TASK CLOSURE TEST helps to determine whether the assumption of population closure can be made from the data read with TASK READ CAPTURES. The only parameter to be specified is OCCASIONS=; this determines which trapping occasions are to be used in the test for closure. The default value OCCASIONS= is all the trapping occasions. For example, suppose that a grid was trapped for 12 days; when TASK CLOSURE TEST is run with the default value, OCCASIONS=1-12. However, if the investigator wants to look at the assumption of closure for only the first 6 days, the input would be TASK CLOSURE TEST OCCASIONS=1-6. Note that there are no embedded blanks around the equal sign, because a blank signifies the end of the specification. The OCCASIONS= parameter, a single-valued parameter, is used in TASK READ CAPTURES to specify the number of trapping occasions. In the TASK CLOSURE TEST and in the rest of the TASKs where OCCASIONS= will be used, a multiple-valued parameter is used to specify the trapping occasions to be analyzed. Hence, a series of the values will be specified with no embedded blanks. Hyphens indicate "through" so that OCCASIONS=1-5 means the numbers 1, 2, 3, 4, and 5. Slashes indicate "by," so that OCCASIONS=1-9/2 means the series 1, 3, 5, 7, and 9, that is, 1 through 9 by 2's. Commas also may be used to separate sequences of numbers, so that OCCASIONS=1-5,9-10,12 means the series 1, 2, 3, 4, 5, 9, 10, and 12.

### **(4) TASK MODEL SELECTION**

This TASK computes the sequence of hypothesis tests described in Chapter 3. It is used to determine which population estimator should be used. The data are those captures read by TASK READ

**TABLE A.7.** Example of TASK READ CAPTURES with the X MATRIX option set. The READ INPUT DATA card signifies to read all the cards up to the next TASK or TITLE = card.

*INSTRUCTIONS read from file CAPTIN*

```
TASK READ CAPTURES X MATRIX OCCASIONS=8
FORMAT='(A4,8F1.0) '
DATA='EXAMPLE INPUT FOR TABLE A.7'
READ INPUT DATA
A00110110110
A00210110110
A00310000011
A00401100111
A00501000000
A00600010000
A00700010100
A00800000111
TASK MODEL SELECTION
```

**CAPTURES.** This TASK also has only one parameter, the OCCASIONS= parameter. The purpose and format for the parameter specification are identical to those described in the TASK CLOSURE TEST.

#### **(5) TASK POPULATION ESTIMATE**

This TASK computes population estimates for data read by TASK READ CAPTURES. The population estimators desired are specified by using any or all of the five key words: NULL, JACKKNIFE, DARROCH, REMOVAL, and ZIPPIN. If all population estimators are desired, the key word ALL may be used. Usually the biologist is unsure of which estimator is appropriate until after he has reviewed the hypotheses testing output. To avoid multiple runs, the key word APPROPRIATE may be used to instruct the program to calculate the estimator selected in TASK MODEL SELECTION. However, the TASK MODEL SELECTION must have been run for the grid being analyzed. Other estimators may be specified with APPROPRIATE, as shown below.

TASK POPULATION ESTIMATE APPROPRIATE NULL

If the NULL estimator is not selected as the appropriate one, two population estimates will be made.

The NULL estimator, derived from Model  $M_0$  in Chapter 3, is described as null because none of the three possible sources of variability is assumed to be operating. The JACKKNIFE estimator is appropriate for Model  $M_{1p}$ , where the probability of capture varies by animal. The DARROCH estimator

is derived from Model  $M_t$ . The REMOVAL estimator is the generalized removal estimator derived from Model  $M_{bh}$ . The ZIPPIN estimator, a special case of the REMOVAL estimator, is derived from Model  $M_b$  in Chapter 3.

This TASK also has available the OCCASIONS= parameter. Use and format of this parameter are identical to those described for TASK CLOSURE TEST. The OCCASIONS= parameter in this TASK is used to look at changes in population during the trapping period.

## (6) TASK UNIFORM DENSITY TEST

This TASK tests the homogeneity of the distribution of captures from the grid read by TASK READ CAPTURES. A matrix of captures by trap station is used to indicate possible trends in density within the grid. Also, the grid is collapsed by rows of traps, a chi-square test is constructed, and the grid is then collapsed by columns. This TASK also has the OCCASIONS= parameter. Its use and format are identical to those described in the TASK CLOSURE TEST. The matrix output from this TASK is difficult to interpret if the upper left trap is not labeled (1,1). No output can be produced when the data are read with NON XY or X MATRIX formats.

## (7) TASK DENSITY ESTIMATE

This TASK computes an estimate of animal density based on the method presented in Chapter 5. An option, several parameters, and additional specifications cards are required.

The option determines the population estimator to be used to estimate the naive density of each grid. The five option key words are NULL, JACKKNIFE, DARROCH, REMOVAL, and ZIPPIN. If all population estimators are desired, the key word ALL may be used. If the estimator selected by TASK MODEL SELECTION is desired, the key word APPROPRIATE may be used.

Two parameters also must be specified on the TASK card. The first parameter, TRAP INTERVAL=, may be shortened to INTERVAL=. It specifies the distance between traps for the grid. For example, if traps are set on a 15-m grid system, TRAP INTERVAL=15 or the shorter form, INTERVAL=15, would be used. The default is INTERVAL=15.

The second parameter, which converts from linear distance to area, is UNITS CONVERSION=, or a shorter form CONVERSION=. For example, if the linear distance between traps is measured in meters, then CONVERSION=1 results in animals/m<sup>2</sup>, whereas CONVERSION=10000 results in animals/ha. To convert from feet to acres, UNITS CONVERSION=43560 would be used; that is, 43 560 ft<sup>2</sup> = 1 A. The default is CONVERSION=10000. As a final example, if traps were placed 30.5 m (100 ft) apart, but density is to be in hectares, the interval is entered in meters, INTERVAL=30.5, and the default of CONVERSION=10000 is used.

Grid definition cards follow the TASK card. Each grid card must specify values for two parameters: X= and Y= determine the range of x- and y-coordinates for the grid, respectively. There can be no embedded blanks in the specification. For example, a card with

X=5-9                    Y=3,8

specifies a 5 by 6 grid with lower left corner at (5,3). Either a hyphen or comma (but not a blank) may be used to separate the values. Labels for the grids punched on the card, such as INNER, MIDDLE, OUTER, and so on, help to interpret the output and will not interfere with the parameter specification. Each grid card has the optional OCCASIONS= parameter, the use and format of which are identical to those described in the TASK CLOSURE TEST.

As many as eight grid cards may be specified. The order in which they appear is not important, although if they are ordered by increasing grid size, the output is easier to interpret. The naive density

estimates are expected to decrease with increasing grid size, and the user can easily note grids that are inconsistent with this pattern if the grid cards are ordered by increasing size.

The last card required is the

#### END OF GRID DEFINITIONS

card, which specifies the end of the input cards required by this TASK. In addition, two optional parameters, DENSITY= and STRIP=, may be set to provide initial values for density and strip width needed to solve the density estimation problem. Initial values should be provided when the user can estimate the value, or when the program has not converged with default values. Default values are calculated from the data, but will not always be close to the final values.

Table A.8 gives an example of input for the TASK DENSITY ESTIMATE. The trap grid is 15 by 15, with 30 ft between traps. Notice the word FEET is placed on the card as a comment, because it is not recognized by the program. To obtain density in acres, CONVERSION=43560 is specified. All five population estimators are to be used. Three nested grids are used, with the largest grid being the total. Default values are for initial values of density and strip width.

### (8) TASK SIMULATE

This task is used to simulate a capture-recapture experiment. As described in Chapter 7, a simulation may be used to determine the sample sizes needed, or the effect that not meeting an assumption would have on an estimator.

Tables 17-19 in *Otis et al. (1978:60-62)*, generated by using TASK SIMULATE, provide the user with an example of the output. This task requires much input. Five parameters may be set on the task card. First, the SEED= parameter provides a random integer used as a starting value to generate random numbers between zero and one. Although this seed usually is somewhat machine specific, a 5- or 7-digit odd integer usually will suffice. The default value is 1234567. The system random number generator is used by the program, so the value of the seed will depend on the type of machine. Therefore the local documentation should be consulted to determine the choice of a seed. A second parameter, POPULATION=, specifies the size of population to be simulated. The default value is 400, with a maximum value of 1000 allowed. OCCASIONS= specifies the number of trapping occasions. The default value is 7, with a maximum of 31 allowed. A third limitation is that POPULATION times OCCASIONS must be less than 4000. REPLICATIONS= specifies the number of experiments (replications) to be simulated. The default value is 50, with no maximum. The number of replications will determine the user's confidence in the output, that is, how precise the estimates are. These parameters may be specified in any order.

**TABLE A.8.** Example input for TASK DENSITY ESTIMATE. TASK READ CAPTURES must have already been executed.

---

*INSTRUCTIONS read from file CAPTIN.*

```
TITLE='EXAMPLE INPUT FOR TABLE A.8'  
TASK DENSITY ESTIMATE INTERVAL=30 FEET CONVERSION=43560 ALL  
INNER GRID X=5-9 Y=5,9  
MIDDLE GRID X=3,11 Y=3-11  
TOTAL GRID X=1,15 Y=1,15  
END OF GRID DEFINITIONS
```

In addition, a PRINT option sets a switch that results in a complete printed output for each experiment. If the user is interested in the MODEL SELECTION output, specifying PRINT will cause it to be printed. Beware, however, of the amount of output that will be printed when the number of REPLICATIONS is large. Do not use PRINT when more than 10 replications are specified. If PRINT is not specified, only the table of summary statistics for the simulations will be printed. This table requires only one page of output, no matter how large the number of REPLICATIONS is. However, more time will be required as the number of REPLICATIONS is increased. We suggest that approximately 100 replications can be expected to provide some useful information. If PRINT is specified, the option X MATRIX may also be specified to have the  $X_{ij}$  matrix listed in the output.

The sixth option that may be set in the TASK SIMULATE card preselects one of the five estimators: NULL, JACKKNIFE, REMOVAL, DARROCH, and ZIPPIN. Only one estimator may be preselected. Normally, TASK SIMULATE selects the appropriate estimator based on the results of TASK MODEL SELECTION. By specifying one of the five estimators, the user is telling TASK SIMULATE not to use TASK MODEL SELECTION, but to go directly to the estimator specified to obtain an estimate. Should the user want to compare two estimators for the same sets of data, two TASK SIMULATE runs must be made, both with the same seed specified.

The most difficult part of the input to TASK SIMULATE is specifying the structure of the probability of captures for the population. Three additional cards may be used for this purpose. The HETEROGENEITY= card specifies a number of individuals and their probability of capture, followed by (optionally) a second number of individuals and their associated probability of capture, and so on. In the following example,

```
TASK SIMULATE POPULATION=150 SEED=4119453 REPLICATIONS=50 OCCASIONS=10
HETEROGENEITY=50,0.5,65,0.3,35,0.1
```

50 animals have 0.5, 65 have 0.3, and 35 have 0.1 probability of capture. This example specifies a total of 150 animals in the population; this value must equal the value specified for POPULATION= on the task card. If only the above card is used to provide capture probabilities, a Model  $M_h$  experiment will be conducted. There are no embedded blanks in the HETEROGENEITY= card.

A Model  $M_{bh}$  experiment is indicated if a BEHAVIOR= card is included with a HETEROGENEITY= card, as in the following example.

```
TASK SIMULATE SEED=4491935 POPULATION=200 OCCASIONS=10 REPLICATIONS=100
HETEROGENEITY=100,0.5,100,0.3
BEHAVIOR=200,1.5
```

In this example, 100 animals have first-capture probability of 0.5, and 100 animals have first-capture probability of 0.3. However, recaptures are influenced by the values on the BEHAVIOR= card. In this example, all 200 animals will have recapture probability of 1.5 times their first-capture probability. If the behavior card had been

```
BEHAVIOR=50,1.5,50,0.5,50,1.5,50,0.5
```

one-half of each of the two groups of animals specified on the HETEROGENEITY= card would have increased recapture probabilities, and one-half would have decreased probabilities. As with the HETEROGENEITY= card, the total number of animals specified must equal the value specified on the TASK card, and no embedded blanks may occur.

A third card for specifying capture probabilities is the TIME= card. The format is different from the above cards. The input

TASK SIMULATE SEED=2288319 OCCASIONS=5 POPULATION=500 REPLICATIONS=30  
TIME=0.9,0.5,0.3,0.5,0.5

specifies that the capture probability on occasion 1 is 0.9, on occasion 2 is 0.5, and so on. This is a Model  $M_t$  experiment. The number of values specified must be equal to the number of occasions specified on the TASK card.

As with the HETEROGENEITY= and BEHAVIOR= cards, the TIME= card will interact with the others through a multiplication process. For example,

TASK SIMULATE POPULATION=200 OCCASIONS=4 REPLICATIONS=100  
TIME=0.5,0.4,0.5,0.4  
BEHAVIOR=100,1.5,100,0.75

results in a first-capture probability of 0.5 for all animals on trapping occasion 1. However on occasion 2, animals not yet captured will have a capture probability of 0.4. Those captured on occasion 1 will have a recapture probability of either  $(1.5)(0.4) = 0.6$ , or  $(0.75)(0.4) = 0.3$ , depending on whether the animal is among the first or second half of the 200 animals in the population. This process continues for the five trapping occasions, providing a Model  $M_{tb}$  experiment.

The last example is one in which all three types of cards are used to simulate a Model  $M_{tbb}$  experiment. The input

TASK SIMULATE POPULATION=200 OCCASIONS=4 REPLICATIONS=50 SEED=459761  
TIME=0.9,0.8,0.9,0.8  
HETEROGENEITY=100,0.9,100,0.5  
BEHAVIOR=50,0.75,50,1.3,50,0.75,50,1.3

means that the initial capture probability on occasion 1 is  $(0.9)(0.9) = 0.81$  for the first 100 animals, and  $(0.9)(0.5) = 0.45$  for the second 100 animals. The BEHAVIOR= card has no effect on capture probabilities on the first occasion because none of the animals are recaptured. However, on occasion 2, the behavior structure is incorporated. If the animal is a recapture its probability will be either  $(0.8)(0.9)(0.75) = 0.54$ , or  $(0.8)(0.5)(1.3) = 0.52$ , depending on whether it is in the first or third group of 50 animals, or in the second and fourth group of 50 animals, respectively. This process continues for the four occasions, and the results are given in Table A.9.

Specifications for Model  $M_o$  can be accomplished in two ways. Both of the following TASKs specify a constant capture probability of 0.5 for the entire population and request that the NULL estimator be selected.

TASK SIMULATE SEED=45763 POPULATION=100 REPLICATIONS=25 &  
OCCASIONS=5 NULL  
HETEROGENEITY=100,0.5  
TASK SIMULATE SEED=45763 POPULATION=100 REPLICATIONS=25 &  
OCCASIONS=5 NULL  
TIME=0.5,0.5,0.5,0.5,0.5.

In addition to the four cards described above, a DATA= card can be used to specify identifying information about the simulation. The format is identical to that given in TASK READ CAPTURES. This card may appear anywhere among or after the three cards used to specify capture probabilities.

## Other Tasks

Several other tasks listed in *White et al. (1978)* are not described in this Appendix. The FORTRAN 77 version of CAPTURE supports them, but we do not encourage their use.

**TABLE A.9.** Capture probabilities for each trapping occasion and capture or recapture status for the example input to TASK SIMULATE.

Animals	First Capture	Recapture
<u>Trapping Occasion 1</u>		
1-50	$(0.9)(0.9)=0.81$	--
51-100	$(0.9)(0.9)=0.81$	--
101-150	$(0.9)(0.5)=0.45$	--
150-200	$(0.9)(0.5)=0.45$	--
<u>Trapping Occasion 2</u>		
1-50	$(0.8)(0.9)=0.72$	$(0.8)(0.9)(0.75)=0.54$
51-100	$(0.8)(0.9)=0.72$	$(0.8)(0.9)(1.3) =0.94$
101-150	$(0.8)(0.5)=0.40$	$(0.8)(0.5)(0.75)=0.30$
151-200	$(0.8)(0.5)=0.40$	$(0.8)(0.5)(1.3) =0.52$
<u>Trapping Occasion 3</u>		
1-50	$(0.9)(0.9)=0.81$	$(0.9)(0.9)(0.75)=0.61$
51-100	$(0.9)(0.9)=0.81$	$(0.9)(0.9)(1.3) =1.05^a$
101-150	$(0.9)(0.5)=0.45$	$(0.9)(0.5)(0.75)=0.34$
151-200	$(0.9)(0.5)=0.45$	$(0.9)(0.5)(1.3) =0.59$
<u>Trapping Occasion 4</u>		
1-50	$(0.8)(0.9)=0.72$	$(0.8)(0.9)(0.75)=0.54$
51-100	$(0.8)(0.9)=0.72$	$(0.8)(0.9)(1.3) =0.94$
101-150	$(0.8)(0.5)=0.40$	$(0.8)(0.5)(0.75)=0.30$
151-200	$(0.8)(0.5)=0.40$	$(0.8)(0.5)(1.3) =0.52$

<sup>a</sup>Program CAPTURE will reduce values greater than 1.0 to a capture probability of 1.0.

# APPENDIX B

# ANSWERS TO QUESTIONS AND EXERCISES

## Chapter 1

1. Yes. Each animal can have its own unique probability of capture on each occasion, and this can change after first capture.
2. Yes. Any summary statistic can be computed from the  $X$  matrix.
3. Model  $M_0$  is not often used in real population work because it makes assumptions that are rarely valid, for example, equal catchability. Moreover, the estimate of  $N$  is poor when these assumptions are violated.
4. No. It is extremely important. The biologists must consider fully both demographic and geographic closure before conducting a trapping program. Without geographic closure,  $N$  is not even defined and, therefore,  $\hat{N}$  is difficult to interpret.
5. a.  $t = 7, j = 1, \dots, t$ .  
 b.  $n_1 = 3$   
 $n_2 = 5$   
 $n_3 = 3$   
 $n_4 = 3$   
 $n_5 = 3$   
 $n_6 = 3$   
 $n_7 = 2$   
 c.  $u_1 = 3 = n_1$  (always).  
 $u_2 = 4$ .  
 d. They are either trap happy (Model  $M_b$ ) or they are animals that have a high capture probability (Model  $M_h$ ). You could argue that animal 6 is trap shy, because it was not captured on any of the 5 nights after its first capture.  
 e.  $M_{t+1} = M_8 = 7$  animals; 7 distinct animals were caught at least once during the study.  
 f.  $M_1 \equiv 0$ ; there are no marked animals at the first trapping occasion.
6. No. This situation is not encompassed by Model  $M_b$ .
7. No. Under closure,  $S \equiv 1$  (no deaths).
8. These sampling methods will provide only an estimate of the parameter  $N$  of interest. Tell him to drain the pond and count the fish if he must know the exact population size.
9. The estimators  $\hat{p}$  and  $\hat{N}$  are closely coupled. The estimates of capture probabilities in the model and the estimate of  $N$  are directly related. A good estimate of population size depends on good estimates of capture probabilities. Finally, if  $p$  is small, few data will be available for analysis. All the methods perform better if  $p$  is large.
10. If you do, you are hallucinating.
11. Very general models usually do not have estimators (for example, Model  $M_{t(bh)}$ ), because of the large number of parameters they require. Furthermore, use of a model that is too general will lack the precision that one desires. For example, if you conducted a tutorial ball and urn experiment, the proper model would be  $M_0$ . The estimator  $\hat{N}$  from this model would have good properties, and the estimated sampling

These are merely column totals in  $X$ :

$$n_j = \sum_{i=1}^N x_{ij} .$$

variance (discussed in detail in Chapter 2) would be small. However, if the ball and urn sample data were analyzed under the more general Model  $M_h$ , the estimator  $\hat{N}$  would still be unbiased, but the sampling variance would be substantially larger.

12. a.  $t = 3$ , the number of trapping occasions.
- b.  $M_{t+1} = 20$ , the number of different animals captured.
- c.  $n_1 = 7$ ,  $n_2 = 11$ , and  $n_3 = 11$ , the sum of each column of the  $X$  matrix, respectively.
- d.  $29 (= 7 + 11 + 11)$ , the total number of captures.
- e.  $u_1 = 7$ ,  $u_2 = 9$ ,  $u_3 = 4$ .
- f.  $f_1 = 13$ ,  $f_2 = 5$ ,  $f_3 = 2$ . (Note that  $M_{t+1} = f_1 + f_2 + f_3 = u_1 + u_2 + u_3 = 20$ .)
- g.  $M_1 = 0$ ,  $M_2 = 7$ ,  $M_3 = 16$ , and  $M_4 = M_{t+1} = 20$ .
- h.  $M = 23$ .
- i.  $m_1 = 0$ ,  $m_2 = 2$ , and  $m_3 = 7$ .
- j.  $m = 9 (= m_1 + m_2 + m_3)$ .

## Chapter 2

1. No. See, for example, the equation for  $\hat{N}$  from Model  $M_t$  in *Otis et al. (1978:106)*.
2. Models form the basis from which estimators of parameters can be derived by providing a mathematical expression of the assumptions in terms of parameters. Some parameters such as annual survival rates, cannot be “observed,” “measured,” or “counted,” and models form a basis for estimating them. Estimation procedures developed without an explicit underlying model are termed *ad hoc*.
3. No. It is still just that—an *ad hoc* approach.
4. This is difficult to say, as these answers depend on many factors. As a rough guide, a cv of 10-15% might be useful for research. Management-oriented studies might provide useful results, if the cv were as large as 20-50%, or even larger in some cases.
5. Study 1, unbiased, precise.  
Study 2, biased, precise.  
Study 3, unbiased, not precise.  
Study 4, biased, not precise.
6.  $T_2$  is preferred, because it will reject a false null hypothesis with probability 0.89.
7. Normal, chi-square, F, t, z.
8.  $H_0$ : the model fits the data.  
 $H_A$ : the model does not fit the data.
9. A true null hypothesis may be rejected (a Type I error) or a false null hypothesis may not be rejected (a Type II error).
10. a. 95% C. I.  $= \hat{\theta} \pm 1.96 \hat{se}(\hat{\theta})$ ,  
 $= 141 \pm 1.96 \times 13.1$ ,  
 $= 115.32 \text{ to } 166.68$ .  
b. It is unlikely that  $\theta = 95$  because this value is far outside the interval.  
c.  $\theta = 135$  is very plausible: it is close to  $\hat{\theta}$  and well within the confidence interval.
11. a. Yes.  $H_0: \theta = 95$ .  $H_A: \theta \neq 95$   
and  $H_0: \theta = 135$   $H_A: \theta \neq 135$ .  
b.  $z = (\theta - \hat{\theta})/\hat{se}(\hat{\theta}) = (141 - 95)/13.1 = 3.51$ ,  
 $z = (\theta - \hat{\theta})/\hat{se}(\hat{\theta}) = (141 - 135)/13.1 = 0.46$ .  
c. The test statistic  $z$  is distributed normally with a mean of zero and a standard deviation of one. If the significance level of the tests is chosen as 0.05, we can see from Fig. 2.11 that the null hypothesis  $\theta = 95$  is rejected, while the null hypothesis  $\theta = 135$  is not rejected.

12. Nothing. Without a measure of precision, nothing can be inferred about the true population sizes of the two areas. Tell your colleague to get his act together.
13. a. Lake bass in Wabo tributary of Lake Powell.  
 b. A census seems impossible—it is better to decide on a sampling method that will provide valid inferences from the sample to the population.  
 c. Population size of “adult” bass—fish capable of breeding.  
 d. Capture-recapture or removal sampling should be considered.  
 e. If  $N_1$  and  $N_2$  are the true population sizes before and after drilling, the hypotheses might be  

$$H_0: N_1 = N_2$$

$$H_A: N_1 > N_2.$$
14. No. This is a common misinterpretation of the meaning of a confidence interval. The correct inference is that if the identical study were repeated a large number of times, 95% of such intervals would cover the true parameter.
15. You would conclude that the null hypothesis is false, which is incorrect.
16. a.  $cv = 0.13, 0.18, 0.22, 0.16,$  and  $0.19,$  respectively.  
 b. Yes,  $cv$ 's of about 20% are reasonable, and each estimate is close to the true parameter value.  
 c. Probably not. The coefficients of variation are fairly large compared to the actual changes in the population.
17. Until computers became widely available about 10 years ago, approximations had to be made so that the estimation could be done on simple calculating machines.
18. No. However, the more that is known, the better the understanding and interpretation can be.
19. The estimator may be biased, the estimated sampling variance may be too small, or the sample size may be too small. (The normality assumption may not be satisfied.)
20. You can conclude that there is strong evidence that  $H_0$  is false, because if the null hypothesis were true you would expect to observe the data you collected only 7 times in 1000 studies. This is very unlikely, so you reject  $H_0$ .
21.  $se(\hat{N}) = \sqrt{\text{var}(\hat{N})} = \sqrt{625} = 25.$
22. No. We find large biases; the average estimate, computed from the estimator under Model  $M_r$ , differs greatly from the parameter  $N$ .

### Chapter 3

1. No. Equal probability of capture is not necessarily achieved by only a high level of trapping effort. The behavior of the animal also is involved.
2. Yes, often very much so.
3. No, although a removal estimator could be used. Unfortunately, a paper was published claiming that for  $t = 2$ ,  $N$  could be estimated by the Petersen-Lincoln method even when there were no captures (*Bell 1974*).
4. No. It is impossible to get enough captures and recaptures to test assumptions and compute reliable estimates of  $N_1$ .
5. Basically, no. If the average capture probability  $\bar{p}_1$  for that trapping occasion is known, then  $\hat{N} = n_1/\bar{p}_1$ , but in practice  $\bar{p}_1$  will not be known.
6. No.
7. Model  $M_r$ . Yes, if sample sizes are sufficient.
8. All except the number of days of trapping.
9. Model  $M_0$  cannot fit these data. Increasing the trapping effort over time will cause average daily capture probabilities to increase, hence to vary with time. Therefore, Model  $M_r$  might be the true model, but neither  $M_h$  nor  $M_p$  can be the true model.

10. (b).
11. No. ML estimates under these models do not exist. One might impose additional assumptions and then obtain an estimator; however, this would change the model. Also, a nonparametric approach might be used to produce an estimator.
12. No. Model  $M_0$  is not at all robust. If sample size is small, the power of tests of assumptions is low, and often  $M_0$  will appear to fit. This situation (a Type II error) is serious, because  $M_0$  is very poor if its assumptions are violated.
13. Because  $\hat{N}_h$  (the estimator under Model  $M_h$ ) is much more robust than  $\hat{N}_0$ .
- 14 No. A completely general statistical test for closure is not possible.

## Chapter 4

1. Your answer should be an emphatic “No.”
2. The most serious defect is that the assumption of constant capture probability cannot be tested, and if that assumption is false the estimator based on it is biased. Even if the constant capture probability model is true, the estimate of  $N$  will be very imprecise unless capture probabilities exceed 0.40.
3. No.
4. Yes. Animals are “removed” from the population by marking them.
5. Yes, because there is additional information from recaptures.
6. We certainly hope not.
7. Closure will fail; that is, animals from outside the grid are often attracted by the “vacuum” left by removed animals.
8. If the removal is accomplished by marking, it may be an acceptable plan if the population is large enough, say 750. If removal is by physical detachment, relocations, etc., the proposal is likely to be politically unacceptable. If the removal involves killing the snails, the biologist is in trouble.
9. No. There is clearly no meaningful decline in the numbers removed over occasions 1 to 5. This study has failed.
10. a. Each row of the  $X$  matrix has exactly one 1 in it, and the remaining entries are zero. In the first 68 rows, the 1 is in column (occasion) 1. In the next 41 rows, the 1 is in column 2. Then there are 25 rows all with a 1 in column 3, and finally 15 rows with a 1 in column 4. The total matrix is 149 rows by 4 columns.
  - b. The study results are acceptable; by looking at the decrease in the removals, we can expect a reasonably precise estimate of  $N$  if the constant capture probability model fits.
  - c. The simplest “quick and dirty” estimate of  $N$  is  $M_3 = 149$ , which, of course, will be low. The next quick estimate is to use occasions 1 and 2 only and Eq. (4.1),

$$\hat{N} = \frac{u_1}{1 - u_2/u_1} = \frac{68}{1 - 41/68} = 171 .$$

This estimate would suffice to satisfy our curiosity while we still were in the field, but a full-blown analysis requires testing assumptions. We leave it to the reader to apply program CAPTURE to these data.

11. No. The results will be garbage. The expected removals are  $E(u_1) = 5$ ,  $E(u_2) = 4.75$ , and  $E(u_3) = 4.51$ .
12. We recommend a capture probability  $p$  of at least 0.2, and  $p \geq 0.3$  is needed to be sure the results will be reliable.
13. Closure has failed after occasion 3; animals not originally in the population are moving into it.
14. It will be worthless—and very misleading if presented without the evidence from the data that closure has failed.

15. There will be  $N - (u_1 + \dots + u_{j-1})$  individuals left in the population on the  $j^{\text{th}}$  removal sampling. The average capture probability of these *remaining* individuals is  $\bar{p}_j$ . If capture probabilities vary in the population, due to innate heterogeneity, then on the first sample individuals with the higher capture probabilities tend to be caught. The individuals remaining (uncaught) on the second removal occasion therefore will have, on the average, lower capture probabilities. Thus, we have  $\bar{p}_1 > \bar{p}_2$ . By the same argument, the even fewer numbers of individuals remaining after the second removal sample again have smaller average capture probability  $\bar{p}_3$ , compared to  $\bar{p}_2$ .

16. Capture probabilities of fish vary greatly by species, and noticeably by fish size, for electrofishing methods. Capture probabilities of small mammals can vary by species, sex, and age; they also can vary due to social dominance and, especially for animals near the edge of the grid, home range size and the number of traps in the home range.

17. Capture probabilities will vary by time, leaving us with an  $M_{tb}$  type of model and making estimation of  $N$  impossible by removal methods. Catch-effort methods could be used, but then the relative effort on each occasion must be known and quantified and the analysis methods are different from those in *Otis et al.* (1978) or in program CAPTURE (cf. *Seber 1973:296-353*).

## Chapter 5

1. No. Nested subgrids could not be constructed, nor would adequate data be obtained.
2. Not necessarily. Animals may not have home ranges that overlap the grid, which implies that  $W = 0$ , but they may still come and go from the grid, thus violating geographic closure.
3. The number of ellipses that intersect or are contained in the grid.
4. The choice depends on animal density. If all traps are expected to be filled on each occasion, the probability of capture may actually be lowered due to nonavailability of traps. However, all traps usually are not filled, and therefore, one trap per station and a larger grid are preferred. Also, placing the traps at half intervals and using the same size grid is preferable to placing two traps per station.
5. No. Subgrids consisting of halves or quarters are biased when a linear gradient in density exists across the grid. Nested subgrids are robust to such a gradient; that is, they will produce an unbiased estimate of the *average* density in the grid.
6. No. Nested subgrids cannot be constructed from one long line of traps.
7. The additional area included in the strip of width  $\hat{W}$  around the grid enlarges the area  $\hat{A}$  to which  $\hat{N}$  applies, so that  $\hat{D} = \hat{N}/\hat{A}$  is reduced.
8. Density is expressed in terms of animals per unit area, whereas population size simply represents an absolute number of animals.
9. The  $X$  matrix does not contain information about capture location. To estimate strip width, and hence density, information concerning the movement of individual animals is obviously required.

## Chapter 7

1. It may be logistically easier to use 100 traps for 8 nights. However, using 200 traps for 4 nights may avoid a closure problem and result in increased capture probabilities.
2. Assume that the home range of the animal is circular, and therefore, the radius is 56.4 m. The formula  $s \leq (\sqrt{2}) W$  gives  $s \leq 80$  m. For  $s \leq W/2$ ,  $s$  should be 28 m. Spacing should be set somewhere within this range, with the actual value depending on the size of the grid and the number of traps available.
3. Model  $M_{tb}$  results, and no estimator is available.
4. The probability of capture is likely to vary by occasion (an unwanted source of variation), because one would not expect an animal to have the same capture probability during the day as during the night.

One could pool the morning and evening captures, if time variation is indicated, or analyze them separately, if enough data are available. Checking traps twice daily is certainly preferable to checking only once per day.

5. Closure will be assured.
6. The MODEL SELECTION procedure lacks power, that is, the ability to identify sources of variation in capture probability, when the probabilities of capture are small. In this population, probabilities generally average less than 0.10, and thus Model  $M_0$  is selected by default because none of the tests reject any of the hypotheses.
7. The capture probabilities on Tuesday night probably would differ from the remaining occasions, resulting in the presence of time variation. Therefore, the study probably should be continued for at least one more night past the planned termination, to avoid models with time variation. During the analysis, the OCCASIONS= option could be used to eliminate the Tuesday data from model selection and estimation.

## Chapter 8

1. a. Geographic closure will be violated. It may be difficult to obtain adequate sample size.  
b. Survival rates and sampling rates may be the only parameters that can be estimated due to the lack of geographic closure.  
c. Catch per unit effort (CPUE) methods, such as *Dupont (1976)*, might be appropriate.
2. Yes. A closed model assumes  $S \equiv 1$ .
3. Yes. More parameters must be estimated. (See *Cormack 1979:241*.)
4. a. Not necessarily, because the estimators  $\hat{S}_j$  and  $\hat{N}_j$  have a high sampling correlation as they are computed from the same data.  
b.  $cv(\hat{N}_4) = 59/422 = 0.14$  or 14%.  
c.  $0.65 \pm 1.96 (0.04)$  or about 0.57 to 0.73 or 57% to 73%.  
d. Yes, by definition.
5. Yes. At least a good approximation can be computed by taking a weighted average of the annual survival rates; a complex iterative procedure is required due to the covariance structure among the estimators. Alternatively, *Jolly (1979)* provides a model for constant survival rate.

# APPENDIX C

## GENERAL READING LIST

The material below presents some guidelines helpful to biologists in getting into the literature. Because the levels of mathematics differ, we have coded the references, to help the reader. The codes are found to the left of each reference, and the code meanings are as follows.

Code	Meaning
C	Conceptual or review paper
S	Specific methodology reported in paper
D	Difficult mathematics (relatively)
L	Little mathematics (relatively)
V	Various levels of mathematics
*	Good initial reading

References to many specific methods derived under the closure assumption are contained in *Otis et al. (1978)*. At this time, no major synthesis has been published for the open-population models; therefore, we provide a number of state-of-the-art references for these models.

### General Reviews

There is very little overlap of intent among the three references.

- CV Cormack, R. M., 1968, The statistics of capture-recapture methods, *Oceanogr. Mar. Biol. Annu. Rev.* 6, pp. 455-506. The first major synthesis and review of existing capture-recapture and removal methods. Several related methods are covered briefly. A readable introduction and critique of the literature that is still useful. Recommended reading.
- CV\* Cormack, R. M., 1979, Models for capture-recapture studies, in R. M. Cormack, G. P. Patil, and D. S. Robson, eds., *Statistical ecology*, Vol. 5, Sampling biological populations, pp. 217-255, International Co-operative Publishing House, Fairland, Maryland. A somewhat more technical review of classes of models and approaches to capture-recapture problems. This is a different type of review from Cormack's 1968 paper, aimed at a slightly more sophisticated audience. Recommended reading.
- CV\* Seber, G. A. F., 1982, *Estimation of animal abundance and related parameters* (2nd ed.), Griffin, London (in press). An in-depth summary of nearly all the literature on capture-recapture and removal methods is found in Chapters 3, 4, 5, 6, and 7 of this edition. Examples of many of the methods are given. Seber's book is certainly recommended reading.

### Estimation Methods for Closed Populations

- SV\* Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson, 1978, Statistical inference from capture data on closed animal populations, *Wildl. Monogr.* 62, 135 pp.

## Estimation Methods for Open Populations

- CL\* Cormack, R. M., 1972, The logic of capture-recapture estimates, *Biometrics* 28(2), pp. 337-343.
- CL\* Cormack, R. M., 1973, Commonsense estimates from capture-recapture studies, *in* M. S. Bartlett and R. W. Hiorns, eds, *The mathematical theory of the dynamics of biological populations*, Academic Press, London, pp. 225-234.
- SD Jolly, G. M., 1965, Explicit estimates from capture-recapture data with both death and immigration—stochastic model, *Biometrika* 52(1/2), pp. 225-247.
- SD Jolly, G. M., 1979, A unified approach to mark-recapture stochastic models, exemplified by a constant survival rate model, *in* R. M. Cormack, G. P. Patil, and D. S. Robson, eds., *Statistical ecology*, Vol. 5, *Sampling biological populations*, International Co-operative Publishing House, Fairland, Maryland, pp. 277 and 282.
- SL Manly, B. F. J., and M. J. Parr, 1968, A new method of estimating population size, survivorship, and birth rate from capture-recapture data, *Trans. Soc. Brit. Ent.* 18, pp. 81-89.
- SD Pollock, K. H., 1975, A K-sample tag-recapture model allowing for unequal survival and catchability: *Biometrika*, 62(4), pp. 577-584.
- CV\* Pollock, K. H., 1978, Building models of capture-recapture experiments. *The Statistician* 25(4), pp. 253-259.

## General Computer Program Packages

- SL Arnason, A. N., and L. Baniuk, 1978, POPAN-2: A data maintenance and analysis system for mark-recapture data, Charles Babbage Res. Cent., St. Pierre, Manitoba, Canada, 269 pp.
- SL White, G. C., K. P. Burnham, D. L. Otis, and D. R. Anderson, 1978, User's manual for program CAPTURE, Utah State University Press, 40 pp.

## Related Reading

- SV Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson, 1978, *Statistical inference from band recovery data*, A handbook, U. S. Fish and Wildlife Service Resource Publication 131, 212 pp.
- SV Paulik, G. J., and D. S. Robson, 1969, Statistical calculations for change-in-ratio estimators of population parameters, *J. Wildl. Manage.* 33(1), pp. 1-27.
- SL Robson, D. S., and D. G. Chapman, 1961, Catch curves and mortality rates, *Trans. Am. Fish. Soc.* 90, pp. 181-189.
- SV Ricker, W. E., 1975, *Computation and interpretation of biological statistics of fish populations*, *Can. Fish. Mar. Serv. Bull.* 191, 382 pp.

# CHAPTER 1

## INTRODUCTION

Interest in estimating the size of populations has had a long history. The crudest methods date back at least to the 17th century and probably long before that. Applied capture-recapture methods with some theoretical basis began to appear in the 1930s and 1940s. The period since then has seen many developments in methods for estimating population size under a wide variety of assumptions and situations. Generally, modern methods can be classified into two groups: capture-recapture and removal methods.

Just what is a capture-recapture study? We will describe a simple example before we discuss more realistic cases (Chapter 3). Suppose that we wish to estimate the size  $N$  of a population in which there is no birth, death, immigration, or emigration over the time period for which we plan to make the estimate. On a first visit, we catch a sample of 1000 animals, mark them so that we can recognize individual animals in the future, and return them to the population where the marked animals mix with unmarked animals. We denote the number of animals in this first sample as  $n_1$ ;  $n_1 = 1000$  in this example. On a second visit, we catch 500 animals (call this  $n_2$ ), of which 450 are unmarked (call this  $u_2$ ) and 50 ( $n_2 - u_2$ ) are marked. In this example, the proportion of the population that is marked, called the "capture probability," is denoted by  $p$  and its estimate by  $\hat{p}$ . The estimate of  $p$  is  $\hat{p} = (n_2 - u_2)/n_2$ , or  $50/500 = 0.10$ . Because the number of marked animals in the population is 1000, an estimate of the total population is  $\hat{N} = n_1/\hat{p}$ , or  $1000/0.10 = 10\ 000$  animals. To improve the precision of  $\hat{N}$ , we can mark the 450 unmarked individuals, return them to the population, and resample. On the third visit we catch more animals (call this  $n_3$ ), of which some are marked and some are not. This procedure can be continued indefinitely as long as the initial assumptions (no births, deaths, immigration, or emigration) hold. Full details of this and other capture-recapture models are given in Chapter 3.

In a removal study, in contrast to the capture-recapture study described above, animals are captured and removed from the population rather than being marked and released. On the second and subsequent visits, more animals ( $u_2, u_3, u_4, \dots$ ) are captured and removed; continued sampling will catch progressively fewer animals on each occasion until eventually none will remain to be captured. The progressive decrease in the first few terms of the series  $u_1, u_2, u_3, \dots$ , is used to estimate  $N$ , the total number of animals. Alternatively, we can mark the captured animals and return them to the population. In this way, they are "removed" from the unmarked population without our having to remove them physically. This kind of removal allows us to view capture-recapture experiments as if they were removal experiments; the advantages of this method will be explained later. Details of removal models are given in Chapter 4.

Capture-recapture and removal sampling are useful methods in some situations. On the whole, however, their usefulness is much overrated in the biological literature. In this primer we explain the methods, the assumptions on which they are based, and their average performance.

We believe a summary of the literature on the state of the art will be helpful. *Cormack (1968, 1979)* and *Seber (1982)* give detailed reviews. Appendix C, the General Reading List, provides an index to the relevant sources. Typically, only special cases of models for capture-recapture and removal experiments have appeared, and these have often involved approximations and simplifications. With some important exceptions, little emphasis has been placed on testing assumptions within or between existing methods. Many methods presented in the literature are *ad hoc* (without a firm basis), and some are demonstrably incorrect. Rarely have new methods been compared with older methods based on the same assumptions.



Richard Cormack

George Seber probably is best known to ecologists for his work on several open-population models and the two editions of his book, *Estimation of Animal Abundance and Related Parameters*. Of tremendous value to biologists and statisticians, his book draws together in a cohesive treatment the literature that had been scattered across various biological and statistical disciplines.

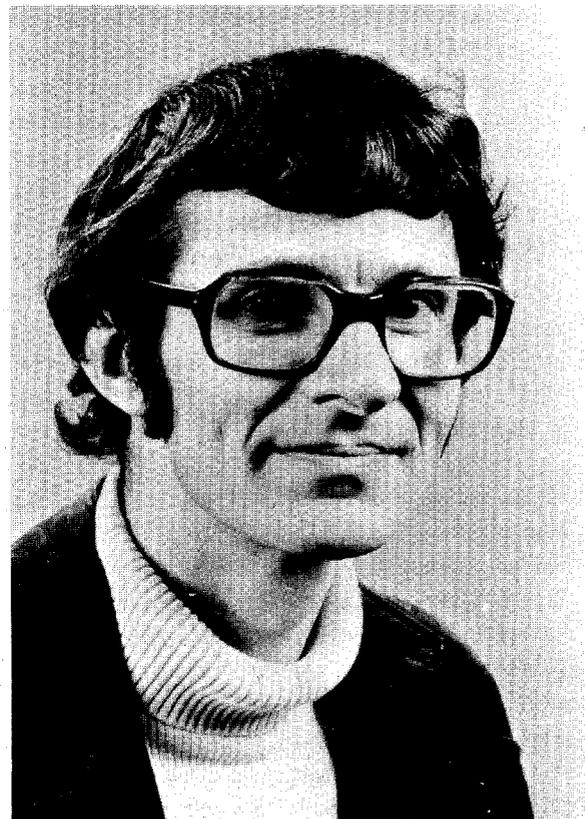
Seber became interested in capture-recapture models less by design than by accident, through his association with J. N. Darroch. His primary interest has been in statistical theory. He has found the quantitative aspects of ecology challenging and interesting from both practical and theoretical standpoints. He believes that mathematical models are important but that they must have a sound biological basis.

Seber took a B.Sc. degree and an M.Sc. degree in mathematics at Auckland University, New Zealand, and a Ph.D. degree from the University of Manchester, England. He served as a statistician at the London School of Economics, as professor of mathematics at Auckland University, and as Director of the Biometrics Unit at Otago University, New Zealand. He now is head of the Mathematics Department at Auckland University. (Recent photograph.)

Richard Cormack has been interested in capture-recapture theory, primarily open-population models, for the past 20 years. He was born and raised in Scotland and educated in England. He took his degrees at Cambridge intending to be a theoretical astronomer, but after graduating he elected to work in the field of mathematical statistics. He spent a year in the United States at the University of Washington, with Douglas Chapman. Cormack worked at the University of Aberdeen for 10 years and at the University of Edinburgh for 6 years before going to the University of Saint Andrews in eastern Scotland, where he holds a chair in statistics.

A review paper in 1968 on capture-recapture methods invited by the *Annual Review of Oceanography and Marine Biology* followed his paper in 1964 on the estimation of survival rates from capture-resighting data. More recently, he has published a series of papers relating primarily to open-population models. He was closely involved with the work of A.N. Arnason and A. D. Carothers while at Edinburgh. Cormack now is studying the use of log-linear methods for the analysis of capture-recapture data and feels this methodology is promising.

Statistical ecology remains his primary interest. He believes that the fascination of statistics resides in its capacity to help other scientists, and that this is best fulfilled by work with others. (Recent photograph by Peter Adamson.)



George A. F. Seber

The results of computer simulations that compare methods or that examine the small-sample properties of a given method have been published only recently. Most estimation methods appear to be very sensitive to the breakdown of certain assumptions: they are not "robust." Little admission has been made of the fact that models developed under the closure assumptions are merely variations on the classic "ball and urn" model (*Feller 1950*). An analogy between this model and real animal populations has not been made, but few biologists seem to be aware of this lack. Finally, in the past decade or so, attention has been focused on the traditional assumption that all members in a population are equally catchable on all occasions. It is now recognized that this assumption rarely holds, and much work has been done in recent years to build models that allow the assumption to be relaxed.

### **Closure: An Important Assumption**

Closure means that the size of a population is constant over the period of investigation: no recruitment (birth or immigration) or losses (death or emigration) occur. This is a strong assumption, and of course it is never completely true in a biological population. For greater generality, we define closure to mean that there are no unknown changes to the initial population. In practice, this means known losses (trap deaths, or deliberate removals) do not violate our definition of closure. If the study is designed properly, closure can be met, at least approximately.

Closure is a very important assumption because all previous capture-recapture and removal models are extensions of ball and urn models (*Feller 1950:45-47*). We subdivide the closure concept into two components:

- (1) "geographic" closure by a boundary, analogous to the sides of an urn, that limits the population.
- (2) "demographic" closure to birth, immigration, death, and emigration.

The distinction between geographic and demographic closure is important because open models (mentioned in Chapter 8) are open only to demographic closure: geographic closure is still a critical assumption.

Unless geographic closure is met, the area relating to the parameter  $N$  is not defined and  $N$  itself has no meaning. For example, geographic closure is met with fish in a small pond, mammals on a small island, or squirrels in an isolated woodlot. Geographic closure is violated when capture-recapture is done with a relatively small grid of traps in a very large field inhabited by small mammals. Similarly, geographic closure probably will be violated when fish are sampled in a large reservoir, unless the whole reservoir can be sampled properly.

The subject of geographic closure arises because models for capture-recapture and removal data are based on ball and urn studies, which assume a three-dimensional container (Figs. 1.1-1.3). In these studies, the samples are drawn from the urn, and the objects are marked and returned to the urn, which is shaken to mix the marked and unmarked objects randomly before the population is sampled again. The concept of geographic closure comes about when models for populations in three-dimensional containers are applied to two-dimensional areas. For example, a biologist may wish to sample a 2- by 3-km island to estimate the size of the meadow vole population. He typically will position a series of traps on a grid or lattice. Unless his grid is 2 by 3 km, he faces problems because he has not met the requirement for geographic closure. This subject is discussed in detail in Chapter 5; see especially Figs. 5.1-5.3.

Sometimes the assumption of demographic closure can be relaxed. *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 instances, the population estimate relates only to the population size at the beginning of the study. However, if both recruitment and mortality occur during the experiment and if both marked and unmarked animals are affected similarly by mortality, the estimate of  $N$  will be too high, on the average, for both initial and final population size (*Robson and Regier 1968*).

Finally, we remark that removal studies often invite failure of the closure assumption. The removal of a significant number of animals may create a "vacuum," and animals outside the area may move into the

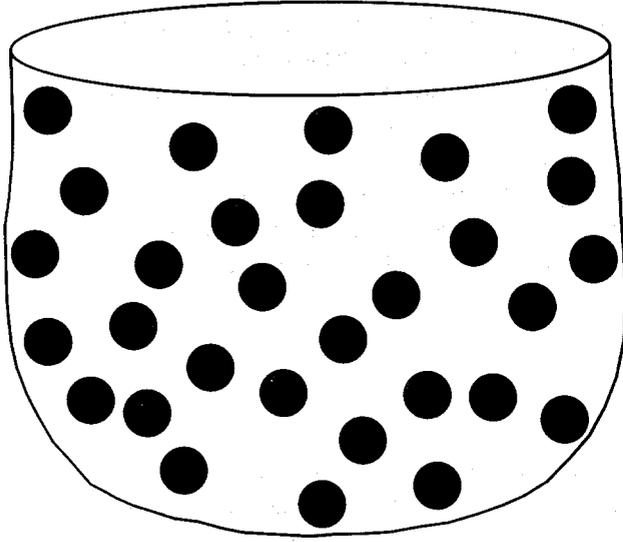


Fig. 1.1. The classic ball and urn experiment is the basis for the ball and urn model (Model  $M_0$ ). Because balls neither die nor give birth, nor do they immigrate or emigrate, the demographic closure assumption is met. The sides of the glass urn limit the population boundaries, and, therefore, ensure geographic closure.  $N$  is well defined and here  $N = 30$  balls.

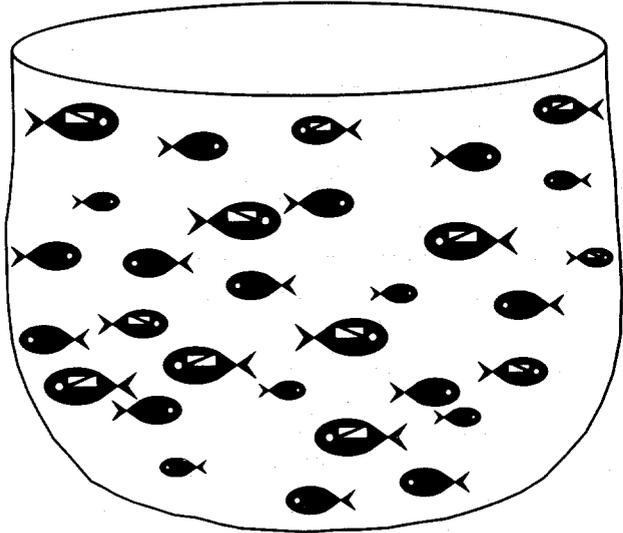


Fig. 1.2. A population of 30 adult fish, of which 12 are marked, in a glass container. Geographic closure is assured because the glass bowl confines the population. The bowl also prohibits immigration and emigration. If the length of a capture-recapture study is short (say 4 days), death can be assumed to be negligible; if it is not, dead fish will be noticed in the container. Any reproduction can be ignored on the basis of the small size of the young. Therefore, demographic closure is valid.

Although the closure assumption is met, the simplistic assumptions of Model  $M_0$  are probably violated. Notice that the larger fish tend to be more prone to capture and, thus, to be marked. Therefore, we might expect Model  $M_h$  to be appropriate. (Models  $M_0$  and  $M_h$  are discussed in this chapter and in Chapter 3.)

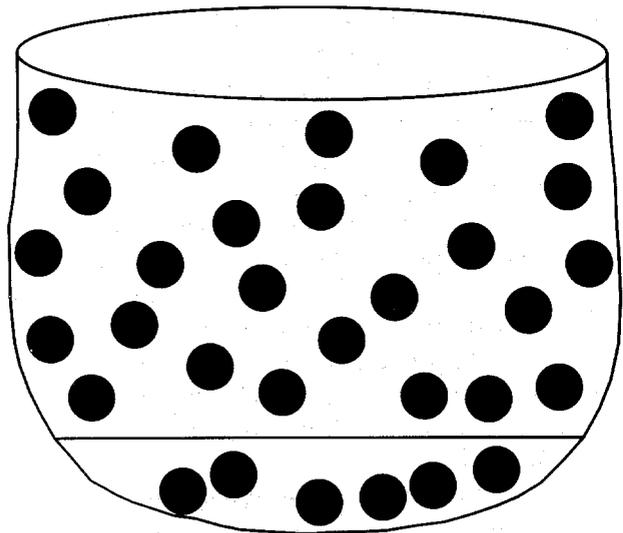


Fig. 1.3. A closed population of  $N = 31$  individuals. The six individuals below the line are not subject to capture; that is, their capture probability  $p$  is zero. When such conditions exist,  $\tilde{N}$  relates only to the catchable population, 25 in this example. Very large fish in a lake, old and wary coyotes, and mosquitoes distant from the nearest trap may be examples of individuals that are essentially untrappable.

trapping area and become subject to capture. In addition, the use of baited traps may induce movement of animals into the trapping area. If capture and marking methods induce mortality, demographic closure is violated. All these examples deal with violation of the closure assumption.

Many models have been developed for "open" populations. The concepts are discussed briefly in Chapter 8. Models for open populations are critically dependent on geographic closure for estimating population size, but they allow mortality + emigration and birth + immigration rates to be estimated.

## Data

In capture-recapture studies, the same individuals are in the population on each trapping occasion,  $j = 1, 2, \dots, t$ , because of the closure assumption. (In removal studies, some individuals are removed on each sampling occasion.) Therefore, we can conceive of the individuals as being numbered from 1, 2,  $\dots$ , to the last individual,  $N$ ; that is,  $i = 1, 2, \dots, N$ . The capture and recapture history of each animal on each sampling occasion can be expressed conveniently in a simple table called the  $X$  matrix and denoted as  $[X_{ij}]$ . Let

$$[X_{ij}] = \begin{bmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1t} \\ X_{21} & X_{22} & X_{23} & \dots & X_{2t} \\ X_{31} & X_{32} & X_{33} & \dots & X_{3t} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ X_{M_{t+1}1} & X_{M_{t+1}2} & X_{M_{t+1}3} & \dots & X_{M_{t+1}t} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ X_{N1} & X_{N2} & X_{N3} & \dots & X_{Nt} \end{bmatrix}$$

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

Note that the first subscript denotes the row, and the second subscript denotes the column. For example,  $X_{31}$  represents the third row, first column.

The  $X$  matrix contains only zeros and ones indicating "not captured" and "captured," respectively. Each column details the history by sampling occasion,  $j = 1, 2, \dots, t$ . The first  $M_{t+1}$  rows relate to the capture and recapture history of each animal that was captured at least once during the study. The remainder of the  $X$  matrix contains all zeros, because these animals were never captured. In real studies, of course, one does not know how many remaining rows there should be.

The following is an example of an X matrix.

$$\begin{array}{r}
 \\
 \\
 \\
 \\
 [X_{ij}] = \\
 \\
 \\
 \\
 \\
 \\
 \\
 \\
 \\
 \end{array}
 \begin{array}{cccc}
 j = 1 & j = 2 & j = 3 & j = 4 = t \\
 \left[ \begin{array}{cccc}
 1 & 0 & 0 & 1 \\
 1 & 0 & 1 & 0 \\
 0 & 1 & 0 & 0 \\
 0 & 0 & 1 & 0 \\
 0 & 0 & 1 & 0 \\
 0 & 0 & 1 & 1 \\
 0 & 0 & 0 & 1 \\
 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0
 \end{array} \right]
 \end{array}$$

In this example, we see that animal 1 was captured on the first occasion and was recaptured on the fourth occasion. Animal 2 also was captured on the first occasion but was recaptured only on the third occasion. Animal 3 was captured on the second occasion and never was recaptured. Seven different animals were captured ( $M_{t+1} = 7$ ) in this 4-day ( $t = 4$ ) study. We can see that two animals were never captured ( $N - M_{t+1} = 9 - 7 = 2$ ) because the two bottom rows, all zeros, indicate that these animals were never caught. Thus it is clear that only a portion of the X matrix is "observed" during a capture-recapture experiment. Of course, in a removal study there are no recaptures, so each row contains, at most, one nonzero entry.

The type of data we collect limits the parameters that we can estimate. In general, we can estimate population size (N), survival (S), and capture or recapture probability (p or c) from a capture-recapture study; however,  $S \equiv 1$  for a closed population. In removal studies, only N and p can be estimated. In some studies, only data on marked animals, such as tag recoveries from dead animals, can be taken. In these studies, S and p can be estimated, but N cannot (see *Brownie et al. 1978*). These restrictions reflect inherent relations between the collected data and the parameters that can be estimated from them.

## Parameters

If we assume population closure, only two parameters are of primary interest:

N = population size (a constant)

and

D = population density (animals per unit area).

Estimation of D is more difficult than estimation of N (see Chapter 5).

The capture probabilities are just as important as the N and D parameters, but are of less biological interest. The nature and number of capture probabilities depend on the assumptions being postulated (the model). For example, when we contrast capture-recapture studies with removal studies, we are interested in the initial capture and recapture probabilities, defined as

p = capture probability, the probability of first capture for an animal,

and

c = recapture probability, the probability that a marked animal will be recaptured.

In "trap happy" populations, we have  $p < c$ , whereas in "trap shy" populations, we have  $p > c$ . When there is no behavioral response to trapping, we have  $p = c$ . As detailed in Chapter 2, we denote estimators of these parameters by the symbol,  $\hat{\phantom{x}}$  called a *caret* or a *hat*. Examples are  $\hat{N}$ ,  $\hat{D}$ ,  $\hat{p}$ , and  $\hat{c}$ .

Estimators of p and c are important because of their close bond with  $\hat{N}$ . In other words, if p or c are estimated poorly and thus show large bias, then  $\hat{N}$  will be affected adversely.

## Statistics

Statistics are entities computed from the data—for example, the X matrix in the context of these sampling studies. Statistics that are used frequently in capture-recapture and removal are defined and discussed below.

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

$n$  = the total number of captures during the study.

$u_j$  = the number of new (unmarked) animals captured in the  $j^{\text{th}}$  sample,  $j = 1, 2, \dots, t$ . Note that  $u_1$  always equals  $n_1$ .

$f_j$  = the capture frequencies = the number of individuals captured exactly  $j$  times in  $t$  days of trapping,  $j = 1, 2, \dots, t$ . For example,  $f_3$  = number of animals captured three times during the  $t$  days of trapping. The term  $f_0$  is used for the number of individuals never captured; obviously,  $f_0$  is not observable.

$M_{t+1}$  = the number of different individuals caught during the experiment. Recall that  $t$  is fixed for a given experiment; this term is merely the number of nonzero rows in the X matrix.

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

$M$  = sum of the  $M_j$ , not including  $M_{t+1}$ .

$m_j$  = the number of marked animals captured in the  $j^{\text{th}}$  sample,  $j = 2, \dots, t$ . Note that  $u_j = n_j - m_j$  and that  $m_1 \equiv 0$ .

$m$  = sum of the  $m_j$ .

The dot notation indicates the summation.

The statistics  $n_j$ ,  $u_j$ ,  $f_j$ ,  $M_j$ , and  $m_j$  may be computed directly from the X matrix; see *Otis et al. (1978:15)* for additional details.

## Fundamental Assumptions

Every modern estimation method is based on a set of well-defined, explicit assumptions. There are three general assumptions for *all* capture-recapture studies.

(1) The population is closed. (Open models allow this assumption to be relaxed, except that geographic closure is still required.)

(2) Animals do not lose their marks during the experiment.

(3) All marks are noted and recorded correctly at each sampling occasion  $j$ .

There are two corresponding assumptions for all removal studies.

(1) The population is closed. (However, see remarks at the end of the *Closure* section in this chapter.)

(2) The number of unmarked animals is counted and recorded correctly at each sampling occasion  $j$ .

The crucial assumption for a model relates to the capture probabilities of the various population members. The modeling of capture probabilities is the key problem in both capture-recapture and removal studies. For example, the earliest assumption was that each animal has a constant and equal probability of capture on each trapping occasion and that capture and marking do not affect subsequent catchability of the animal. This assumption, related directly to the original ball and urn model (Model  $M_0$ ), is unrealistic in capture studies of animal populations. It is now widely recognized that this assumption usually is not met (*Young et al. 1952, Geis 1955, Huber 1962, Swinebroad 1964*). *Edwards and Eberhardt (1967), Nixon et al. (1967), and Carothers (1973a)* provide clear evidence that accurate population estimation usually requires 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 through computer simulation by *Burnham and Overton (1969), Manly (1970), Gilbert (1973), and Carothers (1973b)*. The estimators they studied generally were biased significantly when this assumption was violated.

## Capture-Recapture Models

*Otis et al. (1978)* have presented several methods that allow various relaxations of the assumption of equal catchability. Much of this work started with the results of *Burnham (1972)* and *Pollock (1974)*. Following *Pollock (1974)*, we consider a sequence of models (assumptions) allowing for three major sources of variation in capture probabilities:

- (1) Model  $M_t$  assumes that capture probabilities vary by time or trapping occasion.
- (2) Model  $M_b$  assumes that capture probabilities vary by behavioral responses to capture.
- (3) Model  $M_h$  assumes that capture probabilities vary by individual animal ( $h = \text{heterogeneity among animals}$ ).

The assumptions regarding unequal capture probabilities must be embodied explicitly 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*). Therefore, we discuss the three simplest ways to relax this assumption.

Model  $M_t$  allows capture probabilities to vary by time; that is, to differ on each trapping occasion. This situation may be common even though the number of traps is fixed during the course of a study. For example, a cold rainy period during the study might reduce the probability of capture during this time. If the trapping effort were to vary, Model  $M_t$  might be appropriate. Also, if different capture methods were to be 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 response as a "contagion of catchability." This variation implies that an animal's behavior tends to be altered after its initial capture. For example, if an animal is frightened or hurt during initial capture and marking, it probably will not enter a trap again.

Model  $M_h$  allows capture probabilities to vary by individual animal. This situation has been modeled only with great difficulty and requires that additional assumptions be made about the statistical distribution of the capture probability  $p$ . Individual heterogeneity of capture probability 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 important type of variation has been treated rigorously by *Burnham (1972)* and *Burnham and Overton (1978, 1979)*; their nonparametric approach is presented in Chapter 3.

In addition to Models  $M_t$ ,  $M_b$ , and  $M_h$ , we consider all possible combinations of these three types of unequal capture probabilities; that is, Models  $M_{tb}$ ,  $M_{th}$ ,  $M_{bh}$ , and  $M_{tbb}$ . We also treat Model  $M_o$ , the "null" case in which capture probability is constant with respect to all factors. Model  $M_o$  corresponds to the assumption of equal catchability. For simplicity, we denote estimators of population size for a specific model by using the same subscript notation. For example,  $\hat{N}_o$  denotes the estimator derived from Model  $M_o$ ,  $\hat{N}_t$  denotes the estimator derived from Model  $M_t$ , and  $\hat{N}_{bh}$  denotes the estimator derived from Model  $M_{bh}$ .

## Removal Models

Removal models are discussed in *Otis et al. (1978:44-50)*. The simplest removal model we discuss here was proposed by *Moran (1951)* and *Zippin (1956, 1958)*. It is closely related to Model  $M_b$  for capture-recapture studies. We assume that effort is constant during the study and formulate a sequence of removal models: Models  $M_{R1}$ ,  $M_{R2}$ ,  $M_{R3}$ , . . . In this primer, we do not discuss the models in which effort is deliberately varied; however, a good review of such catch-effort methods can be found in *Seber (1973:296-347)*.

In removal studies, the collected data are the numbers of unmarked animals  $u_j$  captured at each sampling occasion  $j$ ,  $j = 1, 2, \dots, t$ . After the animals are captured, they are "removed" from the



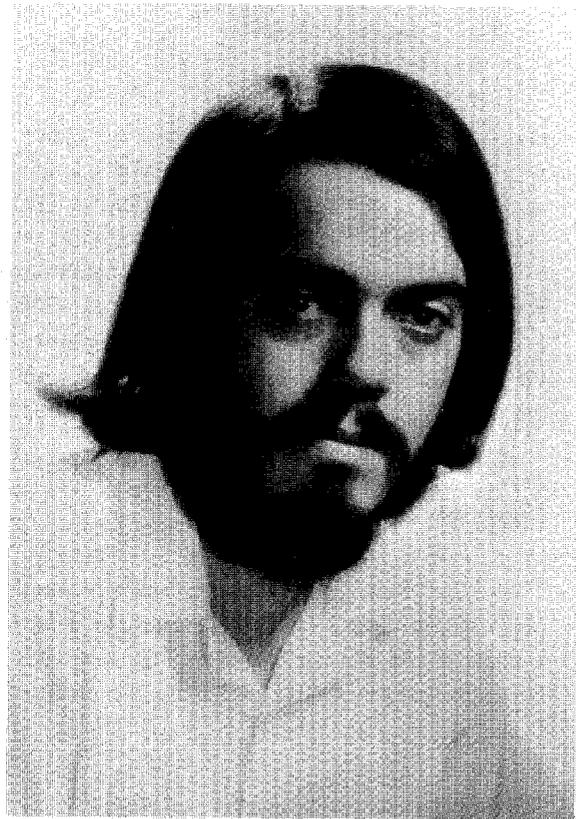
Kenneth P. Burnham

Kenneth Burnham has made a number of important contributions to estimation and testing theory for closed animal population models. He began his education with a B.S. degree in biology at Portland State College, and subsequently took M.S. and Ph.D. degrees in mathematical statistics at Oregon State University under W. S. Overton. His M.S. degree, in the late 1960s, dealt with simulation studies of the robustness of existing estimators of population size. He found that unequal probability of capture (heterogeneity) caused serious bias in the existing estimators. In his Ph.D. work he explored ML estimation for a generalized model for heterogeneity. Having found that it had unacceptable estimators of parameters, he derived a very nontraditional estimator based on the theory of the generalized jackknife. While working in Alaska with C. Cushwa on a rabbit population study, he developed a method to estimate density from capture-recapture sampling. This method allows a parameter to account explicitly for "edge effect"—a problem that has plagued biologists for several decades.

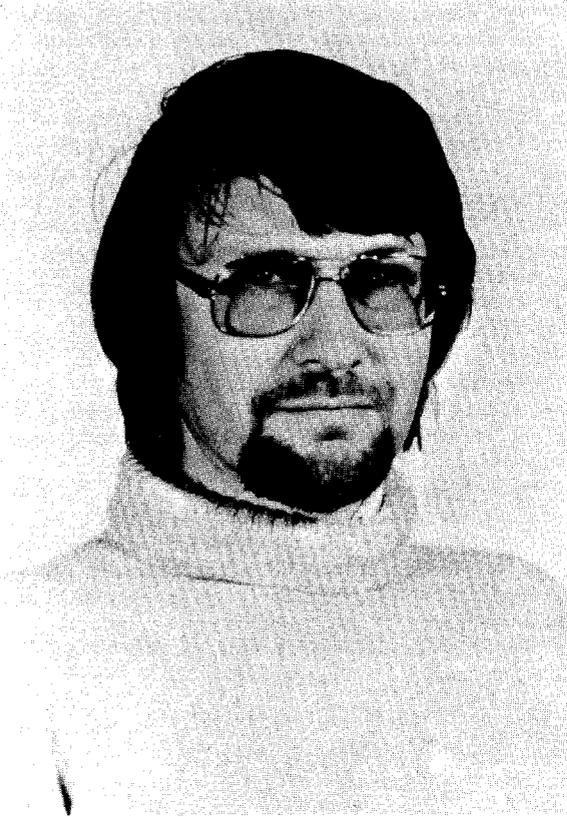
Burnham has worked as a statistician for the U.S. Fish and Wildlife Service since 1973. He believes that more work is needed on the tests of model assumptions, on model selection, and on the derivation of robust estimators. He sees the role of easy-to-use computer algorithms as becoming increasingly important and thinks that practical theory can be advanced as biologists and statisticians work together more closely. (Early 1970s photograph.)

Kenneth Pollock has made major state-of-the-art contributions to the estimation theory for both open and closed capture-recapture models. He was among the first to delve into the problems implicit in the assumption of equal catchability in closed populations. Early in his graduate work, he focused on model-building, estimation, and tests of assumptions for closed-population models. His findings provided an important part of the basis for the monograph by *Otis et al. (1978)*. His recent work has turned to open models, using the multiple hypergeometric approach. He has derived ML estimators for a series of such models for both age-independent and age-dependent populations.

Pollock was born in 1948 in Australia and received a B.S. degree in agriculture from the University of Sydney. He took M.S. and Ph.D. degrees from Cornell University working primarily with D.S. Robson in the Biometrics Unit. Since receiving a Ph.D. in 1974 he has held several professional positions including faculty appointments at the University of Reading and University of California at Davis; he is now with the Department of Experimental Statistics at North Carolina State University. His interests include building realistic statistical models and deriving statistical procedures that are robust to model failure. (Recent photograph.)



Kenneth H. Pollock



Andrew D. Carothers

Andrew Carothers has made contributions to the theory of both open- and closed-population models. Born in Nairobi, Kenya, he took a B.S. degree in Mathematics at the University of Bristol, and, in 1969, an M.S. in Biometry at the University of Reading after teaching for several years in East Africa. He spent 3 years working on capture-recapture theory at the University of Edinburgh in association with R. M. Cormack, G. M. Jolly, and A. N. Arnason. His work focused on the testing of model assumptions and the robustness of estimators, especially in relation to the assumptions of equal catchability. Since 1972, he has worked primarily on the applications of statistics in genetics.

Carothers feels that more emphasis should be given to the power of tests in capture studies and to the relation between test statistics and biased estimators, because he believes there are cases where a model assumption is rejected by a significance test even though the model is still quite useful. (Recent photograph.)

population. Physical removal is the most common application; for example, the fish are transplanted to another pond, or the animals are kept in a holding area until the study is completed. Another possibility is to remove the animals by marking them in some manner; in this instance, the marked animals thus are "removed" from the unmarked population of interest.

Model  $M_{R1}$ , the first removal model, assumes that the capture probability is constant for all trapping periods; that is,  $p_1 = p_2 = p_3 = \dots = p_t$ . (The R stands for removal.) Model  $M_{R2}$  allows the animals caught on the first occasion to have a higher capture probability than on all subsequent occasions; that is,  $p_1 > p_2 = p_3 = \dots = p_t$ . Model  $M_{R3}$  allows  $p_1 > p_2 > p_3 = p_4 = \dots = p_t$  and so on. These models are developed in detail in Chapters 3 and 4.

## Program CAPTURE

Most of the computations necessary in the analysis methods presented here are nearly impossible to perform without a computer. They include computations for estimators of parameters, sampling variances and covariances, test statistics, and model selection. Our philosophy has been to let the computer program CAPTURE do the arithmetic, leaving the biologist free to concentrate on the full

interpretation of results. The computer is far more accurate than a hand calculator, but it requires some adjustments. For example, because most computer printers cannot print lower-case letters and subscripts, the symbol  $n_j$  appears on the computer output as N(J), and the estimators  $\hat{c}$  and  $\hat{p}$  appear as C-HAT and P-HAT, respectively.

Program input follows a simple, free-format style for easy use. A variety of options is available. We advise users to enter the full X matrix to allow a very complete analysis. Alternatively, the user can enter only certain summary statistics, for example,  $n_1, n_2, \dots, n_t$ , and  $M_{t+1}$  for Model  $M_t$ .

The documentation for program CAPTURE is found in *White et al. (1978)*, available without cost from the Utah Cooperative Wildlife Research Unit, Utah State University, Logan, UT 84322. An abbreviated version is presented in Appendix A.

### Questions and Exercises

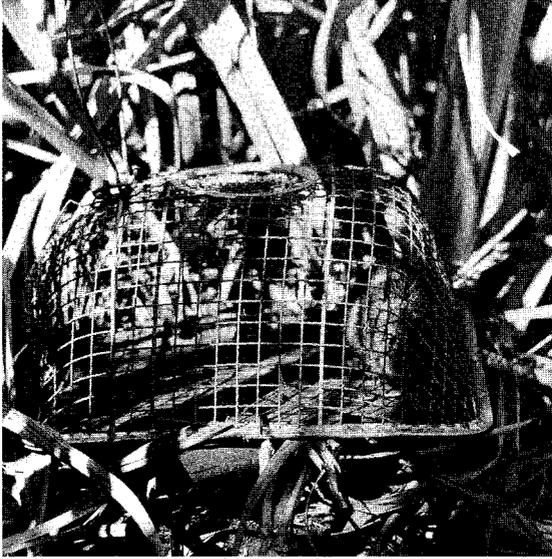
1. Is Model  $M_{t,th}$  the most general model?
2. Can summary statistics, such as  $m_j, u_j, M_j, M_{t+1}$ , and  $m_j$ , be computed from the X matrix?
3. Is Model  $M_o$  generally useful for estimating N in biological populations?
4. Is the closure assumption relatively unimportant? Why? Why not?
5. Consider the following data (X matrix) from a capture-recapture study.

	j = 1	j = 2	j = 3	j = 4	j = 5	j = 6	j = 7
i = 1	1	0	0	0	1	0	0
i = 2	1	1	1	1	1	1	1
i = 3	1	0	1	0	0	0	0
i = 4	0	1	0	1	0	0	0
i = 5	0	1	0	0	0	1	0
i = 6	0	1	0	0	0	0	0
i = 7	0	1	1	1	1	1	1

- (a) What is t (the number of trapping occasions)?
  - (b) What is  $n_j$  for  $j = 1, 2, \dots, t$ ?
  - (c) What is  $u_j$  for  $j = 1, 2$ ?
  - (d) What is noticeable from the data on animals 2 and 7?
  - (e) What is  $M_{t+1}$ ?
  - (f) What is  $M_t$ ?
6. Does Model  $M_b$  allow some animals to be trap happy and others in the same population to be trap shy?
  7. Can survival rates be estimated under the closure assumption?
  8. Assume a fishery biologist must determine the *exact* numbers of fish in a small pond. Would you advise him to use capture-recapture sampling? Removal sampling? What?
  9. Why must biologists be concerned about the capture probabilities?
  10. Do you think that capture-recapture studies are easy to conduct, require only a few traps (say, 25 to 50), and are easy to analyze?
  11. Why not always use a very general model that will not require concern about assumptions involving the type and significance of various types of unequal catchability?
  12. Consider the following X matrix from a capture-recapture study. Only the observed portion is shown; animals that were never captured are not shown. Zero entries are left blank.

	$j = 1$	$j = 2$	$j = 3$
$i = 1$	1		1
$i = 2$	1	1	1
$i = 3$	1		
$i = 4$	1		1
$i = 5$	1	1	1
$i = 6$	1		
$i = 7$	1		
$i = 8$		1	
$i = 9$		1	1
$i = 10$		1	1
$i = 11$		1	
$i = 12$		1	
$i = 13$		1	
$i = 14$		1	
$i = 15$		1	1
$i = 16$		1	
$i = 17$			1
$i = 18$			1
$i = 19$			1
$i = 20$			1

- What is  $t$ ?
- What is  $M_{t+1}$ ?
- What are  $n_1$ ,  $n_2$ , and  $n_3$ ?
- What is  $n$ ?
- What are  $u_1$ ,  $u_2$ , and  $u_3$ ?
- What are  $f_1$ ,  $f_2$ , and  $f_3$ ?
- What are  $M_1$ ,  $M_2$ ,  $M_3$ , and  $M_4$ ?
- What is  $M$ ?
- What are  $m_1$ ,  $m_2$ , and  $m_3$ ?
- What is  $m$ ?



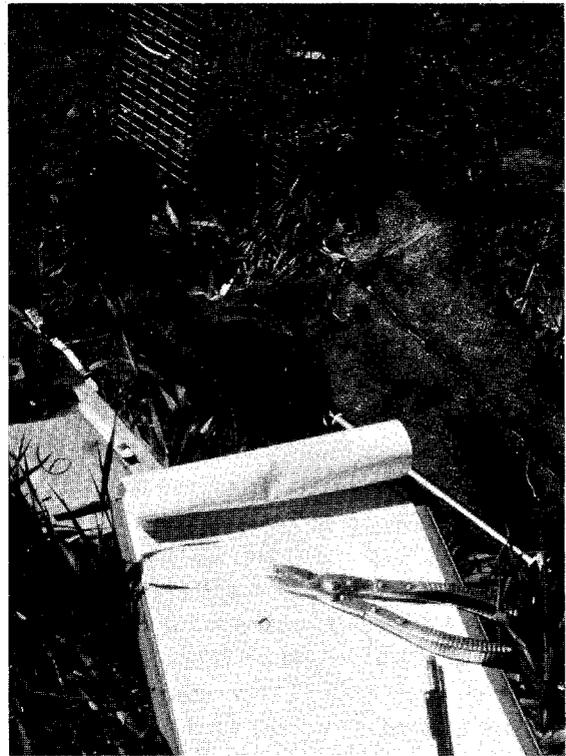
This cotton rat has been captured in a wire mesh Japanese live trap. (Photograph courtesy of Lynn Lefebvre.)



This Richardson's ground squirrel has been captured in a wire live trap. (Photograph courtesy of Kathy A. Fagerstone.)



This Indiana bat (*Myotis sodalis*) has been tagged on the wing with a celluloid ring. Bats tend to show fidelity to roost sites, so that they can be mist-netted in the vicinity of the roost with a good chance the closure assumption will be met. (Photograph courtesy of Richard Clawson.)



The equipment shown is commonly used in capture-recapture studies. The data sheet is in a waterproof notebook, and a pencil, ear tag, and ear tag pliers lie on top of the data sheet. A spring scale lies beside the clipboard. Additional tags are in the envelope. Other miscellaneous equipment and a live trap are in the background. (Photograph courtesy of Kathy A. Fagerstone.)

# CHAPTER 2

## STATISTICAL CONCEPTS

Capture-recapture and removal studies often are not recognized as sampling methods because they are quite unlike those used in much of the usual sampling theory (see standard texts such as *Cochran 1977*). For example, in capture-recapture and removal studies the sampling probabilities are not known and cannot be pre-established. No sampling "frame" is available, and the investigator has relatively little control over the situation.

Progress in science usually is made through experimentation—data are collected and analyzed, and conclusions are drawn. The conclusions drawn from the sample data are meant to go beyond the particular study. Biologists often wish to generalize from a particular experiment to the class of all similar experiments. This type of generalization is termed "inductive inference."

Let us examine two idealized case studies to illustrate some elementary concepts of sampling and inductive inference. First, consider the task of a quality control specialist who must estimate the proportion of newly manufactured light bulbs that are defective. Because 5 million bulbs are manufactured each month, he cannot test each light bulb (make a complete census); therefore, he must sample the bulbs and test the sample for defects. If he randomly samples 1000 bulbs from the population of 5 million, tests them, and finds only 3 defective, he might conclude that about 0.3% of the bulbs manufactured during his experiment are defective. That is, he makes an inference about the population from a sample.

Second, consider a biologist faced with estimating the number of mice on a large tract of land in south-central Wyoming in June of a given year. Because total enumeration (a census) is impossible, he might sample the area by establishing several 20 by 20 trapping grids located randomly throughout the area. (He may, in fact, want to stratify the sample by vegetative type, but he will avoid such considerations for the moment.) If he performs a capture study for six nights at each area, he can estimate the density (the number per unit area) for each of the study areas. The density estimates could be averaged over the areas, and inferences could be made about the density of the population, based on the sample data collected from the grids.

Both of these samples involve sampling a defined population, acquisition of data from the sampling process, and finally estimation and conclusions about the population rather than conclusions about only the sample. A theorem of logic tells us that there is uncertainty in inductive inference and, therefore, that we cannot make perfectly certain generalizations about a population by studying only a sample. However, we can make uncertain inferences, and we can measure the degree of uncertainty if the experiment has been performed in accordance with certain scientific principles (*Mood et al. 1974*). One function of the science of statistics is to provide techniques for making inductive inferences and for measuring their degree of uncertainty (*Ostle 1963:1-16*).

With respect to the subject of statistics, many think of statistics in terms of simple *t* and chi-square tests, analysis of variance, regression, and other such methods. In fact, the field is far broader than is suggested by the methods for data analysis to which people are exposed in the first two or three courses on statistics. Statistics is not a branch of mathematics, but it is an area of science concerned with the development of a practical theory of information. It involves sampling, design of experiments, analysis of information, estimation of parameters, and testing of hypotheses. It is the basis for inductive inference, and it is an integral part of what is termed the Scientific Method. The following sections introduce basic statistical concepts that are needed for an understanding of the following chapters.

## Theory, Reality, and Models

It is essential to understand the difference between theory or theoretical statistical models and reality. The methods presented here and in *Otis et al. (1978)* are approximations or models of reality. No model gives an exact explanation of a real biological or physical phenomenon. A "good" model, however, can be very useful to our understanding of a process.

In this primer we are concerned with the statistical theory of animal trapping experiments designed to enable the estimation of population size or density, or both. We postulate theoretical probability models for sampling animal populations, apply a theory of probability and inference based on rigorous mathematical foundations (see *Otis et al. 1978*), and present several theoretical models for use in the collection and analysis of information in real biological populations. The models are not exact representations of nature. Their utility is measured by the extent to which they assist us in understanding the dynamics of animal populations.

For our purposes we think of a model as a mathematical representation of a postulated set of assumptions concerning a capture-recapture or removal experiment. Such models are stochastic because they allow for the fact that the data arise from a random process. In a stochastic process, the outcome (data) is not completely predictable. Stochastic processes are common in everyday life, and they represent the rule rather than the exception. (This topic will be discussed further later in this chapter.)

Although the biologist need not understand the details or derivation of stochastic models, he should be able to see how these models help to achieve the goal of estimating population parameters. Their role is illustrated in Fig. 2.1. The model is the link between the data and the procedure used to estimate the population parameters contained in the model. Thus, whether or not the fact is stated explicitly, all statistical estimation procedures are based on a model of the sampling experiment or, stated differently,

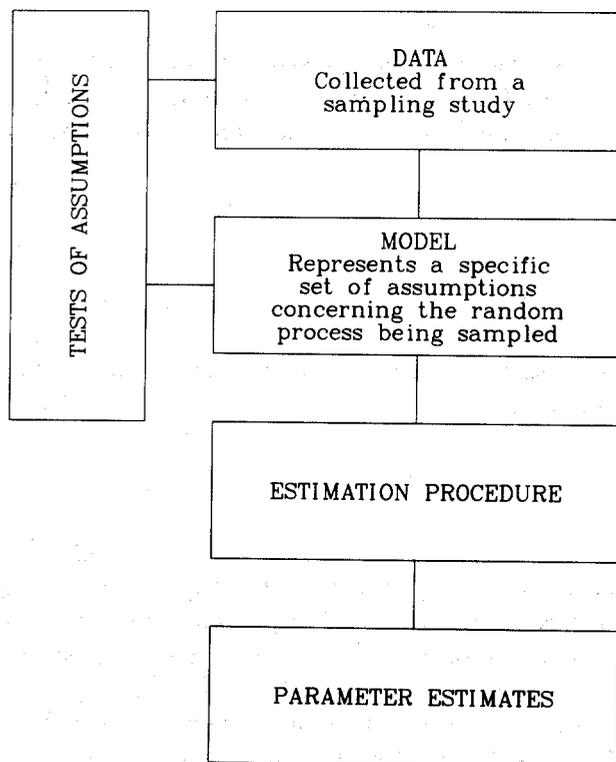


Fig. 2.1. The roles of data, models, and estimation procedures and their relationships in producing estimates of parameters.

estimation procedures are based on a specific set of assumptions concerning the sampling experiment. In Chapter 3, we deal more directly with models of capture-recapture experiments and the assumptions they represent.

Education in the biological sciences often has not included adequate explanation of either the scientific method or the theory of inference. The books by *Baker and Allen (1968)* and *Goldstein and Goldstein (1978)* provide an introduction to both subjects, and those by *Popper (1962)* and *Medawar (1969)* present more technical discussions. The use of the scientific method represents a broad philosophy concerning rigorous inference. We might ask, "What justifies a conclusion?" The answer to this question always involves "valid methodology."

Valid methodology is a package of essential ingredients: proper hypothesis formulation, design of data collection, conduct of the experiment or sample, rigorous analysis of the data to test the hypothesis, and inference (a conclusion) to reject or support, but never to "accept" the hypothesis. Inference depends critically on study design and data analysis.

## Estimation

In the discussion of models, we frequently referred to parameters and estimation procedures (or estimators). These terms are discussed below. [See *Kendall and Buckland (1970)* for related material].

**Parameter.** A parameter is the true population value of interest, expressed as a number. In capture-recapture studies the parameter of interest is either population size  $N$ , the total number of animals in the population, or population density  $D$ , the number of animals per unit of area. Examples of other important parameters in biological work are annual survival rate, average clutch size, average number of corpora lutea, and the proportion of males in a population.

**Estimator.** An estimator is a mathematical expression that indicates how to calculate an estimate of a parameter from the sample data. Estimators are necessary because we almost never know the value of the population parameter. The following formula for calculating a mean is the estimator most commonly used by biologists.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (2.1)$$

This formula for the sample mean is an estimator of the population mean  $\mu$ . The Petersen-Lincoln estimator of population size is simply

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

where  $n_1$  and  $n_2$  are the total number of animals captured on the first and second sampling occasions, respectively, and  $m_2$  is the number of marked animals captured on the second occasion (*Seber 1973:59*). In general, an *estimator* is shown with a "hat" over the parameter to indicate clearly that it is an estimator, rather than the true parameter. For example,  $\hat{N}$  and  $\hat{D}$  are estimators of the parameters  $N$  and  $D$ . An *estimate* is the numerical value resulting from substituting the sample data into the estimator. For example, the data set {4, 2, 7, 3, 4}, when substituted into the estimator given in Eq. (2.1), produces the estimate  $20 \div 5 = 4$ .

In most practical situations, a "proper" estimator is not obvious. In other words, intuition is often of little help in deriving a good estimator of a parameter. Without a model, we can only guess at valid



Carl George Johannes Petersen

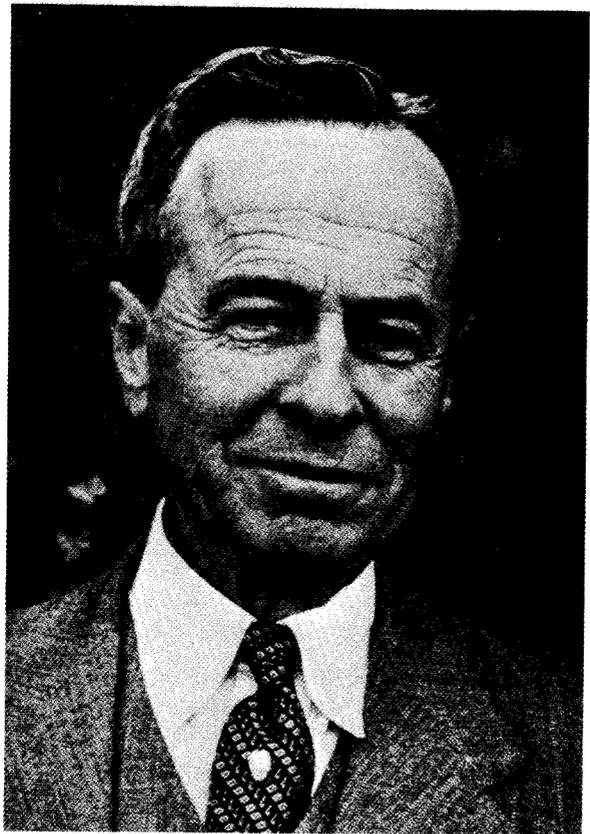
Nearly every terrestrial ecologist is aware of the Lincoln Index, a simple estimator developed by Frederick Lincoln to estimate the size of the waterfowl population in North America from banding and recovery data. The same method was derived before the turn of the century by C. G. J. Petersen for fishery problems and as far back as the 17th century by P.S. LaPlace for human population problems. In the past half century the Petersen-Lincoln method has found many uses in a variety of disciplines.

Frederick Lincoln, born in Denver, Colorado, in 1892, spent his life in the study of birds. He joined the Biological Survey (now the U.S. Fish and Wildlife Service) in 1920 and was responsible for the bird banding program, a cooperative program among Canada, the United States, and Mexico. He developed the flyway concept for management of migratory waterfowl and was the leading authority on the distribution and migration of birds. He devoted much of his energy to developing better methods of trapping and banding birds and to developing procedures for recording, reporting, and analyzing banding data. He was awarded an honorary Doctor of Science degree by the University of Colorado in 1956. Additional information concerning his life appears in *Auk* 79:494-499, written after his death. (Photograph with permission of the American Ornithologists Union.)

The practice of estimating population size from capturing, marking, and recapturing both marked and unmarked animals had its beginning with the Petersen estimate. Carl George Johannes Petersen was born in Denmark in 1860 and was a famous fishery scientist and the director of the Danish Biological Station from its founding in 1889 until his retirement in 1926.

The estimation method now bearing his name was published in 1896 and stemmed from his work on plaice. He invented a brass tag that he attached to the fish, to study their migrations. When one-third of the marked fish in his study were recaptured by fisherman, Petersen recognized that this information constituted a basis for estimating population size.

Petersen was awarded the LL.D. degree *honoris causa* in 1912 from the University of St. Andrews, Scotland. Much additional information about his work can be found in *J. Du Conseil* 1928, 3(2):135-138 and in the *Report of the Danish Biological Station*, 1940, Copenhagen.



Frederick C. Lincoln

estimators. Such guesses are typically poor (if not incorrect), and no estimates of precision can be made without a model. However, with a proper model relating the data, assumptions, and parameters of interest, we can derive valid estimates of parameters in the model routinely, by very general, available methods. The principal method used in statistical estimation over the past half century has been the method of maximum likelihood (ML), which is discussed later in this section.

Our goal is to use good estimators to produce estimates of the parameters of interest. To evaluate estimators, we need criteria by which to judge them. In statistical theory, two essential criteria arise from the concepts of bias and precision.

**Accuracy.** Accuracy is defined as “exact conformity to truth” or “freedom from error or defect.” This ideal is unattainable in sampling studies and inductive inference; therefore, we rely on the concepts of bias and precision (defined below) as aids in making good inductive inference.

**Bias.** Ideally, an estimator should be free of bias. That is, if we were to repeat a sampling experiment under the same conditions on a very large number of occasions, each time computing an estimate from the sample data, the average of the estimates should equal the parameter being estimated. Frequently, we denote the “average” value of an estimator  $\hat{N}$  by  $E(\hat{N})$ , read as the “expected” value of  $\hat{N}$  or the “average” value of  $\hat{N}$  over a very large number of repetitions. Thus if  $E(\hat{N}) = N$ , we say that the estimator  $\hat{N}$  is unbiased. Note that bias is a conceptual quantity because usually we have only one set of data and can compute only one value of  $\hat{N}$  from the data. Bias relates strictly to the *average* performance of an estimator.

It is often convenient to discuss the subject of bias or biased estimators in two classes—small-sample bias and model bias. These terms are not well established in the literature, but the distinction between them is important for biologists. *Small-sample bias* is often of negligible importance to a biologist in the analysis of one or only a few data sets. This type of bias decreases as sample size increases. Biologists frequently encounter a “biased” estimator for the first time when estimating the variance from a random sample. One learns that the ML estimator

$$\begin{aligned} \text{variance} &= s_1^2 \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \end{aligned}$$

is biased. However, we find that the expression

$$\begin{aligned} \text{variance} &= s_2^2 \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \end{aligned}$$

is unbiased. When  $n$  is 30 to 40 or more, the difference between  $s_1^2$  and  $s_2^2$  becomes negligible. This is an example of small-sample bias. Another example relates to the Petersen-Lincoln estimator.

$$\hat{N}_1 = \frac{n_1 n_2}{m_2}$$

is biased, but

$$\hat{N}_2 = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1$$

is unbiased in some instances and virtually unbiased in others (see *Seber 1973:60* for details). However, the difference between the two estimators is negligible if the sample size is reasonably large. For example, compare using  $n_1 = 60$ ,  $n_2 = 50$ , and  $m_2 = 30$ .

A much more serious problem deals with *model bias*. This problem arises when important assumptions such as equal catchability, made in creating the model, are incorrect for a particular situation. An incorrect assumption can cause large bias even when sample sizes are quite large because it is theoretically independent of sample size. An example will illustrate this important concept. Consider a theoretical population of animals in which individuals become trap shy after they have been captured the first time. Given that the population is composed of 400 animals ( $N = 400$ ), the probability of first capture is 0.20 ( $p = 0.20$ ), and the probability of recapture drops to 0.05 ( $c = 0.05$ ), we can ask what bias could be expected if we (incorrectly) assume that the population is equally catchable (that  $p = c$ ) and estimate the population size under this assumption. Using simulation procedures, we find that the expected value of the estimator  $E(\hat{N})$  is about 1071, which illustrates that model bias can be very substantial. Formally, if  $E(\hat{N}) = 1071$ , the bias is  $E(\hat{N}) - N = 1071 - 400 = 671$ . Interested readers can find this example and others in *Otis et al. (1978:127)*.

Expressing bias as a percentage is often useful; called "percent relative bias," the expression is defined as

$$\text{PRB} = \frac{E(\hat{N}) - N}{N} \times 100 .$$

In the example above,  $\text{PRB} = 168$ ; that is,  $[(1071 - 400)/400] \times 100 = 168$ .

**Precision.** Precision relates to the repeatability of a result. If, for example, a sample is drawn and the total population size is estimated to be 10 700, will the next sample yield an estimate of 400, or 31 900, or will it be near 10 700? Repeatability is an integral concept in science. The precision of an estimator is measured by the sampling variance and its square root, called the standard error of the estimate (Fig. 2.2). People not familiar with the concept of repeatability, with numerical quantities plotted as histograms (bar graphs), or with theoretical probability functions overlying the histograms should examine Fig. 2.3.

In this context, our concern is an estimate of the sampling variance of the estimator  $\hat{N}$ , denoted  $\text{var}(\hat{N})$ , as a measure of precision or repeatability. The sampling variance and standard error [ $\text{se}(\hat{N}) = \sqrt{\text{var}(\hat{N})}$ ] are measures of the variability of the individual estimates around their expected or average value over different samples. Of course, we would prefer to have an estimation procedure that would give very similar estimates from different samples.

The concepts of bias and precision are illustrated in Figs. 2.4 and 2.5. The information in Fig. 2.4, adapted from *Overton and Davis (1969)*, shows a series of targets and shot patterns. If we make an analogy by considering each shot as an estimate (made from sample data by using a given estimator), we can illustrate the concepts in terms of standard frequency diagrams as shown in Fig. 2.5.

Research scientists and managers always prefer the unbiased and precise estimator illustrated in Figs. 2.4a and 2.5a to the very precise, incorrect estimate depicted in Figs. 2.4c and 2.5c, which is considered especially undesirable. Unfortunately, situations like those depicted in Figs. 2.4b-d and 2.5b-d are probably very common in attempts to estimate population size from the data for capture-recapture studies. In this primer we attempt to improve the accuracy of such estimates by emphasizing study design, increased sample size, and improved methods of analysis.

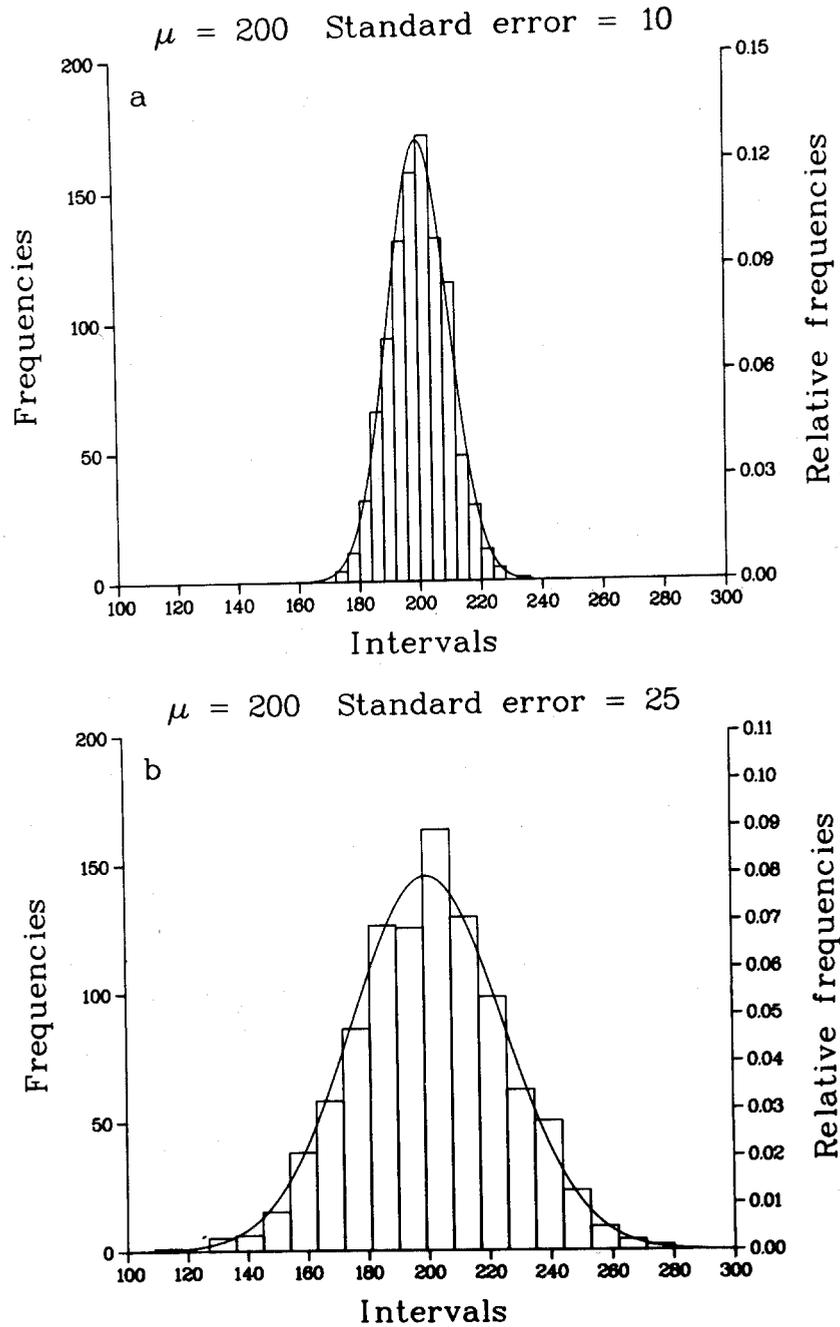


Fig. 2.2. Histograms of two data sets with differing amounts of variability. In a, the data are clustered fairly closely around the population parameter of 200. The spread of the data around the mean is measured by the standard error, which is 10 in this example (variance =  $10^2 = 100$ ). For example, the range of the data is about 170 to 230 for 1000 data points. In contrast, the data shown in b are much more variable, as reflected by the larger standard error of 25 (variance =  $25^2 = 625$ ). Here the range is from 130 to 280, also based on 1000 data points. In each instance, a normal curve has been fitted.

**Stochastic Processes and Models.** Processes that are not completely predictable (deterministic) are termed stochastic. Examples include coin-flipping, card games, all forms of gambling, weather patterns, stock market fluctuations, and, most important in the context here, all sampling data.

It follows that stochastic models are appropriate for data that arise from a stochastic process. Such is the case in capture-recapture and removal-sampling studies. The biologist need not understand the details

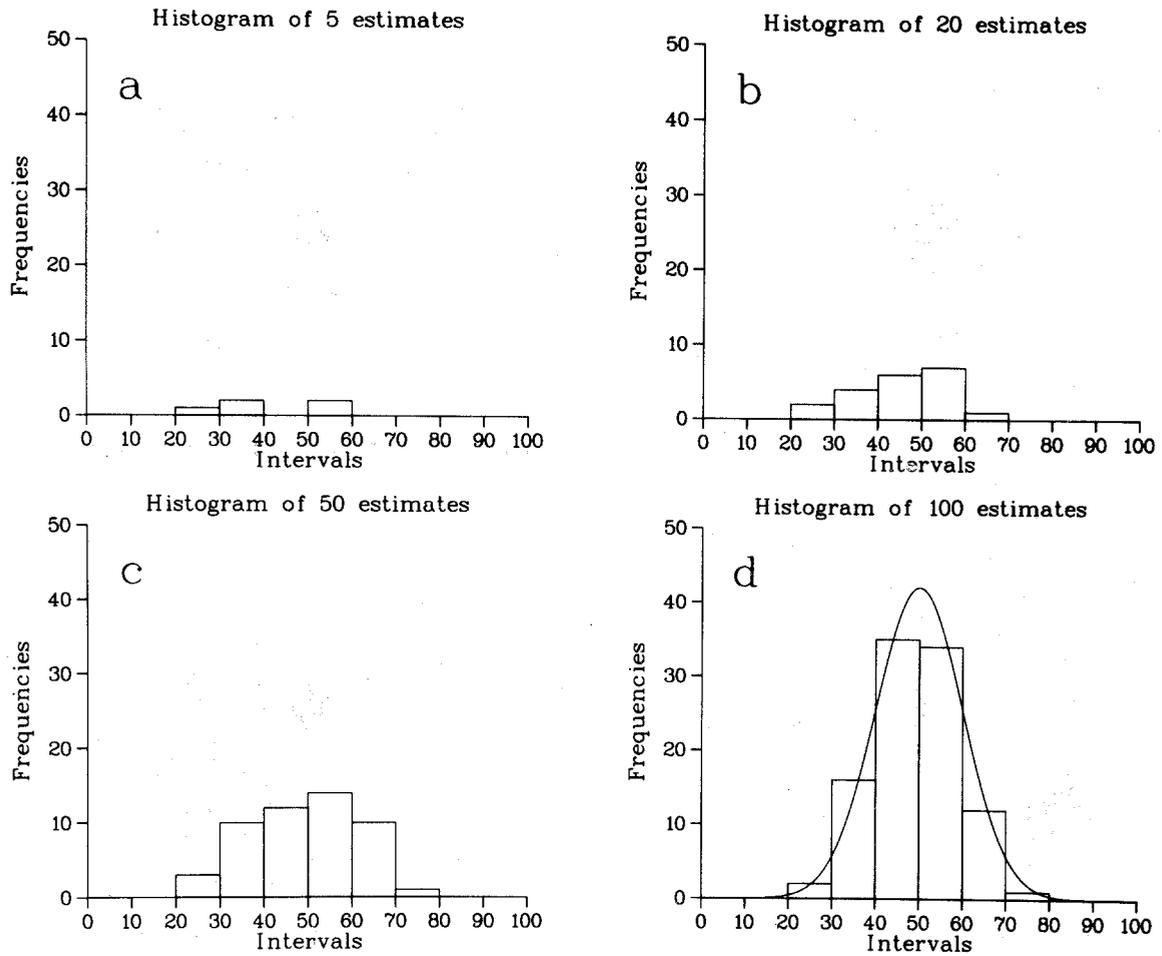


Fig. 2.3. Histograms of estimates based on samples of sizes 5, 20, 50, and 100. Computer simulation studies are useful in assessing whether an estimator has a normal sampling distribution for a given sample size. In the case shown, the estimator is distributed approximately normally as indicated in d.

of stochastic models, model building, or estimation theory to use the methodology presented in this primer; our purpose is to concentrate on concepts rather than on mathematical or statistical derivations.

## Variation

Important variation is found everywhere in the biological sciences. It is crucial to understand clearly the two distinctly different types of variation in capture-recapture and removal studies.

**Spatial and Temporal Variation.** First, there is the obvious variation in space and time in the real world. Neither animal density nor plant cover is uniform over the State of Utah; both are clear examples of spatial variation. Animal numbers fluctuate over time; these changes constitute temporal variation. As another example consider a 20-km portion of stream divided equally into 20 numbered segments. Assume that we know the exact number of fish in each segment. In other words, we know the

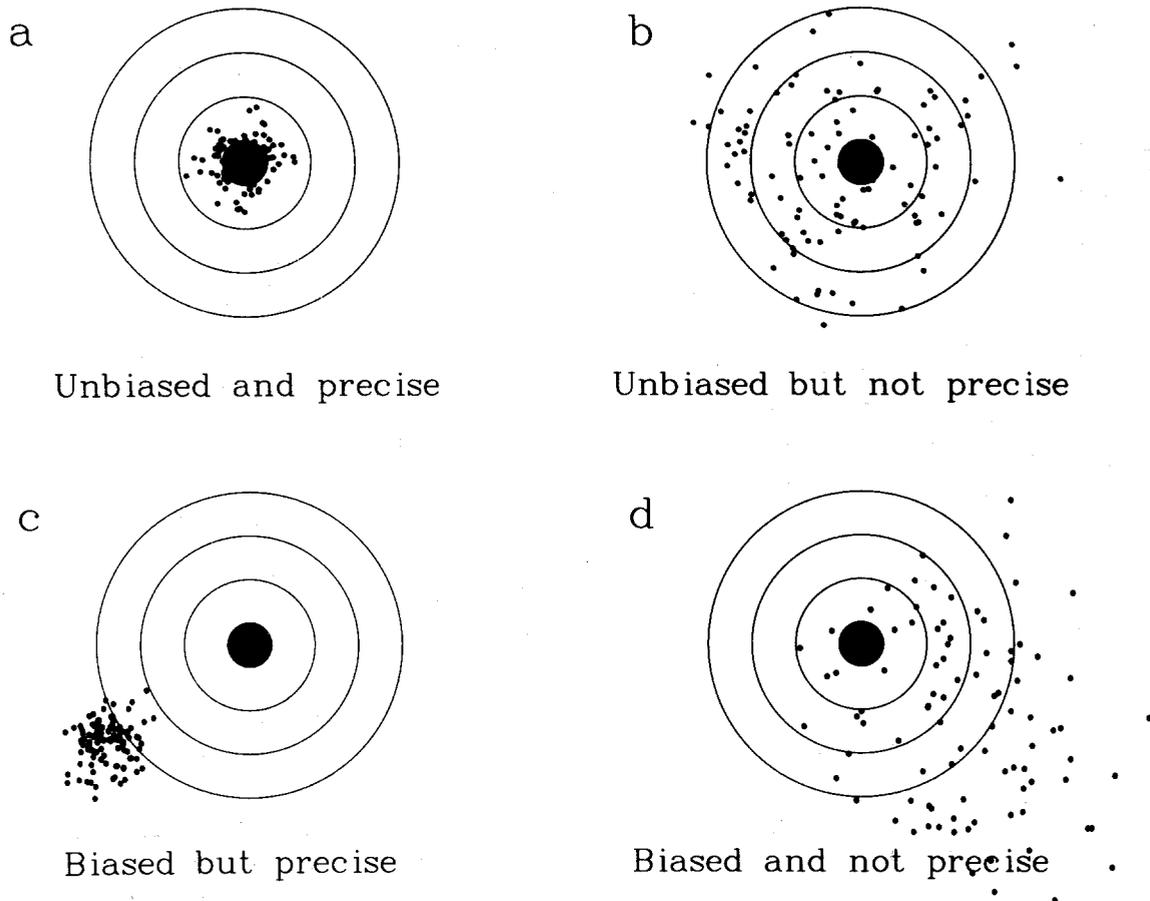


Fig. 2.4. Targets and shot patterns illustrate the concepts of bias and precision. The goal of the marksman is shown in a. Although c might not be too bad for a marksman, who merely needs to adjust his sights, it is the worst case for the biologist attempting to estimate population abundance. It is a highly precise, incorrect estimate. Furthermore, after completing a shooting session, the marksman can approach the target and compare his shot pattern with the bull's eye (true parameter). In contrast, the biologist usually will never know the true parameters; therefore, his inductive inferences from the sample data about the true parameter must be made carefully.

true parameters  $N_1, N_2, N_3, \dots, N_{20}$ , and these parameters vary. As a measure of this variation, the population variance  $\sigma^2$  could be computed by the usual definition,

$$\sigma^2 = \frac{\sum_{i=1}^{20} (N_i - \bar{N})^2}{20} \quad (2.2)$$

This quantity measures the variation in population size over space (20 segments of the stream). This type of variation is most frequently studied in basic statistics courses, where "sampling error" or "measurement error" is ignored.

**Stochastic Variation.** Second, there is stochastic variation of basically unpredictable events such as coin flipping, success of a nest, or time of death of an animal. This kind of variation is somewhat more difficult to understand. We will use the same example of the 20-km stream and consider segment 3 of the

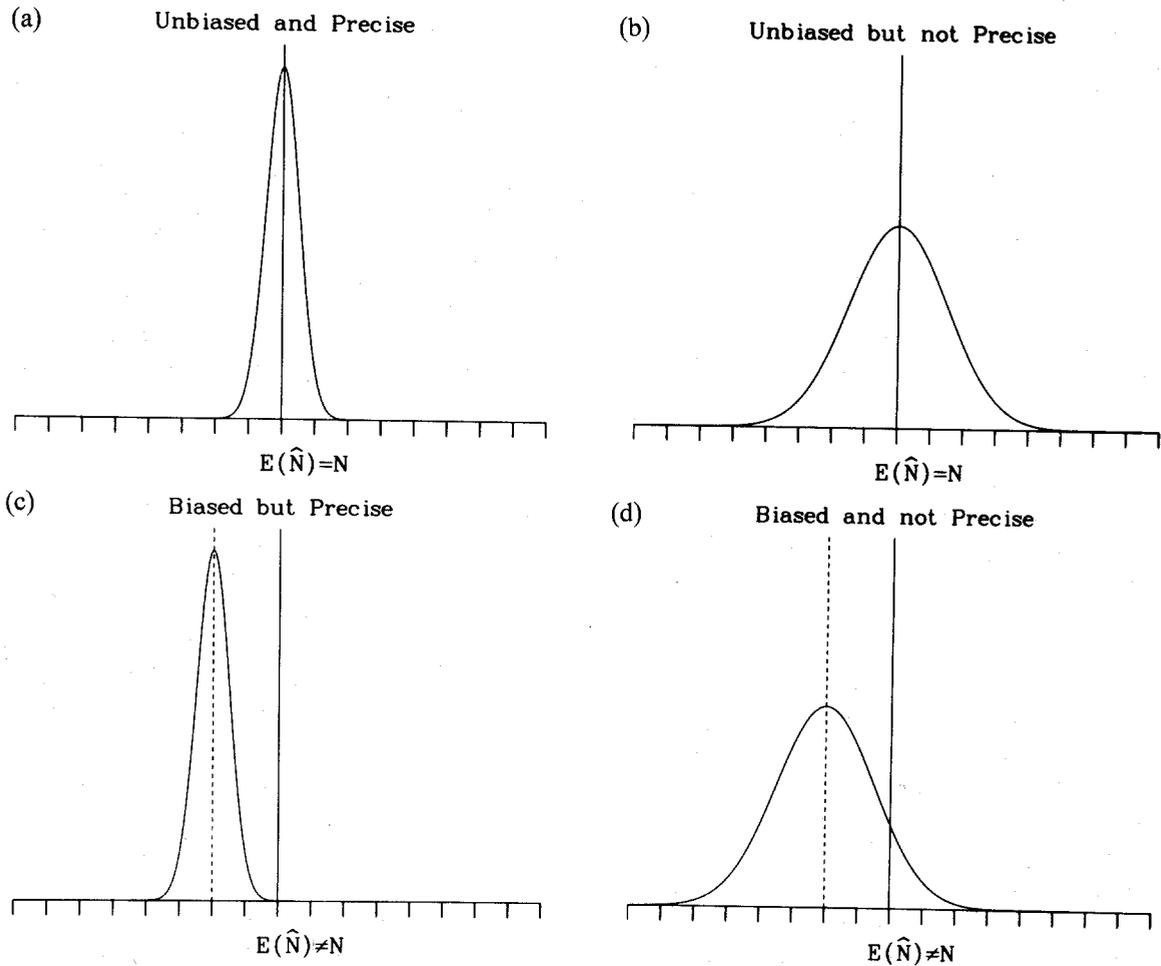


Fig. 2.5. An illustration of the concepts of bias and precision. Here, the information from a large number of samples is used to compute individual values of  $\hat{N}$ , which then are graphed as frequency distribution. (Alternatively, they could have been shown as simple histograms.) We desire an estimator with properties shown in a, in which the average value of  $\hat{N}$  is equal to the parameter  $N$  and the variation of the estimates around  $N$  is small. Precision is measured by the sampling variance of the estimator  $[\text{var}(\hat{N})]$  and relates directly to the spread of the frequency distribution. (Compare these spreads with those shown in Figs. 2.4.a and b.)

stream. In this instance, we have to estimate the number of fish in the 1-km segment. (Of course, a complete census would give us the parameter  $N_3$  directly, but we will assume that such a census is not feasible.) A removal experiment using electrofishing methods on 3-5 occasions and analysis of the resulting sample data would provide an estimate  $\hat{N}_3$  of the parameter  $N_3$ . If the survey were repeated, we would get a second estimate  $\hat{N}_3$  of  $N_3$ , and so on. It is this variation among the estimates that we consider in this primer, but it is important to recognize the existence of spatial and temporal variation.

Catching fish, however, is not a perfectly predictable event—it is, instead, a stochastic process. Therefore, we obtain  $\hat{N}_3 = N_3 \pm \epsilon_1$ , where  $\epsilon$  is sampling error or measurement error. (Note, if  $\hat{N}_3$  is an unbiased estimator, then the average or expected value of  $\epsilon_1$  must be zero.) Sampling variation occurs whenever we sample from a defined population ( $N_3$  in this instance) and attempt to estimate a parameter associated with the population. We denote sample variation associated with the estimators  $\hat{N}$  and  $\hat{p}$  as  $\text{var}(\hat{N})$  and  $\text{var}(\hat{p})$ , respectively. Usually sampling variances themselves are only estimated, and we employ the notation  $\hat{\text{var}}(\hat{N})$  and  $\hat{\text{var}}(\hat{p})$  to signify that these are estimators.

**Standard Errors and Sample Size.** The specific meaning and the concept of the terms “sampling variation” and “standard errors” are sometimes difficult to grasp. Expressions such as  $\text{var}(\hat{N})$  or  $\text{se}(\hat{N})$  are recognized as measures of precision, but the underlying concepts often are not understood clearly. The following example illustrates these concepts.

Consider a small island off the coast of Alaska with a population of colonial sea birds nesting on the island’s rocky edges. There are exactly 11 000 nests on the island in a certain year, and in this species the female lays only 1 egg. An investigator must find the proportion of the nests that are successful. In other words, he wants to know the parameter  $p$ . We know that  $p$  in this example is 0.70, that exactly 70% of the nests were successful; however, this fact is unknown to the investigator. He could contemplate a complete census of all nests and, by classifying the nests as successful ( $p$ ) or unsuccessful ( $1 - p$ ), arrive at the exact value of  $p$ . This approach is impossible, however, because the time and expense would be prohibitive. Therefore, he must sample a fraction of the total nests and estimate the parameter using the estimator  $\hat{p}$ .

Let us assume that the investigator decides to randomly select a sample  $n$  of size 25 nests. If  $s$  is the number of successful nests, then an estimator of  $p$  is computed as  $\hat{p} = s/n$ . He conducts the survey, finds the random 25 nests, and observes that 15 were successful and 10 failed. Thus the estimate of the proportion of successful nests is  $15/25 = 0.60$ . We know that  $p = 0.70$  in this example, and that this estimate of  $p$  is not too bad. The investigator, however, has no idea at this point how good his estimate is. In other words, he does not know how close his estimate is to the true, unknown parameter.

The investigator then decides to run a second survey of 25 different nests to check his first estimate. This survey yields  $s = 20$  and an estimate of  $\hat{p} = 20/25 = 0.80$  or 80% successful. Now his confidence is shaken a bit and he decides to conduct eight more surveys. The results, including the first two estimates, are as follows: 0.60, 0.80, 0.69, 0.52, 0.88, 0.69, 0.76, 0.64, 0.56, and 0.76. All 10 values are estimates of the same parameter  $p$ . The fact that the estimates vary represents sampling variation.

Sampling variation occurs when we sample a population (in this example, 11 000 nests) and *estimate* the parameter  $p$ , rather than making a complete census and *computing* the parameter  $p$  exactly. It is clear, once we consider the matter, that the variation in the 10 estimates would have been much smaller if a larger sample had been taken (if  $n = 500$  instead of 25) in each of the 10 surveys. Now we can ask, How do I measure how much variation to expect in the estimates for a sample of a certain size? The answer to this question is in the realm of mathematical statistics. Theory exists to enable calculation of the variance of the estimate,  $\text{var}(\hat{p})$  in this example, which is a measure of the sampling variation we can expect. The formula for estimating the variance of a proportion is given by  $\hat{\text{var}}(\hat{p}) = [\hat{p}(1 - \hat{p})]/n$ . In the first sample of 25 nests,  $\hat{p} = 0.60$ ; therefore,  $\hat{\text{var}}(\hat{p}) = (0.60)(0.40)/25 = 0.0096$ .

The square root of the variance, a more useful quantity, is called the “standard error.” That is,  $\text{se}(\hat{p}) = \sqrt{\text{var}(\hat{p})}$  or, in the above example,  $\hat{\text{se}}(\hat{p}) = \sqrt{0.0096} = 0.098$ . The standard error is used in calculating confidence intervals and coefficients of variation. The estimated coefficient of variation ( $cv$ ) of an estimate is defined as

$$cv(\hat{\theta}) = \frac{\hat{\text{se}}(\hat{\theta})}{\hat{\theta}},$$

where  $\hat{\theta}$  is, in general, an estimate of some parameter  $\theta$ . In our example from above,

$$\begin{aligned} cv(\hat{p}) &= \frac{\hat{\text{se}}(\hat{p})}{\hat{p}} \\ &= \frac{\sqrt{[\hat{p}(1 - \hat{p})]/n}}{\hat{p}} \\ &= \frac{0.098}{0.60} \\ &= 0.16 . \end{aligned}$$

In biological studies, a coefficient of variation (of an estimator) of 0.10 or less is considered good, so we see that the estimate from the first survey of 25 nests has only fair precision. The most effective means of increasing the precision of our estimate is to increase the sample size.

A 95% confidence interval in our example is computed as  $\hat{p} \pm 1.96se(\hat{p})$  or  $0.60 \pm 1.96 \times 0.098 = 0.60 \pm 0.192$ ; hence the interval is (0.41 to 0.79). If the investigator had run a large number of independent surveys, each of a random sample of 25 nests, 95% of the confidence intervals would be expected to include the true parameter (which, in our example, is 0.70). Clearly a confidence interval as wide as (0.41 to 0.79) is not of much use. The important concept here is that expressions like  $var(\hat{p})$ ,  $se(\hat{p})$ , and  $cv(\hat{p})$  are measures of precision (repeatability) or sampling variation.

A final point will illustrate the advantage and importance of sample size. Assume that the 10 surveys of nests were pooled and the effort is considered as 1 survey, with a sample size of 250 (10 surveys  $\times$  25 nests per survey = 250). Then  $\hat{p} = 0.69$ , which is very close to the true parameter of 0.70. The sampling variance is then  $\hat{var}(\hat{p}) = [0.69 \times (1 - 0.69)]/250 = 0.000855$ , and  $\hat{se}(\hat{p}) = 0.029$ . The coefficient of variation of the estimate is only 0.04. The 95% confidence interval is much narrower (0.63 to 0.75), a good indication that the variation in the estimates of  $p$  would not vary much from survey to survey. In other words, the repeatability is good, which allows much stronger inference about the parameter of interest. Repeatability is an important part of inductive inference.

**A Further Example of Variation.** Now let us assume we have five islands, each similar to the others, and each supporting a colony of birds of the same species. The situation is summarized as follows.

Island Number	Unknown Parameter <sup>a</sup>	Estimate	Sampling Variance
1	$p_1$	$\hat{p}_1 = p_1 \pm \epsilon_1$	$\hat{var}(\hat{p}_1)$
2	$p_2$	$\hat{p}_2 = p_2 \pm \epsilon_2$	$\hat{var}(\hat{p}_2)$
3	$p_3$	$\hat{p}_3 = p_3 \pm \epsilon_3$	$\hat{var}(\hat{p}_3)$
4	$p_4$	$\hat{p}_4 = p_4 \pm \epsilon_4$	$\hat{var}(\hat{p}_4)$
5	$p_5$	$\hat{p}_5 = p_5 \pm \epsilon_5$	$\hat{var}(\hat{p}_5)$

<sup>a</sup>The proportion of successful nests.

If we average the five estimates, we see that both spatial and sampling variation are involved. For example, define  $\bar{p}$  as the average of the five estimates.

$$\bar{p} = \frac{\sum_{i=1}^5 \hat{p}_i}{5}$$

The variance of  $\bar{p}$  is

$$\hat{var}(\bar{p}) = \frac{\sum_{i=1}^5 (\hat{p}_i - \bar{p})^2}{4} \quad (2.3)$$

It should be clear that the  $var(\bar{p})$  has two components of variation: sampling variation (given in Column 4 of the table above) and spatial variation among the islands [similar to the expression in Eq. (2.1)]. Spatial variation enters the computation of  $\bar{p}$  because physical and biological factors may cause differences among the islands. The separation of the sources of variation in expressions like Eq. (2.3) is a difficult subject known as “variance components” in statistics. We will not explore the subject here because it would take us too far afield. In this example, both spatial and sampling variation are likely to be quite important, and biologists should keep the two sources in mind.

As a second example, consider the sunfish in all the small ponds (potholes) in a particular county in Minnesota. There are 89 ponds, varying in size from 0.1 to 1.8 ha. Each pond is assigned an identification number, 1, 2, . . . , 89.

Spatial variation arises because the actual population size  $N$  of sunfish will differ among the 89 small ponds. Some may not have any sunfish ( $N = 0$ ), whereas others may have large populations. We also might suspect that population size could vary with pond size. Temporal variation relates to the actual population size  $N$  in a particular pond as it changes over time, because of births and deaths. Temporal variation is often seasonal. Note that both spatial and temporal variation relate to changes in  $N$ .

The stochastic component is encountered in many studies of animal populations because we usually cannot count each member of a population to determine  $N$ . Instead, a sampling procedure, such as capture-recapture, must be used to *estimate*  $N$ . The sample data (the  $X$  matrix or some summary of it) are used with an estimator, such as  $\hat{N} = (n_1 n_2) / m_2$ , to compute an estimate of the population size for a given point in time and space. The estimates vary with each sample drawn (see Fig. 2.3). It is this variation that we call sampling variance; it is a measurement error, caused by the stochastic nature of the sampling and capturing process and denoted as  $\text{var}(\hat{N})$ . Sampling variation can be illustrated by looking at pond 32. Because the pond is small, electrofishing was used to sample the population of sunfish each day for 4 days. The fish were returned to the pond after each sample day. An appropriate estimator was used to estimate  $N$  from the first 4-day sampling study, and the estimate was 243 sunfish. A second 4-day sampling study was conducted, and the estimate was 202 sunfish. A total of five 4-day sampling studies yielded estimates of the parameter  $N$  as 243, 202, 157, 231, and 192. This variation is called sampling variation. Fortunately, it can be estimated without having to conduct replicate samples. Sampling variation is a measure of the precision among the estimates.

Many practical problems necessarily involve both types of variation. However, this primer treats the estimation problems for a given point in time and space, such as a single large trapping grid within a larger area, and thus spatial and temporal variation are not relevant here.

### Properties of a Good Estimator

Because estimators are functions of random variables (the sample), they possess probability (sampling) distributions. Estimators, therefore, must be derived from probabilistic (for stochastic) models. A good estimator

- (1) is robust to crucial assumptions: it is not very sensitive to the failure of some important assumptions. It is robust to model bias.
- (2) exhibits minimum variance: it is the most precise estimator possible. It makes full use of all the information in the sample.
- (3) is distributed normally: for the sample sizes usually encountered, the distribution of the estimator is normal. If not normal, the distribution should be known, at least approximately.
- (4) is unbiased, given the assumptions: at least the small-sample bias is zero when sample size is large.

All of these properties are more complex than we have indicated; however, those listed should provide a working basis for an understanding of the material that follows. [The interested reader should consult a text on mathematical statistics, such as *Lindgren (1968:266-278)*.]

A word of caution is appropriate here because most capture-recapture and removal analysis methods fall somewhat short of our expectations for a good estimator. For example, most estimators have a slight small-sample bias, most are nonnormal (skewed to the right) for the sample sizes typically encountered, and most are not robust to the failure of certain assumptions. Poor coverage of confidence intervals sometimes is due to nonnormality. However, estimators are derived by using the ML method, which guarantees that they will be minimum variance estimators, at least asymptotically. Many methods described in the literature are *ad hoc* methods, and their properties are generally unknown. Such deficiencies call for careful design, field work, and analysis. These needs are the central focus of the material to follow.

### Estimation Methods

The data from capture-recapture or removal studies are collected from samples, thus requiring a probabilistic treatment of the data to derive good estimation and inference procedures. As we have shown, 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 express the assumptions quantitatively. A probability model is a form of mathematical representation of the observed data under a specific set of assumptions and, as such, it provides a basis for quantitatively and explicitly incorporating the specific assumptions about closure and capture probabilities and for developing the point and interval estimators by rigorous statistical estimation techniques. Most parameter estimators used here were derived by using the maximum likelihood (ML) method.

Estimators derived by this method are optimal, at least for large samples. This general method of estimating parameters was derived in the early 1920s by the famous statistician and geneticist, Sir Ronald A. Fisher, and it has been the backbone of statistical estimation theory for more than 50 years. Alternative estimation procedures, such as method of moments and minimum chi-square, have been developed, but the ML method is generally accepted as the best. The interested reader is referred to any book on mathematical statistics for additional material on the ML method; one example is *Kempthorne and Folks (1971:242)*.

**Random Sampling.** Many authors state that random sampling is required in capture studies. This assumption stems from ball and urn experiments (Fig. 1.1), in which marked and unmarked balls are shaken completely, and a random sample is taken at the end of each sampling occasion.

Traditional sampling methods include procedures for drawing random samples. Use of the procedures requires knowledge of the sampling probabilities (for finite populations). Deliberate control over the elements to be sampled is required. However, such control is clearly absent in capture studies of animal populations. The sampled animals are not selected by the investigator; the capture probabilities are not preset, nor are they even fixed during the course of the study. It is unrealistic to think that an animal captured in one corner of the trapping grid may be captured subsequently in the opposite corner. There is simply no basis for thinking that samples are drawn randomly in capture studies of animal populations. *Mendenhall et al. (1971:187)* and *Johnson and Kotz (1977:248:250)* present a different view, although they recognize some practical problems. *Feller (1950:45)* gives an example of the capture-recapture method for a hypothetical fish population (essentially the Petersen-Lincoln estimate) and mentions in a footnote that the method is used widely in practice.

The concept of random sampling does not apply to situations assumed in Models  $M_b$ ,  $M_{th}$ ,  $M_{bh}$ ,  $M_{tb}$ ,  $M_{th}$ , and  $M_{tbb}$ . The goal of these models is to provide an analysis of the sampled data in the face of behavioral response and heterogeneity, both of which are contrary to the traditional role of random sampling.

**Robustness of an Estimator.** In addition to the important properties of bias and precision, an estimator also may be judged by its robustness to the failure of certain assumptions. In the previous section, we described how every estimation procedure is based on a model that represents a specific set of assumptions concerning the population or process being sampled. The concept of robustness relates to the question, How well does the estimator perform if one (or more) of the assumptions on which it is based is false? If the performance of an estimator is little affected by the failure of an assumption, it is said to be robust to the particular assumption (also see *Otis et al. 1978:15*). As an example, recall the discussion of model bias in the previous section, where we considered a capture-recapture experiment in which the probability of first capture and recapture were 0.20 and 0.05, respectively. We found that if we incorrectly assumed that probability of first capture and probability of recapture were equal, the estimation procedure based on these assumptions was very biased and therefore performed poorly. Thus the estimator is not robust to the assumption that the probability of recapture and first capture are equal. Unfortunately, most methods for the analysis of capture-recapture data are not very robust. In particular, the assumption of equal catchability is important, because most traditional methods (the Petersen-Lincoln and Schnabel estimators and the Zippin removal method) are not robust to failure of this assumption (*Burnham and Overton 1969; Otis et al. 1978:123-133*).

**Closed-Form Solutions to the ML Estimator.** In general, ML estimators of unknown parameters, such as  $N$  or  $D$ , are found by using calculus techniques on a function closely related to the stochastic model and called the "likelihood function." In many cases, the ML estimator can be written in



Zoe Emily Schnabel

Calvin Zippin became interested in removal sampling while at Johns Hopkins University in the early 1950s, when he began consulting with a group in vertebrate ecology. The group was involved in trapping rats in the Baltimore area to estimate population size. The removal method was known, but its statistical properties, including the standard error formulas for population size and capture probabilities, had not been explored thoroughly. This consulting work developed into his doctoral dissertation at Johns Hopkins.

Zippin joined the faculty of the University of California, San Francisco, in 1953, and he has worked there since then concentrating on biometry and the epidemiology of cancer. He remains interested in capture-recapture and removal sampling.

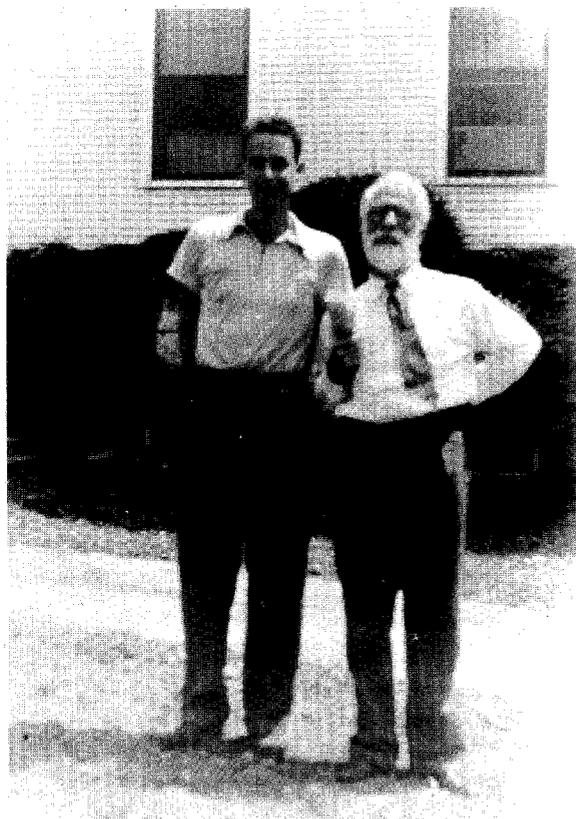
Zippin, shown at the left, is standing with Sir Ronald A. Fisher. Fisher, an outstanding pioneer in statistical theory and practice, worked with several scientists involved in capture-recapture studies. Some of his theoretical developments in mathematical statistics form the basis for much of what we now call the field of statistics. (The photograph was taken in the early 1950s in North Carolina.)

The "Schnabel estimate" has been the backbone of population size estimation, assuming closure, for the past 40 years. It provides an easy-to-compute method for estimating population size in the case where animals are captured, marked, and recaptured over  $t$  occasions. Before Schnabel's work, only  $t = 2$  occasions could be handled by the Petersen-Lincoln method.

Zoe Emily Schnabel (Mrs. George S. Albert since the late 1930s) completed an A.B. degree at Oberlin College and an M. A. degree in mathematics at the University of Wisconsin (1937). She taught mathematics and biometry at Ohio State University and mathematics and statistics at the University of Tennessee before retiring in 1969.

Schnabel's work on capture-recapture studies, however, was done between 1936 and 1938, when she was a graduate assistant in the Computing Laboratory of the Mathematics Department in Madison, Wisconsin. The Laboratory had been established in the early 1930s to assist university researchers with the statistical analysis of data. Members of the mathematics faculty served as consultants, and Schnabel and others assisted with the computations using calculating machines of the era.

The Schnabel estimate was an outgrowth of the work done by Schnabel, E. Hull, and M. Ingraham in the Mathematics Department and D. Juday, a limnologist in the Biology Department. Schnabel concluded her paper with an observation that, unfortunately, many have disregarded: "It should be emphasized, however, that none of the solutions can be expected to provide more than an estimate of the general order of magnitude of the total population." (Photograph taken about 1937-1938.)



Calvin Zippin

a simple form that is easy to use. For example, the Petersen-Lincoln estimator is the ML estimator of  $N$  for a special case of one of the models discussed in Chapter 3. The formula is written as

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

where  $n_1$ ,  $n_2$  are the total number of animals captured on the first and second sampling occasions, respectively, and  $m_2$  is the number of marked animals captured on the second occasion. We say the estimator “exists in closed form.”

Many closed-form estimators found in the published literature are only approximations to the exact ML estimator; examples are the *Schnabel (1938)* and *Zippin (1956)* estimators.

**Numerical Solutions to the ML Estimator.** In capture-recapture models we rarely find that the exact ML estimators exist as simple formulas (as shown above). To illustrate this, consider the model developed by *Darroch (1958)* in which there are four sampling occasions and the capture probabilities



John N. Darroch

In many ways, John Darroch’s work represents a cornerstone in capture-recapture theory. He studied optimal estimation in the model underlying the Schnabel method for closed populations, laid the foundations for the fully stochastic open model developed later and independently by George Jolly and George Seber, and developed the theory for stratified populations—a subject that later captured the attention of Neil Arnason. Also, he supervised the Ph.D. program for Seber at Manchester University.

Dr. Darroch received his undergraduate and early graduate training in mathematics and statistics at Cambridge University in England. He took a lectureship in mathematical statistics at the University of Cape Town, South Africa, in 1955. There, he became interested in the problem of estimating the number of species in a marine environment. This interest led to three papers on capture-recapture, which were published in *Biometrika* and were accepted as a Ph.D. thesis at the University of Cape Town. He returned to England for 3 years before going to Australia in the early 1960s. He is now at the Flinders University in South Australia. (Photograph taken in late 1950s.)

are assumed to vary only by time. The approximate ML estimator for  $N$  for this model is the unique value of  $N$  that satisfies

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

where  $M_5 = M_{t+1}$  = number of individuals caught during the study and  $t$  equals the number of sampling occasions. For example, the total caught on each of four occasions might be  $n_1 = 30$ ,  $n_2 = 15$ ,  $n_3 = 22$ , and  $n_4 = 45$ , and the total individual animals caught at least once,  $M_5$ , might be 79. Then, the ML estimate of  $N$  is the solution of the equation

$$\left(1 - \frac{79}{N}\right) = \left(1 - \frac{30}{N}\right) \left(1 - \frac{15}{N}\right) \left(1 - \frac{22}{N}\right) \left(1 - \frac{45}{N}\right) .$$

There are efficient numerical methods to solve such equations. However, simple trial and error and a little patience will solve an equation as simple as this one. In general, the ML estimator for Darroch's model is derived by solving 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 an iterative basis, by using a sophisticated trial and error numerical procedure. We say the equation does not have a simple, closed-form solution. Complex probability models often do not have simple estimators; nonetheless, complex models appear necessary to describe many capture-recapture studies adequately.

Although we cannot show simple closed-form estimators for most of the models to be discussed, it is the ML concept that is important and we leave it to the computer to do the arithmetic. The numerical methods employed are given in detail in *Otis et al. (1978:103-114)*. The concept that is so important here involves the notion of a likelihood function.

**Likelihood Function.** Formally, the likelihood function is the joint probability density function of the sample data. In the context here, it is a function of the integer-valued parameter  $N$  and the real-valued parameter  $p$  (the vector containing all the probability parameters necessary to the model), given the discrete sample data contained in the  $\underline{X}$  matrix (a matrix of zeros and ones).

The notation is not as complex as it may seem. For example,  $\mathcal{L}(N, p | \mathbf{X})$  denotes the likelihood function of the unknown parameters  $N$  and  $p$ , given a specific set of sample data contained in the  $\mathbf{X}$  matrix. As we will see in the next chapter, this is notation for the likelihood function for Model  $M_0$  (Fig. 2.6). Two more examples will be given. For Model  $M_t$ , we have capture probabilities that may vary among sampling occasions; that is,  $p_1, p_2, p_3, \dots, p_t$ . If we let these values be denoted as the vector  $\mathbf{p}$ , then the likelihood function for Model  $M_t$  can be denoted as  $\mathcal{L}(N, \mathbf{p} | \mathbf{X})$  (Otis *et al.* 1978:106). In Model  $M_b$ , the parameter  $p$  is the probability of first capture, the parameter  $c$  is the probability of recapture, and the likelihood function for Model  $M_b$  is denoted as  $\mathcal{L}(N, p, c | \mathbf{X})$ .

The likelihood function is a formal way to express quantitatively the relative "likeliness" of several values that may be considered as candidates for  $\hat{N}$ . The ML method selects as the value for  $\hat{N}$  the most likely one, on the basis

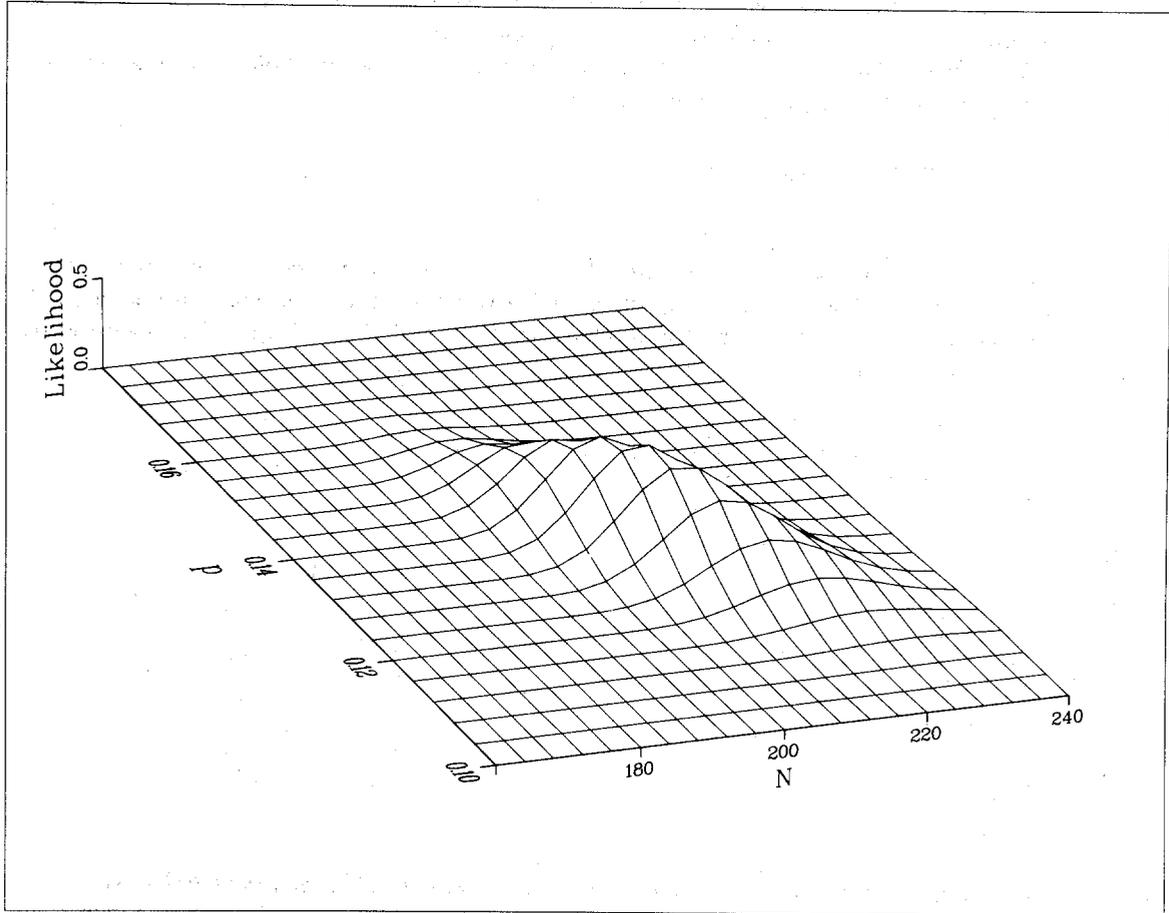


Fig. 2.6. A three-dimensional graph, showing the likelihood function for a given data set under Model  $M_0$ . This model involves only two unknown parameters:  $N$  = population size and  $p$  = constant capture probability. Given the data, the ML estimates of  $N$  and  $p$  are found as those values that maximize the likelihood function. The idea is that we are trying to find those values of  $N$  and  $p$  that make our data seem "most likely." Graphically, we see that values of  $N = 208$  and  $p = 0.136$  are approximately the ML estimates in this example. Of course, with different data, the likelihood function would be of a somewhat different shape and the values of  $N$  and  $p$  that maximize the function also would be different.

of the available data (hence the name, maximum likelihood). The use of likelihood theory and the ML method extends our intuition and ability to make inductive inference. To understand these concepts, consider a 2-sample capture experiment in which 40 animals are captured, marked, and released ( $n_1 = 40$ ) in the first sample. In the second sample, 50 animals are captured; of these 25 (half) have marks from the first occasion ( $n_2 = 50, m_2 = 25$ ). Before computing the ML estimate from the Petersen-Lincoln estimator, let us use our intuition and see what can be inferred from the data about population size. First, there must be at least 50 animals, because we caught that number on the second sample. In fact, at least 65 animals must be present (40 from the first occasion plus the 25 unmarked from the second occasion). However it seems unlikely that  $N = 1000$  because half the animals caught on the second occasion were marked. Moreover, it seems fairly unlikely that the population could be as large as 500 or even 400. For example, if  $N = 400$ , only  $40/400 = 10\%$  of the population would have been marked before the second sample; if  $m_2 = 50$  we would have expected only 5 marked animals in the second sample. Intuitively, we have reason to believe that  $N$  is at least as large as 65 and probably well below 400. The ML estimate is the value selected as the most likely, given the data we observed;  $\hat{N} = (n_1 n_2) / m_2 = 80$ .

The likelihood function is difficult to deal with directly because it involves products of often complicated terms. The likelihood function for Model  $M_0$  (see Chapter 3) is

$$\mathcal{L}(N, p | \underline{X}) = \frac{N!}{(N - M_{t+1})!} p^n (1 - p)^{tN - n},$$

where  $n$  is the total number of captures and recaptures. By taking the natural logarithm of the likelihood function, we can deal with the sums of the terms; dealing with sums is nearly always more desirable than dealing with products of terms (see *Larson 1969:224-226*). This function, denoted as  $\ln \mathcal{L}(N, p | \underline{X})$ , is called "the log likelihood function." The log likelihood function for Model  $M_0$  is

$$\ln \mathcal{L}(N, p | \underline{X}) = \ln \left[ \frac{N!}{(N - M_{t+1})!} \right] + (n) \ln(p) + (tN - n) \ln(1 - p).$$

The term

$$\ln \left[ \frac{N!}{(N - M_{t+1})!} \right]$$

can be written more simply as

$$\sum_{j=N-M_{t+1}+1}^N \ln(j).$$

Details of likelihoods for some capture-recapture models are given in *Otis et al. (1978:102-114)*.

### Basis for Rigorous Inference

Often, capture-recapture data are analyzed, and conclusions are drawn from them by *ad hoc* procedures. For example,  $M_{t+1}$  is used frequently as an "index to abundance." Another index used frequently is the number of animals captured per 100 trap nights. However, the use of indexes in science is to be discouraged because indexes lack the basic factors (Fig. 2.1) required for making inferences about parameters based on data. Indexes are useful only when they have been calibrated with the parameter of interest by using, for example, the theory of double sampling (*Cochran 1977*).

Initially, we must know what assumptions may be needed and which of them seem realistic. (See the previous section on Theory and Reality.) These assumptions should be built into a stochastic model that deliberately relates the sample data to the unknown parameters of interest. Then, a good estimation

procedure is required. This procedure (Fig. 2.1) is essential for making inferences from data. A final, integral step is to test and evaluate the assumptions. This step is especially critical in capture-recapture and removal studies because most estimators are not robust to certain assumptions about capture and recapture probabilities. Tests of model assumptions are computed in program CAPTURE.

### Confidence Intervals

An estimate without both a measure of precision (the sampling variance) and an assessment of the relevant assumptions is not trustworthy and must be regarded as scientifically invalid. A single estimate of  $N$  is not meaningful without a measure of the sampling variation in the estimator. While the variance, standard error, and coefficient of variation are measures of sampling variation (or precision), the construction of a confidence interval for the parameter of interest represents a much stronger inferential statement. A confidence interval usually is written as

$$P[a \leq N \leq b] = 1 - \alpha ,$$

where  $a$  and  $b$  are the lower and upper bounds calculated from the sample data,  $N$  is the parameter of interest, and  $1 - \alpha$  is the significance level. The value of  $\alpha$  is frequently chosen to be 0.05. The bounds for the interval are constructed from a given formula, depending upon the distributional assumptions made about the estimator. For example, confidence intervals for the mean  $\mu$  of a normal population are familiar to most biologists (see *Bliss 1967:186-204*). In this example, the bounds for the 95% confidence interval are computed as

$$\bar{y} \pm 1.96 \text{ se}(\bar{y}) ,$$

where  $\text{se}$  is the standard error. The confidence interval statement implies that if one repeatedly drew a random sample from the population, computed the estimate  $\bar{y}$  of  $\mu$ , and computed the 95% confidence interval, then 95% of the intervals would cover the parameter  $\mu$  (Fig. 2.7).

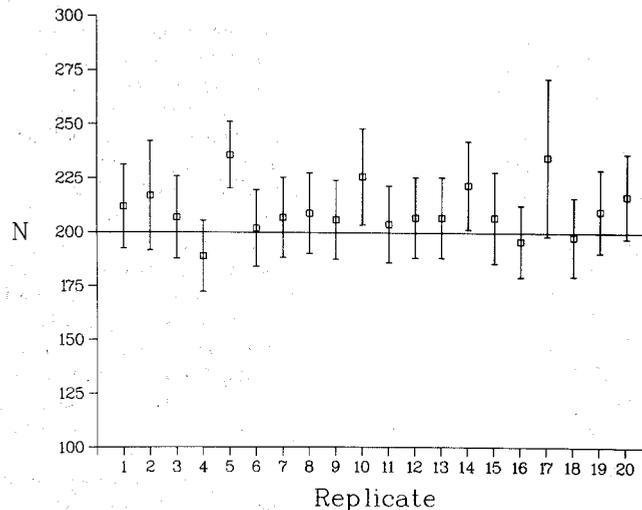


Fig. 2.7. Estimates and 95% confidence intervals plotted for 20 independent surveys. In each simulated case, the true parameter is  $N = 200$ . Note that only 17 of the 20 (85%) intervals cover the parameter. The average value of  $\hat{N}$  appears to be biased high. The bias, of course, could explain why the coverage may be a little below the nominal 95% level. In several capture-recapture models, the actual coverage, unfortunately, is substantially less than 95%.

Most methods in the statistical literature, and all methods in this primer, for constructing confidence intervals are based on the assumption that estimators are distributed as normal random variables; it is thus very desirable for an estimator to be distributed normally. Distribution of ML estimators becomes approximately normal as the sample size increases. Unfortunately, significant nonnormality (positive skewness) often occurs with capture-recapture estimators, partly because sample sizes are too small (see *Otis et al. 1978:133-135*).

Thus, in capture-recapture sampling it is often difficult to find a procedure to compute the lower and upper bounds (not always symmetric around the estimate) of the confidence interval so that, in fact, they include the parameter  $N$  95% of the time. In other words, the actual coverage is often less than 95%. Poor coverage also is often due to poor estimates of the sampling variances, to biased estimators, or to nonnormal distribution of the estimators. (Refer to *Otis et al. 1978:126* and *133-135* for examples and discussion of these problems.) In this primer, 95% confidence intervals are constructed as  $\hat{N} \pm 1.96 [\hat{se}(\hat{N})]$ .

## Tests of Hypotheses

A statistical hypothesis is a statement about one or more parameters of the population of interest. A decision concerning the validity of the hypothesis is made based on the value of a test statistic calculated from the sample data. The test statistic frequently has a common distribution, such as chi-square, normal,  $t$ , or  $F$  (see *Mendenhall and Scheaffer 1973:325-365*). Mathematical statistics is employed to derive the



C. H. N. Jackson

C. H. N. Jackson, D.Sc., made several theoretical contributions to the analysis of capture-recapture data in a series of papers published during the 1930s and early 1940s. His work stemmed from his life-long interest in tsetse flies in the Tanganyika region (now Tanzania). Jackson, an Englishman, consulted with R. A. Fisher on statistical questions during the 1930s.

Jackson proposed several methods based on a variety of assumptions. He gave point estimators and, usually, sampling variance estimators for his methods. In most respects, his work was far advanced for the time.

Jackson was born in 1900 and was awarded Ph.D. and D. Sc. degrees from Cambridge University, the latter for his population studies on tsetse flies. He was awarded the Order of the British Empire not long before he died.

Capture-recapture studies were merely a small part of Jackson's long professional career; his publications cover the period 1927-1955. He was a distinguished entomologist working near Old Shinyanga with the Tsetse Research Center, now the Uganda Trypanosomiasis Research Organization in Tororo, Uganda. Those wishing to gain further insight into this famous entomologist should read the paper by Potts and Jackson (1952); "The Shinyanga game destruction experiment," *Bull. Entomol. Res.*, 43(2);363-374. (Photograph shows Jackson apparently at middle-age, perhaps in the 1930s or 1940s; courtesy of P. M. Mwambu.)

theoretical distribution of the test statistic if the null hypothesis ( $H_0$ ) is true. From such distributions we obtain critical values—numbers that are compared to the value of the test statistic to decide whether  $H_0$  is rejected. For every significance level ( $\alpha$  value) there is a corresponding critical value. Usually  $\alpha$  is set at 0.05 or 0.01. However, the experimenter is free to set the significance level of the test at any value, although very rarely does this value exceed 0.10. Thus, in a sense, the user is specifying the chance that a true null hypothesis will be rejected (Type I error). At this point, an interesting tradeoff is made. If the user is willing to increase the chance of making a Type I error from, say, 0.04 to 0.10, then the corresponding chance of accepting a null hypothesis that is not true (Type II error) is decreased. This result of statistical testing theory explains why different significance levels may be used in different testing situations. It is the responsibility of the experimenter to decide which type of error is the more serious in a specific situation.

To test one hypothesis (specifically, the null hypothesis  $H_0$ ) against another (termed the alternative hypothesis  $H_A$ ), a study is designed, data are collected and analyzed, and if the results are unlikely under this hypothesis, the null hypothesis is rejected. If the results seem probable under the null hypothesis, there is no reason to reject it. The test statistic measures the degree to which the results conform to the null hypothesis.

For example, we might test the null hypothesis  $H_0$ , stating that a penny is fair (that 50% of all tosses will be heads and 50% tails), against the alternative hypothesis  $H_A$ , stating that a penny is not fair, by flipping the penny 500 times and observing the outcome. Intuitively, if we observed 50 heads in 500 tosses we would consider that the result was improbable under the null hypothesis and that  $H_0$  should be rejected in favor of  $H_A$ . We would conclude that the penny is not fair. On the other hand, 248 heads would be considered a likely result, very close to the 250 we would expect, and we would have no reason to reject  $H_0$ . In this intuitive example, we have used the “number of heads obtained” as the value of our test statistic.

When results are obvious, the decision to reject or not reject is clear. In actual practice, however, experimental results are not always so clear and intuition may be of little help. Statistical theory then provides objective methods for making inductive decisions and for evaluating the goodness of the inferential procedures. Simply stated, the decision is whether to reject  $H_0$ .

A basic philosophy of science is that the truth of a null hypothesis cannot be proven. We can reject a null hypothesis on the basis of data from a proper experiment. If the experiment is replicated several times and each time  $H_0$  is clearly rejected, the evidence becomes very convincing that  $H_0$  is false. Conversely, if in repeated experiments, properly conducted, we fail to reject  $H_0$ , we continue to entertain the possibility that  $H_0$  is true. We can never truly “accept”  $H_0$ , but repeated failure to reject it adds to its authenticity.

**Error Types and Distributions under the Null Hypothesis.** In hypothesis testing, two types of errors can be made.

- A Type I error is the rejection of a null hypothesis ( $H_0$ ) that is true. The probability of a Type I error is denoted as  $\alpha$  (the significance level).
- A Type II error is the acceptance of a null hypothesis ( $H_0$ ) that is false. The probability of a Type II error is denoted as  $\beta$ .

The possible outcomes of hypothesis testing are illustrated in Fig. 2.8. Commonly,  $100\alpha$  (in percent) is referred to as the significance level of the test (for example, 5% and 1% are frequently used).

Nearly all of the relevant tests in *Otis et al. (1978)* and this primer are distributed as chi-square ( $\chi^2$ ) variables if sample size is large. For these tests,  $H_0$  is rejected if the test statistic is larger than the critical value. Various chi-square distributions are shown in Fig. 2.9, with rejection (significance) regions. The concept that a test statistic, such as chi-square test, has a distribution is difficult to understand. In many cases, a test statistic follows a chi-square distribution if the null hypothesis  $H_0$  is true. If such a test is replicated 5, 25, 50, and 100 times, a strong tendency toward a chi-square distribution is observed (Fig.

Decision	Null Hypothesis $H_0$	
	True	False
Reject $H_0$	Type I Error ( $\alpha$ )	No Error
Do Not Reject $H_0$	No Error	Type II Error ( $\beta$ )

Fig. 2.8. These are four possible outcomes of a statistical test of hypotheses and their associated errors.

2.10). Again, the concept of repeated samples is the basis for the theory that indicates distribution of a particular test is chi-square under the null hypothesis.

The test of closure has a test statistic with a standard normal distribution (mean = 0, standard deviation = 1). The test, shown in Fig. 2.11, is one-sided. Rejection of  $H_0$  (closure) is based only on negative values of the test statistic (see *Otis et al. 1978:120*).

The power of the test (in percent), defined as  $(1 - \beta)100$ , relates to the ability of the test to reject  $H_0$  if it is false. If a test routinely fails to reject a false hypothesis, we say it lacks power. The power of a test

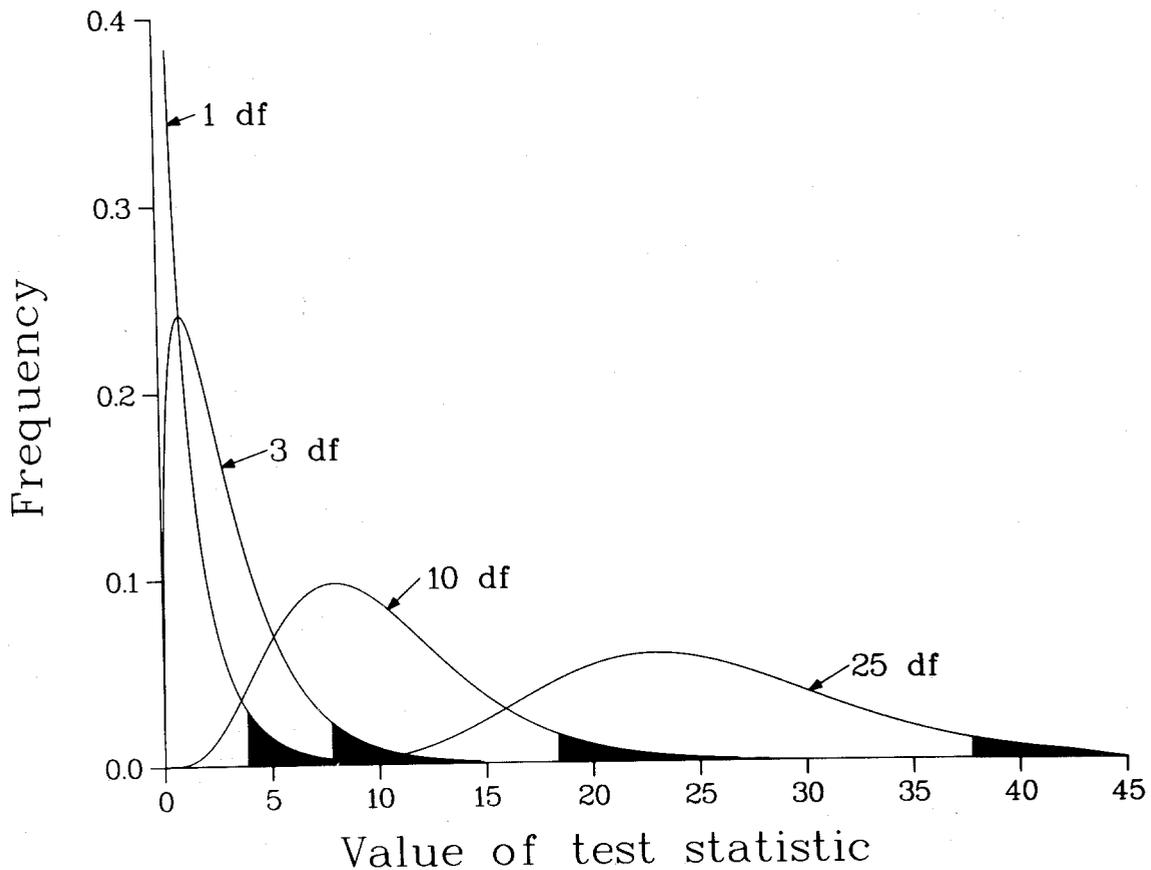


Fig. 2.9. The chi-square distribution for 1, 3, 10, and 25 degrees of freedom (df). In each case, the 0.05 rejection region is shown as a shaded area. All seven test statistics in *Otis et al. (1978:115-119)* are distributed as chi-square under the null hypothesis  $H_0$ . The interpretation of a test statistic that is distributed as chi-square is simple. Suppose that a test statistic for a particular  $H_0$  is distributed as chi-square with 25 df [written as  $\chi^2_{(25)}$ ]. If the computed value of the test statistic exceeds about 35.5, we will reject  $H_0$  (at the 0.05 significance level). The concept is that a value as large as 35.5 is very unlikely if the test statistic is, in fact, distributed as  $\chi^2_{(25)}$ . We consider it sufficiently unlikely, so we decide to reject  $H_0$ . The output from program CAPTURE in *Otis et al. (1978:92)* gives various test statistics that are distributed as chi-square.

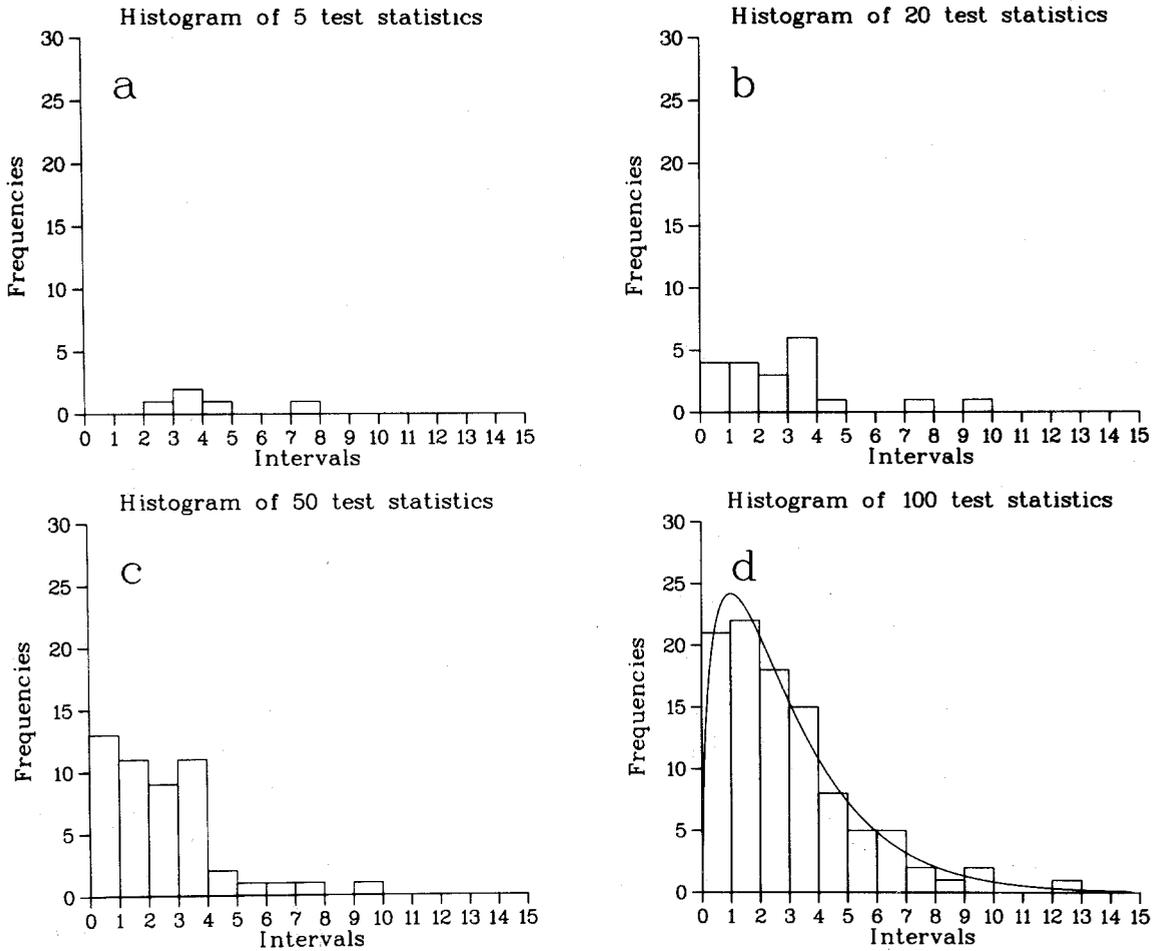


Fig. 2.10. Histograms of test statistics based on 5, 20, 50, and 100 sample data sets. The test statistics are distributed as chi-square under the null hypothesis. In d, the observed value of the test statistic at 12-13 represents a Type I error.

can be computed based on the theory of statistics. For now, only the concept of power is important (Figs. 2.12 and 2.13).

**Error Types and Distributions under the Alternative Hypothesis.** We have discussed and illustrated the concept of the test statistic distribution under a null hypothesis  $H_0$  and how this distribution actually determines the test statistic values that cause rejection of  $H_0$ , for a given significance level ( $\alpha$ ) of Type I error. A test statistic also has a distribution under the alternative hypothesis  $H_A$ . The shape of this distribution determines the power of the test or, equivalently, the size of the Type II error. Figs. 2.12 and 2.13 illustrate this concept.

**Hypothesis Testing in Capture-Recapture and Removal Studies.** In capture-recapture and removal studies we encounter two basic hypothesis-testing situations. To illustrate them, let us suppose that we are considering two models of a capture-recapture study, Model  $M_0$  and Model  $M_t$ . The first test is for “goodness of fit.” An example of this test is

$H_0$ : Model  $M_t$  fits the data

$H_A$ : Model  $M_t$  does not fit the data.

The question asked here is whether Model  $M_t$  is an adequate representation of the data.

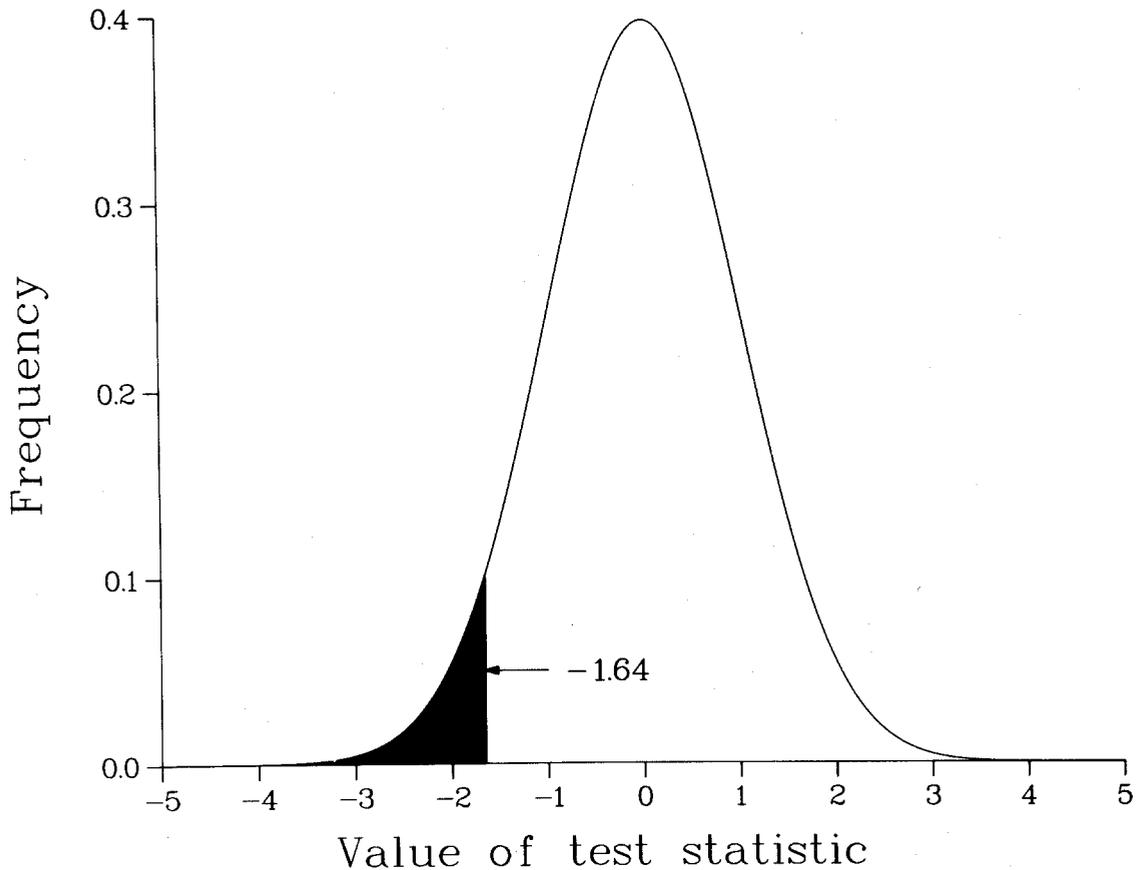


Fig. 2.11. Test statistics frequently are distributed as a standard normal distribution with a mean of 0 and a standard deviation of 1. Such a standard normal distribution is shown, with a one-tailed rejection region for  $\alpha = 0.05$  (shaded). The closure test (Otis et al. 1978:120-121) has this distribution and a one-sided (negative) rejection region. If the test statistic under  $H_0$  (closure) is less than  $-1.64$ ,  $H_0$  is rejected at the  $\alpha = 0.05$  level.

The second test might be termed a “simple alternative” test. An example of this test is

$H_0$ :  $M_0$  fits the data as well as  $M_1$

$H_A$ :  $M_1$  fits significantly better than  $M_0$ .

In the simple alternative tests discussed in this primer, the model under  $H_A$  is more general than the model under  $H_0$ . This test is a comparison of the two models. Thus, the question asked here is, Does the more general model (Model  $M_1$ ) fit the data significantly better than the simpler model (Model  $M_0$ ), or does the simpler model do as well?

The fundamental difference between these two tests is that the first is concerned with the question of whether one specific model provides a good fit to the data, whereas the second compares a specific model to a more general model to see which provides the better fit to the data. *Begon (1979:55-75)* presents a section on testing assumptions in capture-recapture models that is easy to understand.

Capture-recapture and removal studies represent not only very difficult testing problems, but also difficult modeling and estimation problems. The many technical reasons for these difficulties are beyond the scope of this primer. We will, however, mention four causes of testing problems.

(1) With small samples, the distribution of a test statistic may not follow the theoretical (large-sample) distribution very well (\*Fig. 2.14).

(2) The test may have poor power and thus may make rejecting the null hypothesis difficult when, in fact,  $H_0$  is false (\*Fig. 2.15).

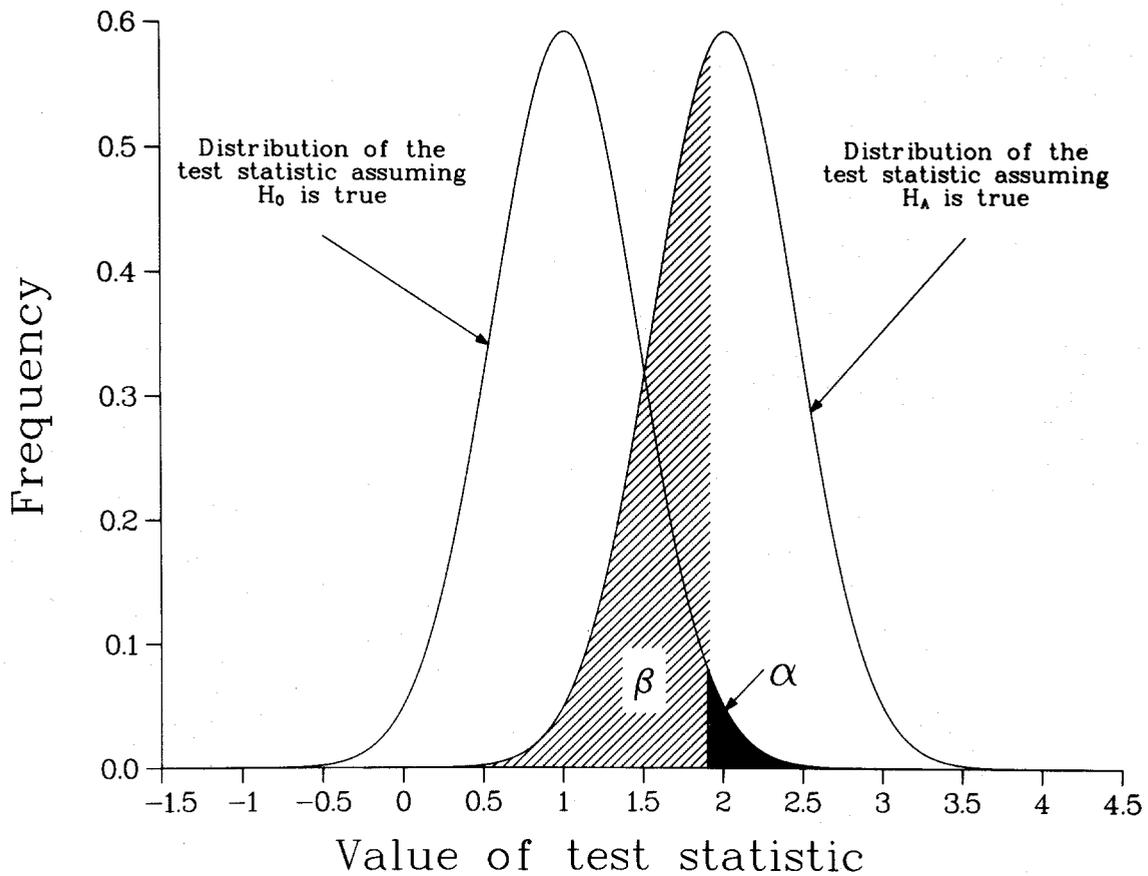


Fig. 2.12. An example of the distribution of a test statistic under the null hypothesis  $H_0$  and under the alternative hypothesis  $H_A$ . The probability of a Type I error, denoted as  $\alpha$ , is black; the probability of a Type II error, denoted as  $\beta$ , is crosshatched. In this example,  $\alpha = 0.05$  and  $\beta = 0.45$ ; the power of the test of the null hypothesis is  $1 - \beta$  and is considered to be fairly poor in this example.

(3) The battery of tests to be described (see *Otis et al. 1978:115-120*) are very dependent on each other. Dependence, which is to be expected because all the tests are computed with roughly the same data, makes interpretation difficult. (The difficulty is alleviated, in part, by the automatic model selection algorithm in CAPTURE.)

(4) Certain tests cannot be computed unless substantial amounts of data are available for analysis; *Leslie's (1958)* test is an example.

In dealing with capture-recapture and removal methods, the user of this primer need not be concerned with the derivation of the test or how the test statistic is distributed. Program CAPTURE computes the value of the required test statistics and the observed significance levels. Biologists, therefore, need only to interpret the results.

## Simulation Methods

Since the late 1960s, computer simulation has been used to study the performance of various estimators of population size from capture studies. Simulated populations provide a set of essential features: (1) the primary parameter  $N$  is known exactly; (2) the capture probabilities are known and can be manipulated at the will of the investigator; (3) the assumptions can be deliberately met or violated; (4)



Patrick H. Leslie

Patrick Leslie, “George” to many friends and colleagues, is probably best known (in the context of the subject here) for his work in estimating population size and death rates in populations of small mammals. During the late 1940s and early 1950s he collaborated on a series of papers about voles with Dennis and Helen Chitty while working in the Bureau of Animal Population with Charles Elton. Before this work, he had developed several regression-based methods. His most important work dealt with the population projection matrix methods—the “Leslie matrix.”

Leslie took his undergraduate training in physiology in 1921 and earned the Doctor of Science degree at Oxford. He joined the Bureau of Animal Population Research at Oxford in 1935 and continued there as Senior Research Officer until his retirement in 1967. Incredibly, in view of his accomplishments, Leslie had no formal training in advanced mathematics; his talent for applying mathematical approaches to ecological problems did not become apparent until after he was 35 years old. Leslie’s career is reminiscent of that of Sir Ronald Fisher, in that they both maintained contact with the real problems of colleagues in other fields and loved to explore real data. Further information can be found in *Nature* 239(5373):477-478, in an article written after his death. (Photograph taken in 1949 by D. A. Kempson, Bureau of Animal Population, Oxford University.)

the proper stochastic sampling variation can be emulated; and (5) the simulated study can be repeated exactly, if necessary (*Bishop and Sheppard 1973*).

Simulated populations are very useful in answering a host of questions concerning the small-sample properties of an estimator under data from its model (bias), or under other models (robustness); the confidence interval coverage; the power of hypothesis tests; and other important issues. Simulated population data are inexpensive to generate and useful for many purposes. For example, Tables 17-19 of *Otis et al (1978:60-62)* involve the analyses of 2400 simulated data sets.

We use simulation to generate a data matrix  $X$  with known properties. Let us start by considering the first animal on the first trapping occasion. How can we determine if it is to be captured or not? We begin by looking at the capture probability (a parameter) for this animal. Let us say that this is 0.3; that is,  $p = 0.3$ . Of course, an animal is either captured or not—it cannot be 0.3 caught. In addition, the data must be simulated to preserve the chance (or stochastic) element in sampling studies. Therefore a uniform random number between 0 and 1 is generated by the computer. This number is compared to 0.3. If the random number is less than 0.3, the animal is “captured” and the first element of the  $X$  matrix is set to one, indicating the animal was captured. The sample procedure is performed for animal 2, and so on. Input values for the capture probabilities  $p_{ij}$  are available within the computer. For example, if animals are assumed to be trap happy, then animal 1 will have  $p > 0.3$  on subsequent trapping occasions.

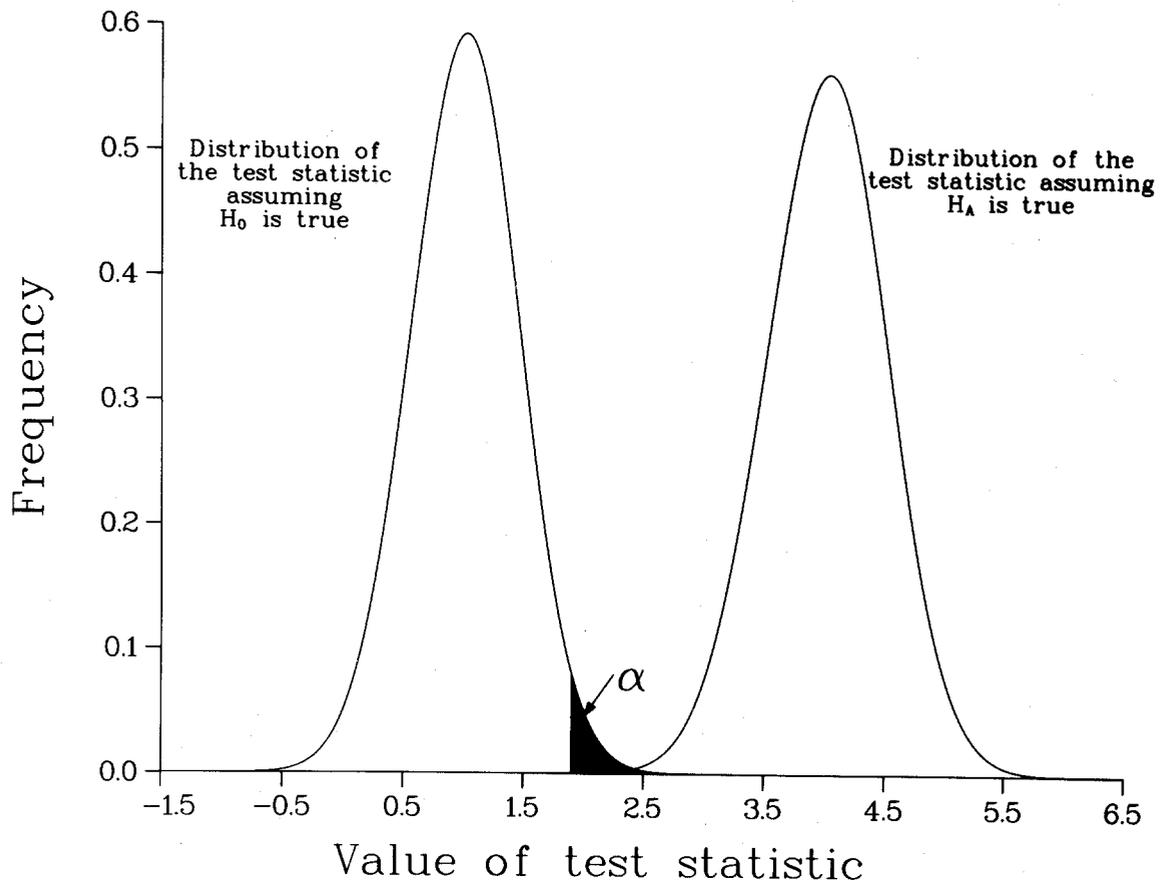
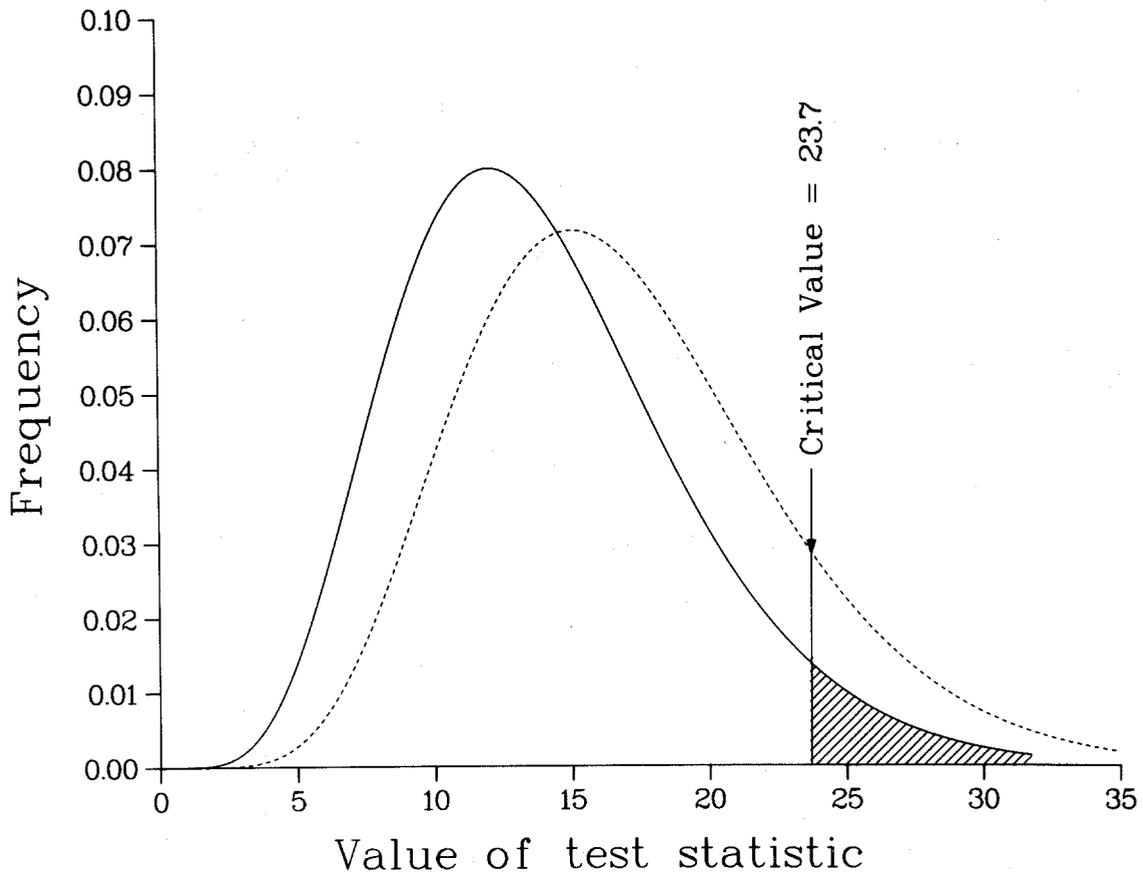


Fig. 2.13. A second example of the distribution of a test statistic under the null hypothesis  $H_0$  and the alternative hypothesis  $H_A$ . The probability ( $\alpha$ ) of a Type I error is shown in black; the probability ( $\beta$ ) of a Type II error is essentially zero. The power of the test  $1 - \beta$  is nearly 1, indicating a very good and powerful test of  $H_0$ . (Compare with Fig. 2.12.)

### Summary

1. Sampling a population enables valid inferences to be made about various parameters if proper procedures are used in the field and during the analysis.
2. A mathematical model is required to link the sample data with the necessary assumptions and to provide a basis for parameter estimation. Stochastic models are needed because the sample data result from processes with a strong random component.
3. Estimators for capture-recapture and removal studies should be unbiased and precise. Proper consideration of basic principles enables estimators to have these properties. Estimates of population parameters must have a measure of precision to be of value in making valid inductive inference.
4. Variation is everywhere in capture-recapture and removal experiments. The two types of variation (spatial and temporal, and stochastic) must be recognized in biological work.
5. Adequate sample size and the magnitude of the capture probabilities are critical elements to consider in the design of a study.
6. Random sampling is inappropriate to most capture-type studies, and the methods discussed here make no use of this assumption.
7. Only the estimator for Model  $M_h$  can be computed easily by hand. The comprehensive computer program CAPTURE is required for essentially all of the analyses described here.
8. Tests of hypothesis are important to assess the tentative assumptions and to select the best model.

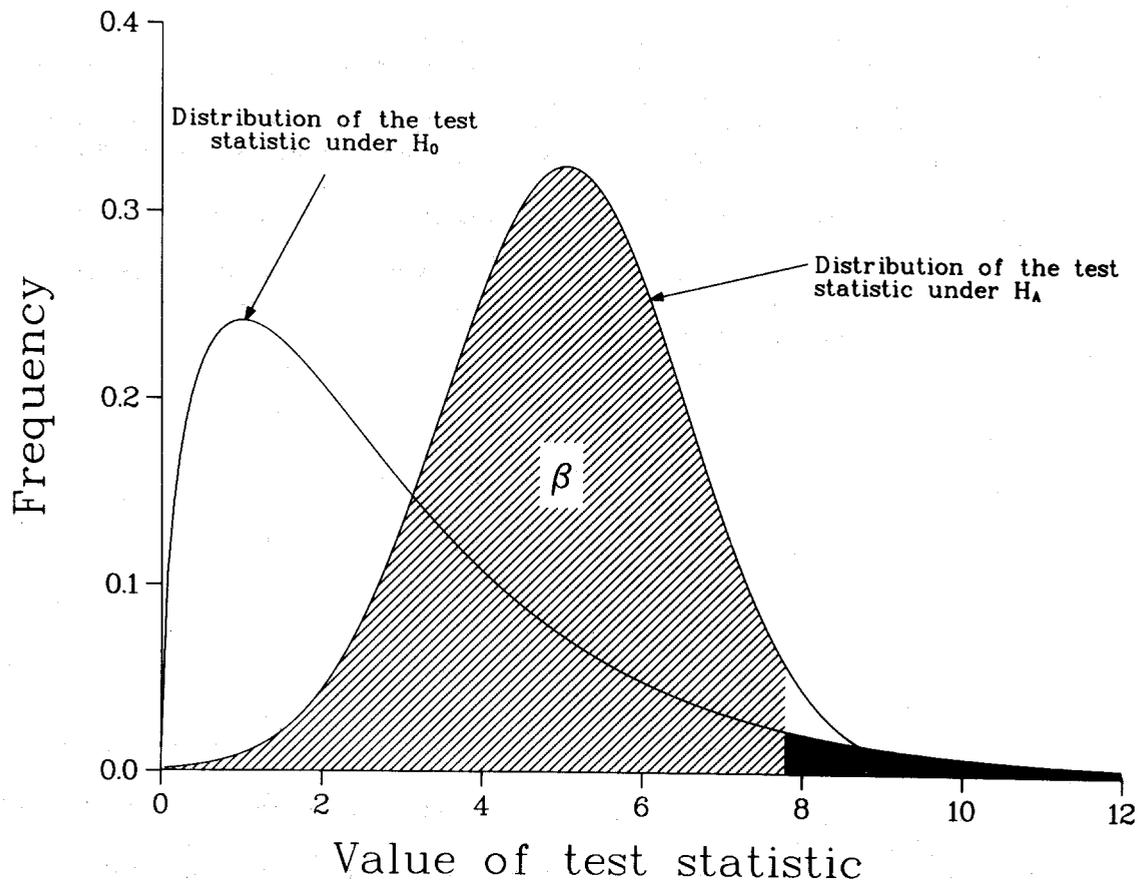


\*Fig. 2.14. The fact that a test statistic may not follow the theoretical distribution if sample size is small can present problems. Shown here is the distribution of a test statistic that asymptotically follows a chi-square distribution with 14 df (solid line). However, with the sample sizes encountered in this series of experiments, the distribution was approximated poorly by the chi-square distribution. The rejection region for  $\alpha = 0.05$  for the chi-square distribution is shown in the crosshatched area. Note, however, that the critical value 23.7 defines a significance level on the actual sample distribution that is much larger than 0.05. Because the sample distribution is not well approximated by the chi-square distribution, we would reject the null hypothesis in this example when it is, in fact, true more often than 5% of the time (because more than 0.05 of the area of the actual distribution is to the right of 23.7).

### Questions and Exercises

1. Are simple formulas available that enable biologists to compute ML estimates of  $N$  for most models?
2. Why is a model needed to estimate parameters from data?
3. If an *ad hoc* method provides  $\hat{N}$  in close agreement with an estimate from a method with a rigorous underlying theoretical basis, does this agreement provide substantial support for the *ad hoc* method?
4. What might be a reasonable coefficient of variation for  $\hat{N}$  for research studies? For management-oriented studies?
5. Compute the average, say  $\bar{N}$ , its standard error  $\hat{se}(\bar{N})$ , and its coefficient of variation (cv) for each study below. If  $N = 20$ , which estimate is precise, which is biased, and which is both precise and biased or neither? (See Figs. 2.4 and 2.5.)

Study 1	20	21	17	19	22
Study 2	25	26	22	24	27
Study 3	5	8	24	30	33
Study 4	25	38	46	50	57



\*Fig. 2.15. Distribution of a test statistic with very poor power (power =  $1 - \beta$ ). The very large  $\beta$  region, which is the probability of accepting  $H_0$  given  $H_A$  is true, is shown in the crosshatched area.

6. Consider  $T_1$  and  $T_2$ , two test statistics of a specific hypothesis;  $T_1$  has power 0.13 and  $T_2$  has power 0.89. Which would you prefer?
7. Name some statistical distributions that test statistics commonly follow, if sample size is large.
8. What are the null and alternative hypotheses for a goodness of fit test?
9. What are the two ways in which a hypothesis test can fail to give the "correct" result?
10. Based on a large sample, you compute an estimate of a parameter  $\theta$ , as  $\hat{\theta} = 141$ , with  $\hat{se}(\hat{\theta}) = 13.1$ .
  - a. What is the 95% confidence interval?
  - b. Is a true value of  $\theta = 95$  a reasonable value?
  - c. Similarly, is  $\theta = 135$  plausible?
11. The answers to 10b and 10c were somewhat intuitive.
  - a. Can formal hypotheses be formed for 10b and 10c?
  - b. What is the form of the test?
  - c. Compute and interpret the test statistics.
12. A colleague shows you an estimate of population size for snails in geographic areas A and B. In area A,  $\hat{N} = 4306$ , and in area B,  $\hat{N} = 3911$ . What can be inferred from these estimates?
13. You work for an agency and your supervisor tells you of the agency's concern for the Wabo tributary of the huge Lake Powell. The concern is over the possible reduction in the number of spawning lake bass in this area caused by oil drilling and exploration in this area of the lake. You are told to "find two technicians and get out there and find out what we need to know." The following questions should occur to you. How would you answer them?
  - a. What is the population of interest?
  - b. Is a sample or a census called for?
  - c. What is the parameter of interest?

- d. What sampling methods might be useful if a sample is required?
  - e. Suppose a good estimate of population size before exploration and drilling is available. Formulate a formal null and alternative hypothesis of interest.
14. You see in the literature that a certain parameter estimator had a 95% confidence interval of (31 to 91) for a given sample. Does the true parameter lie within this specific interval with probability 0.95? Why?
  15. If you tested a null hypothesis and made a Type I error, what would you conclude? Is your conclusion correct?
  16. Consider the results of a 5-year study of mice in an old field in Wisconsin. Grid trapping was done with live traps for 7 days and the data for each year were analyzed carefully. The estimates of population size ( $\hat{N}$ ) appear below, along with the true parameters ( $N$ ).

Year	$\hat{N}$ ( $\pm se$ )	$N$
1	115 ( $\pm 15$ )	100
2	170 ( $\pm 30$ )	150
3	150 ( $\pm 33$ )	200
4	256 ( $\pm 40$ )	225
5	42 ( $\pm 8$ )	50

- a. What is the cv for each estimate?
  - b. Are the individual estimates fairly good?
  - c. Can good inferences be made from the five estimates about the actual population changes over the 5 years?
17. Why have exact ML estimators for many of the capture-recapture and removal models not appeared in the literature until recently?
  18. Is it necessary to know the details concerning likelihood functions and estimation theory before using some of the analysis methods presented here and in *Otis et al. (1978)*?
  19. Give two or three reasons why a stated 95% confidence interval may cover the true parameter less than 95% of the time.
  20. You have defined a null hypothesis, collected appropriate data, computed a proper test statistic, and found the observed significance level is 0.007. What can you conclude? Why?
  21. If  $var(\hat{N}) = 625$ , what is  $se(\hat{N})$ ?
  22. Examine Table N.3.6 in *Otis et al. (1978:127)*. Is the estimator for Model  $M_t$  robust to trap-happy and trap-shy populations?

# CHAPTER 3

## CAPTURE-RECAPTURE MODELS

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

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

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

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

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

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

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

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

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

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

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

## Modeling Capture Probabilities

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

unmarked balls in the sample. The basic assumption is that on the average, that is, in terms of statistical expectations, the ratio of marked balls to total balls in the population will be the same as the ratio of marked balls in the sample— $(n_1/N) = (m_2/n_2) \equiv M_2/N$ .

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

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

or

$$\begin{aligned}\hat{N} &= \frac{30 \times 36}{10} \\ &= 108 .\end{aligned}$$

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

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

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

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

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

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

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

handful of balls from an urn) on a series of separate occasions; the capture probabilities of our models reflect this aspect of the process.

### Model $M_0$ , Constant Capture Probabilities

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

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

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

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

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

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

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

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

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

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

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

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

NUMBER OF TRAPPING OCCASIONS HAS	6
NUMBER OF ANIMALS CAPTURED, M(T+1), HAS	47
TOTAL NUMBER OF CAPTURES, N., HAS	88
ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.2654	
POPULATION ESTIMATE IS	55 WITH STANDARD ERROR 4.1570
APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL	46 TO 64

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

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

The estimated (average) capture probability  $\hat{p}$  gives some insight into the reliability of results as do, the standard error of the estimate of  $N$ , and the resultant confidence interval. With typical studies of populations in the size range from 50 to several hundred, results are not reliable unless the true average capture probability is at least 0.1 and preferably at least 0.2. In Model  $M_0$ , the estimate of  $p$  is the same as the estimated average capture probability. A  $\hat{p}$  of 0.27 and the small standard error on  $\hat{N}$  of 4.2 indicate the results are trustworthy.

There is no fail-safe way to be sure that the true average capture probability  $\bar{p}$  for a study exceeds 0.10, because we will have only an estimate of this value. We have not presented confidence intervals about  $p$  itself in Model  $M_0$  (or about  $\bar{p}$  in other models). Such intervals would be reliable only if the model were true. Our reason for trying to judge whether  $\bar{p}$  is at least 0.10 is to provide a basis for deciding whether we can reasonably trust the model selection procedure and the resultant estimate of  $N$ . If the data are poor, the estimate  $\hat{p}$  can be quite a bit greater than 0.10, yet the true  $\bar{p}$  can be less than 0.10. Thus, the problem of judging the data's reliability from the data themselves is one of circularity. If the study is reliable (that is, if the capture probabilities are high enough and the true population size is sufficiently large), one probably will conclude this from the data. But if the data are poor, especially because of very low capture probabilities, this fact cannot always be determined from analysis of the data.

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

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

In the example of Fig. 3.2 we have

$$\begin{aligned} cv(\hat{N}) &= \frac{4.157}{55} \\ &= 0.0756 \text{ or } 7.56\% . \end{aligned}$$

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

We have illustrated Model  $M_0$  with the first simulation results of 10 repetitions, using  $N = 50$ ,  $p = 0.3$ , and  $t = 6$ . Summary results of all 10 simulations for this "study," given in Table 3.3, illustrate natural sampling variability. For example,  $M_t$  varies from 41 to 47,  $n_t$  varies from 81 to 99, and  $\hat{N}$  varies from 45

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

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

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

### Model $M_t$ , Variation by Time

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

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

$$E(n_1) = 150.0 \times 0.20 = 30$$

$$E(n_2) = 150.0 \times 0.40 = 60$$

$$E(n_3) = 150.0 \times 0.30 = 45$$

$$E(n_4) = 150.0 \times 0.35 = 52.5$$

$$E(n_5) = 150.0 \times 0.25 = 37.5$$

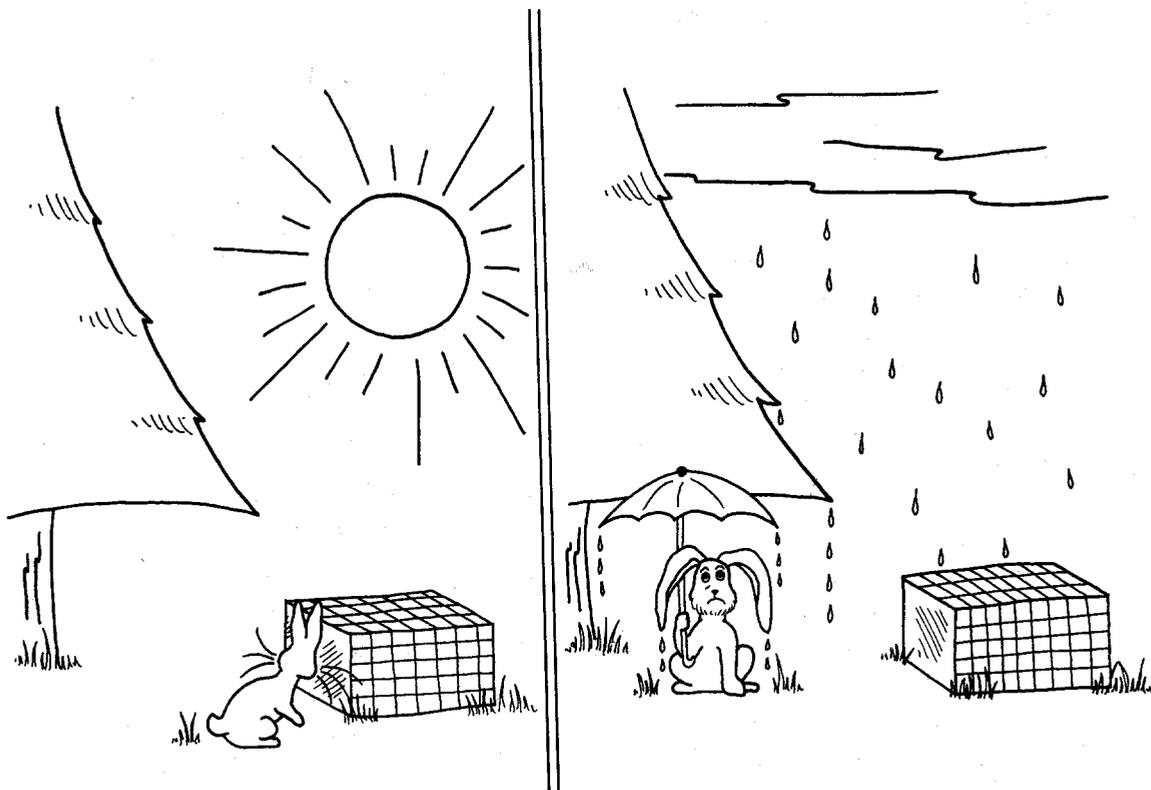
The observed daily captures  $n_j$  shown in Fig. 3.3 reflect these expected numbers.

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

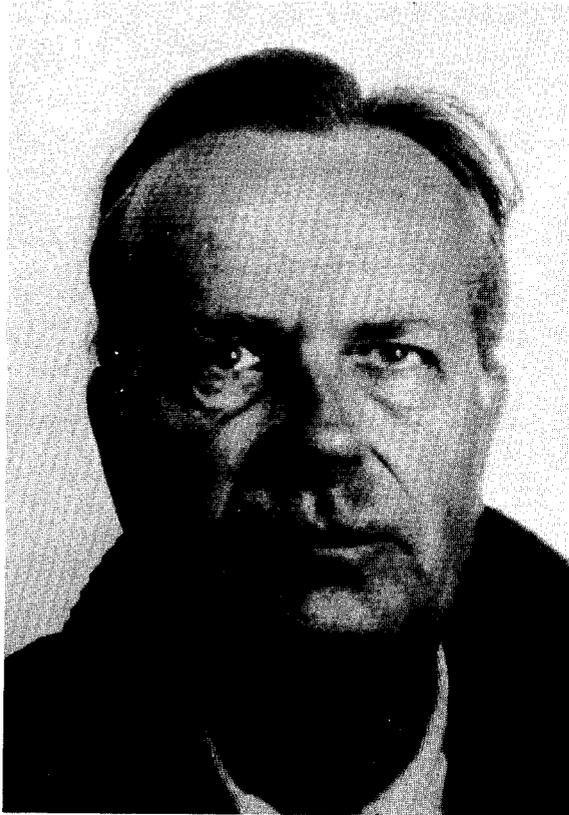
Fig. 3.3. Basic summary statistics for the first of 10 simulations of Model  $M_t$ , with  $N = 150$ ,  $t = 5$  occasions, and daily capture probabilities  $p_1 = 0.20$ ,  $p_2 = 0.40$ ,  $p_3 = 0.30$ ,  $p_4 = 0.35$ , and  $p_5 = 0.25$ . The average capture probability is  $\bar{p} = 0.30$ . Observe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

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

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



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



Norman T. J. Bailey

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

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

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

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

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

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

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



Douglas G. Chapman

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

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

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

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

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

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

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

```

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

TOTAL ANIMALS CAPTURED      127

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

POPULATION ESTIMATE IS      151 WITH STANDARD ERROR      6.9754
APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL      137 TO      165

HISTOGRAM OF N(J)

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

  60
  54
  48
  42
  36
  30
  24
  18
  12
   6

```

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

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

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

## Model $M_b$ , Behavioral Response

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

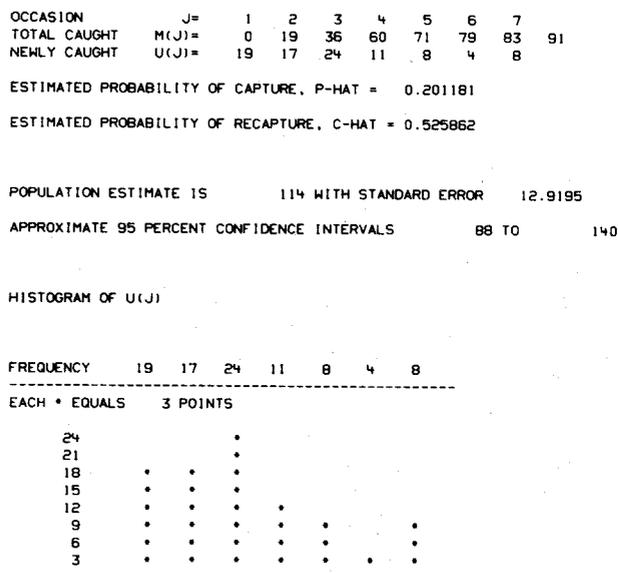
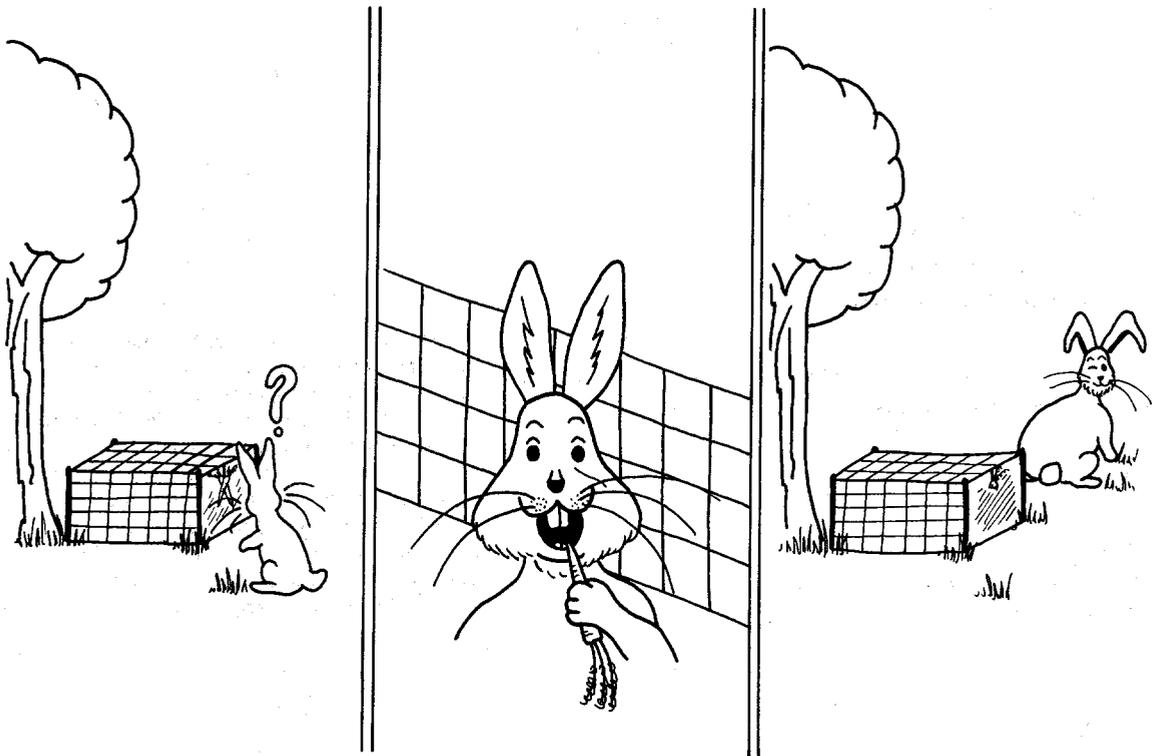


Fig. 3.6. Results of estimating  $N$ ,  $p$ , and  $c$  under Model  $M_b$  from the data in Fig. 3.5. The ML estimate of  $N$  is 114, with an estimated standard error of 12.9. The approximate 95% confidence interval is computed to be 88 to 140; however, the lower limit must be replaced by 91 when the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

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

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



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

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

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

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

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

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

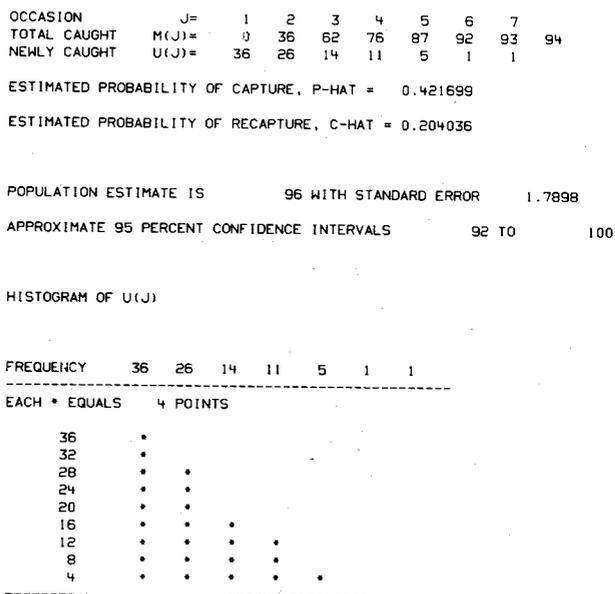
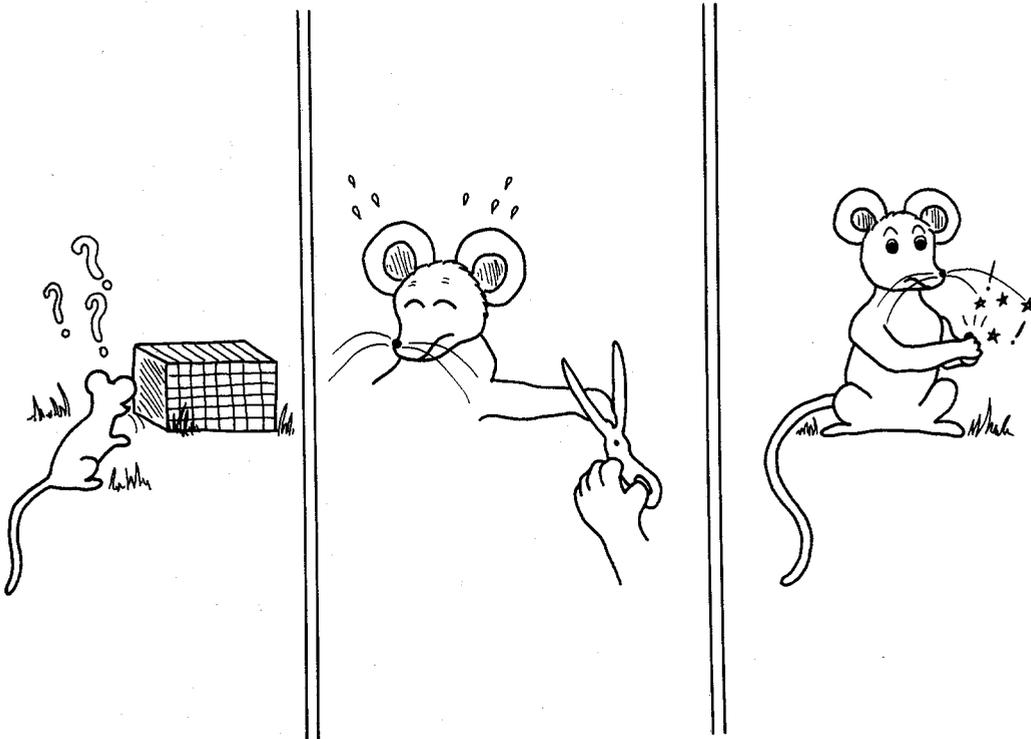


Fig. 3.8. Results of estimating  $N$ ,  $p$ , and  $c$  under Model  $M_b$  from data in Fig. 3.7. The ML estimate of  $N$  is 96, with an estimated standard error of 1.8. The confidence interval should be taken as 94 to 100. The true parameter values underlying these Model  $M_b$  simulated data are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

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

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

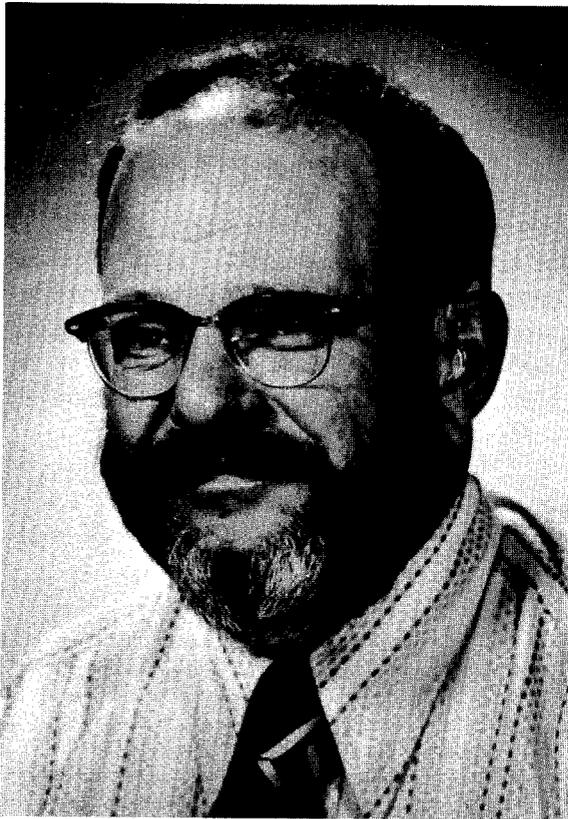


Trap-shy behavioral changes in capture probabilities may result from an unfavorable first-capture experience.

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

## Model $M_n$ , Heterogeneity

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



L. Lee Eberhardt

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

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

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

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

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

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

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

Estimation of  $N$  under Model  $M_h$  is based on the capture frequency data. For the example in Fig. 3.9, these data are  $f_1 = 50, f_2 = 46, f_3 = 35, f_4 = 24, f_5 = 14, f_6 = 5, \text{ and } f_7 = 0$ . Thus, 50 animals were caught only once, but 5 animals were caught on 6 of the 7 capture occasions. The jackknife estimator of  $N$  is computed as a linear combination of these capture frequencies:  $\hat{N} = a_1 f_1 + a_2 f_2 + \dots + a_t f_t$ . The key to this estimator is the derivation of the coefficients  $a_i$ . In keeping with our emphasis on concepts, we will not delve into the mathematics behind the jackknife estimator of  $N$ . For those details see *Burnham and Overton (1978, 1979)* and *Otis et al. (1978:33-37, 108-109)*.

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

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

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

NUMBER OF TRAPPING OCCASIONS WAS 7  
 NUMBER OF ANIMALS CAPTURED,  $M(t+1)$ , WAS 174  
 TOTAL NUMBER OF CAPTURES,  $N$ , WAS 443

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

COMPUTED JACKKNIFE COEFFICIENTS

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

THE RESULTS OF THE JACKKNIFE COMPUTATIONS

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

AVERAGE  $\bar{p}$  = 0.2944

INTERPOLATED POPULATION ESTIMATE IS 215 WITH STANDARD ERROR 8.6844

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 198 TO 233.

HISTOGRAM OF  $F(i)$

FREQUENCY	50	46	35	24	14	5	0
EACH * EQUALS	5 POINTS						
50	*						
45	*	*					
40	*	*					
35	*	*	*				
30	*	*	*				
25	*	*	*	*			
20	*	*	*	*	*		
15	*	*	*	*	*	*	
10	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*

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

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

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

The jackknife estimator is not an ML estimator. We cannot derive a useful ML estimator for Model  $M_h$  because of the many parameters. This model is mathematically very difficult to deal with, but it probably is very realistic for many studies. Consequently, having an estimator for it is important, even though the jackknife estimator does not perform well under some patterns of heterogeneity. Specifically, if many animals have very small capture probabilities (say, less than 0.05), the jackknife estimator will

underestimate  $N$ , as will every other known estimator. There is no mathematical solution to this problem; if some animals are essentially uncachable, no estimation method can estimate  $N$  properly.

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

### Model $M_{bh}$ , Behavioral Response and Heterogeneity

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

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

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

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

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

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

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

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

The results of applying the Model  $M_{bh}$  estimator to the first-capture data of Fig. 3.11 are shown in Fig. 3.12. The computed estimate is 192 with an estimated standard error of 5.0. The approximate 95% confidence interval covers the true value of  $N = 200$ , and the lower limit of the interval (182) is not less than the number of different individuals caught ( $M_3 = 181$ ). In addition to the estimate of  $N$ , the print-out gives information on the estimation process and provides a histogram of the "removal" data,  $u_1, \dots, u_8$ . To explain the main body of numbers in Fig. 3.12, we must discuss the ideas behind estimation of population size with Model  $M_{bh}$ .

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

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	N(J)=	64	47	54	39	40	50	46	39
TOTAL CAUGHT	M(J)=	0	64	96	128	142	155	168	175
NEWLY CAUGHT	U(J)=	64	32	32	14	14	12	7	6
FREQUENCIES	F(J)=	74	50	30	21	5	1	0	0

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

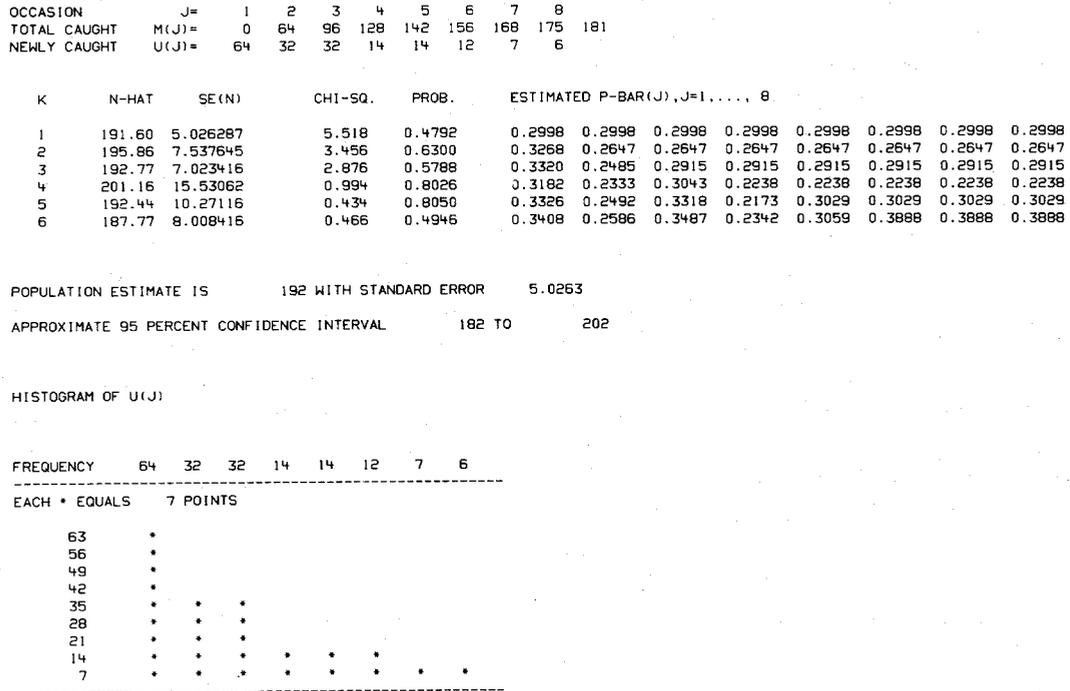
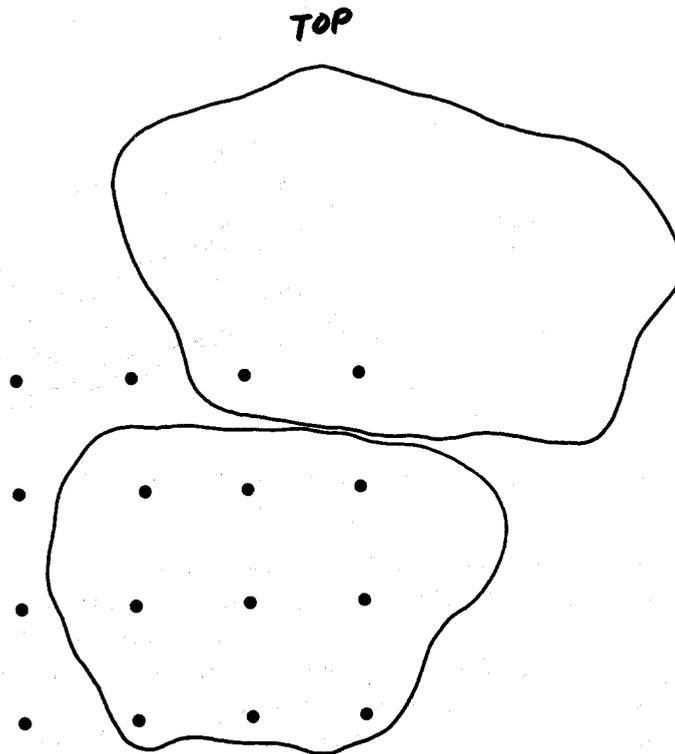


Fig. 3.12. Results of estimation of N from the simulation data of Fig. 3.11; the true population size is 200, and the true model is  $M_{bh}$ . The estimator under Model  $M_{bh}$  computed by CAPTURE is 192, with an estimated standard error of 5.03. The confidence interval covers N and does not overlap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

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

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

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



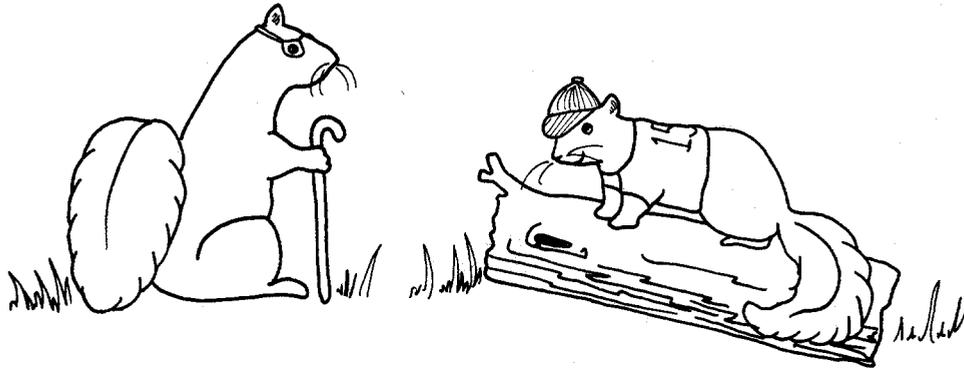
Capture probabilities can be affected by the number of traps in an individual's home range. The two traps in the range above versus the nine in the range below are a source of heterogeneity.

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

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

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

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



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

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

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

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

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

## Model $M_{th}$ , Time Effects and Heterogeneity

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

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

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

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

## Model $M_{tb}$ , Time Effects and Behavioral Response

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

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

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

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

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

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

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

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

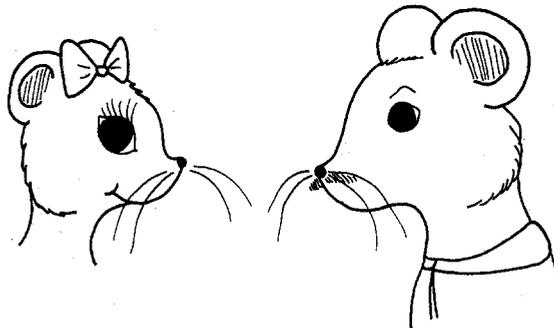
**TABLE 3.11.** A representation of the capture probabilities applicable to different capture histories of animals in the simulation example of Model  $M_{tb}$ . If an animal is captured on day 1, the first row of the table gives its capture probabilities. On days 2 through 5, the animal is subject to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

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

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

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

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



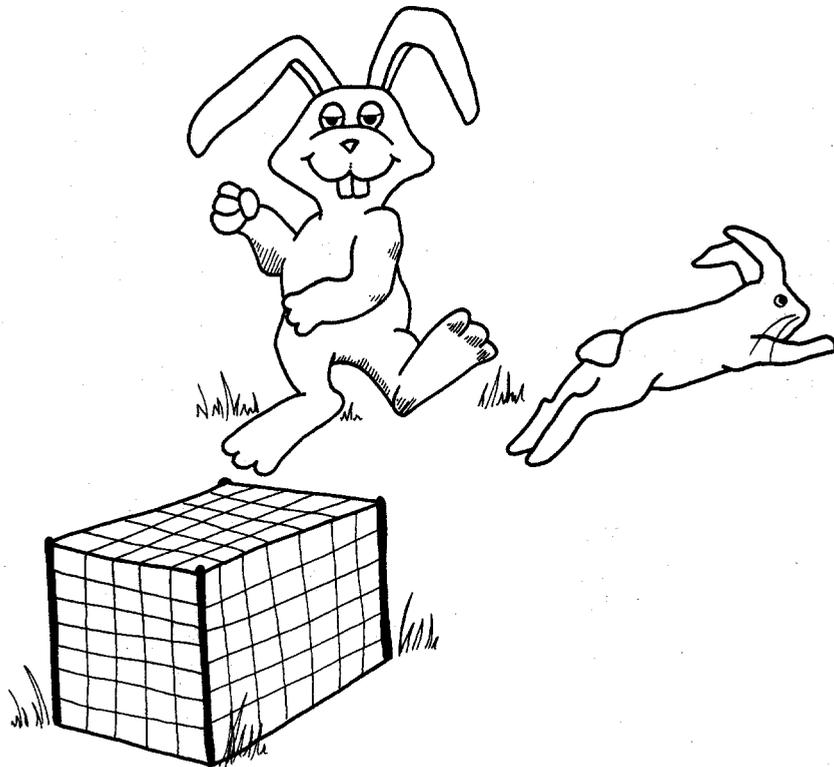
Heterogeneity of individual capture probabilities can be caused partly by differences in sex-specific capture rates.

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

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

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

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



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

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

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

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

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

## Summary of Models

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

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

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

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

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

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

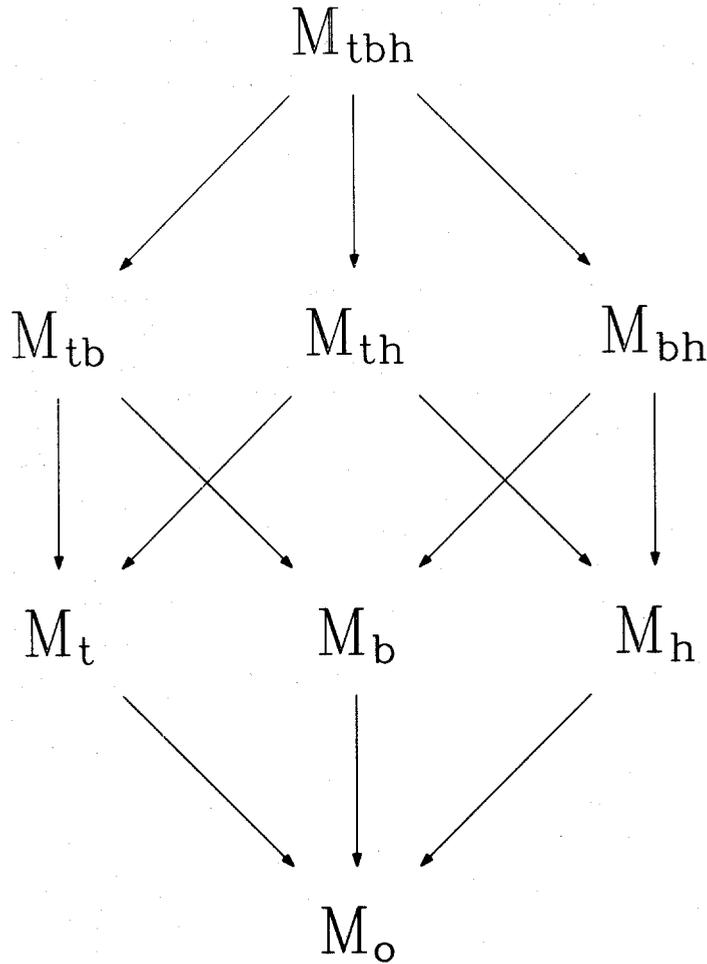


Fig. 3.16. Relations among the models. Each arrow points from one model to an immediate special case of that model. Tests of assumptions are based partly on these relations among the models.

## Testing Model Assumptions and Model Selection

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

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

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

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
CHI-SQUARE VALUE = 3.524 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.31767
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
CHI-SQUARE VALUE = 0.002 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.96170
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
CHI-SQUARE VALUE = 2.670 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.75075
4. GOODNESS OF FIT TEST OF MODEL M(H)  
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
CHI-SQUARE VALUE = 2.485 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.77878  
TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE  
(FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)  
NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY  
-----  
1 3.824 5 0.57509  
2 2.614 5 0.75929
5. GOODNESS OF FIT TEST OF MODEL M(B)  
NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
CHI-SQUARE VALUE = 6.433 DEGREES OF FREEDOM = 8 PROBABILITY OF LARGER VALUE = 0.59886  
5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
CHI-SQUARE VALUE = 3.183 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.52774  
5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
CHI-SQUARE VALUE = 3.250 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.51686
6. GOODNESS OF FIT TEST OF MODEL M(T)  
NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
EXPECTED VALUES TOO SMALL. TEST NOT PERFORMED.
7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
CHI-SQUARE VALUE = 5.410 DEGREES OF FREEDOM = 11 PROBABILITY OF LARGER VALUE = 0.90970

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

Fig. 3.17. The seven tests of assumptions used in model selection applied to the simulated data from Model  $M_0$  with true  $N = 50$  and  $p = 0.30$ . The underlying data are exactly the same as those used for Figs. 3.1 and 3.2 and given in Table 3.2.

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

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

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

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

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

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

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

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

**Test 5, Goodness of Fit of Model  $M_b$ .** The null hypothesis for test 5 is that Model  $M_b$  fits the data versus the alternative that Model  $M_b$  does not fit the data. This null hypothesis can be broken into two parts: (1) first-capture probabilities are constant over time, and (2) recapture probabilities are

constant over time. The parts constitute tests 5a and 5b, respectively. The sum of the two chi-square test statistics gives the overall goodness of fit statistic for Model  $M_b$ .

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

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

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

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

## The Model Selection Procedure

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(0) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 1.460 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.69161

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(0) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 1.654 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.19842

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(0) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 23.479 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00010

4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 20.405 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00042

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

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

5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 21.235 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00166

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 14.827 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00197

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 6.408 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.09336

6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 CHI-SQUARE VALUE = 56.082 DEGREES OF FREEDOM = 68 PROBABILITY OF LARGER VALUE = 0.84807

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 21.904 DEGREES OF FREEDOM = 10 PROBABILITY OF LARGER VALUE = 0.01567

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	M(0)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
	0.13	0.00	0.10	0.12	1.00	0.75	0.49	0.26

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

Fig. 3.18. The seven tests of assumptions used in model selection applied to the simulated data from Model M<sub>1</sub> with true N = 150. The underlying data are exactly the same as those used for Figs. 3.3 and 3.4. (See Table 3.4 and the discussion of the simulation of Model M<sub>1</sub> for the capture probability parameters.)

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

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

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

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

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

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

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

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
CHI-SQUARE VALUE = 13.877 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.01641
  
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
CHI-SQUARE VALUE = 38.965 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00000
  
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
CHI-SQUARE VALUE = 47.445 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00000
  
4. GOODNESS OF FIT TEST OF MODEL M(H)  
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
CHI-SQUARE VALUE = 40.018 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00000  
  
TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE  
(FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)  
  
NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY  
-----  
1 17.176 6 0.00866  
2 12.000 6 0.06197  
3 23.194 6 0.00073
  
5. GOODNESS OF FIT TEST OF MODEL M(B)  
NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
CHI-SQUARE VALUE = 12.468 DEGREES OF FREEDOM = 10 PROBABILITY OF LARGER VALUE = 0.25497
  
- 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
CHI-SQUARE VALUE = 9.269 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.09879
  
- 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
CHI-SQUARE VALUE = 3.198 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.66943
  
6. GOODNESS OF FIT TEST OF MODEL M(T)  
NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
CHI-SQUARE VALUE = 23.572 DEGREES OF FREEDOM = 23 PROBABILITY OF LARGER VALUE = 0.42781
  
7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
CHI-SQUARE VALUE = 74.257 DEGREES OF FREEDOM = 19 PROBABILITY OF LARGER VALUE = 0.00000

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)

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

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)

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

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)

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

4. GOODNESS OF FIT TEST OF MODEL M(H)

NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)

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

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

NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	17.586	6	0.00735
2	5.219	6	0.51610
3	8.833	6	0.18317

5. GOODNESS OF FIT TEST OF MODEL M(B)

NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)

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

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

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

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

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

6. GOODNESS OF FIT TEST OF MODEL M(T)

NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)

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

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.

NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)

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

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)

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

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)

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

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)

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

4. GOODNESS OF FIT TEST OF MODEL M(H)

NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)

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

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

NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	4.040	6	0.67126
2	4.096	6	0.66373
3	6.200	6	0.40116
4	6.333	6	0.38690
5	5.400	6	0.49362

5. GOODNESS OF FIT TEST OF MODEL M(B)

NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)

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

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

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

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

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

6. GOODNESS OF FIT TEST OF MODEL M(T)

NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)

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

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.

NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)

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

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

Fig. 3.21. The seven tests of assumptions used in model selection applied to the simulated data from Model  $M_{1h}$  with true  $N = 200$ . The underlying data are exactly the same as those used for Figs. 3.9 and 3.10. (See Table 3.8 and the discussion of the simulation of Model  $M_{1h}$  for the capture probability parameters.)

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H).  
CHI-SQUARE VALUE = 11.852 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.01849
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B).  
CHI-SQUARE VALUE = 8.073 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00449
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T).  
CHI-SQUARE VALUE = 13.570 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.05938
4. GOODNESS OF FIT TEST OF MODEL M(H)  
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
CHI-SQUARE VALUE = 14.735 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.03956  
TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE  
(FREQUENCIES LESS THAN 21 ARE NOT CALCULATED.)  

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	9.459	7	0.22133
2	9.520	7	0.21745
3	17.609	7	0.01387
4	5.667	7	0.57917
5. GOODNESS OF FIT TEST OF MODEL M(B)  
NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
CHI-SQUARE VALUE = 8.698 DEGREES OF FREEDOM = 12 PROBABILITY OF LARGER VALUE = 0.72846
  - 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
CHI-SQUARE VALUE = 5.518 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.47924
  - 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
CHI-SQUARE VALUE = 3.180 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.78595
6. GOODNESS OF FIT TEST OF MODEL M(T)  
NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
CHI-SQUARE VALUE = 162.378 DEGREES OF FREEDOM = 125 PROBABILITY OF LARGER VALUE = 0.01251
7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
CHI-SQUARE VALUE = 23.398 DEGREES OF FREEDOM = 21 PROBABILITY OF LARGER VALUE = 0.32314

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)

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

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)

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

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.

NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)

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

4. GOODNESS OF FIT TEST OF MODEL M(H)

NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)

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

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

NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY

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

5. GOODNESS OF FIT TEST OF MODEL M(B)

NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)

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

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

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

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

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

6. GOODNESS OF FIT TEST OF MODEL M(T)

NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)

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

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.

NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)

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

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 0.605 DEGREES OF FREEDOM = 2 PROBABILITY OF LARGER VALUE = 0.73897

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 12.829 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00034

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 39.776 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 34.454 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

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

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	25.494	4	0.00004
2	9.591	4	0.04790

5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 29.390 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00005

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 19.987 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00017

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 9.403 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.02439

6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 CHI-SQUARE VALUE = 47.145 DEGREES OF FREEDOM = 49 PROBABILITY OF LARGER VALUE = 0.54860

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 29.927 DEGREES OF FREEDOM = 8 PROBABILITY OF LARGER VALUE = 0.00022

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

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

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

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 14.763 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00522

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 18.741 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00002

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 116.194 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.00000

4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 104.054 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.00000

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

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	16.333	5	0.00596
2	24.500	5	0.00018
3	50.794	5	0.00000
4	18.125	5	0.00280
5	18.647	5	0.00224

5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 93.254 DEGREES OF FREEDOM = 8 PROBABILITY OF LARGER VALUE = 0.00000

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 23.418 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00010

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 69.836 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 CHI-SQUARE VALUE = 176.095 DEGREES OF FREEDOM = 124 PROBABILITY OF LARGER VALUE = 0.00114

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 50.998 DEGREES OF FREEDOM = 15 PROBABILITY OF LARGER VALUE = 0.00001

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

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

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

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

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

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

## A Comprehensive Look at the Simulation Examples

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

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

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

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

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

True Model	True N	Estimator (Model) Used				
		Null (M <sub>0</sub> )	Darroch (M <sub>t</sub> )	Zippin (M <sub>b</sub> )	Jackknife (M <sub>h</sub> )	Generalized Removal
M <sub>0</sub>	50	49.1(1.1)	49.0(1.0)	48.5(1.1)	53.4(0.9)	49.7(1.8)
M <sub>t</sub>	150	147.1(1.7)	146.0(1.7)	155.8(2.4)	168.1(3.0)	136.7(2.7)
M <sub>b</sub> <sup>a/</sup>	100	86.9(1.6)	86.8(1.6)	102.4(3.9)	94.5(2.4)	105.0(4.6)
M <sub>b</sub> <sup>b/</sup>	100	114.2(1.2)	113.3(1.1)	99.2(0.8)	129.1(3.4)	100.0(1.2)
M <sub>h</sub>	200	180.4(2.3)	180.4(2.3)	182.2(2.5)	209.4(2.9)	182.3(2.5)
M <sub>bh</sub>	200	201.3(2.4)	200.7(2.4)	186.8(2.1)	239.4(5.0)	192.6(7.0)
M <sub>th</sub>	200	182.2(1.9)	178.8(1.9)	178.9(1.7)	216.8(3.2)	161.3(1.3)
M <sub>tb</sub>	150	212.8(7.3)	208.1(7.1)	160.5(5.3)	252.3(5.3)	131.9(2.4)
M <sub>tbh</sub>	200	176.2(1.0)	174.8(0.9)	190.3(2.0)	201.0(2.1)	181.9(4.4)

<sup>a/</sup> Trap-happy case.

<sup>b/</sup> Trap-shy case.

In terms of their robustness we rank the five estimators in the order  $\hat{N}_h$  (jackknife, Model M<sub>h</sub>);  $\hat{N}_{bh}$  (generalized removal, Model M<sub>bh</sub>);  $\hat{N}_b$  (Zippin, Model M<sub>b</sub>);  $\hat{N}_t$  (Darroch, Model M<sub>t</sub>); and  $\hat{N}_0$  (null, Model M<sub>0</sub>). Darroch's estimator is always valid when Model M<sub>0</sub> is true; moreover, very little precision is lost by using  $\hat{N}_t$  when Model M<sub>0</sub> is indicated. Also,  $\hat{N}_b$  is a special case of  $\hat{N}_{bh}$ . Thus, we can reduce the choices to three estimators:  $\hat{N}_h$ ,  $\hat{N}_{bh}$ , and  $\hat{N}_t$ . Certainly, when the selection procedure suggests Model M<sub>tb</sub>, M<sub>th</sub>, or M<sub>tbh</sub>, and estimation is necessary, only one of these three estimators should be considered for use. Of the three,  $\hat{N}_h$  will generally be the best choice, although the particular set of data, circumstances, and model selection criteria may cause one to select either  $\hat{N}_{bh}$  or  $\hat{N}_t$ .

**Model Selection and Estimation.** In practice, we attempt to select the correct model before estimating N. Thus, the real test of the methods we recommend first involves model selection, then estimation based on the most appropriate model and concurrent evaluation of whether that model fits the data. These results, for each simulated case, are shown in Table 3.15. For example, the model selection procedure correctly selected Model M<sub>0</sub> in 8 of the 10 simulation cases. In the other two cases, Model M<sub>h</sub> was used as the basis for estimation, even though M<sub>th</sub> was chosen once. Results of point estimation, confidence interval coverage, and model goodness of fit were all very good when Model M<sub>0</sub> was the true model. (A significance level of 5% was used to judge model fit.)

Results from Table 3.15 are also good for Models M<sub>t</sub>, M<sub>b</sub>, and M<sub>h</sub>. The model selection procedure in these examples generally led to use of the correct model. For Model M<sub>t</sub>, the correct model was selected only 7 (of 10) times, but when Model M<sub>th</sub> was selected, the selection criteria clearly indicated that Model M<sub>t</sub> (not M<sub>h</sub>) should be used for estimation.

Results for Model M<sub>bh</sub> are good, but not as good as for the four simple (one-factor) models (M<sub>0</sub>, M<sub>t</sub>, M<sub>b</sub>, and M<sub>h</sub>). In general, the results when Model M<sub>bh</sub> is true tend to underestimate N; the presence of heterogeneity "causes" this tendency toward a negative bias.

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

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

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

**TABLE 3.15.** (cont)

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

**TABLE 3.15. (cont)**

Model $M_b$ , (trap-happy case), $N = 100$					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	$M_b$	114	Yes	$M_b$	Yes
2	$M_b$	97	Yes	$M_b$	No
3	$M_b$	85	No	$M_b$	Yes
4	$M_b$	112	Yes	$M_b$	Yes
5	$M_b$	98	Yes	$M_b$	Yes
6	$M_b$	105	Yes	$M_b$	Yes
7	$M_b$	83	No	$M_b$	Yes
8	$M_b$	97	Yes	$M_b$	Yes
9	$M_b$	114	Yes	$M_b$	Yes
10	$M_b$	119	Yes	$M_b$	Yes
Mean		102.4			

**TABLE 3.15. (cont)**

Model $M_b$ , (trap-shy case), $N = 100$					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	$M_b$	99	Yes	$M_b$	Yes
2	$M_b$	101	Yes	$M_b$	Yes
3	$M_b$	99	Yes	$M_b$	Yes
4	$M_{tb}$	103	Yes	$M_{bh}$	Yes
5	$M_b$	96	Yes	$M_b$	Yes
6	$M_b$	99	Yes	$M_b$	Yes
7	$M_b$	98	Yes	$M_b$	Yes
8	$M_b$	96	Yes	$M_b$	Yes
9	$M_b$	99	Yes	$M_b$	Yes
10	$M_{bh}$	104	Yes	$M_{bh}$	Yes
Mean		99.4			

**TABLE 3.15.** (cont)

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

**TABLE 3.15.** (cont)

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

**TABLE 3.15. (cont)**

Model $M_{th}$ , N = 200					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	$M_t$	177	No	$M_t$	Yes
2	$M_t$	179	No	$M_t$	Yes
3	$M_t$	179	No	$M_t$	Yes
4	$M_{th}$	188	Yes	$M_t$	No
5	$M_{th}$	176	No	$M_t$	Yes
6	$M_{th}$	175	No	$M_t$	Yes
7	$M_t$	170	No	$M_t$	Yes
8	$M_t$	176	No	$M_t$	Yes
9	$M_{th}$	190	Yes	$M_t$	Yes
10	$M_t$	178	No	$M_t$	Yes
Mean		178.8			

**TABLE 3.15. (cont)**

Model $M_{tb}$ , N = 150					
Repetition	Model Selected	$\hat{N}$	Coverage	Model Used For Estimation	Goodness Of Fit
1	$M_t$	207	No	$M_t$	Yes
2	$M_t$	208	No	$M_t$	Yes
3	$M_{tb}$	192	No	$M_t$	Yes
4	$M_{tb}$	221	No	$M_t$	Yes
5	$M_{tb}$	145	Yes	$M_b$	Yes
6	$M_{th}$	207	No	$M_t$	Yes
7	$M_{tb}$	149	Yes	$M_b$	No
8	$M_{th}$	182	Yes	$M_t$	Yes
9	$M_{tb}$	236	No	$M_t$	Yes
10	$M_{th}$	182	No	$M_t$	Yes
Mean		192.8			

**TABLE 3.15.** (cont)

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

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

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

### Testing for Closure

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

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

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

OVERALL TEST RESULTS --  
 Z-VALUE -0.328  
 PROBABILITY OF A SMALLER VALUE 0.37145

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

NUMBER OF CAPTURES	Z-VALUE	PROBABILITY
2	0.285	0.61214

Fig. 3.26. Results of the closure test for the first simulation case of Model  $M_0$ . Under the null hypothesis of either Model  $M_0$  or  $M_h$  (only) and closure, the computed z-value has a standard normal distribution. The program gives the observed significance level of the test,  $P = 0.37145$  in this example.

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

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

NUMBER OF CAPTURES	Z-VALUE	PROBABILITY
2	-1.494	0.06758
3	-1.988	0.02341
4	-1.246	0.10647
5	-1.215	0.11216

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

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

In all, for examples in this chapter, we simulated 10 repetitions of 9 models (two cases of Model  $M_h$ ). For Models  $M_0$  and  $M_h$ , the closure test did not reject the model even once at the 5% significance level. This result is not strange, given that only 20 tests were made and that the test is valid for these two models. But for the 70 remaining simulations, the closure test rejected the model 23 times at the 5% significance level; of course, closure was really true in all these cases. The number of rejections corresponds to a 33% rejection rate, when it should be 5%, and illustrates that this closure test is invalid whenever time or behavior affects capture probabilities. We emphasize that this problem is fundamental; no valid statistical test of closure can be constructed on the basis of only the capture-recapture data.

## Summary

1. There are three critical considerations in constructing capture-recapture models: what population size  $N$  means (this relates to geographic closure), whether the model should be demographically closed or open, and how to model capture probabilities.
2. Ball and urn models have motivated most of the thinking about capture-recapture models. However, in real populations, capture probabilities vary and there is no analogy to the sides of the urn.
3. We have not dealt with the case of  $t = 2$  because we cannot test any assumptions in this case, and it is covered adequately in the literature.
4. Models are based on the concept of capture probabilities:  $p_{ij}$  = the probability of capturing the  $i^{\text{th}}$  animal (in the population at risk of capture) on the  $j^{\text{th}}$  sampling occasion.
5. Three factors can affect capture probabilities: time effects, behavioral response to capture, and innate heterogeneity (that is, variations among individuals in capture probabilities). On the basis of the three factors, eight different basic models for closed-population capture-recapture studies are presented (see

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

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

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

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

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

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

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

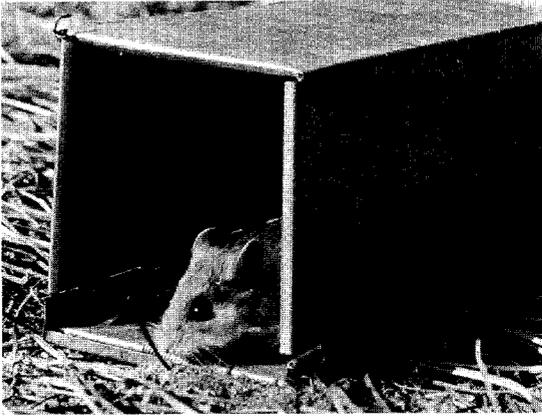
12. The full-blown procedure of doing all seven tests, and then selecting the model leads to a robust estimation procedure for good capture-recapture data on closed populations.

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

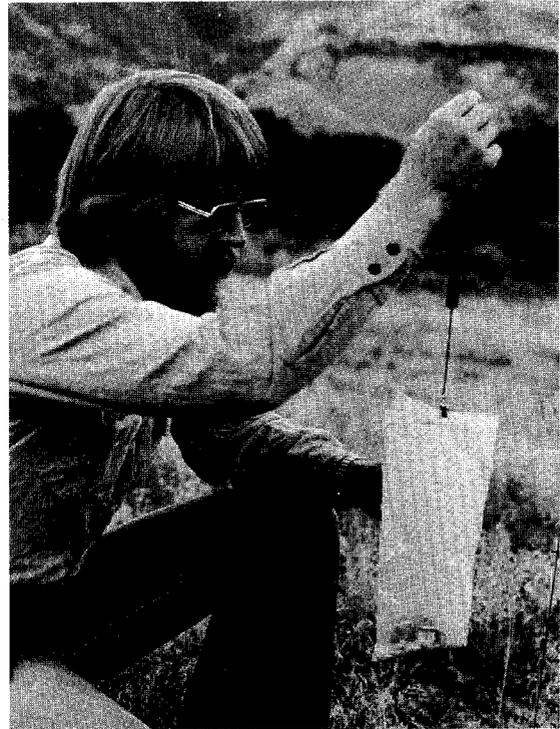
### Questions and Exercises

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

9. If you trap within the same fixed area with 100 traps on days 1, 2, and 3, then with 150 traps on days 4, 5, and 6, and finally with 200 traps on days 7, 8, and 9, can Model  $M_0$  fit the data? Can Model  $M_h$ ,  $M_b$ , or  $M_t$  fit the data?
10. The capture probability  $p$  (say, in Model  $M_0$ ) is which of the following?
  - a. The probability that a trap will catch an animal,
  - b. The probability that an individual animal will be caught on a given trapping occasion,
  - c. The probability that an individual animal will be caught at least once during the study.
11. Will ML estimators of  $N$  be developed for Models  $M_{th}$ ,  $M_{tb}$ , and  $M_{tth}$ ?
12. Is Model  $M_0$  robust in trap-shy populations?
13. If the model selection criteria are the same for Models  $M_0$  and  $M_h$ , why do we recommend selecting Model  $M_h$  as the basis for estimating  $N$ ?
14. Is it just a matter of time until a general, completely valid test for closure is developed?



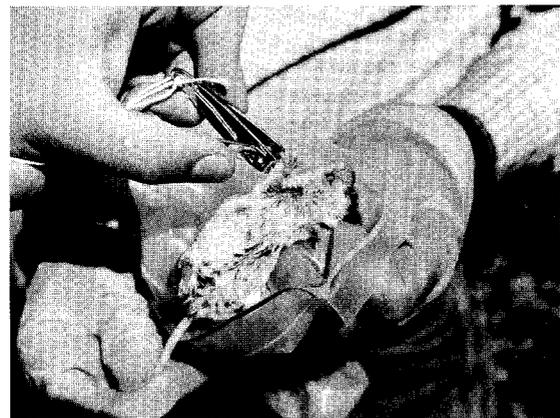
This deer mouse (*Peromyscus maniculatus*) is about to be captured in a Sherman live trap. (Photograph by Gary White.)



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



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



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

## CHAPTER 4

# REMOVAL METHODS

When our goal is to estimate population size, it is logical to want to do a one-time only, 100% count of the population. In theory, a 100% count can be achieved by a removal study; that is, as each individual is caught it is removed from the population, either permanently or for the duration of the study. Thus, additional captures must represent different individuals. By extending the capture effort until no more individuals are caught, we should obtain a 100% count. Unfortunately, there are practical problems with this approach, just as there are with the capture-recapture method. In this chapter we discuss the modeling, data analysis, and advantages and disadvantages of the constant-effort removal method.

Electrofishing in ponds, small lakes, or small streams with a flow of only 0.1 to 1 cubic meters per second is the most common example of the removal method. For an introduction to the literature on this subject see *Friedman (1974)*, and *Seber (1982)*. Usually, the fish are not killed; rather, they are removed when they are caught, held until the end of the several removal passes, and then returned to the pond, lake, or stream. One removal occasion usually consists of a complete pass, going first upstream and then downstream in the study area. At least two such passes should be made, and we maintain that three or more are necessary unless the efficiency of the gear is very high—that is, unless the capture probability is 0.8 or more on each pass. Equal effort is required on each pass for the models presented here. In the optimal situation, 100% of the fish are removed in the first pass; the second pass is made to verify that all fish have been counted. In practice, capture probabilities as high as 0.8 are uncommon (although this may be a reflection of the electrofishing gear in use), and significant numbers of fish will be caught on the second and subsequent passes.

Removal sampling of small mammals is usually accomplished with snap traps. However, simply removing live-trapped animals from the area would suffice. In fact, neither removal from the area nor kill trapping is necessary to apply removal analysis methods to trapping data for small mammals. Initially, all animals in the population at risk of capture are unmarked, or their marks are known to be from an earlier trapping effort. Animals can be captured, marked, and released, and we can think of them as removed from the unmarked population segment; *Seber (1973:323)* has referred to this approach as “removal by marking.” Thus, by using only first captures in a live-trapping study, we can analyze the data as a removal study.

The removal method circumvents any behavioral response to trapping during the study. (We define behavioral response as a change in capture probability after, and as a result of, first capture.) In fact, as mentioned in Chapter 3, the data from capture-recapture models  $M_b$  and  $M_{bh}$  are properly analyzed as removal studies because of this dominant behavior effect. If killing the animal is tolerable, a removal study on small mammals is quicker and cheaper than a live-trapping study. No tags are needed, kill traps are cheaper, handling the animals is easier, and on successive occasions fewer animals are caught so less field time is required.

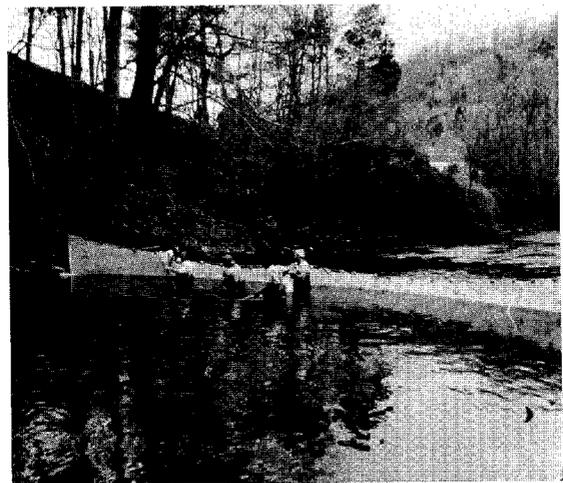
The immediate and major disadvantage of physically removing the animals is that geographic closure is usually violated—after a few days, individuals originally too far from the grid to have been caught will start moving into the area vacated by the removals. Also, there is the obvious disruption of the local population, so the method is not suitable for studies to be repeated in the near future.



Electrofishing in a Virginia stream. Notice the generator in the boat and the lead to an electrode. (Photograph courtesy of R. F. Raleigh.)



Working the two electrodes during electrofishing in a Virginia stream. (Photograph courtesy of R. F. Raleigh.)



A net is used to close off the lower end of the sampled stream during electrofishing in a Virginia stream. (Photograph courtesy of R. F. Raleigh.)

## The Removal Model, Assuming Constant Capture Probability

**The Model and Its Assumptions.** In the simplest removal model, every individual has an equal and constant probability of capture on all removal, or capture, occasions. This is the model most often encountered in the literature (*Hayne 1949a; Moran 1951; Zippin 1956, 1958; Seber and White 1970; Otis et al. 1978:28-32, 107-108; and Seber 1982*). To illustrate the model, we consider a population of size  $N = 625$ , with every individual having an independent, unchanging probability of capture of  $p = 0.4$ . On the first capture occasion, the expected number of captures is 250 ( $= 625 \times 0.4$ ), which leaves 60% of the population (375) uncaught. On the second occasion, the expected number of captures is 40% of the remaining 375 animals, or 150 ( $= 375 \times 0.4$ ). The expected (average) data from four capture occasions shown in Table 4.1 continue the argument. Even after four occasions, we can expect an average of 81 individuals to remain uncaught.

This model has only two parameters: population size  $N$  and capture probability  $p$ . The only data from removal sampling are the numbers of captures on each of  $t$  capture occasions; we let  $u_1, \dots, u_t$  represent these data. If  $E(u_j)$ ,  $j = 1, \dots, t$  represents the expected (average) number of captures on occasion  $j$ , the constant-effort removal model for  $E(u_j)$  is

$$E(u_j) = N(1 - p)^{j-1}p, \quad j = 1, \dots, t.$$

Thus,  $E(u_1) = Np$  for occasion 1;  $E(u_2) = N(1 - p)p$  for occasion 2; and  $E(u_t) = N(1 - p)^{t-1}p$  for the final occasion. Our understanding of this model can be tested by computing the elements of Table 4.1. For example,  $E(u_4) = 625(1 - 0.4)^3(0.4) = 625 \times 0.216 \times 0.4 = 54$ .



P. A. P. Moran

The contributions made to removal sampling and analysis methods by P. A. P. Moran were somewhat fortuitous. After World War II, he worked at the Institute of Statistics at Oxford. The Institute was housed in the same quarters as the Bureau of Animal Population under Charles Elton. Moran, through friendships with both Elton and P. H. Leslie, became interested in animal populations and estimation in removal experiments.

Moran was born in 1917 and educated at the Universities of Sydney and Cambridge. His interests over the past 30 years have turned in other directions. He is currently in the Statistics Department at the Australian National University, Canberra, Australian Capital Territory. (Recent photograph.)

**TABLE 4.1.** Expected (average) results of a removal study on a population of size  $N = 625$  with equal, independent capture probability of  $p = 0.4$  for all individuals and  $t = 4$  removal occasions.

Removal Occasion, $j$	Population Size At Start Of Occasion	Expected Number Captured, $E(j)$	Number Not Yet Caught
1	625	250	375
2	375	150	225
3	225	90	135
4	135	54	81

The assumption of equal effort is necessary if the capture probability is to be the same on every capture occasion. In electrofishing studies, the equal effort requirement demands the use of a standardized technique on each occasion. It is especially important to prevent a letdown of effort after the first pass. In trapping small mammals, the same number of operating traps is required on each occasion, and the same length of time must intervene between setting the traps and checking them on each occasion.

The usual assumption for removal studies does not allow heterogeneity of capture probabilities. That is, there can be no innate differences in capture probabilities. However, the assumption fails to some extent in electrofishing because capture probabilities are related to the size of the fish. In trapping small mammals, capture probabilities may vary by species, age, or sex of animals. In both types of studies, a solution to the problem is to separate the data by size-class of fish (within broad limits) or by species, age, or sex of animal and to analyze each data subset separately. An example showing a partitioning of data by the sex of the animal appears in Chapter 6.

**Some Simulated Data.** In real applications, the data are subject to sampling variations; therefore, the numbers of animals captured will not decrease as smoothly as those shown in Table 4.1. An actual application is more likely to produce results like those shown in Table 4.2, where simulated data are given, based on the same model as Table 4.1 ( $N = 625$  and  $p = 0.4$ ). The numbers caught are not calculated in a deterministic manner; rather, they are the realization of a random process. For example, on the first occasion the number caught is a random (binomial) variable with an expected value of 250 ( $= N \times p$ ) and a sampling variance of 150 [ $N \times p \times (1 - p)$ ]. Hence, the captures on the first occasion have a sampling standard deviation of 12.2 ( $= \sqrt{150}$ ). Thus, for reasons rooted in mathematical statistics, we expect the number of individuals caught on the first occasion to fall between about 225 and

**TABLE 4.2.** A simulated instance of removal data on a population of  $N = 625$  with equal, independent capture probability of  $p = 0.4$  for all individuals and  $t = 4$  removal occasions.

Removal Occasion, $j$	Population Size At Start Of Occasion	Number Captured, $u_j$	Number Not Yet Caught
1	625	260	365
2	365	141	224
3	224	97	127
4	127	50	77

275. With a more detailed analysis of this statistical model, we can predict the properties of the data for any values of  $N$ ,  $p$ , and  $t$ .

**Estimation for the Case  $t = 2$ .** Given removal data like those shown in Table 4.2, we need an estimator of  $N$ . An acceptable, simple, closed-form estimator is available only for data from two ( $t = 2$ ) occasions; it is not, however, the maximum likelihood (ML) estimator. For three or more occasions, efficient estimation becomes more complicated. We recommend using the exact ML estimator of  $N$  for any value of  $t$ . We also strongly recommend at least three sampling occasions, because the model's basic assumption (equal capture probability on every occasion) cannot be tested if  $t = 2$ . Of course, if capture probabilities are around 0.95, or even 0.9, two occasions will suffice, but such high capture probabilities are rarely achieved.

Because the ML estimator does not exist in closed form, numerical iterative techniques are required to compute it, even for  $t = 2$ . Program CAPTURE is available to perform these computations; we recommend that biologists use it to obtain the exact ML estimates for use with removal data.

For  $t = 2$ , a good approximation to the exact ML estimator is

$$\hat{N} = \frac{u_1}{1 - (u_2/u_1)} \quad (4.1)$$

Using the "data" of Table 4.1 to illustrate this formula, we have  $u_1 = 250$ ,  $u_2 = 150$ , and hence

$$\begin{aligned} \hat{N} &= \frac{250}{1 - (150/250)} \\ &= \frac{250}{1 - 0.6} \\ &= \frac{250}{0.4} \\ &= 625 \end{aligned}$$

In this example we obtain the exact result of 625 because the data are the expected (average) numbers of removals under the model with  $N = 625$  and  $p = 0.4$ . That is, they are not real data, subject to random variation. Using the simulated data of Table 4.2, we have  $u_1 = 260$ ,  $u_2 = 141$ , and hence

$$\begin{aligned} \hat{N} &= \frac{260}{1 - (141/260)} \\ &= \frac{260}{1 - 0.542} \\ &= \frac{260}{0.458} \\ &= 568 \end{aligned}$$

A formula for estimating the sampling variance of the estimator in Eq. (4.1) is given in *Otis et al. (1978:108)*. This estimate is computed automatically by program CAPTURE for  $t \geq 2$ . For the above example, the estimated sampling variance of  $\hat{N}$  is 2687.2 [=  $\hat{\text{var}}(\hat{N})$ ]. The estimated standard error of  $\hat{N}$  is thus  $\hat{\text{se}}(\hat{N}) = 51.8$  ( $= \sqrt{2687.2}$ ). Finally, for this example, an approximate 95% confidence interval on the true population size  $N$  is  $568 \pm 1.96 \times 51.8$ , or 466 to 670. The parameter  $p$  also has a simple estimator for the case  $t = 2$ . In fact,  $\hat{p} = 1 - (u_2/u_1)$ ; thus, in the above example,  $\hat{p} = 0.458$ .

Figure 4.1 shows part of the output from program CAPTURE, giving the ML estimate for the same example with  $t = 2$ . In the figure, the ML estimates are  $\hat{p} = 0.463605$  and  $\hat{N} = 562$ . (Two significant digits suffice in reporting such results.) These results are only slightly different from those of the closed-form formulas (Eq. 4.1). In the same figure, the estimated sampling standard error of the ML estimate of  $N$  is 50.0; in comparison, the estimated sampling error of the estimate of  $N$  derived from Eq. (4.1) is 51.8. The same formula is used to estimate the standard errors of both the  $\hat{N}$  given by Eq. (4.1) and the ML estimate of  $N$  computed by CAPTURE. These sampling variance estimates are different only because the estimates of  $N$  and  $p$  differ slightly when the exact ML estimate is used instead of the closed-form formula.

We do not deal with the instance of only two removals in greater detail because we recommend using three or more capture occasions for two reasons: (1) three or more occasions are required to test the assumption of constant capture probability (that is, to perform a goodness of fit test of the model), and (2) a large gain in precision is achieved by making the extra effort of one or two additional passes. The case of two removals has been studied in considerable detail (see for example, *Seber and Whale 1970; Seber 1982*).

### Estimation and Goodness of Fit Testing for Three or More Capture Occasions

We can recommend no closed-form estimators of  $N$  and  $p$  for  $t \geq 3$ . Figure 4.2 presents output from CAPTURE giving the ML estimates for the complete removal data of Table 4.2; that is,  $t = 4$  rather than  $t = 2$ . In Fig. 4.2,  $\hat{p} = 0.413609$ ,  $\hat{N} = 621$ , and the estimated standard error for  $\hat{N}$  is 16.9401. Contrast these results with those for  $t = 2$  in Fig. 4.1, especially the substantial improvement in precision when  $t = 4$ . From Fig. 4.1, we see that  $\hat{se}(\hat{N}) = 49.95$  for  $t = 2$ , whereas from Fig. 4.2, we find that  $\hat{se}(\hat{N}) = 16.94$  for  $t = 4$ . This change is a threefold improvement in precision of the estimate of  $N$  and is apparent in the point estimates of  $N$ . The true value of  $N$  is 625. For  $t = 2$ ,  $\hat{N} = 562$  and the error is  $-63$ , but for  $t = 4$ ,  $\hat{N} = 621$  and the error is only  $-4$ . In both examples, the approximate 95% confidence interval includes the true value of 625. The gain in precision with four capture occasions is substantial when capture probability is less than (about) 0.7, as this example illustrates.

If three or more capture occasions are used, there is a goodness of fit test for this constant capture probability model. It is a chi-square test of form  $(\text{Observed} - \text{Expected})^2 / \text{Expected}$ , and Program CAPTURE computes this test. For the full data of Table 4.2, the test statistic value is 1.568 (based on 2 df); if the model is true, the probability of a value this large or larger is 0.4567. We conclude that the constant capture probability model fits the data of Table 4.2.

```

OCCASION      J=      1      2
TOTAL CAUGHT  M(J)=    0 260 401
NEWLY CAUGHT  U(J)=   260 141

ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.463605

POPULATION ESTIMATE IS      562 WITH STANDARD ERROR 49.9536
APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS      464 TO      660

OCCASION      J=      1      2      3      4
TOTAL CAUGHT  M(J)=    0 260 401 498 548
NEWLY CAUGHT  U(J)=   260 141  97  50

ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.413609

POPULATION ESTIMATE IS      621 WITH STANDARD ERROR 16.9401
APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS      587 TO      655

```

Fig. 4.1. Exact ML estimate of  $N$  as computed by CAPTURE by using the first two capture occasions in Table 4.2. The data are simulated from a true population of size  $N = 625$  and  $p = 0.4$ .

Fig. 4.2. Exact ML estimate of  $N$  as computed by using all four capture occasions in Table 4.2. The data are simulated from a true population of size  $N = 625$  and  $p = 0.4$ .

Given the output of Fig. 4.2, we can compute the goodness of fit test on a hand calculator; we illustrate the computation as follows. First, we compute the estimated expected number of captures on each occasion; the formula is

$$\hat{E}(u_1) = \hat{N}\hat{p}$$

and

$$\hat{E}(u_j) = \hat{N}(1 - \hat{p})^{j-1}\hat{p}, \quad \text{for } j > 1,$$

where  $\hat{N}$  and  $\hat{p}$  are the ML estimates of  $N$  and  $p$ . From Fig. 4.2,  $\hat{N} = 621$  and  $\hat{p} = 0.414$ . (Three-digit accuracy is sufficient for these computations.) For this example,

$$\hat{E}(u_1) = 621 \times 0.414 = 257.1,$$

$$\hat{E}(u_2) = 621 \times (1 - 0.414) \times 0.414 = 150.7,$$

$$\hat{E}(u_3) = 621 \times (1 - 0.414)^2 \times 0.414 = 88.3,$$

and

$$\hat{E}(u_4) = 621 \times (1 - 0.414)^3 \times 0.414 = 51.7,$$

The quantity  $[u_j - \hat{E}(u_j)]^2/\hat{E}(u_j)$  is then computed for each capture occasion ( $j = 1, 2, 3, 4$  in this example). For example, for  $j = 1$  we have  $(260 - 257.1)^2/257.1 = 0.0327$ . For  $j = 2, 3,$  and  $4$  the corresponding results are 0.6244, 0.8572, and 0.0559. The sum of these four values is the goodness of fit test statistic. For this example, the sum is 1.570. The test statistic is printed by CAPTURE, along with other information, as part of the "generalized removal method" discussed below, for reasons related to the overall structure and logic of program CAPTURE.

If the constant capture probability model fits the data perfectly, the test statistic value will be zero. The poorer the fit of the model to the data, the larger the test statistic value. To judge the fit or lack of fit, we need to know the probability that the test statistic value will be as large as, or larger than, the observed value (such as 1.570) given that the constant capture probability model is true. This observed "significance level" is computed by program CAPTURE (see Fig. 2.9). In the example given above from Fig. 4.2, the probability of an observed test statistic value exceeding 1.570 is 0.4567. Thus, this test statistic is not unusually large, and we conclude that the model adequately fits the data. For philosophical reasons, we do not conclude that the model is a true representation of the real world underlying this study, only that the model fits the data.

### Another Simulation Example

To illustrate the sampling variability of data and parameter estimates under even this simple, constant capture probability model, we generated 10 independent replications for the situation  $N = 200$ ,  $p = 0.20$ , and  $t = 5$ . Table 4.3 shows the actual captures, the resultant ML estimates of  $N$  and  $p$ , and the (approximate) 95% confidence interval limits.

Table 4.3 shows that the estimates of  $N$  are quite variable (as are the actual removals from replicate to replicate):  $\hat{N}$  ranges from 169 to 249. The average  $\hat{N}$  from these 10 replications is 206, which is close to the true value of  $N = 200$ . Of the 10 confidence intervals, only the one for replicate 9 does not cover the true value of  $N$ . These results illustrate the variable nature of the data and estimators, even when the underlying model is known to be true.

**TABLE 4.3.** Results of 10 independent replications of the constant capture probability removal model with  $N = 200$ ,  $p = 0.20$ , and  $t = 5$ . Notice that  $\hat{N}$  varies from 169 to 249; this sampling variation is to be expected under this model. Notice also that all the confidence intervals except replicate 9 cover the true value of  $N$ .

Replicate	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$\hat{N}(\hat{se}(\hat{N}))$	$\hat{p}$	95% Confidence Interval Limits	
1	37	36	31	15	18	200(30.7)	0.205	139	261
2	38	29	31	22	20	249(61.1)	0.152	129	369
3	43	33	30	20	18	209(30.6)	0.208	149	269
4	33	29	34	18	16	215(46.2)	0.169	124	306
5	48	34	25	21	22	219(31.9)	0.206	156	282
6	45	36	28	19	16	191(20.2)	0.243	149	233
7	37	25	32	23	14	209(41.4)	0.178	127	291
8	42	25	31	19	20	220(42.9)	0.177	135	305
9	55	30	29	15	15	179(12.6)	0.300	154	204
10	39	26	27	14	16	169(23.0)	0.225	123	215
Averages	41.7	30.3	29.8	18.6	17.5	206.0	0.206	137.9	272.9

In practice there is only one repetition. It is important to recognize that innate sampling variability will occur and that the results of any study are just one of many possible outcomes. Thus the computed estimate  $\hat{N}$  cannot be expected to equal the true  $N$ . The sampling variability of  $\hat{N}$  must be recognized and estimated. Moreover, it is also important to present a confidence interval on  $N$  that will express a likely range of values for the true population size.

There is another point to be made. In this example, we used a capture probability of 0.2. This value is at the low end of the scale of acceptable capture probabilities required to get precise and unbiased population estimates from removal studies. The larger the capture probability, the more precise will be  $\hat{N}$ ; that is, the lower will be the sampling variability. We recommend  $p \geq 0.4$  for really good results in the typical removal study, having  $N$  of a few hundred and  $t = 3$  to 6.

### Failure of Removal Experiments

The basis for estimating population size in removal experiments is that the population size is reduced significantly on each sampling occasion. The number of animals caught should tend to diminish:  $u_1 > u_2 > u_3$ , and so on. If no depletion is realized, the experiment is said to "fail"; that is, unless the number caught tends to decrease, the model parameters cannot be estimated. *Seber and White (1970)* showed that  $N$  and  $p$  can be estimated from data when the following "failure criterion" is satisfied:

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

If the criterion is just barely satisfied, the point estimators and their sampling variance will have very poor properties. Failure of removal experiments can be avoided by conducting the sampling so that  $p$  is as large as possible.

## Generalized Removal Method

**The Basic Idea.** The constant capture probability model does not always fit removal data. An example from *Otis et al. (1978:46-48)* illustrates the problem. The removal data, given in Fig. 4.3, are from *Andrezejewski and Jezierski (1966)*. They were estimating population densities of the European hare by using drive trapping. From Fig. 4.3 we see that  $u_1 = 722$ , then  $u_2$  drops off sharply to 191. After that, the drop-off in removals is much less severe. The pattern suggests that the constant capture probability model does not fit these data. The chi-square goodness of fit test confirms this suspicion; from Fig. 4.3 the goodness of fit test value for the constant capture probability model is 13.150, with a significance level of 0.0014. This significance level means that there is only a probability of 0.0014 that a chi-square test statistic value this large, or larger, would result if in fact the constant capture probability model were the true model for these data. Therefore, we are justified in rejecting this model and concluding that it is not valid for these data.

Having rejected the constant capture probability model as providing a good fit, we must either base the estimate of population size on a different model or accept the estimate from a model that does not fit. Sometimes the point estimate of  $N$  is useful, even when the model does not fit, but we should never feel comfortable in those circumstances. The sampling variance will be poor and usually will be underestimated.

If the constant capture probability model does not fit, we must generalize it—that is, make it more flexible by allowing some degree of unequal capture probabilities. The most general removal model would allow a different (average) capture probability to apply to the uncaught animals still in the population on each capture occasion. Such a model would have  $t + 1$  parameters:  $N$  = population size and  $p_j$  = capture probability applicable to the  $N - (u_1 + \dots + u_{j-1})$  remaining individuals at the start of occasion  $j$  for  $j = 1$  to  $t$ . However, with only  $t$  pieces of data ( $u_1, \dots, u_t$ ) we cannot estimate  $t + 1$  parameters. This model is too general to be of any use. Some reduction in the number of parameters is necessary—for example, by assuming a relation between the capture probabilities  $p_1, \dots, p_t$ . The constant capture probability model assumes  $p_1 = p_2 = \dots = p_t = p$ .

The two possible generic sources of variation in the capture probabilities  $p_1, \dots, p_t$  are variation associated with time (or capture occasion) and variation intrinsic in the capture probabilities of individuals. The latter source we call “heterogeneity” (Chapter 1). Both types have been discussed with reference to capture-recapture studies (Chapter 3). Note that in a removal study, there can be no behavioral variation in capture probabilities of the type defined and discussed in Chapter 3. If the capture probabilities vary because of heterogeneity, a generalized removal model is possible; however, we know of no way to deal with time variation in a removal study. Therefore, we must conduct the study so as to minimize any time variation in capture probabilities. In electrofishing studies, a standardized method

OCCASION	J=	1	2	3	4	
TOTAL CAUGHT	M(J)=	0	722	913	982	1018
NEWLY CAUGHT	U(J)=	722	191	69	36	

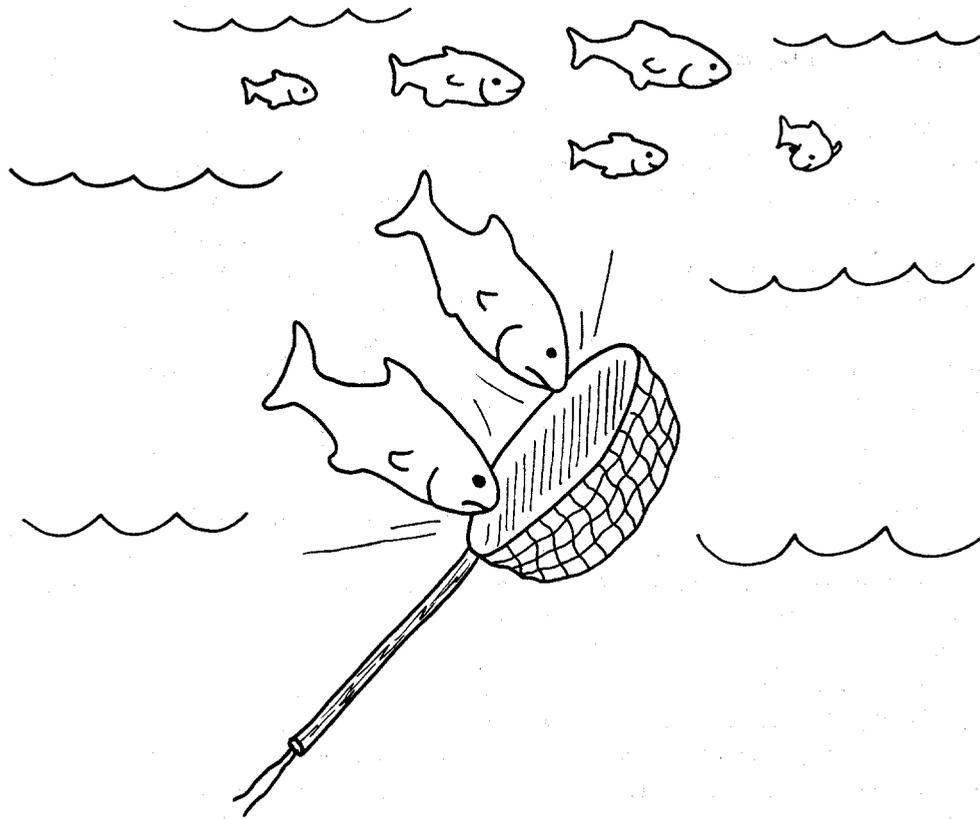
  

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 4			
1	1028.21	3.789703	13.150	0.0014	0.6806	0.6806	0.6806	0.6806
2	1039.10	7.658751	1.528	0.2164	0.6948	0.5916	0.5916	0.5916

POPULATION ESTIMATE IS	1039	WITH STANDARD ERROR	7.6588
APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL	1023	TO	1055

Fig. 4.3. Removal data on hares from Andrezejewski and Jezierski (1966). The data are from an actual field study with four removal occasions; true  $N$  is not known. The generalized removal estimator has been applied. For  $k = 1$  removal occasions, the model is the constant capture probability model. In the generalized removal method,  $k$  represents the number of different capture probability parameters in the model.



In electrofishing, larger fish have greater capture probabilities.

must be used on each occasion. In small-mammal trapping, the method must be standardized and trapping must be done during constant environmental conditions.

Intrinsic variation in capture probabilities has been demonstrated for electrofishing by *Cross and Scott (1975)* and *Bohlin and Sundstrom (1977)*. This means that capture probabilities vary even within size or age classes of fish. Intrinsic variation also is well documented in trapping studies of small mammals.

As an extreme, hypothetical example of heterogeneity, imagine a population of 200 in which 50 individuals have a capture probability of 1.0, and 150 individuals have a capture probability of 0.5. On the first occasion, all 50 individuals with a capture probability of 1.0 will be caught, that is, removed. Also, on the average, 75 of the 150 individuals with a capture probability of 0.5 will be caught, leaving 75 individuals each with a capture probability of 0.5. Thus, on and after occasion 2, the constant capture probability model becomes valid. But on the first occasion the average capture probability for the population is 0.625  $[= (1 \times 50 + 0.5 \times 150)/200]$ . If the constant capture probability model is applied to removal data from such a population,  $N$  will be underestimated. (Numerous empirical studies of the removal method on ponds and small lakes have shown that  $N$  usually is underestimated.)

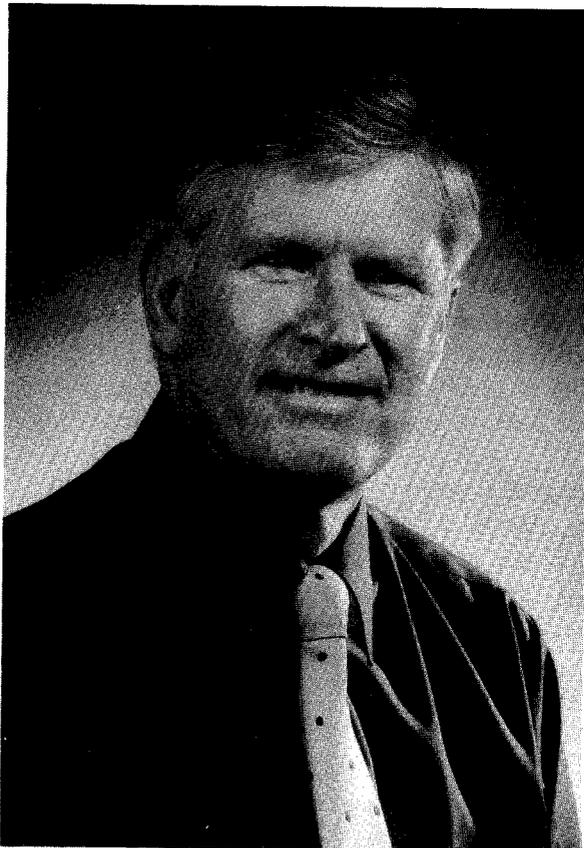
If heterogeneity is present, there is a tendency for the individuals with higher capture probabilities to be caught before the individuals with lower capture probabilities. This tendency causes a decrease in successive capture probabilities. Thus, we expect that  $p_1 > p_2 > p_3 > \dots > p_t$ . In the above extreme example,  $p_1 = 0.6250$ , but  $p_2 = p_3 = \dots = p_t = 0.5$ . A less extreme example is a population size of 200, composed of four 50-member subpopulations, with the respective capture probabilities 0.8, 0.6, 0.4, and 0.2. For five removal occasions, we would have  $p_1 = 0.5000$ ,  $p_2 = 0.4000$ ,  $p_3 = 0.3333$ ,  $p_4 = 0.2920$ , and  $p_5 = 0.2655$ . This example illustrates the expected sequence of removal probabilities applicable to each occasion under the assumption of heterogeneity within the population. The  $p_i$ 's decrease, but tend to level off. Thus, although the difference between  $p_1$  and  $p_2$  is 0.1, the difference between  $p_4$  and  $p_5$  is only 0.0265. In the absence of specific information about the values of these capture probabilities, a logical

modeling approach is to allow the first few  $p_j$ 's to differ, but thereafter to assume that  $p_{j+1} = p_{j+2} = \dots = p_t$ . In so doing, we seek only an adequate fit to the data, not a perfect model, and we make no assumptions about how the  $p_j$ 's decrease.

**Examples of the Generalized Removal Model.** *Otis et al. (1978:44-50)* presented a generalized removal method that implements the ideas presented here (see also *Skalski and Robson 1979*). The method involves fitting successively more general removal models, starting with the constant capture probability model, until an adequate fit to the data is found. (The method has some similarities to step-wise regression.) The second model of the sequence allows the capture probability for the first capture occasion to differ from the capture probabilities on the second and subsequent occasions. Thus, the parameters for the second model are  $N$ ,  $p_1$ , and  $p$  ( $= p_2 = \dots = p_t$ ). Specifically, this model has two capture probability parameters.

The results of the second model are presented in Fig. 4.3 in the  $k = 2$  row, and the results for the constant capture probability model are presented in the  $k = 1$  row. (The value of  $k$  is the number of different capture probability parameters in the model.) For  $k = 1$ , the constant capture probability model does not fit ( $P = 0.0014$ ), but for  $k = 2$  the chi-square goodness of fit statistic is 1.528, with  $P = 0.2164$ . We judge that the second model fits these data adequately, and hence we use it as the basis for estimating  $N$ . Consequently, we take  $\hat{N} = 1039$ , with a standard error of 7.66 and a confidence interval of 1023 to 1055.

The estimator of  $N$  used with this generalized removal model ( $k = 2$  in the above example) is the ML estimator. Details of its theory and computation are found in *Otis et al. (1978:40-43,45-50, 112-114)*.



Douglas S. Robson

Douglas Robson has been interested in capture-recapture and removal sampling since the early 1950s and has published more than 20 papers on many aspects in this area. Born in North Dakota, he took a B.S. degree in statistics from Iowa State University, received M.S. and Ph.D. degrees from Cornell University in statistics, and has had professional experience at the University of Washington, Princeton University, and Colorado State University. He has been Professor of Biological Statistics at Cornell since 1963.

Robson's contributions to capture-recapture have ranged from fairly theoretical to very applied work, for example, sample size estimation. His most recent interests lie in the estimation of survival rates in open-population models and in removal experiments. Robson certainly has been one of the leading contributors to the development of testing and estimation techniques in capture-type studies. He has worked with a host of other people contributing to the general theory, including C. Brownie, D. G. Chapman, R. M. Cormack, G. J. Paulik, K. H. Pollock, and H. A. Reiger. (Recent photograph.)



Larger mammals can be captured in wire live traps such as the one shown. (Photograph courtesy of Raymond Greenwood and Alan Sargeant.)



Two types of traps can be used at the same station to catch small mouse-sized mammals (left trap), and larger chipmunk-sized mammals (right trap). (Photograph courtesy of Robert Streeter.)

The basic assumption behind this generalized removal model is that of heterogeneity of capture probabilities in the sampled population. This should result in  $p_2$  being less than  $p_1$ . Program CAPTURE prints the ML estimator of the distinct capture probabilities in each generalized removal model. From Fig. 4.3 for row  $k = 1$ , there is only one distinct capture probability, thus  $\hat{p}_1 = \hat{p}_2 = \hat{p}_3 = \hat{p}_4 = \hat{p} = 0.6806$ . However, for the next more general model (on row  $k = 2$ ),  $\hat{p}_1 = 0.6948$  and  $\hat{p}_2 = \hat{p}_3 = \hat{p}_4 = \hat{p} = 0.5916$ . Thus, after the first capture occasion the estimated average capture probability of hares still uncaptured is 0.6, rather than 0.7. We interpret this change to be a result of an innate heterogeneity existing in the population and the tendency for more catchable animals to be caught first.

To demonstrate the generalized removal method further, we present more real data, this time from the use of removal sampling to estimate stream bottom benthos (see *Carle 1976*; *Carle and Strube 1978*). Figure 4.4 presents the numbers of Ephemerellidae found in five removals from one site on the Cache la Poudre River in Colorado;\* the field methodology was essentially that of *Carle (1976)*. Just a look at the removal data (310, 26, 14, 7, 6) shows that the constant capture probability model does not fit. Indeed, the goodness of fit test to that model has a chi-square test statistic value of 77.118; taken to four decimal places, the probability of a value this large, if the constant probability model is true, is 0.0000. In Fig. 4.4, rows  $k = 2$  and  $k = 3$  show the results of fitting the two- and three-parameter generalized removal models. The two point estimators of  $N$  hardly differ because apparently almost all individuals were caught.

From Fig. 4.4, we see that the model with two capture parameters ( $k = 2$ ) provides an adequate fit to these data. For that model,  $\hat{N} = 368$ , with a standard error of 4.2. Perhaps more interesting are the estimates of capture probabilities. For the two-parameter model,  $\hat{p}_1 = 0.8425$  and the subsequent estimated capture probabilities are  $\hat{p} = 0.4460$ . The drop in capture probabilities is dramatic. It is consistent with a model in which the capture probabilities are fairly high for most individuals, but are moderate to low for a few individuals.

The generalized removal model with three parameters also fits the data of Fig. 4.4 (row  $k = 3$ ), as it should if the less general model with two parameters fits. In this model, the three capture probabilities are  $p_1, p_2,$  and  $p_3 = p_4 = p_5 = p$ . We see that  $\hat{p}_1 = 0.8418$  is almost the same as  $\hat{p}_1 = 0.8425$  for the two-parameter model. Also,  $\hat{p}_2 = 0.4464$  and  $\hat{p}_3 = \hat{p}_4 = \hat{p}_5 = \hat{p} = 0.4375$  are only slightly different from the

\*This information was provided by R. F. Raleigh, US Fish and Wildlife Service, Ft. Collins, CO 80526.

OCCASION	J=	1	2	3	4	5					
TOTAL CAUGHT	M(J)=	0	310	336	350	357	363				
NEWLY CAUGHT	U(J)=	310	26	14	7	6					
K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 5						
1	363.00	0.4118716	77.118	0.0000	0.7857	0.7857	0.7857	0.7857	0.7857	0.7857	
2	367.96	4.208916	0.703	0.7035	0.8425	0.4460	0.4460	0.4460	0.4460	0.4460	
3	368.24	6.008948	0.640	0.4237	0.8418	0.4464	0.4375	0.4375	0.4375	0.4375	
POPULATION ESTIMATE IS			368	WITH STANDARD ERROR	4.2089						
APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL					359	TO	377				

Fig. 4.4. Removal data on stream bottom benthos, Ephemerelellidae, from one site on the Cache la Poudre River, Colorado. The circular depletion sampler described in Carle (1976) and Carle and Strube (1978) was used to obtain these data.

OCCASION	J=	1	2				
TOTAL CAUGHT	M(J)=	0	310	336			
NEWLY CAUGHT	U(J)=	310	26				
ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.919746							
POPULATION ESTIMATE IS			338	WITH STANDARD ERROR	1.7377		
APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS					334	TO	342

Fig. 4.5. Application of the removal estimator to only the first two occasions of the stream bottom benthos data shown in Fig. 4.4. Of necessity, the constant capture probability model is used here. The results are misleading; compare them with those in Fig. 4.4 for the entire data set ( $t = 5$ ).

two-parameter model values. Again, the result is consistent with the adequacy of the two-parameter model for these data.

What if only two capture occasions had been used for the example in Fig. 4.4? The results would be those shown in Fig. 4.5. First, because no goodness of fit test is possible, we have no clue that the constant capture probability model does not fit. The estimated capture probability of 0.92 is not only grossly in error, but is also misleading, because it suggests that virtually all individuals have been caught after two removals. The estimate of  $N$  is 338, with an estimated approximate 95% confidence interval of 336 to 342. Remember that if the computed lower limit is less than the total number of individuals seen, the lower limit should be replaced by the number of individuals actually seen. In fact, at least 363 organisms were present; thus the estimate of 338 is clearly too small. More than two removals are required to detect such problems as these.

Such comparisons are also interesting for the data on hares presented in Fig. 4.3. Figure 4.6 presents the results of analyzing only the first two removals of the *Andrzejewski and Jezierski (1966)* data. Again, there is no clue that the point estimate of  $\hat{N} = 980$  is biased low. The computed confidence interval has an upper limit of 1009, which is less than the 1010 hares removed on four occasions.

Finally, we return to the simulated data first given in Table 4.2. Applying the generalized removal method to these data, we get the results shown in Fig. 4.7. As discussed above, this goodness of fit test shows that the model fits, so  $\hat{N}$  is taken as 621. Note that for the model in which  $p_1$  is allowed to differ from  $p_2 = p_3 = p_4 = p$ , we get  $\hat{p}_1 = 0.4132$  and  $\hat{p} = 0.3952$ , which is to be expected because all  $p_j$  are the same (0.4).

**An Example Using Fish Removal Data.** The constant capture probability model does not fit all fish removal data. Figure 4.8 is an example of data from a major study of the removal method applied to fish in small streams (*Mahon 1980*). Mahon conducted more than a dozen removal experiments; he obtained well over 100 data sets, with the data partitioned by fish species and size. After each experiment, consisting of 4 to 8 removal occasions, he collected the fish remaining in the stream segment by using rotenone. Thus, he knew the true population sizes. We illustrate the generalized removal method for fish

OCCASION            J=     1     2  
TOTAL CAUGHT    M(J)=   0 722 913  
NEWLY CAUGHT    U(J)=   722 191

ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.737208

POPULATION ESTIMATE IS            980 WITH STANDARD ERROR    14.6068

APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS            951 TO            1009

Fig. 4.6. Application of the removal estimator to only the first two removal occasions of the Andrezejewski and Jezierski (1966) study shown in Fig. 4.3. Of necessity, the constant capture probability model is used here. The results are again misleading; compare them with those in Fig. 4.3.

OCCASION	J=	1	2	3	4	
TOTAL CAUGHT	M(J)=	0	260	401	498	548
NEWLY CAUGHT	U(J)=	260	141	97	50	

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 4			
1	620.98	16.94006	1.568	0.4567	0.4136	0.4136	0.4136	0.4136
2	629.28	26.25738	1.246	0.2643	0.4132	0.3952	0.3952	0.3952

POPULATION ESTIMATE IS            621 WITH STANDARD ERROR    16.9401

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL            587 TO            655

Fig. 4.7. Results of applying the generalized removal method to the simulated data given in Table 4.2. With four removal occasions, the first two removal models are fitted to the data;  $k = 1$  is the constant capture probability model. For  $k = 2$ ,  $p_1$  is allowed to be different from  $p_2 = p_3 = p_4 = p$ . Because the constant capture probability model fits the data, the results,  $\hat{p}_1 = 0.4132$  and  $\hat{p} = 0.3952$  for  $k = 2$  are very close both to each other and to the true value,  $p = 0.4$ . Compare these results with those in Fig. 4.2.

with data from one of Mahon's study stations. The data are for fantail darter (*Etheostoma flabellare*) longer than 35 mm. The true population size for individuals of this species and size was  $N = 1151$ .

Figure 4.8 shows the removals for the seven sampling occasions ( $t = 7$ ) used at this station. The 666 fantail darters caught were only 58% of the true population. We see from the figure that the constant capture probability model produced an estimate of  $\hat{N} = 900(\pm 48)$ . The chi-square goodness of fit test statistic for this model was 9.071, which has an observed significance level of  $P = 0.106$ . We reject a removal model, for lack of fit, if this observed significance level is less than 0.20 because we believe it is better to sacrifice some precision to minimize bias. This example illustrates such a tradeoff. The generalized removal model with  $k = 2$  produces  $\hat{N} = 1025$ , with an estimated standard error of 105. The (approximate) 95% confidence interval on  $N$  is 819 to 1231; it covers  $N (= 1151)$ . By contrast, for  $k = 1$  (the constant capture probability model), the interval of 805 to 995 does not come close to covering the true value of  $N$ .

## Regression, or Catch Per Unit Effort Methods

The wildlife and fisheries literature presents several methods for estimation of  $N$  based on the constant capture probability removal model. These methods are of two types: (1) regression methods or (2) the maximum likelihood method and its modifications or approximations (see Zippin 1956; Carle and Strube 1978). The regression methods often are referred to as methods based on "catch per unit effort" (CPUE). This acronym derives from the fisheries literature, especially concerning commercial fisheries, where fishing provides catch statistics but where units of fishing effort vary in each time period. In wildlife studies, effort usually can be nearly equal on each capture occasion. In removal trapping, keeping the same number of traps working on each occasion is considered to provide equal effort. In electrofishing, standardization of the technique provides equal effort. We do not deal here with CPUE methods for varying effort on each occasion, but both Ricker (1975) and Seber (1982) give an introduction to the methodology.

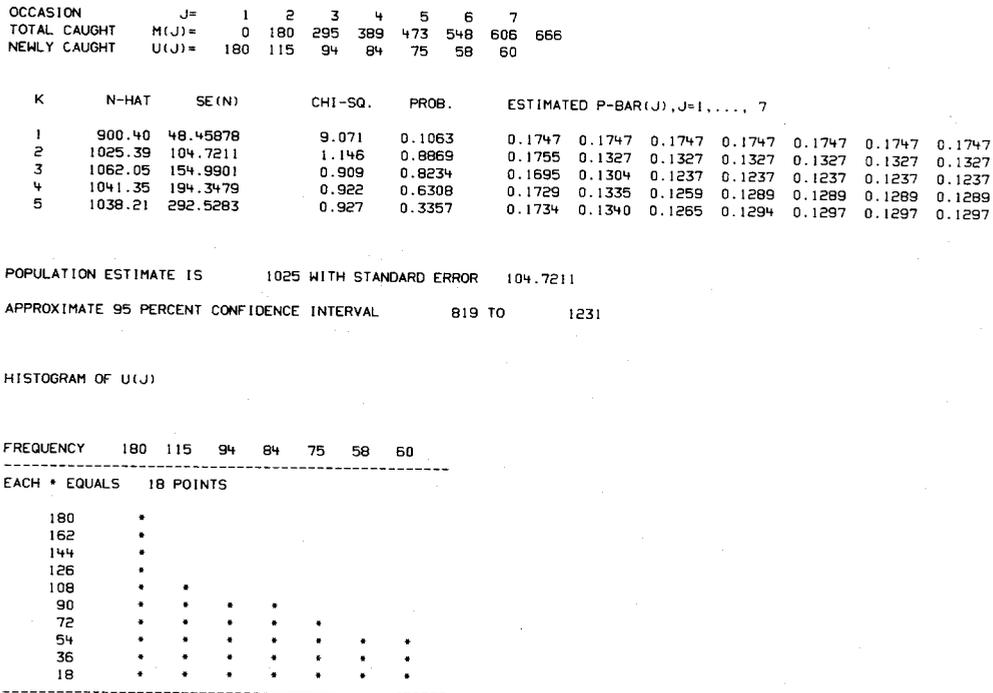


Fig. 4.8. Results of applying the generalized removal method to data on fantail darters longer than 35 mm (from Mahon 1980). The true value of  $N = 1151$  was known because after the seven removals, rotenone was used to collect the remaining fish.

Under the model of constant capture probability, all regression-based CPUE methods are inferior to the ML estimation method. We will discuss one regression-based method primarily because it appears so frequently in the literature. First, however, we must define another symbol in addition to  $u_j$ :  $M_j$  = the number of individuals marked and hence removed from the population before the  $j^{\text{th}}$  capture occasion. Thus  $M_1 = 0$ , because before the first capture occasion no individuals have been marked and removed from the population. Before the second capture occasion  $u_1$  individuals have been marked and removed, hence  $M_2 = u_1$ . After the second capture occasion,  $u_1 + u_2$  individuals have been marked and removed, so that  $M_3 = u_1 + u_2$ . In general,  $M_j = u_1 + \dots + u_{j-1}$ . Also defined (but not used in the regression method) is  $M_{t+1} = u_1 + u_2 + \dots + u_t$ , which is the total number of individuals marked and removed and hence known to be in the population. Any admissible estimator of  $N$  must be at least as large as  $M_{t+1}$ .

Under the constant capture probability model, a linear structural relation between the two variables  $E(u_j)$  and  $E(M_j)$  is, specifically,

$$E(u_j) = pN - pE(M_j) \quad (4.2)$$

or, equivalently,

$$E(M_j) = N - \left( \frac{1}{p} \right) E(u_j) \quad (4.3)$$

Based on Eq. (4.2), Hayne (1949a) published the following regression method of estimating  $N$ : regress  $u_j$  on  $M_j$ ,  $j = 1, \dots, t$ , to get an estimate of  $p$  (as the slope of the regression equation) and an estimate of  $-pN$  (the intercept of the regression equation). Symbolically, the fitted regression is

$$u_j = \hat{a} + \hat{b} M_j, \quad j = 1, \dots, t, \quad (4.4)$$

where  $\hat{a} = p\hat{N}$  and  $\hat{b} = -\hat{p}$ , so that

$$\begin{aligned} \frac{\hat{a}}{-\hat{b}} &= \frac{p\hat{N}}{-(-\hat{p})} \\ &= \hat{N}. \end{aligned} \tag{4.5}$$

Basing the regression on Eq. (4.3) rather than on Eq. (4.4) has statistical (theoretical) advantages, but commonly it has been based on the form of Eqs. (4.2) and (4.4) in conjunction with a plot of the data as  $u_j$  (on the y-axis) versus  $M_j$  (on the x-axis). In such a plot, the point where the line intercepts the x-axis is the estimate of  $N$ , the mathematical equivalent to Eq. (4.5).

From Fig. 4.2 we have the pair of variables  $u_j$  and  $M_j$ , as

$j$	$u_j$	$M_j$
1	260	0
2	141	260
3	97	401
4	50	498

A plot of these data and the results of applying the simple linear regression of  $u_j$  on  $M_j$ , using Eq. (4.4), are shown in Fig. 4.9;  $\hat{a}$  and  $\hat{b}$  are explained above, and  $r$  is the correlation coefficient of  $u_j$  and  $M_j$ . From this regression, the estimate of  $p$  is  $\hat{p} = -\hat{b} = 0.41488$  and the estimate of  $N$  is  $\hat{N} = (257.2/0.42488) = 620$ . Note also that  $M = 620$  is the intercept of the line with the x-axis.

These point estimates of  $N$  and  $p$  compare well with the ML estimates in Fig. 4.2. However, it is very difficult to obtain a valid estimate of the sampling variance of this regression estimator of  $N$ . Even if we could obtain a valid variance estimate, statistical theory assures us that the ML estimate of  $N$  has a smaller sampling variance under the assumed model and hence is the better estimator.

The usual model for simple linear regression (for a sample size of  $n$  pairs of variables  $y_j, x_j$ ) is, symbolically,

$$y_j = a + bx_j + \varepsilon_j, \quad j = 1, \dots, n,$$

under the assumption that the  $n$  pairs of variables  $(y_j, x_j)$  are independent of each other. An alternative expression of this is that the "errors"  $\varepsilon_1, \dots, \varepsilon_n$  are assumed to be independent. Moreover, it is assumed that the sampling variance of  $\varepsilon_j$  is a constant, say  $\sigma^2$ , for all  $j = 1, \dots, n$ .

These usual assumptions of simple linear regression are violated severely by the application in Eq. (4.4). Specifically, the different pairs of the variables  $(u_j, M_j)$  are highly correlated, so that the assumption of independence of the errors is violated. Also, the sampling variances of the errors are not equal; instead, they are proportional to  $u_j$ . Consequently, the usual formulas for the sampling variances of  $\hat{a}$  and  $\hat{b}$  are invalid.

For the simulated data that fit the constant capture probability model, the correlation of  $M_j$  and  $u_j$  is  $r = -0.99761$ . Applied to the hare data of *Andrezejewski and Jezierski (1966)*, this same regression method produces  $r = -0.99926$ , which is a stronger correlation (Fig. 4.10). Yet we know from the results in Fig. 4.3 that the hare data do not fit the constant capture probability model. The disparity demonstrates that the correlation coefficient is useless as a measure of how well this model fits the data. Similarly, visual examination of the plotted data, as in Figs. 4.9 and 4.10, does not serve as a goodness of fit test of the model to the data. Only the chi-square goodness of fit test provides an adequate test of the model fit.

The regression method can, and often does, produce estimates of  $N$  below the total numbers of animals caught. For example, when we applied the regression method to the hare data of Fig. 4.3 (Fig. 4.10), we

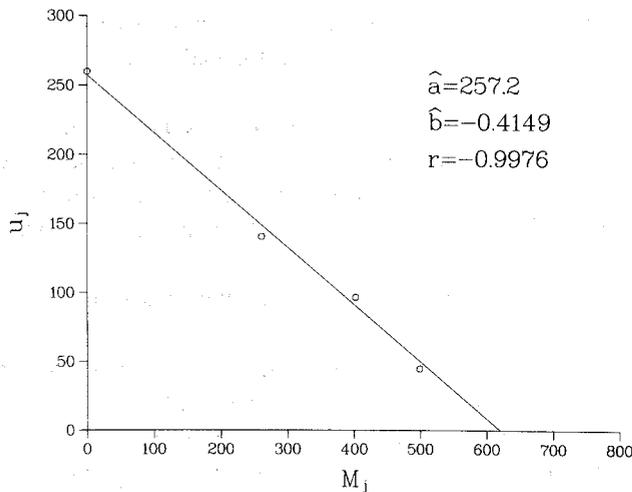


Fig. 4.9. A plot of the  $u_j, M_j$  data shown in Table 4.2 and used in Fig. 4.2. The fitted regression line, Eq. (4.4), also is shown. Notice that the intercept of the line and the x-axis is  $\hat{N} = 620$ .

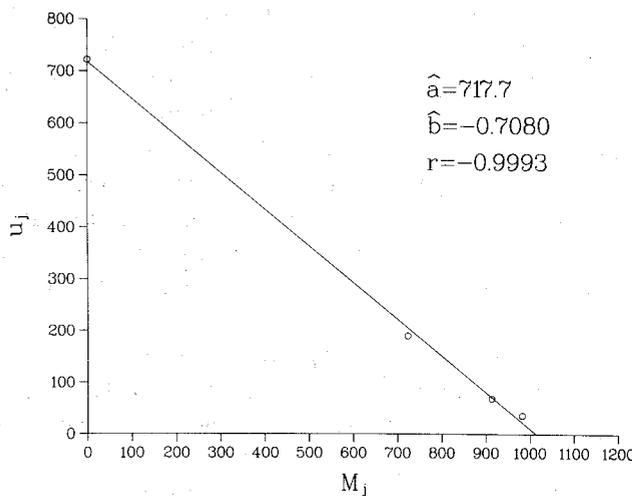


Fig. 4.10. The regression estimator applied to the hare data of Andrezejewski and Jeziewski (1966). The correlation is very high,  $r = -0.99926$ , and the plotted data apparently are linear, yet these removal data do not fit the constant capture probability model. (See Fig. 4.3.)

obtained  $\hat{N} = 717.7/0.70830 = 1014$ , even though the total number of hares caught was 1018. Of course, such an estimate is ridiculous. The problem is not hypothetical. *Andrezejewski and Jeziewski (1966)* used exactly this regression method and reported in their paper the point estimate of 1010. They got this number rather than 1014 by rounding the slope estimate to 0.71, computing  $\hat{N} = 717.7/0.71$ , and truncating the result to 1010. Their variance estimate is also quite invalid; they estimated  $\hat{se}(\hat{N}) = 38$ .

In contrast, the generalized removal method applied to the hare data of Fig. 4.3 produced  $\hat{N} = 1039$  with  $\hat{se}(\hat{N}) = 7.7$ . Thus, an approximate 95% confidence interval  $N$  is 1023 to 1055.

## Summary

1. In true removal sampling, each animal is caught only once. The most common applications are snap trapping of small mammals and electrofishing in small streams. In snap trapping the animals are killed. In electrofishing the captured fish usually are held in buckets until the removal passes have been completed, and then the fish are returned to the stream.

2. The assumption of closure is critical to the results of a removal study. If closure is violated, the estimate of (original) population size will be biased—often badly so. If closure fails drastically, the study

becomes a total waste of time and money. If animals leave the study area between capture occasions, the population size  $N$  will be underestimated. If they enter the area,  $N$  will be overestimated—possibly badly so.

3. Removal studies require a minimum of two capture occasions, and even two occasions do not allow a test of the constant capture probability assumption. If all animals have an average capture probability of at least 0.8, two occasions will suffice because failure of constant capture probability will not matter.

4. The common model for analysis of removal data assumes an equal and constant capture probability ( $p$ ) for all individuals.

5. If the assumption of equal and constant capture probability is to be tested, there must be at least three capture occasions.

6. To obtain useful results with the removal method, capture probabilities should be at least 0.2, and if they are as low as 0.2, even six capture occasions will give only marginally reliable results.

7. The removal method is not suitable for populations exceeding a few thousand individuals.

8. The removal method will fail to give reliable results unless there is an obvious decrease in number of animals caught over the  $t$  capture occasions.

9. We recommend use of the ML estimator of  $N$ , although it does not exist in closed form. Its sampling variance is smaller than those of other estimators under the constant  $p$  or the generalized removal model. In particular, we note that the ML estimator of  $N$  will never be less than  $M_{t+1}$  (the number of individuals removed). All commonly used regression methods are based on the same assumptions as the ML estimator, and their estimators of  $N$  can be less than  $M_{t+1}$  even when the assumptions are true.

10. The assumption of constant capture probability can be relaxed to allow for some degree of heterogeneity. The assumption of heterogeneity leads to the generalized removal method, which fits successively more general models to the data until an acceptable fit is found. For example, the first more general model allows one capture probability,  $p_1$ , on occasion 1 and different (but equal) capture probabilities on occasions 2, 3, . . . ,  $t$ . If this generalized model fits the data, it often does so with just two parameters.

11. Common sources of differences in capture probabilities (heterogeneity) are differences by species, sex, or age. In electrofishing, fish size (within wide limits) influences capture probability—smaller fish are less catchable. If the data are sufficient, we recommend partitioning all data by species and sex, and electrofishing data by two or three broad categories of size.

12. Although some degree of heterogeneity can be allowed for, time variation in capture probabilities totally destroys the usefulness of the removal method. No estimator is possible for a removal study with time variation in the capture probabilities, unless the number of capture occasions is extended until all animals are caught.

13. The estimator of  $N$  under the generalized removal is the ML method. It does not exist in closed form.

### Questions and Exercises

1. Would you accept a removal study with only one ( $t = 1$ ) capture occasion?
2. What are the potential defects of removal studies with  $t = 2$  occasions?
3. Can you have an open-model removal study?
4. Can you use removal estimation methods on live-trapping data?
5. In general, do capture-recapture studies yield more information about  $N$  than removal studies?
6. Would you accept, as a point estimate of  $N$ , a value less than  $M_{t+1}$ , the total number of animals removed?
7. Why is kill trapping over a long time period likely to give a poor estimate of  $N$ ?
8. A biologist, asked to estimate the population size of the endangered lion snail, suggests a removal study over five occasions. Assess his proposal.
9. Are the following removal data useful?  $u_1 = 100$ ,  $u_2 = 99$ ,  $u_3 = 105$ ,  $u_4 = 98$ , and  $u_5 = 95$ , with  $t = 5$ .

10. A snap trap study of small mammals conducted for 4 days gives  $u_1 = 68$ ,  $u_2 = 41$ ,  $u_3 = 25$ , and  $u_4 = 15$ .
- Describe the X matrix.
  - Was the study acceptable; that is, do you expect a reasonable estimate of N to result?
  - Give a rough estimate of N, then if possible, use program CAPTURE to obtain the exact ML estimate.
11. If true population size is  $N = 100$  and a three-occasion removal study has the constant capture probability  $p = 0.05$ , will the results be useful? Compute the expected numbers removed on the study's three occasions.
12. In studies of small populations (N of at most 200), what capture probability should you have to achieve reliable results if  $t = 4$ ?
13. You conduct a removal study, with carefully controlled equal effort, and find  $u_1 = 20$ ,  $u_2 = 15$ ,  $u_3 = 8$ ,  $u_4 = 17$ ,  $u_5 = 29$ , and  $u_6 = 36$ . What is happening?
14. You assess your experimental situation and conclude that closure will be violated significantly; you decide to try a removal study anyway. What can be said about  $\hat{N}$  as an estimate of N?
15. Let  $\bar{p}_j$  be the average capture probability on removal occasion j. What does  $\bar{p}_j$  mean? Why do the  $\bar{p}_j$  decrease under model  $M_{bh}$ ?
16. What sources of heterogeneity are likely to be present in electrofishing? In kill trapping?
17. What happens in a removal study if equal effort fails to be true?

# CHAPTER 5

## DENSITY ESTIMATION

Previous chapters in this primer have dealt only with the problem of estimating population size  $N$ . The only assumption made about the area to which  $N$  applies is that its size is finite.

Density  $D$  is defined as the number of animals per unit of area. Density estimation extends population size estimation to include an estimate of the area to which the population estimate  $\hat{N}$  relates. To estimate density, the assumption of demographic closure still must be met. That is,  $N$  must be reasonably constant over the time interval in which capture-recapture data are collected. Thus, animals cannot be immigrating to or emigrating from the area, and no death or recruitment is allowed. These assumptions concerning closure have been made throughout this primer. In the present discussion, we also assume that traps have been placed in a square or rectangular grid design.

Density estimation is not as simple as dividing  $\hat{N}$  by the area ( $A$ ) of the trapping grid, because of the phenomenon known as "edge effect." For example, animals at the edge of a grid will not spend all of their time on the grid, because the grid area contains only part of their home range. Thus, the effective area trapped is somewhat larger than the grid area (Fig. 5.1), because the area to which  $\hat{N}$  applies includes the entire home ranges of such animals. Hence, the naive density estimate  $\hat{D} = \hat{N}/A$  will tend to overestimate  $D$ , because  $A$  does not include the additional area around the grid boundary, called the boundary strip. In Fig. 5.2, the grid is very large relative to the home range, and thus the area of the edge effect is almost negligible relative to the size of the grid. In contrast, Fig. 5.3 represents the opposite situation: the grid is very small relative to the home range. In this situation, a valid density estimate probably cannot be obtained.

Biologists have been aware of this potential source of bias for many decades (*Dice 1938, 1941; Stickel 1954*) and have proposed several approaches for handling the problem. Perhaps the most frequently used approach is based on Dice's suggestion that the width of the boundary strip ( $W$ ) be taken as one-half the average diameter of the home range of the species. Suggestions for estimating this diameter from the trapping data have been made by several authors (*Hayne 1949b; Stickel 1954; Tanaka 1972; Otis et al. 1978:72-73*), but the approaches are all subject to difficulties. For example, the estimates depend on trap spacing and numbers of recaptures. A second approach involves the use of assessment lines (*M. H. Smith et al. 1971, 1975; H. D. Smith et al. 1972; Swift and Steinhorst 1976; O'Farrell et al. 1977*). Although this approach has produced good results, the method can become quite complex and is heavily dependent on the design of the trap layout. Our discussion here focuses on a third approach, which depends on jointly estimating density and boundary strip width by using data from selected subgrids (*Otis et al. 1978:67-74*).

### Theory

The key concept in this approach to estimating the boundary strip  $W$  of a trapping grid is that the importance of the boundary strip decreases as the size of the grid increases. To illustrate, consider a square with each side 5 units long. If a strip 1 unit wide is added all around (with corners rounded to quarter circles), the area is increased by 93%. However, if the grid is 50 units on each side, the increase in area owing to the addition of a 1-unit strip is only 8% (Fig. 5.4). Hence, the naive estimate of density for a species whose home range radius is on the order of 1 unit would have a large bias for the 5 by 5 grid, but the same estimate would have a small bias for a 50 by 50 grid.



Michael Smith

Michael Smith received his Ph.D. degree from the University of Florida in 1966. He worked on the population biology of mice in old fields, in the sandy soils of the southeastern United States, where the mice can be dug out of burrows easily. This technique allowed independent estimates of population characteristics in addition to those derived from classic trapping approaches. The realization that strong biases are inherent in trapping data has influenced his approach to study of the population biology of vertebrates since that time.

Smith joined the staff of the Savannah River Ecology Laboratory at the University of Georgia in 1966 and shortly thereafter became involved in the International Biological Program (IBP). Estimates of density, not just population numbers, were a critical part of this program because calculations of energy flow and elemental cycling require density values. Most of his work for the IBP, conducted in conjunction with John B. Gentry, involved the use of techniques to estimate the size of the sampling area. Assessment lines and "radioactive trap nights" were investigated in efforts to determine the best methods for obtaining density estimates and confidence intervals for small mammal populations.

Smith's current interests lie with the genetic regulation of population processes. Animal population density is one of the most important variables in this field, as it is in most fields that have occupied his interests. Smith is now Director of the Savannah River Ecology Laboratory. (Recent photograph.)

Program CAPTURE estimates  $W$  and  $D$  from a single trapping grid by constructing a set of nested grids (Fig. 5.5). The naive estimate of density for each grid is

$$\hat{Y}_i = \hat{N}_i / A_i, \quad i = 1, \dots, k;$$

that is, the population estimate for the  $i^{\text{th}}$  grid,  $\hat{N}_i$ , divided by the area of the  $i^{\text{th}}$  grid,  $A_i$ . Clearly,  $\hat{Y}_1$ , which corresponds to the innermost grid, has the most bias, and  $\hat{Y}_k$  probably has the least, because of the decrease in importance of strip width as grid size increases.

To demonstrate this phenomenon, we will consider the (unpublished) data of Burnham and Cushwa,\* discussed in *Otis et al. (1978:36-37)*. Showshoe hares (*Lepus americanus*) were trapped in a black spruce forest 48 km north of Fairbanks, Alaska. A 10 by 10 live-trapping grid was used, with traps spaced 200 ft (61 m) apart. Because the traps were not baited for the first 3 days, only the data from the last 6 days appear here. The model selection procedure chose Model  $M_h$  as the appropriate model, so the jackknife estimator is used.

The area  $A_1$  of the inner, 4 by 4 grid is 360 000 ft<sup>2</sup>, and the population estimate is 31 animals. Thus, the naive estimate of density  $\hat{Y}_1$  is 3.8 hares per acre. Similarly, for the next larger, 6 by 6 grid, the population estimate is 44 hares and the naive estimate of density  $\hat{Y}_2$  is 1.9 hares per acre. The naive estimates of density for the 8 by 8 and 10 by 10 grids are  $\hat{Y}_3 = 1.6$  and  $\hat{Y}_4 = 1.2$  hares per acre, respectively. The estimates trend downward (from 3.8 to 1.2) because the additional strip width around the grid becomes less and less important as the grid size increases.

\*K. P. Burnham and C. Cushwa, US Fish and Wildlife Service.

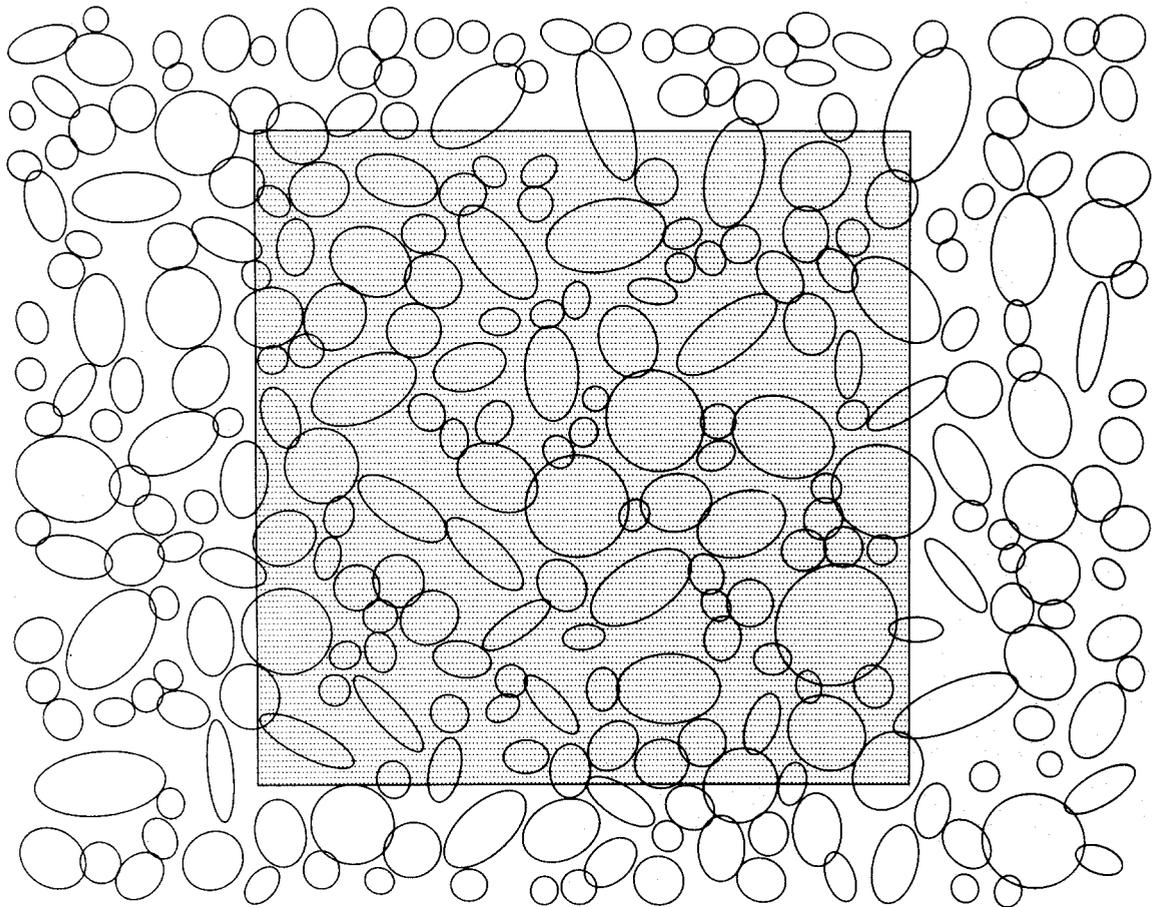


Fig. 5.1. The trapping grid (shaded area) includes all, part, or none of the home ranges (ellipses) of the animals in the vicinity of the grid. Some of the animals whose home ranges overlap the grid will be captured, and thus the effective area of the trapping grid is larger than its physical size, necessitating a technique to estimate the effective grid area and density at the same time. Note that some of the home ranges do not overlap the trapping grid, and hence the animals belonging to these ranges are not captured.

The statistical treatment of the estimation method is given in *Otis et al. (1978:69-72, 121-122)*. Briefly, it is assumed that some fixed but unknown strip width  $W$  surrounds each of the  $k$  subgrids (Fig. 5.6), and an equation is derived that relates the naive density estimates  $\hat{Y}_i$  to  $D$  and  $W$ . A complicated statistical approach known as the generalized nonlinear least-squares method uses this basic relation to derive the estimates. The estimates  $\hat{D}$  and  $\hat{W}$  are derived so that the sum of squared differences between the naive estimates  $Y_i$  and the predicted values is minimized.

### Example

This illustration of the density estimation procedure is based on a set of capture data\* collected on Richardson's ground squirrel (*Spermophilus richardsoni*) inhabiting rangeland near Kremmling, Colorado. A 10 by 10 grid with 10-m trap spacing was established in June and trapped for 6 consecutive days. Figures 5.7a and b give the numbers of captures at each trap during this period and the results of three simple chi-square tests, each of which tests a hypothesis concerning the uniformity of ground

\*The data, unpublished, are from K. Fagerstone, US Fish and Wildlife Service.

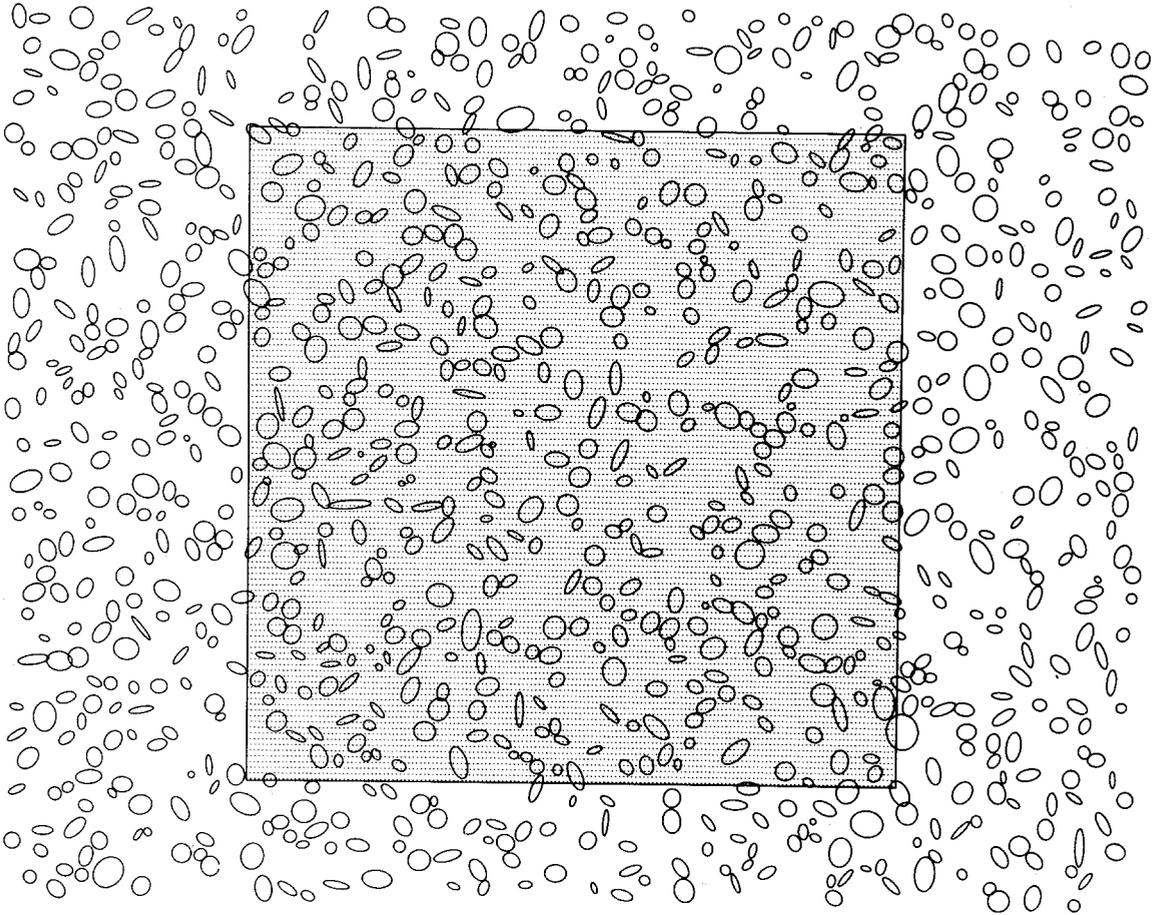


Fig. 5.2. Here the trapping grid (shaded area) is much larger relative to home range size than in Fig. 5.1. Therefore, the effective area of the trapping grid probably is little, if any, larger than its physical size.

squirrel density over the grid. A uniform density has been assumed in the derivation of the estimation method and, therefore, we do not wish to see a rejection of the uniformity hypothesis in any of the tests. The data set gives no evidence of nonuniformity, because none of the tests approach significance. We are not yet sure how severely the estimation method would be affected if the uniform density condition were not met (see Chapter 6, Example 7), but we can assume safely that accuracy would be decreased. If density is not uniform, defining the concept of parameter D becomes difficult.

Figures 5.7c-f give the results of population estimation for each of the four subgrids. The estimator for Model  $M_{bh}$  has been used in each case because this is the model chosen by the model selection procedure for the entire data set. Although relatively few squirrels were captured in the inner and middle inner grids, the high capture success indicated by the  $p_i$  estimates provides sufficient confidence in the reliability of the population estimates for these grids. Furthermore, in all four subgrids, the model provided a good fit to the data, as indicated by the chi-square and probability columns in the output.

Figure 5.7g gives the basic results of the density estimation procedure. Listed first are the values of necessary parameters and the starting values for density D and strip width W. (*Otis et al. 1978:121-122* gives the derivation of the starting values.) The next information of interest compares the naive estimate of density for each subgrid (that is, the population estimate divided by the area of the subgrid,  $\hat{Y}_i = \hat{N}_i/A_i$ ) with the values predicted by the fitted nonlinear least-squares regression equation. This output gives the experimenter a feel for how well the regression model fits the data, and the multiple correlation coefficient provides a quantitative measure of the model's adequacy. The coefficient's value in this example,  $R =$

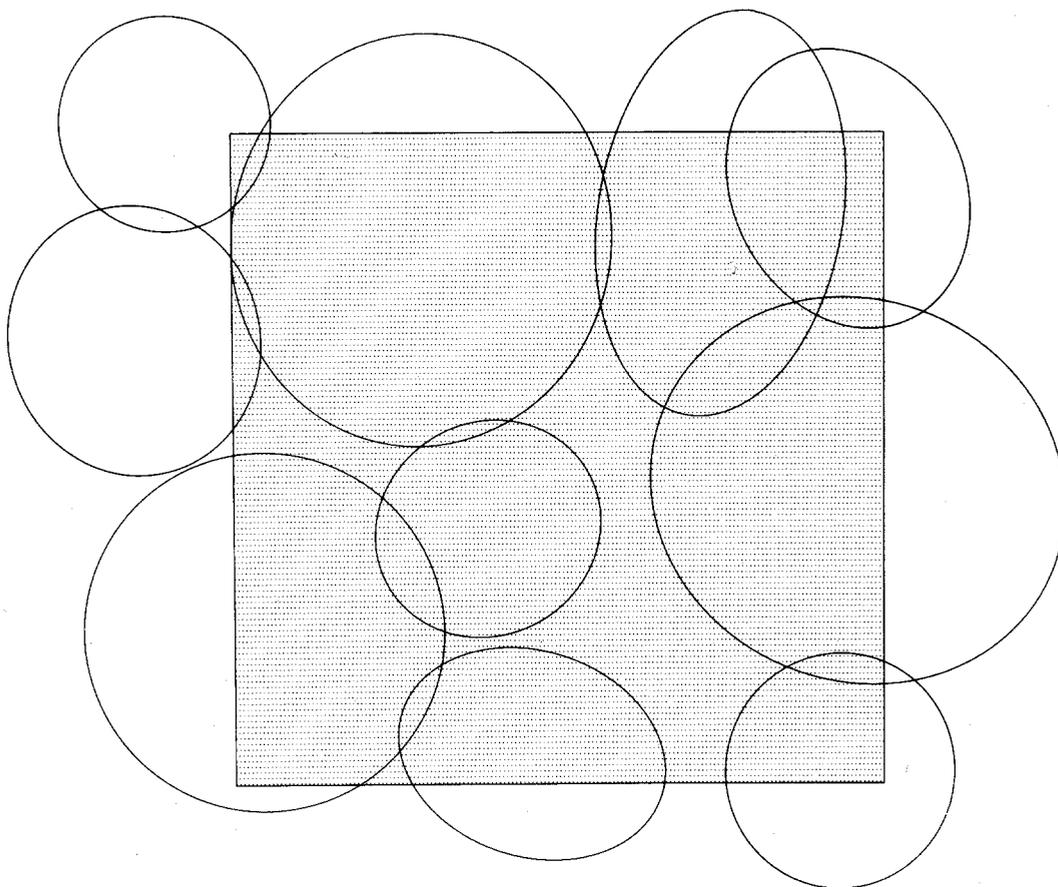


Fig. 5.3. Because almost all of the home ranges (ellipses) include some area outside the trapping grid (shaded area), the grid's effective area is much larger than its physical area. At best, a very poor density estimate would be achieved under these circumstances.

0.93169, indicates that the regression equation has accounted for about 93% of the variation in the naive estimates  $\hat{Y}_i$ . Next, estimates of the parameters  $D$  and  $W$  and their standard errors are given. From this capture-recapture experiment, program CAPTURE estimates that  $D$  equals about 44 ground squirrels per hectare. The estimate depends on the estimate of  $W$ , that is, on the fact that a strip of estimated width 12.4 m must be added to all four sides of the physical grid area to account for the larger effective grid area. The output also provides an estimate of the correlation between the estimates of  $D$  and  $W$ . Intuitively, we expect a high negative correlation because the larger the strip width (the effective area trapped), the lower the density estimate should be. Our estimate of  $-0.9805$  indicates such a relation. Finally, program CAPTURE calculates a simple test of the hypothesis that  $W$  is significantly different from zero—that the data justify the addition of a strip width. Our results show that such an addition is indeed necessary.

### Design Requirements

The density estimation procedure requires that the following assumptions be met if the data are to be analyzed by program CAPTURE.

1. All the assumptions made earlier in this text continue to be met, because a population estimate must be made for each of the nested grids.

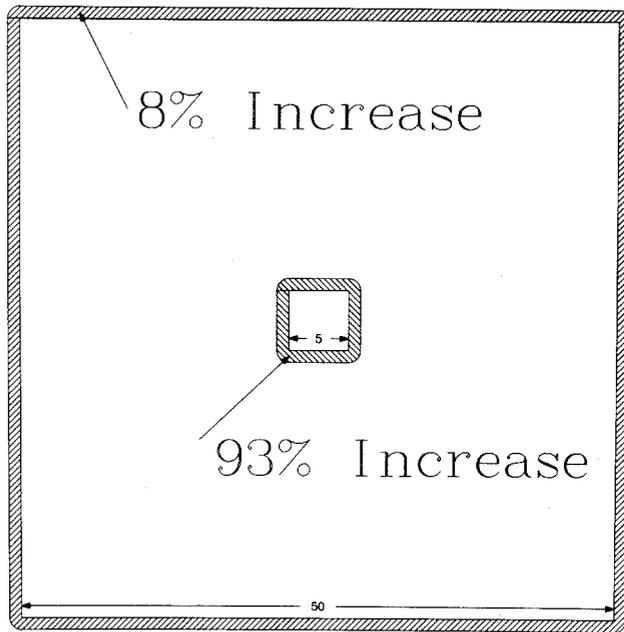


Fig. 5.4. The relative importance of a fixed-width boundary strip decreases as the grid size increases. Thus, the addition of a 1-unit-wide boundary strip to a 5- by 5-unit trapping grid adds 93% to the grid's effective size. In contrast, the addition of the same width boundary strip to a 50- by 50-unit trapping grid adds only 8% to the larger grid's effective size.

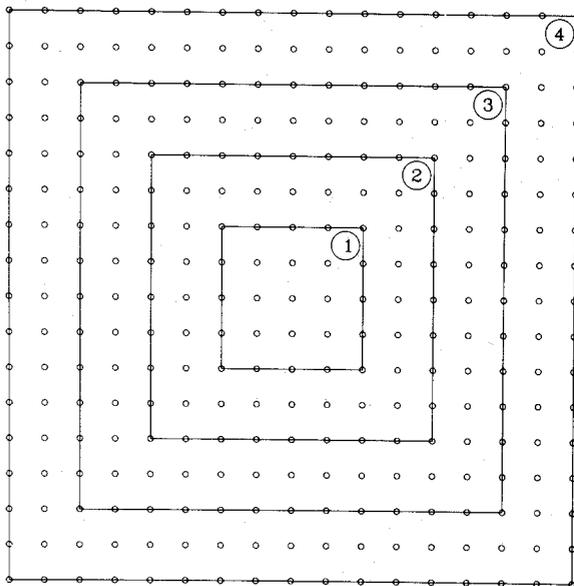


Fig. 5.5. The effect of boundary strip width on density is determined by program CAPTURE from nested subgrids within the entire grid. Thus, the outer grid (4) consists of all the traps, with  $x = 1 - 17$  and  $y = 1 - 17$ , and the smallest inner grid (1) consists of only the traps  $x = 7 - 11$  and  $y = 7 - 11$ . Because the impact of the strip width decreases with increasing grid size (Fig. 5.4), the nested grid structure provides a means of estimating strip width. The naive density estimates should decrease as the grid size increases; that is, the naive estimate of density ( $\hat{Y}_1$ ) for Grid 1 is expected to be larger than the naive estimate of density ( $\hat{Y}_4$ ) for Grid 4.

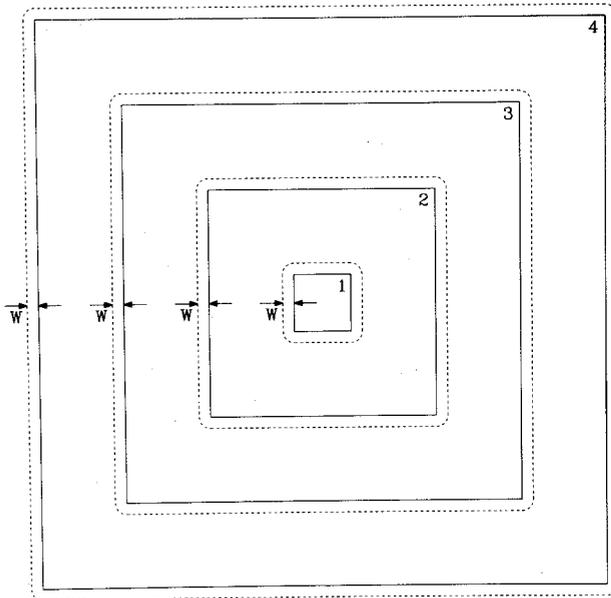


Fig. 5.6. The boundary strip width  $W$  is assumed to be constant, regardless of the grid size. The nested subgrids are used to construct an estimate of  $W$  in program CAPTURE.

MARK-RECAPTURE POPULATION AND DENSITY ESTIMATION PROGRAM DEVELOPED BY THE UTAH COOPERATIVE WILDLIFE RESEARCH UNIT.  
 CAPTURE RECAPTURE WORKBOOK EXAMPLES  
 PROGRAM VERSION OF MAY 07, 1980

TEST FOR UNIFORM DENSITY. SEE THIS SECTION OF THE MONOGRAPH FOR DETAILS.  
 CHAPTER 5 DENSITY ESTIMATION EXAMPLE

PAGE 701  
 80/07/26

MATRIX OF CAPTURES PER TRAP STATION.

COLUMNS	1	2	3	4	5	6	7	8	9	10
ROW 1	0	0	1	1	0	1	0	1	3	1
ROW 2	0	0	3	5	1	0	0	1	0	2
ROW 3	0	2	3	0	0	0	5	0	0	0
ROW 4	1	1	0	1	0	0	1	1	2	0
ROW 5	1	2	2	0	0	0	2	2	1	0
ROW 6	0	0	1	3	0	0	3	1	2	0
ROW 7	3	0	0	3	3	2	0	2	0	2
ROW 8	2	3	0	0	4	3	0	1	3	0
ROW 9	0	3	0	4	5	0	0	1	0	1
ROW 10	2	2	0	5	0	4	3	2	0	0

Grid Number

- ← 4 Entire grid  $x=1-10$   $y=1-10$
- ← 3 Middle outer grid  $x=2-9$   $y=2-9$
- ← 2 Middle Inner grid  $x=3-8$   $y=3-8$
- ← 1 Inner grid  $x=4-7$   $y=4-7$

IN THE ABOVE MATRIX, TRAP COORDINATES ARE ROUNDED TO THE NEAREST WHOLE INTEGER.  
 IN THE FOLLOWING GOODNESS OF FIT TESTS, TRAP COORDINATES THAT ARE NOT INTEGERS  
 AND NON-RECTANGULAR TRAPPING GRIDS WILL CAUSE SPURIOUS RESULTS.

Fig. 5.7a. Nested subgrids used in the density estimation procedure with data on Richardson's ground squirrels. In the matrix, each entry is the number of ground squirrels caught at a particular trap station.

TEST FOR UNIFORM DENSITY. SEE THIS SECTION OF THE MONOGRAPH FOR DETAILS.  
 CHAPTER 5 DENSITY ESTIMATION EXAMPLE

CHI-SQUARE TEST OF UNIFORM DENSITY BY ROWS.

ROW	1	2	3	4	5	6	7	8	9	10
OBSERVED	8	12	10	7	10	10	15	16	14	18
EXPECTED	12.000	12.000	12.000	12.000	12.000	12.000	12.000	12.000	12.000	12.000
CHI-SQUARE	1.333	0.000	0.333	2.083	0.333	0.333	0.750	1.333	0.333	3.000

TOTAL CHI-SQUARE = 9.83 WITH 9 DEGREES OF FREEDOM. PROBABILITY OF LARGER VALUE = 0.3641

CHI-SQUARE TEST OF UNIFORM DENSITY BY COLUMNS.

COLUMN	1	2	3	4	5	6	7	8	9	10
OBSERVED	9	13	10	22	13	10	14	12	11	6
EXPECTED	12.000	12.000	12.000	12.000	12.000	12.000	12.000	12.000	12.000	12.000
CHI-SQUARE	0.750	0.083	0.333	8.333	0.083	0.333	0.333	0.000	0.083	3.000

TOTAL CHI-SQUARE = 13.33 WITH 9 DEGREES OF FREEDOM. PROBABILITY OF LARGER VALUE = 0.1481

CHI-SQUARE TEST OF UNIFORM DENSITY BY RINGS (OUTER RING IS NUMBER 1).

RING	1	2	3	4	5
OBSERVED	38	39	25	18	0
EXPECTED	43.200	33.600	24.000	14.400	4.800
CHI-SQUARE	0.626	0.868	0.042	0.900	4.800

TOTAL CHI-SQUARE = 7.24 WITH 4 DEGREES OF FREEDOM. PROBABILITY OF LARGER VALUE = 0.1240

None of the circled significance values are less than 0.05

Fig. 5.7b. Chi-square tests of uniform density with data on Richardson's ground squirrels. If any of the three probability levels are small ( $p < 0.05$ ), the investigator must question whether the assumption of uniform density across the grid is met.

- The trapping grid is square or rectangular with the intertrap distance the same in both "row" and "column" directions. Note in Fig. 5.5 that the traps are equally spaced in a lattice.
- The total trapping grid is large enough to contain at least three, but preferably four, subgrids (including the total grid), as shown in Fig. 5.5. It is not absolutely necessary that each subgrid be increased by two rings of traps as shown, but the greater the relative increase in area to the boundary strip, the better the estimation procedure will work. If additional grids were inserted in Fig. 5.5, the estimate probably would be more precise.
- The population density is approximately constant in the area of trapping. That is, there is no marked trend in density across the grid. The presence of a trend breaks down the concept of a constant strip width around each nested grid, shown in Fig. 5.6.
- Trapping success is high. By implicit assumption, a large proportion of the animals are captured and recaptured, because the population estimates  $\hat{N}_i$  from each of the subgrids must be reasonably reliable if the method is to work well.

In addition, the coordinates of the trap location of each capture and recapture must be recorded. The (x, y) coordinates of the traps must be known so that population estimates for the various subgrids can be calculated. Thus, if an animal is captured at three locations, but only two of these locations are in the inner subgrid, only these two captures can be used for estimation in the inner subgrid. Additional details about data recording are given in Chapter 7, Study Design.

POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 INNER GRID X=4-7 Y=4-7

OCCASION	J=	1	2	3	4	5	6
TOTAL CAUGHT	M(J)=	0	3	4	10	13	13
NEWLY CAUGHT	U(J)=	3	1	6	3	0	0

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 6					
1	13.81	2.01375	12.974	0.0047	0.3261	0.3261	0.3261	0.3261	0.3261	0.3261
2	13.00	1.012297	10.912	0.0122	0.2308	0.4545	0.4545	0.4545	0.4545	0.4545
3	13.00	0.2028692	1.631	0.4424	0.2308	0.1000	0.7500	0.7500	0.7500	0.7500
4	13.00	0.4201734E-05	0.001	0.9814	0.2308	0.1000	0.6667	0.9998	0.9998	0.9998

POPULATION ESTIMATE IS 13 WITH STANDARD ERROR 0.2029  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 12 TO 14

HISTOGRAM OF U(J)

FREQUENCY	3	1	6	3	0	0
6			*			
5			*			
4			*			
3	*		*	*		
2	*		*	*	*	
1	*	*	*	*	*	*

$$Y_1 = 13 / (30 \text{ m} \cdot 30 \text{ m} \cdot 1 \text{ ha} / 10000 \text{ m}^2) = 144.44 \text{ animals/ha}$$

Fig. 5.7c. An example of population estimation using the generalized removal estimator for Model  $M_{bh}$  with data on Richardson's ground squirrels taken from Grid 1 (inner grid) of Fig. 5.7a ( $x = 4 - 7, y = 4 - 7$ ). The first of four naive density estimates is produced from the population estimate for this grid.

POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 MIDDLE INNER GRID X=3-8 Y=3-8

OCCASION	J=	1	2	3	4	5	6
TOTAL CAUGHT	M(J)=	0	9	13	21	25	27
NEWLY CAUGHT	U(J)=	9	4	8	4	2	1

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 6					
1	30.85	3.390958	4.084	0.3948	0.3109	0.3109	0.3109	0.3109	0.3109	0.3109
2	30.19	3.284894	3.976	0.2640	0.2981	0.3395	0.3395	0.3395	0.3395	0.3395
3	28.00	0.9337696	0.393	0.8216	0.3214	0.2105	0.5769	0.5769	0.5769	0.5769
4	28.00	0.8693015	0.422	0.5159	0.3214	0.2105	0.5333	0.6363	0.6363	0.6363

POPULATION ESTIMATE IS 31 WITH STANDARD ERROR 3.3910  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 24 TO 38

HISTOGRAM OF U(J)

FREQUENCY	9	4	8	4	2	1
9	*					
8	*		*			
7	*		*			
6	*		*			
5	*		*			
4	*	*	*	*		
3	*	*	*	*	*	
2	*	*	*	*	*	*
1	*	*	*	*	*	*

$$Y_2 = 31 / (50 \text{ m} \cdot 50 \text{ m} \cdot 1 \text{ ha} / 10000 \text{ m}^2) = 124 \text{ animals/ha}$$

Fig. 5.7d. An example of population using the generalized removal estimator for Model  $M_{bh}$  with data on Richardson's ground squirrels taken from Grid 2 (middle inner grid) of Fig. 5.7a ( $x = 3 - 8, y = 3 - 8$ ). The second naive density estimate is produced from the population estimate for this grid.

POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 MIDDLE OUTER GRID X=2-9 Y=2-9

OCCASION	J=	1	2	3	4	5	6
TOTAL CAUGHT	M(J)=	0	17	27	35	39	43
NEWLY CAUGHT	U(J)=	17	10	8	4	4	1

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 6					
1	46.14	2.446583	1.324	0.8573	0.3799	0.3799	0.3799	0.3799	0.3799	0.3799
2	45.81	2.548772	1.423	0.7000	0.3711	0.3969	0.3969	0.3969	0.3969	0.3969
3	44.68	1.766119	1.458	0.4823	0.3805	0.3612	0.4895	0.4895	0.4895	0.4895
4	44.06	1.288259	1.862	0.1723	0.3858	0.3695	0.4688	0.5925	0.5925	0.5925

POPULATION ESTIMATE IS 46 WITH STANDARD ERROR 2.4466  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 41 TO 51

HISTOGRAM OF U(J)

FREQUENCY	17	10	8	4	4	1
-----						
EACH * EQUALS	2 POINTS					
18	.					
16	.					
14	.					
12	.	.				
10	.	.	.			
8	.	.	.	.		
6	.	.	.	.	.	
4	.	.	.	.	.	.
2	.	.	.	.	.	.

$$Y_3 = 46 / (70m \cdot 70m \cdot 1ha / 10000 m^2) = 93.88 \text{ animals/ha}$$

Fig. 5.7e. An example of population estimation using the generalized removal estimator for Model  $M_{bh}$  with data on Richardson's ground squirrels taken from Grid 3 (middle outer grid) of Fig. 5.7a ( $x = 2 - 9, y = 2 - 9$ ). The third naive density estimate is produced from the population estimate for this grid.

POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 TOTAL GRID X=1-10 Y=1-10

OCCASION	J=	1	2	3	4	5	6
TOTAL CAUGHT	M(J)=	0	24	38	48	53	57
NEWLY CAUGHT	U(J)=	24	14	10	5	4	1

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 6					
1	59.86	2.124398	1.036	0.9043	0.4169	0.4169	0.4169	0.4169	0.4169	0.4169
2	59.44	2.097676	1.120	0.7721	0.4038	0.4404	0.4404	0.4404	0.4404	0.4404
3	58.51	1.513301	1.103	0.5762	0.4102	0.4057	0.5258	0.5258	0.5258	0.5258
4	58.00	1.113654	1.437	0.2307	0.4138	0.4118	0.5000	0.6250	0.6250	0.6250

POPULATION ESTIMATE IS 60 WITH STANDARD ERROR 2.1244  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 55 TO 65

HISTOGRAM OF U(J)

FREQUENCY	24	14	10	5	4	1
-----						
EACH * EQUALS	3 POINTS					
24	.					
21	.					
18	.	.				
15	.	.	.			
12	.	.	.	.		
9	.	.	.	.	.	
6	.	.	.	.	.	.
3	.	.	.	.	.	.

$$Y_4 = 60 / (90m \cdot 90m \cdot 1ha / 10000 m^2) = 74.07 \text{ animals/ha}$$

Fig. 5.7f. An example of population using the generalized removal estimator for Model  $M_{bh}$  with data on Richardson's ground squirrels taken from Grid 4 (entire grid) of Fig. 5.7a ( $x = 1 - 10, y = 1 - 10$ ). The fourth naive density estimate is produced from the population estimate for this grid.

JOINT ESTIMATION OF DENSITY AND BOUNDARY STRIP WIDTH FROM CAPTURE DATA. SEE THIS SECTION OF THE MONOGRAPH FOR DETAILS.  
 CHAPTER 5 DENSITY ESTIMATION EXAMPLE

STARTING VALUES FOR DENSITY ESTIMATION--  
 NUMBER OF GRIDS 4  
 TRAP INTERVAL 10.00  
 UNITS CONVERSION 10000.00  
 INITIAL DENSITY ESTIMATE 66.0122  
 INITIAL STRIP WIDTH ESTIMATE 7.1501

GRID I	NAIVE DENSITY	PERIMETER/AREA	PI/AREA	STARTING COVARIANCE MATRIX			
	Y(I)	A(I)	B(I)				
1	144.4444	0.1333333	0.3491E-02	5.08			
2	124.0000	0.800000E-01	0.1257E-02	14.5	184.		
3	93.87755	0.5714286E-01	0.6411E-03	3.11	39.4	24.9	
4	74.07407	0.4444444E-01	0.3879E-03	1.07	13.5	8.55	6.88

RESULTS OF ITERATIONS  
 FUNCTION EVALUATIONS REQUIRED 54  
 ESTIMATED SIGNIFICANT DIGITS OF PARAMETER VALUES 6

Note the decrease in the naive density estimates with increasing grid size

FITTED MODEL COMPARED TO THE DATA

GRID(I)	Y(I)	F(I)
1	144.444	142.398
2	124.000	97.929
3	93.878	80.648
4	74.074	71.879

MULTIPLE CORRELATION COEFFICIENT IS 0.93169

ESTIMATED DENSITY= 44.612 ± 2.998 = ITS STANDARD ERROR  
 ESTIMATED STRIP WIDTH= 12.409 ± 1.0139 = ITS STANDARD ERROR  
 CORRELATION OF ESTIMATORS - .9805

The final density estimate is even smaller than  $Y_4$

TEST OF ESTIMATED STRIP WIDTH GREATER THAN ZERO.  
 Z-VALUE = 12.2389 PROBABILITY OF LARGER VALUE = 0.0000

FINAL COVARIANCE MATRIX

5.081				
16.41	184.0			
3.762	42.17	24.93		
1.348	15.11	8.930	6.879	

Fig. 5.7g. An example of the joint estimation of density  $D$  and strip width  $W$  with data on Richardson's ground squirrels,  $\hat{D} = 44.612 \pm 2.998$  and  $\hat{W} = 12.409 \pm 1.014$ . The boundary strip model discussed in the text explains 93% of the variance of the  $Y(I)$  (the naive estimates of density from the nested subgrids). Note that the naive density estimate for the entire grid ( $\hat{Y}_4 = 74.07$  squirrels/ha), which is the best of the four estimates, is reduced almost by half for the final estimate of  $D$ .

## Summary

1. Density  $D$  is the number of animals per unit of area.
2. Because a portion of the animals trapped on a grid actually have some home range area outside the grid, the effective area of a grid is larger than the physical area.
3. To correct the physical area of a trapping grid to its effective area, a constant strip width,  $W$ , is added around the grid perimeter.
4. The importance of the strip width  $W$  is much greater for a small grid than for a large grid, and thus a set of nested subgrids partitioned from the entire grid is used by program CAPTURE to estimate  $D$  and  $W$ , simultaneously.
5. The methods used in program CAPTURE to estimate density make the same assumptions as those used for the population estimators in Chapter 3; in addition, density is assumed to be constant in the area of the trap grid.
6. Because  $\hat{D}$  and  $\hat{W}$  are highly correlated and because a high variability is associated with density estimation experiments, a large proportion of the population must be captured and recaptured to achieve a reliable estimate of either  $D$  or  $W$ .

## Questions and Exercises

1. Would a 3 by 3 trapping grid be satisfactory for estimating density?
2. If the strip width  $W$  is equal to zero, is the population geographically closed? (See Chapter 1 for a discussion of geographical closure.)
3. In Fig. 5.1, what is the population (number of animals) at risk of capture?
4. Which would you recommend, one trap per station and twice the grid size, or two traps per station and half the grid size? Why?
5. If there is a linear gradient in density across the grid, would you suggest using subgrids that consist of halves or quarters of the entire grid? Why?
6. Can density be estimated when 100 traps are laid out in one long line instead of in a 10 by 10 grid? If so, how?
7. Why is the naive density estimate of 74.07 for the entire grid in Fig. 5.7g reduced to the final estimate of 44.61?
8. How do the units of density  $D$  differ from those of population  $N$ ?
9. Why does the estimation of  $D$  require more than the  $X$  matrix defined in Chapter 1?

# CHAPTER 6 EXAMPLES

When the reader has become familiar with the concepts presented in Chapters 2 to 5, he will need some experience in their application. In this chapter, we provide an opportunity to perform extensive analysis of several examples. We have included annotated computer output so that the reader can become familiar with the output from program CAPTURE, although in some cases only the output relevant to the particular example has been included.

Note that notation in the computer output differs from notation in the text, because most computer printers cannot provide subscripts. Thus, for example, the text notation  $M_t$  corresponds to the computer notation  $M(T)$ . Parentheses are used to denote subscripts. Additional examples can be found in *Otis et al.* (1978:81-96).

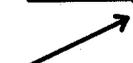
## Example 1. Interpreting the Data

For this example, we have created a series of simulated experiments to help the reader develop insight into the catchability structure of the population under study by examining some key summary statistics of the capture data. In all of these experiments, the population size is 300 and there are 7 trapping occasions.

Figure 6.1a gives the summary statistics for a population having no variation in capture probabilities among its members (Model  $M_o$ ,  $p = 0.20$ ). The most significant aspect of the statistics is that there is nothing unusual about them: daily captures remain relatively constant and the frequencies of capture show a steady decline. Contrast this vector of frequencies with those displayed in Fig. 6.1b. The statistics in Fig. 6.1b were generated from a heterogeneous population of individuals in which 3 groups of 100 animals have capture probabilities equal to 0.1, 0.2, and 0.3, respectively (Model  $M_h$ ). Many more animals have been caught three or more times in this experiment than in the previous one; the increase indicates that some members of the population are very susceptible to trapping. Indeed, nearly 35% of the animals captured are captured more than twice, as opposed to only 15% in the  $M_o$  experiment. Also, the daily captures have remained relatively stable. Thus, the best clue the researcher has for detecting heterogeneity arises from irregularities in the frequencies of capture. As we shall see below, other sources of unequal catchability tend to produce more obvious changes in the statistics than does individual heterogeneity.

CHAPTER 6 M(O) PART OF M(TBH) EXAMPLE 1

OCCASION	J=	1	2	3	4	5	6	7	
ANIMALS CAUGHT	N(J)=	58	58	61	66	61	57	51	
TOTAL CAUGHT	M(J)=	0	58	104	146	175	205	229	240
NEWLY CAUGHT	U(J)=	58	46	42	29	30	24	11	
FREQUENCIES	F(J)=	115	89	27	7	2	0	0	


  
 Only 36 animals were captured  
 more than twice in 7 occasions

Daily captures are fairly constant

CHAPTER 6 M(H) PART OF M(TBH) EXAMPLE 1

OCCASION	J=	1	2	3	4	5	6	7
ANIMALS CAUGHT	N(J)=	60	66	59	57	57	48	67
TOTAL CAUGHT	M(J)=	0	60	108	143	166	180	194
NEWLY CAUGHT	U(J)=	60	48	35	23	14	14	17
FREQUENCIES	F(J)=	97	55	37	16	4	2	0

During the course of the experiment 59 animals were captured more than twice

Figure 6.1c contains statistics generated from trapping a population with the heterogeneity structure described in the  $M_h$  experiment and an added factor: the overall trapping success on the first two occasions was roughly 25% greater than that experienced on the remaining five occasions. Thus, the probability structure of this population corresponds to Model  $M_{th}$ . The time factor has caused a much larger number of captures on the first two occasions relative to the other occasions. A large variation of this type among the total animals caught on each occasion is the best indicator that time variation is playing a role. Again, the percentage of animals captured more than twice is large (almost 30%), as was the case in Fig. 6.1b.

Finally, consider the statistics presented in Fig. 6.1d. The most interesting feature of these data is that the daily captures exhibit a U-shape—that is, the number of animals caught starts out high, gradually declines, and then gradually builds up again, peaking on the last occasion. This pattern is caused by the addition of a behavioral response in the population to the same sources of unequal catchability (heterogeneity and time variation) contained in the previous population. (Animals have a 30% greater chance of recapture than of first capture; they become trap happy.) Thus, we are dealing with a population in which all three sources of unequal catchability are operating (Model  $M_{tbh}$ ). As before, the large number of first captures is caused by a greater overall trapping success (time variation) on those occasions. The gradual buildup in the number of animals caught beginning on the fifth occasion is caused by the trap happiness of the population. That is, after an animal is caught, it has a 30% greater chance of being recaptured; thus, as more and more individuals are caught, the overall susceptibility of the population to trapping increases, and more and more individuals are caught. Of course, the variation among daily captures could have been caused by time variation alone. However, the fact that captures show a steady increase with time is a signal that perhaps behavioral variation is responsible. Conversely, we might suspect that animals were trap shy if captures showed a steady decline from some relatively constant level.

This series of experiments represents only a sample of the literally infinite number of situations that can arise in the real world. The reader must think about what kinds of probability structure could be causing the pattern discerned in the statistics generated by the experiment. The potential complexity of the probability structures of real data necessitates rigorous testing of the data; this testing is the function of the model selection procedure in program CAPTURE.

There is a big drop in capture success after the second occasion

CHAPTER 6 M(TBH) PART OF M(TBH) EXAMPLE 1

OCCASION	J=	1	2	3	4	5	6	7
ANIMALS CAUGHT	N(J)=	82	84	59	57	57	48	67
TOTAL CAUGHT	M(J)=	0	82	135	166	186	199	210
NEWLY CAUGHT	U(J)=	82	53	31	20	13	11	15
FREQUENCIES	F(J)=	98	61	42	14	8	2	0

Fig. 6.1c. The summary statistics from a simulated experiment on a Model  $M_{th}$  population.

Because heterogeneity is still present many animals are captured more than twice

CHAPTER 6 M(TBH) PART OF M(TBH) EXAMPLE 1

OCCASION	J=	1	2	3	4	5	6	7
ANIMALS CAUGHT	N(J)=	82	80	70	70	73	83	95
TOTAL CAUGHT	M(J)=	0	82	128	158	191	210	228
NEWLY CAUGHT	U(J)=	82	46	30	33	19	18	13
FREQUENCIES	F(J)=	73	77	54	22	14	1	0

Fig. 6.1d. The summary statistics from a simulated experiment on a Model  $M_{tbh}$  population.

The effects of behavioral response on daily captures appear in the last few occasions

### Example 2. Trap Happy or Trap Shy—No Difference?

In a Florida sugar cane field, 76 traps were placed along 6 parallel transects and baited with apples. Traps were placed 15.4 m apart on a transect, transects were an average 80 m apart, and trapping was done for 8 consecutive days. The species under study was the cotton rat (*Sigmodon hispidus*). Trapping results are summarized in Fig. 6.2a. The model selected as appropriate for population estimation was the behavioral response Model  $M_b$  (Fig. 6.2b). The results given in Fig. 6.2c show that the rats evidently are becoming trap happy—the probability of recapture is 0.38 as opposed to 0.23 for first capture. To illustrate a point, we manipulated this data set so that the sequence of newly captured animals remains unchanged, but rats now tend to avoid recapture (Fig. 6.2d). The model selection procedure still selects the behavioral model as appropriate, and the population estimation results are as given in Fig. 6.2e. The probability of first capture is the same as before (0.23), but now probability of recapture is 0.07. The population estimate and its estimated standard error, however, remain the same! The lack of change in the estimate illustrates the point that, in behavioral response models, recaptures have no effect on the bias or precision of the estimate. This fact should affect construction of the study design if, on the basis of previous knowledge, the researcher has decided that the behavioral response model probably will be used to analyze the data. See Chapter 7 (Study Design) for a discussion of ways to make such a priori decisions. The emphasis in the design should be placed on capturing as many different animals as

CHAPTER 6 EXAMPLE 2

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	N(J)=	19	26	33	27	33	37	27	28
TOTAL CAUGHT	M(J)=	0	19	36	52	60	66	74	81
NEWLY CAUGHT	U(J)=	19	17	16	8	6	8	7	1
FREQUENCIES	F(J)=	24	21	11	13	5	3	4	1

t = 8



82

↑

$M_{t+1} = M_9 = 82 = \text{Total number of different animals captured}$

possible, possibly at the sacrifice of recaptures; that is, the probability of first capture  $p$  should be made as large as possible. One way to achieve this goal is to change the location of the traps between trapping occasions to increase the number of different areas of activity that are trapped effectively.

### Example 3. Closure?

Perusal of the sample statistics given in Fig. 6.3a, created from another trapping experiment in the Florida sugar cane field, reveals a serious problem with respect to the assumption of closure. During the first 5 days of trapping, the number of newly caught animals steadily declines, as one would expect when trapping a population not subject to immigration. On the sixth day, however, the number of newly caught animals jumps to 20, and on the seventh and eighth days as well, the numbers of unmarked animals caught are significant. This phenomenon should suggest an influx of new animals into the study area to the investigator, for it is not likely that such a severe jump in newly captured animals could occur without immigration. Unfortunately, the researcher cannot depend on the results of the closure test (Fig. 6.3b) to alert him to this possibility, because the test gives no indication that the closure assumption has been violated. This is not surprising, because the model selected as most appropriate for this data set is the behavioral response Model  $M_b$  (Fig. 6.3c), and as has been pointed out (*Otis et al. 1978:66*), the closure test is not reliable in the presence of behavioral response. A warning signal exists, however, in the huge standard error of the population estimate, resulting in a coefficient of variation of 68%. In view of the suspected lack of closure, the researcher cannot present the estimate of 285 animals (Fig. 6.3d) as a valid estimate of the population size at the beginning of the experiment. Eliminating the last 3 days of trapping and reanalyzing the data probably would provide a better estimate of size at that time.

### Example 4. Separating the Sexes

In Chapter 7, we point out that one potential method for eliminating individual heterogeneity in capture probabilities is to stratify the data into groups based on age, sex, or any other factor that the researcher may suspect as the cause of heterogeneity. The requirement that each subgroup have adequate data often prevents the use of this approach, but in a capture experiment in our Florida sugar cane field, ample data were collected and we were able to stratify them by sex. In this example, the objective was not so much to eliminate heterogeneity and thus improve the estimate of total population size as it was to satisfy a

Strong evidence of behavioral response

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 38.216 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 9.892 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00166
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 12.195 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.09434
4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 12.828 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.07641  
 TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE  
 (FREQUENCIES LESS THAN 21 ARE NOT CALCULATED.)  

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	14.000	7	0.05118
2	7.444	7	0.38412
5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 14.600 DEGREES OF FREEDOM = 12 PROBABILITY OF LARGER VALUE = 0.26404
- 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 5.795 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.44652
- 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 8.805 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.18485
6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 EXPECTED VALUES TOO SMALL. TEST NOT PERFORMED.
7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 32.358 DEGREES OF FREEDOM = 19 PROBABILITY OF LARGER VALUE = 0.02847

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
CRITERIA	0.74	0.77	1.00	0.93	0.00	0.37	0.76	0.79

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

Evidence of trap response even in the presence of individual heterogeneity

Fig. 6.2b. Output from the model selection procedure.

Probability of capture increases  
after first capture - animals  
are evidently trap happy

CHAPTER 6 EXAMPLE 2

OCCASION	J=	1	2	3	4	5	6	7	8
TOTAL CAUGHT	M(J)=	0	19	36	52	60	66	74	81
NEWLY CAUGHT	U(J)=	19	17	16	8	6	8	7	1

ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.230132  
ESTIMATED PROBABILITY OF RECAPTURE, C-HAT = 0.381443

POPULATION ESTIMATE IS 93 WITH STANDARD ERROR 6.6865  
APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS 79 TO 107

HISTOGRAM OF U(J)

FREQUENCY	19	17	16	8	6	8	7	1
EACH * EQUALS	2 POINTS							
20	*							
18	*	*						
16	*	*	*					
14	*	*	*					
12	*	*	*					
10	*	*	*					
8	*	*	*	*		*	*	
6	*	*	*	*	*	*	*	
4	*	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*	*

Fig. 6.2c. Estimates of population size and capture probabilities produced by the Zippin procedure for the original data.

CHAPTER 6 EXAMPLE 2 - MANIPULATED

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	N(J)=	19	17	18	12	11	13	12	8
TOTAL CAUGHT	M(J)=	0	19	36	52	60	66	74	81
NEWLY CAUGHT	U(J)=	19	17	16	8	6	8	7	1
FREQUENCIES	F(J)=	67	7	3	5	0	0	0	0

The sequence of first captures is  
the same as before, but total daily  
captures have decreased because  
animals have been made trap shy

Fig. 6.2d. The summary statistics for the manipulated trap-shy data set.

CHAPTER 6 EXAMPLE 2 - MANIPULATED

OCCASION	J=	1	2	3	4	5	6	7	8
TOTAL CAUGHT	M(J)=	0	19	36	52	60	66	74	81
NEWLY CAUGHT	U(J)=	19	17	16	8	6	8	7	1

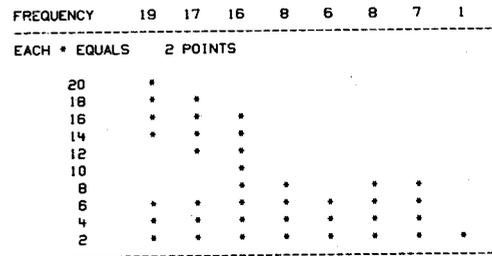
ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.230132

ESTIMATED PROBABILITY OF RECAPTURE, C-HAT = 0.072165

POPULATION ESTIMATE IS 93 WITH STANDARD ERROR 6.6865

APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS 79 TO 107

HISTOGRAM OF U(J)



Estimates of  $p$  and  $N$  have remained the same, even though the probability of recapture has fallen from 0.38 to 0.07!

Fig. 6.2e. Estimates of population size and capture probabilities for the manipulated data set.

CHAPTER 6 EXAMPLE 3

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	N(J)=	17	16	16	17	12	29	29	25
TOTAL CAUGHT	M(J)=	0	17	30	41	51	58	78	91
NEWLY CAUGHT	U(J)=	17	13	11	10	7	20	13	7
FREQUENCIES	F(J)=	55	31	7	3	1	1	0	0

Evidence of immigration

# The closure test fails to detect evidence of the apparent breakdown of closure

CHAPTER 6 EXAMPLE 3

OVERALL TEST RESULTS --  
 Z-VALUE 0.824  
 PROBABILITY OF A SMALLER VALUE 0.79498

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

NUMBER OF CAPTURES Z-VALUE PROBABILITY

NUMBER OF CAPTURES	Z-VALUE	PROBABILITY
2	-.207	0.41785

See Figure 2.9 in Chapter 2

Fig. 6.3b. The formal test for population closure.

curiosity about (1) the relative sizes of populations by sex, (2) the relative catchability of the sexes, and (3) whether individual differences in catchability could be ascribed to sex. Figure 6.4a shows a summary of the entire data set, and Fig. 6.4b reveals that analysis of the data results in an excellent fit to the heterogeneity Model  $M_h$ . The population estimate of 391 (S.E. = 36) is entirely satisfactory in terms of standard error (Fig. 6.4c). The analyses for males and females are given in Figs. 6.4d-f and 6.4g-i, respectively. With respect to the males, we notice that Models  $M_0$  and  $M_h$  both receive values of 1.00 in the model selection procedure, and that the Model  $M_h$  is chosen for the estimation. This choice has been built into program CAPTURE because Model  $M_h$  has the more robust estimator; that is, we believe that it is the "safer" model to use when the selection process produces a tie. The estimate of the male population size is 211 (S.E. = 21), and the estimate of average capture probability of males is 0.1013, a figure that is very close to the corresponding estimate of 0.1087 for the entire population. Similar results are produced for females. Model  $M_h$  is chosen and proves to be a much more solid choice than it was for males. The estimate of female population size is 148 (S.E. = 15), and the estimate of 0.1427 for the average probability of capture is again close to the corresponding figure for the entire population.

Have the analyses satisfied our earlier curiosities? First, we should test to see if the population estimates for each sex are different. A simple method is the z-test, which assumes that the estimates of  $N$  are distributed normally, with known variance. These conditions should hold at least approximately when sample size is large. We assume that our sample meets these assumptions. The null hypothesis is  $H_0: N_\delta = N_\phi$ . The two-tailed z-test is calculated as

$$z = \frac{\hat{N}_\delta - \hat{N}_\phi}{\sqrt{\hat{\text{Var}}(\hat{N}_\delta) + \hat{\text{Var}}(\hat{N}_\phi)}}$$

CHAPTER 6 EXAMPLE 3

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 8.808 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.03196

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 6.769 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00928

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 16.849 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.01840

4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 17.982 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.01205

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

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	11.473	7	0.11929
2	11.366	7	0.12345

5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 14.705 DEGREES OF FREEDOM = 12 PROBABILITY OF LARGER VALUE = 0.25796

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 11.598 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.07157

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 3.108 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.79522

6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 EXPECTED VALUES TOO SMALL. TEST NOT PERFORMED.

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 35.797 DEGREES OF FREEDOM = 14 PROBABILITY OF LARGER VALUE = 0.00112

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
	0.48	0.41	1.00	0.70	0.00	0.27	0.73	0.48

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

Model  $M_b$  gives a good fit to the data; therefore it is not surprising that it is chosen as the best model

Fig. 6.3c. The results of tests of capture probability structure and the model selection procedure.

Width of the confidence interval reflects the total unreliability of the population estimate. This is the best evidence the biologist has of the fact that lack of closure has prevented production of a good estimate

CHAPTER 6 EXAMPLE 3

OCCASION	J=	1	2	3	4	5	6	7	8
TOTAL CAUGHT	M(J)=	0	17	30	41	51	58	78	91
NEWLY CAUGHT	U(J)=	17	13	11	10	7	20	13	7

ESTIMATED PROBABILITY OF CAPTURE, P-HAT = 0.051280

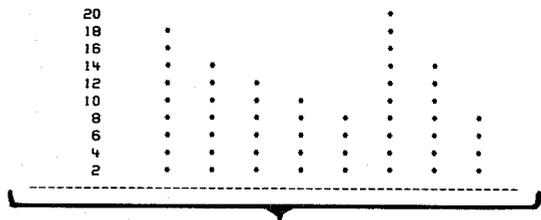
ESTIMATED PROBABILITY OF RECAPTURE, C-HAT = 0.172131

POPULATION ESTIMATE IS 285 WITH STANDARD ERROR 194.1353

APPROXIMATE 95 PERCENT CONFIDENCE INTERVALS -95 TO 666

HISTOGRAM OF U(J)

FREQUENCY	17	13	11	10	7	20	13	7
EACH * EQUALS	2 POINTS							



Visual representation of the apparent influx of new animals

Fig. 6.3d. Population estimation using Model  $M_b$ .

CHAPTER 6 EXAMPLE 4 - SEXES COMBINED

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	N(ij)=	44	38	44	47	45	44	45	33
TOTAL CAUGHT	M(ij)=	0	44	74	102	129	152	172	185
NEWLY CAUGHT	U(ij)=	44	30	28	27	23	20	13	9
FREQUENCIES	F(ij)=	109	47	26	5	3	4	0	0

↑  
 A total of 109 animals were captured only once, and 47 animals were captured exactly twice

Fig. 6.4a. An example of stratification of the trapping data by sex, using results of an experiment on the cotton rat. The summary statistics are for the entire data set.

Substituting the numbers from Figs. 6.4f and 6.4i gives

$$z = \frac{211 - 148}{\sqrt{20.87^2 + 15.21^2}} = 2.44 .$$

The probability of a larger z value (taken from a z or standard normal deviate table) is 0.0146. We thus conclude that the populations of the two sexes differ in size and, because  $\hat{N}_\delta > \hat{N}_\phi$ , that there are more males than females.

The same approach could be used to test for differences in average capture probability between the sexes, but the procedure is not straightforward because of the difficulty in estimating the variances of these estimates. Even without such a test at our disposal, we can say that no biologically significant differences exist between the two parameters; the estimates are 0.10 for males and 0.14 for females. Furthermore, we can say that sex is not the cause of heterogeneous capture probabilities in the population.

Thus, we have seen that the population size appears to be weighted in favor of males and that, on the basis of subjective evaluation, the average male is about as catchable as the average female. Furthermore, we must conclude that individual heterogeneity in capture probabilities is present within each sex, although we find indications that such differences may not be quite as large among males. Finally, we point out that the estimates from the two sexes add up to 359 animals, a figure that is close to the estimate of 391 produced from the entire data set. In general, the estimate obtained by summing the individual subgroup estimates will not equal the estimate produced from the entire data set; in fact, the two can be quite different, particularly if different models are selected in the analyses. In this instance, however, close agreement between the two estimates is reached, probably because of the consistent use of the jackknife estimator and the similarity in capture probabilities between the two sexes.

### Example 5. Time Is of the Essence

We have described eight different probability models for capture-recapture experiments and discovered that three of them ( $M_{tb}$ ,  $M_{th}$ , and  $M_{tbb}$ ) do not have associated estimators. In each of these models, only the time factor appears consistently, and all the corresponding models without this factor ( $M_b$ ,  $M_h$ , and  $M_{bb}$ ) have associated estimators. Thus, if time is not a factor affecting capture probabilities, the

CHAPTER 6 EXAMPLE 4 - SEXES COMBINED

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.

NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(H)$

CHI-SQUARE VALUE = 35.250 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00000

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.

NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(B)$

CHI-SQUARE VALUE = 0.000 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.99596

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.

NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(T)$

CHI-SQUARE VALUE = 4.798 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.68463

4. GOODNESS OF FIT TEST OF MODEL  $M(H)$

NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(H)$

CHI-SQUARE VALUE = 4.444 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.72740

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

NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	7.330	7	0.39532
2	7.099	7	0.41862
3	1.651	7	0.97661

5. GOODNESS OF FIT TEST OF MODEL  $M(B)$

NULL HYPOTHESIS OF MODEL  $M(B)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(B)$

CHI-SQUARE VALUE = 7.095 DEGREES OF FREEDOM = 12 PROBABILITY OF LARGER VALUE = 0.85129

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

CHI-SQUARE VALUE = 2.844 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.82810

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

CHI-SQUARE VALUE = 4.250 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.64286

6. GOODNESS OF FIT TEST OF MODEL  $M(T)$

NULL HYPOTHESIS OF MODEL  $M(T)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(T)$

CHI-SQUARE VALUE = 204.825 DEGREES OF FREEDOM = 147 PROBABILITY OF LARGER VALUE = 0.00090

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.

NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(BH)$

CHI-SQUARE VALUE = 23.447 DEGREES OF FREEDOM = 19 PROBABILITY OF LARGER VALUE = 0.21822

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	$M(O)$	$M(H)$	$M(B)$	$M(BH)$	$M(T)$	$M(TH)$	$M(TB)$	$M(TBH)$
	0.63	1.00	0.15	0.37	0.00	0.24	0.22	0.41

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

↑  
Model  $M_t$  is a very poor model for these data

Fig. 6.4b. The results of the model selection procedure, using data from both sexes.

CHAPTER 6 EXAMPLE 4 - SEXES COMBINED

NUMBER OF TRAPPING OCCASIONS HAS 8  
 NUMBER OF ANIMALS CAPTURED,  $N(T+1)$ , WAS 194  
 TOTAL NUMBER OF CAPTURES,  $N$ , WAS 340

FREQUENCIES OF CAPTURE,  $F(I)$   
 $I = 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8$   
 $F(I) = 109 \ 47 \ 26 \ 5 \ 3 \ 4 \ 0 \ 0$

COMPUTED JACKKNIFE COEFFICIENTS

	N(1)	N(2)	N(3)	N(4)	N(5)
1	1.875	2.625	3.250	3.750	4.125
2	1.000	0.357	-.625	-1.696	-2.661
3	1.000	1.000	1.372	2.098	2.964
4	1.000	1.000	1.000	0.848	0.535
5	1.000	1.000	1.000	1.000	1.036

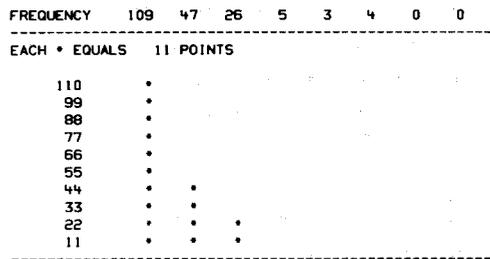
THE RESULTS OF THE JACKKNIFE COMPUTATIONS

J	N(I)	SE(I)	0.95 CONF. LIMITS		TEST OF $N(I+1)$ VS. $N(I)$
0	194				CHI-SQUARE(1 D.F.)
1	289.4	13.37	263.2	315.6	39.410
2	340.9	21.31	299.1	382.7	11.531
3	372.5	29.29	315.1	430.0	5.331
4	394.8	37.39	321.5	468.1	3.540
5	411.4	44.87	323.5	499.4	0.000

AVERAGE P-HAT = 0.1087

INTERPOLATED POPULATION ESTIMATE IS 391 WITH STANDARD ERROR 35.9808  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 320 TO 462

HISTOGRAM OF  $F(I)$



Average catchability of the population is estimated to be 0.11 - fairly low

Fig. 6.4c. Population estimation of the total population, using the Model  $M_h$  procedure.

CHAPTER 6 EXAMPLE 4 - MALES

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	$N(J) =$	18	17	18	28	24	25	26	15
TOTAL CAUGHT	$M(J) =$	0	18	33	46	67	82	93	102
NEWLY CAUGHT	$U(J) =$	18	15	13	21	15	11	9	6
FREQUENCIES	$F(J) =$	66	27	11	3	0	1	0	0

Fig. 6.4d. The summary statistics for the males.

There were 5 (=18-13) recaptures on the third trapping occasion

CHAPTER 6 EXAMPLE 4 - MALES

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(H)$   
 CHI-SQUARE VALUE = 3.844 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.27886
  
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(B)$   
 CHI-SQUARE VALUE = 0.511 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.47478
  
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(T)$   
 CHI-SQUARE VALUE = 9.256 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.23480
  
4. GOODNESS OF FIT TEST OF MODEL  $M(H)$   
 NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(H)$   
 CHI-SQUARE VALUE = 9.299 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.23192  
 TEST OF MODEL  $M(H)$  BY FREQUENCY OF CAPTURE  
 (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)  

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	6.970	7	0.43204
2	11.321	7	0.12522
  
5. GOODNESS OF FIT TEST OF MODEL  $M(B)$   
 NULL HYPOTHESIS OF MODEL  $M(B)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(B)$   
 CHI-SQUARE VALUE = 10.600 DEGREES OF FREEDOM = 12 PROBABILITY OF LARGER VALUE = 0.56346
  
- 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 6.142 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.40743
  
- 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 4.458 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.61497
  
6. GOODNESS OF FIT TEST OF MODEL  $M(T)$   
 NULL HYPOTHESIS OF MODEL  $M(T)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(T)$   
 CHI-SQUARE VALUE = 26.084 DEGREES OF FREEDOM = 20 PROBABILITY OF LARGER VALUE = 0.16305
  
7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(BH)$   
 CHI-SQUARE VALUE = 12.078 DEGREES OF FREEDOM = 14 PROBABILITY OF LARGER VALUE = 0.60001

No evidence of significant variation among daily captures

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	$M(O)$	$M(H)$	$M(B)$	$M(BH)$	$M(T)$	$M(TH)$	$M(TB)$	$M(TBH)$
	1.00	1.00	0.37	0.60	0.00	0.53	0.41	0.72

APPROPRIATE MODEL PROBABLY IS  $M(H)$  OR  $M(O)$   
 SUGGESTED ESTIMATOR IS JACKKNIFE.

These two models have tied for best model - The jackknife (Model  $M_H$ ) estimator is chosen because it is a more robust (see Chapter 2) estimator

Fig. 6.4e. The results of the model selection procedure for males.

# A total of 108 males were captured during the experiment

CHAPTER 6 EXAMPLE 4 - MALES

NUMBER OF TRAPPING OCCASIONS WAS 8  
 NUMBER OF ANIMALS CAPTURED,  $M(T+1)$ , WAS 108  
 TOTAL NUMBER OF CAPTURES,  $N$ , WAS 171

FREQUENCIES OF CAPTURE,  $F(I)$   
 $I = 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8$   
 $F(I) = 66 \ 27 \ 11 \ 3 \ 0 \ 1 \ 0 \ 0$

COMPUTED JACKKNIFE COEFFICIENTS

	N(1)	N(2)	N(3)	N(4)	N(5)
1	1.875	2.625	3.250	3.750	4.125
2	1.000	0.357	-.625	-1.696	-2.661
3	1.000	1.000	1.372	2.098	2.964
4	1.000	1.000	1.000	0.848	0.535
5	1.000	1.000	1.000	1.000	1.036

THE RESULTS OF THE JACKKNIFE COMPUTATIONS

I	N(I)	SE(I)	0.95 CONF. LIMITS		TEST OF $N(I+1)$ VS. $N(I)$ CHI-SQUARE(1 D.F.)
0	108				
1	165.8	10.41	145.4	186.1	26.438
2	197.9	16.59	165.4	230.4	7.012
3	216.7	22.71	172.2	261.2	2.559
4	228.3	28.79	171.9	284.8	1.245
5	235.6	34.31	168.4	302.9	0.000

AVERAGE P-HAT = 0.1013

INTERPOLATED POPULATION ESTIMATE IS 211 WITH STANDARD ERROR 20.8709

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 170 TO 253

HISTOGRAM OF  $F(I)$

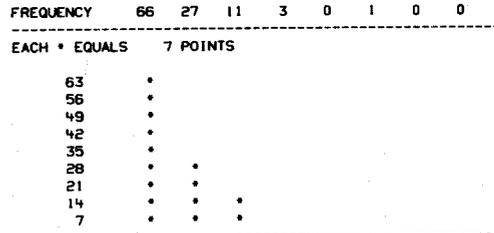


Fig. 6.4f. Population estimation of the male population, using the Model  $M_h$  procedure.

CHAPTER 6 EXAMPLE 4 - FEMALES

OCCASION	J=	1	2	3	4	5	6	7	8
ANIMALS CAUGHT	$N(J)=$	26	21	26	19	21	19	19	18
TOTAL CAUGHT	$M(J)=$	0	26	41	56	62	70	79	83
NEWLY CAUGHT	$U(J)=$	26	15	15	6	8	9	4	3
FREQUENCIES	$F(J)=$	43	20	15	2	3	3	0	0

Fig. 6.4g. The summary statistics for the females.

CHAPTER 6 EXAMPLE 4 - FEMALES

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 26.293 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00001
  
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 1.360 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.24354
  
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 3.966 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.78365
  
4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 4.515 DEGREES OF FREEDOM = 7 PROBABILITY OF LARGER VALUE = 0.71887  
 TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE  
 (FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)  
 NUMBER OF CAPTURES CHI-SQUARE D.F. PROBABILITY  
 -----  

1	8.535	7	0.28780
2	3.733	7	0.80993
  
5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 7.977 DEGREES OF FREEDOM = 12 PROBABILITY OF LARGER VALUE = 0.78693
  
- 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 4.548 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.60292
  
- 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 3.429 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.75342
  
6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 CHI-SQUARE VALUE = 47.103 DEGREES OF FREEDOM = 25 PROBABILITY OF LARGER VALUE = 0.00477
  
7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 14.969 DEGREES OF FREEDOM = 16 PROBABILITY OF LARGER VALUE = 0.52689

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
CRITERIA	0.58	1.00	0.20	0.48	0.00	0.25	0.20	0.50

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

Fig. 6.4h. The results of the model selection procedure for females.

No females were captured every day

CHAPTER 6 EXAMPLE 4 - FEMALES

NUMBER OF TRAPPING OCCASIONS WAS 8  
 NUMBER OF ANIMALS CAPTURED,  $M(T+1)$ , WAS 86  
 TOTAL NUMBER OF CAPTURES,  $N$ , WAS 169

FREQUENCIES OF CAPTURE,  $F(I)$   
 I= 1 2 3 4 5 6 7 8  
 $F(I)$ = 43 20 15 2 3 3 0 0

COMPUTED JACKKNIFE COEFFICIENTS

	N(1)	N(2)	N(3)	N(4)	N(5)
1	1.875	2.625	3.250	3.750	4.125
2	1.000	0.357	-1.625	-1.696	-2.661
3	1.000	1.000	1.372	2.098	2.964
4	1.000	1.000	1.000	0.848	0.535
5	1.000	1.000	1.000	1.000	1.036

THE RESULTS OF THE JACKKNIFE COMPUTATIONS

I	N(I)	SE(I)	0.95 CONF. LIMITS		TEST OF $N(I+1)$ VS. $N(I)$ CHI-SQUARE (1 D.F.)
0	86				
1	123.6	8.40	107.2	140.1	13.238
2	143.0	13.37	116.8	169.2	4.475
3	155.8	18.50	119.6	192.1	2.784
4	166.5	23.86	119.7	213.3	2.444
5	175.8	28.91	119.1	232.5	0.000

AVERAGE P-HAT = 0.1427

INTERPOLATED POPULATION ESTIMATE IS 148 WITH STANDARD ERROR 15.2086  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 118 TO 178

HISTOGRAM OF  $F(I)$

FREQUENCY	43	20	15	2	3	3	0	0
EACH * EQUALS	5 POINTS							
45	*							
40	*							
35	*							
30	*							
25	*							
20	*	*						
15	*	*	*					
10	*	*	*	*	*	*		
5	*	*	*	*	*	*	*	*

A coefficient of variation of  $\approx 10\%$  - not bad!

Fig. 6.4i. Population estimation of the female population, using the Model  $M_h$  procedure.

researcher is assured that the most appropriate model selected for the data will produce an estimate. This fact has important implications for study design. Clearly, the experimenter should do everything possible to avoid the situation in which one of these “no estimator” models is most appropriate for the experiment.

Usually we think of weather changes as the cause of differential trap success among trapping occasions, or significant time variation. The reader may ask whether we are suggesting that uniform weather conditions be designed into the experiment. The answer is no. We realize that such designs are difficult to come by. However, another parameter of the experiment is controllable and also is a potential cause of time variation. This parameter is effort; it can be measured both in terms of the number of traps used and the frequency with which they are checked. Thus, to help achieve uniform trapping success, the same number of traps should be used on each trapping occasion (usually a day), and the traps should be checked the same number of times each day. To illustrate this point, let us consider the results of an experiment on Richardson’s ground squirrels (*Spermophilus richardsoni*) in an abandoned wheat field in the rangeland of southwestern Montana.

In the experiment, 93 livetraps were placed by active burrow holes in the 1.5-ha (3.8-acre) study plot, baited with rolled oats, and checked for 6 consecutive days. Although the entire group of traps was checked more than once daily, the trapping occasion was defined as 1 day, and therefore the captures resulting from separate trap checks within the day were pooled.

On the first day of trapping, traps were checked twice, but on each of the remaining days, they were checked three times. The reason for this is simple—on the first day, there was not time to check the traps three times because trapping success was high and every animal had to be handled and tagged. On subsequent days, recaptures (which demand less time than new captures) increased, and the researchers became more efficient. In the summary statistics for the trapping experiment (Fig. 6.5a) the daily captures reflect almost perfectly the effect of unequal trapping effort. The number caught on the first day is only about two-thirds of the numbers caught on each remaining day. Thus, as the model selection procedure (Fig. 6.5b) indicates, significant time variation is included in the selected Model  $M_{tbb}$ . However, Figs. 6.5c-e reveal the excellent results produced when we use only the data from the last 5 days of trapping, when effort was equal. The lesson here is plain: devote equal effort to each trapping occasion or risk the chance that trapping effort on some occasions may be wasted. Equal effort on each trapping occasion is a fundamental assumption of all eight models described in Chapter 3.

### Example 6. How Many Are Not Enough?

We frequently emphasize that we do not recommend using capture-recapture estimation procedures when the population to be trapped is expected to be “small.” The researcher often responds by asking, “Why not, when I have used capture-recapture before in such situations and obtained reasonable-looking population estimates with small standard errors?” Such a question is to be expected, because of data sets like the one presented in Fig. 6.6a. The data were collected from a study site in Idaho, which had been clearcut in the 1960s and recently replanted with lodgepole pine (*Pinus contorta*) seedlings. A 10 by 10

CHAPTER 6 EXAMPLE 5 - ALL OCCASIONS

OCCASION	J=	1	2	3	4	5	6
ANIMALS CAUGHT	N(J)=	127	193	180	176	173	193
TOTAL CAUGHT	M(J)=	0	127	254	319	356	385
NEWLY CAUGHT	U(J)=	127	127	65	37	29	21
FREQUENCIES	F(J)=	121	102	74	68	23	18

Fig. 6.5a. The summary statistics for an experiment on Richardson’s ground squirrels, illustrating the importance of equal effort on all trapping occasions.

On Day 2, 66 more animals were captured than on Day 1

CHAPTER 6 EXAMPLE 5 - ALL OCCASIONS

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(H)  
 CHI-SQUARE VALUE = 181.481 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(B)  
 CHI-SQUARE VALUE = 19.438 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00001

3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
 NULL HYPOTHESIS OF MODEL M(O) VS. ALTERNATE HYPOTHESIS OF MODEL M(T)  
 CHI-SQUARE VALUE = 31.975 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.00001

4. GOODNESS OF FIT TEST OF MODEL M(H)  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(H)  
 CHI-SQUARE VALUE = 32.462 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.00001

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

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	6.587	5	0.25323
2	6.912	5	0.22729
3	20.450	5	0.00103
4	13.199	5	0.02159
5	9.087	5	0.10565

Strong evidence of time variation

5. GOODNESS OF FIT TEST OF MODEL M(B)  
 NULL HYPOTHESIS OF MODEL M(B) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(B)  
 CHI-SQUARE VALUE = 17.554 DEGREES OF FREEDOM = 8 PROBABILITY OF LARGER VALUE = 0.02483

5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
 CHI-SQUARE VALUE = 12.212 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.01585

5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
 CHI-SQUARE VALUE = 5.342 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.25394

6. GOODNESS OF FIT TEST OF MODEL M(T)  
 NULL HYPOTHESIS OF MODEL M(T) VS. ALTERNATE HYPOTHESIS OF NOT MODEL M(T)  
 CHI-SQUARE VALUE = 519.588 DEGREES OF FREEDOM = 316 PROBABILITY OF LARGER VALUE = 0.00000

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
 NULL HYPOTHESIS OF MODEL M(H) VS. ALTERNATE HYPOTHESIS OF MODEL M(BH)  
 CHI-SQUARE VALUE = 39.514 DEGREES OF FREEDOM = 15 PROBABILITY OF LARGER VALUE = 0.00054

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
	0.57	0.49	0.46	0.87	0.00	0.46	0.49	1.00

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

Selection of Model  $M_{TBH}$  is to be expected because all null hypotheses tested above (with the exception of Test 5b) were rejected

Fig. 6.5b. Output from the model selection procedure.

CHAPTER 6 EXAMPLE 5 - OCCASION 1 ELIMINATED

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	193	180	176	173	193	
TOTAL CAUGHT	M(J)=	0	193	294	336	367	393
NEWLY CAUGHT	U(J)=	193	101	42	31	26	
FREQUENCIES	F(J)=	132	102	82	52	25	

Fig. 6.5c. The summary statistics for the data from only the last 5 days of trapping.

grid of 100 livetraps spaced at 15-m intervals was used to trap the area for 5 consecutive days, to estimate the size of the small mammal population. The traps, baited with a mixture of rolled oats and peanut butter, were checked in the morning and evening. (The data in Fig. 6.6a have been pooled to obtain daily captures.)

The results of the model selection and estimation tasks are given in Figs. 6.6b and 6.6c. Only Test 7 indicates variation in capture probabilities, and samples are too small to compute two of the tests. Thus, the choice of  $M_0$ , the simplest model, as most appropriate is not surprising. Although the estimated population size is the same as the number of animals captured, this estimate seems reasonable and appears to be very precise.

Why then do we recommend that results obtained from experiments in which very few animals are captured not be trusted? Let us view the results of this example in light of a small simulation study that was conducted to evaluate model selection and estimation procedures when the population is small and trapping success moderate. The study generated 100 sets of data from each of the 8 models. Population size was 50, the average probability of capture was about 0.20 in each model, and trapping was done on 6 occasions.

The following relevant conclusions arose from the study.

1. The procedures performed satisfactorily when the data were generated from Model  $M_0$  or Model  $M_t$ .
2. Model  $M_0$  was chosen as the most appropriate model in nearly 75% of the data sets generated from Models  $M_b$ ,  $M_h$ ,  $M_{bh}$ , and  $M_{th}$ . In these instances, the estimate averaged 42.5, and the confidence interval coverage was about 50%. The correct model was selected only about 8% of the time.
3. The correct model was selected in 62 of the 100 data sets generated from  $M_{tb}$ ;  $M_0$  was selected only 3 times, and these 3 estimates averaged 72.3.
4. The correct model was selected in only 1 of the 100 data sets generated from  $M_{tth}$ ;  $M_0$  was selected 4 times, and these estimates averaged 59.5.

How do these results affect our interpretation of the real world example? The second point above has particular relevance for us; it says that the chances are very good that Model  $M_0$  is not a good model for the data, that the estimate therefore is biased, and that the presumed 95% confidence interval is probably closer to a 50% confidence interval.

Taken as a whole, this example illustrates one central point: unless both behavioral response and heterogeneity are absent, the model selection and estimation procedures are likely to produce misleading results when small populations are involved. The answer will be precise, with a small confidence interval, but misleading.

### Example 7. Density Estimation

The method described in Chapter 5 for estimating density is again illustrated here, by a data set collected from a population of Richardson's ground squirrels. Trapping was done on a 10 by 10 grid with one trap per station and a 10-m trap spacing, for five trapping occasions. The number of animals captured per station and the definition of the four subgrids to be used to estimate density are given in Fig. 6.7a.

The results of three tests of a null hypothesis concerning uniform density over the grid are given in Fig. 6.7b. Although there is no evidence of a gradient in the direction of columns or with respect to distance

Time specific variation has been eliminated by excluding the first day of trapping

CHAPTER 6 EXAMPLE 5 - OCCASION 1 ELIMINATED

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(H)$   
CHI-SQUARE VALUE = 87.624 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.00000
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(B)$   
CHI-SQUARE VALUE = 0.078 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.78022
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(T)$   
CHI-SQUARE VALUE = 3.499 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.47806

4. GOODNESS OF FIT TEST OF MODEL  $M(H)$   
NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(H)$   
CHI-SQUARE VALUE = 3.891 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.42092  
TEST OF MODEL  $M(H)$  BY FREQUENCY OF CAPTURE  
(FREQUENCIES LESS THAN 2T ARE NOT CALCULATED.)

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	4.061	4	0.39787
2	1.203	4	0.87767
3	6.455	4	0.16763
4	15.308	4	0.00410

The data are now fit nicely by Model  $M_h$

5. GOODNESS OF FIT TEST OF MODEL  $M(B)$   
NULL HYPOTHESIS OF MODEL  $M(B)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(B)$   
CHI-SQUARE VALUE = 10.248 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.11458
- 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
CHI-SQUARE VALUE = 8.473 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.03719
- 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
CHI-SQUARE VALUE = 1.776 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.62028
6. GOODNESS OF FIT TEST OF MODEL  $M(T)$   
NULL HYPOTHESIS OF MODEL  $M(T)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(T)$   
CHI-SQUARE VALUE = 455.905 DEGREES OF FREEDOM = 292 PROBABILITY OF LARGER VALUE = 0.00000

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(BH)$   
CHI-SQUARE VALUE = 11.025 DEGREES OF FREEDOM = 10 PROBABILITY OF LARGER VALUE = 0.35554

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	$M(O)$	$M(H)$	$M(B)$	$M(BH)$	$M(T)$	$M(TH)$	$M(TB)$	$M(TBH)$
	0.69	1.00	0.29	0.41	0.00	0.34	0.24	0.56

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

Model  $M_t$  continues to be a poor model for these data

Fig. 6.5d. Output from the model selection procedure when Day 1 is eliminated.

CHAPTER 6 EXAMPLE 5 - OCCASION 1 ELIMINATED

NUMBER OF TRAPPING OCCASIONS WAS 5  
 NUMBER OF ANIMALS CAPTURED,  $M(T+1)$ , WAS 393  
 TOTAL NUMBER OF CAPTURES,  $N$ , WAS 915

FREQUENCIES OF CAPTURE,  $F(I)$   
 $I = 1 \ 2 \ 3 \ 4 \ 5$   
 $F(I) = 132 \ 102 \ 82 \ 52 \ 25$

COMPUTED JACKKNIFE COEFFICIENTS

	N(1)	N(2)	N(3)	N(4)	N(5)
1	1.600	2.400	2.800	3.000	3.000
2	1.000	0.550	0.050	-.250	-.250
3	1.000	1.000	1.133	1.250	1.250
4	1.000	1.000	1.000	0.992	0.992
5	1.000	1.000	1.000	1.000	1.000

THE RESULTS OF THE JACKKNIFE COMPUTATIONS

I	N(I)	SE(I)	0.95 CONF. LIMITS		TEST OF $N(I+1)$ VS. $N(I)$ CHI-SQUARE(1 D.F.)
0	393				16.924
1	498.6	13.79	471.6	525.6	3.393
2	531.9	20.45	491.8	572.0	1.564
3	544.6	25.94	493.8	595.5	0.000
4	549.6	29.14	492.5	606.7	0.000
5	549.6	29.14	492.5	606.7	0.000

AVERAGE P-HAT = 0.3446

INTERPOLATED POPULATION ESTIMATE IS 531 WITH STANDARD ERROR 20.2128  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 491 TO 571

HISTOGRAM OF  $F(I)$

FREQUENCY	132	102	82	52	25
EACH * EQUALS	14 POINTS				
126	.				
112	.				
98	.	.			
84	.	.	.		
70	.	.	.	.	
56	.	.	.	.	.
42	.	.	.	.	.
28	.	.	.	.	.
14	.	.	.	.	.

High average probability of capture,  
 and a corresponding small standard  
 error are two qualities the  
 experimenter should strive for

Fig. 6.5e. Estimation of population size, using the jackknife estimator (Model  $M_h$ ) on data from Days 2-6.

CHAPTER 6 EXAMPLE 6

OCCASION	J=	1	2	3	4	5	
ANIMALS CAUGHT	N(J)=	13	16	17	18	15	
TOTAL CAUGHT	M(J)=	0	13	20	23	24	24
NEWLY CAUGHT	U(J)=	13	7	3	1	0	
FREQUENCIES	F(J)=	4	3	4	8	5	

Fig. 6.6a. The summary statistics illustrating a trapping experiment on small mammals in which few animals were captured.

Only 24 different animals were captured in 5 days of trapping

from the center of the grid, there is evidence ( $P < 0.05$ ) that density is greater on the bottom half of the grid. In this instance, however, the estimation procedure should not be biased with respect to estimating average density on the grid, because of the configuration of the nested subgrids. The area of every subgrid is divided evenly between the “high” and “low” density areas. This advantage of the nested design also should hold if the density pattern in the direction of columns is similar.

Estimates of population size on each of the subgrids are presented in Figs. 6.7c-6.7f. Previous analysis of the entire data set, using the model selection procedure, indicated that the Model  $M_{bh}$  estimator provided the most reliable estimates, and therefore this estimator was chosen for use on all subgrids. In general, adequate fits to the data were obtained for all four data sets.

Figure 6.7g first gives the starting values for parameters essential to the actual density estimation procedure. Next are the naive estimates (no edge effect assumed) of density, produced by dividing the population size estimate by the subgrid area, and the values of parameters necessary to the regression analysis. The estimates of size increase as the grids become larger, but the naive density estimates decrease because the importance of the boundary strip diminishes as the grid size increases. However, we will see that a boundary strip is necessary even with the entire grid. Next, a list of comparisons indicates good agreement between the naive estimates and estimates predicted from the nonlinear regression routine; it is followed by estimates of density and boundary strip width. The estimated density of 64.9 squirrels/ha, which is much smaller than the smallest naive estimate from the total grid, suggests a need to adjust for the edge effect. Such an adjustment, here the addition of a strip 11.8 m wide around the grid, is justified further by the significance of the test for nonzero strip width ( $P < 0.01$ ). From the estimated standard errors for the density and strip width estimates also provided, we see that the density estimate is not especially precise ( $cv = 16\%$ ), nor is the approximately 95% confidence interval of 44.5-85.4 very narrow. Use of a larger grid probably would have increased the precision. (See Chapter 7 for recommendations regarding grid size.)

Biologists interested in running program CAPTURE to obtain density estimates may be interested in the following information. The data matrix  $X_{ij}$  was read by using TASK READ CAPTURES. (Only XY REDUCED or XY COMPLETE data formats are allowed because the trap coordinates are necessary.) TASK UNIFORM DENSITY TEST produced Fig. 6.7b.

Model  $M_{bh}$  was selected by using the output from TASK MODEL SELECTION, although Model  $M_{tth}$  was ranked highest because of an apparently high probability of capture on day 3.

The density estimation method is called by TASK DENSITY ESTIMATE and starts with the four naive density estimates (Figs. 6.7c-6.7f), the covariance matrix expressing the degree of overlap among the nested grids, the coefficients A(I) and B(I), and the starting values provided by the user. In this example, 59 iterations were required to maximize the multiple correlation coefficient to within a small tolerance value. The procedure simultaneously estimates D and W. It can not be performed on a hand calculator because of the large amount of computation required. Additional details on program CAPTURE are given in *White et al. (1978)*.

The data set is too inadequate for the performance of the test

CHAPTER 6 EXAMPLE 6

1. TEST FOR HETEROGENEITY OF TRAPPING PROBABILITIES IN POPULATION.  
NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(H)$   
EXPECTED VALUES TOO SMALL. TEST NOT PERFORMED.
2. TEST FOR BEHAVIORAL RESPONSE AFTER INITIAL CAPTURE.  
NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(B)$   
CHI-SQUARE VALUE = 0.722 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.39557
3. TEST FOR TIME SPECIFIC VARIATION IN TRAPPING PROBABILITIES.  
NULL HYPOTHESIS OF MODEL  $M(O)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(T)$   
CHI-SQUARE VALUE = 2.777 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.59572
4. GOODNESS OF FIT TEST OF MODEL  $M(H)$   
NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(H)$   
CHI-SQUARE VALUE = 3.289 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.51069

The test statistics may not be closely approximated by chi-square distributions due to small sample size (see Fig. 2.12)

5. GOODNESS OF FIT TEST OF MODEL  $M(B)$   
NULL HYPOTHESIS OF MODEL  $M(B)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(B)$   
CHI-SQUARE VALUE = 1.416 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.92254
- 5A. CONTRIBUTION OF TEST OF HOMOGENEITY OF FIRST CAPTURE PROBABILITY ACROSS TIME  
CHI-SQUARE VALUE = 0.678 DEGREES OF FREEDOM = 2 PROBABILITY OF LARGER VALUE = 0.71233
- 5B. CONTRIBUTION OF TEST OF HOMOGENEITY OF RECAPTURE PROBABILITIES ACROSS TIME  
CHI-SQUARE VALUE = 0.738 DEGREES OF FREEDOM = 3 PROBABILITY OF LARGER VALUE = 0.86431
6. GOODNESS OF FIT TEST OF MODEL  $M(T)$   
NULL HYPOTHESIS OF MODEL  $M(T)$  VS. ALTERNATE HYPOTHESIS OF NOT MODEL  $M(T)$   
EXPECTED VALUES TOO SMALL. TEST NOT PERFORMED.

7. TEST FOR BEHAVIORAL RESPONSE IN PRESENCE OF HETEROGENEITY.  
NULL HYPOTHESIS OF MODEL  $M(H)$  VS. ALTERNATE HYPOTHESIS OF MODEL  $M(BH)$   
CHI-SQUARE VALUE = 6.125 DEGREES OF FREEDOM = 2 PROBABILITY OF LARGER VALUE = 0.04677

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL CRITERIA	$M(O)$	$M(H)$	$M(B)$	$M(BH)$	$M(T)$	$M(TH)$	$M(TB)$	$M(TBH)$
	1.00	0.81	0.37	0.64	0.00	0.17	0.35	0.73

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

The data are again inadequate for computation of the test

Due to small sample size, the tests do not have much power for detecting the sources of unequal catchability, and thus the model selection procedure defaults to Model  $M_0$

Fig. 6.6b. Output from the model selection procedure.

CHAPTER 6 EXAMPLE 6

NUMBER OF TRAPPING OCCASIONS WAS 5  
 NUMBER OF ANIMALS CAPTURED,  $M(T+1)$ , WAS 24  
 TOTAL NUMBER OF CAPTURES,  $N$ , WAS 79  
 ESTIMATED PROBABILITY OF CAPTURE,  $P\text{-HAT} = 0.6583$

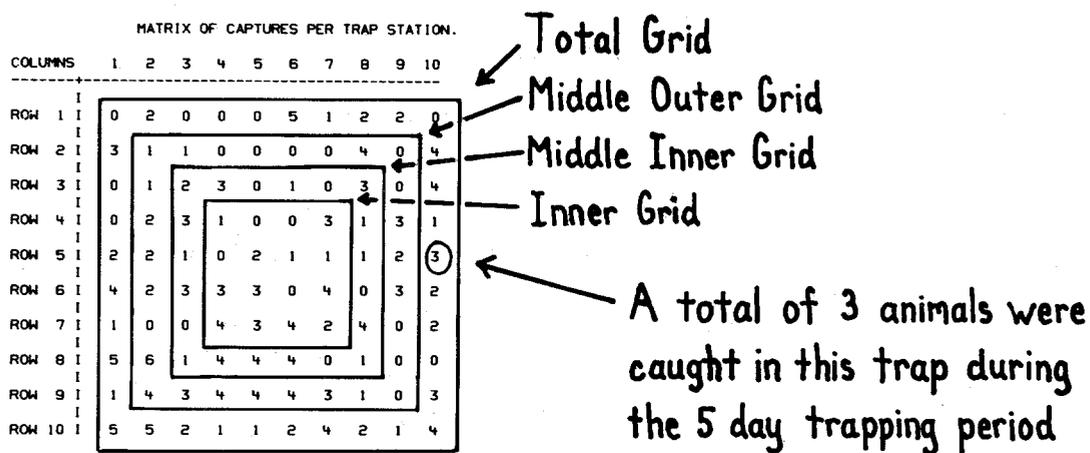
POPULATION ESTIMATE IS 24 WITH STANDARD ERROR 0.3429  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 23 TO 25

A deceptively narrow confidence interval often results from studies in which very few animals are captured, and confidence levels are very much less than the stated level of 95 percent

Fig. 6.6c. Estimation of population size using the null estimator of Model  $M_0$ .

MARK-RECAPTURE POPULATION AND DENSITY ESTIMATION PROGRAM DEVELOPED BY THE UTAH COOPERATIVE WILDLIFE RESEARCH UNIT. PAGE 881  
 CAPTURE RECAPTURE WORKBOOK EXAMPLES PROGRAM VERSION OF MAY 07, 1980 80/07/26.

TEST FOR UNIFORM DENSITY. SEE THIS SECTION OF THE MONOGRAPH FOR DETAILS.  
 CHAPTER 6 EXAMPLE 7



IN THE ABOVE MATRIX, TRAP COORDINATES ARE ROUNDED TO THE NEAREST WHOLE INTEGER.  
 IN THE FOLLOWING GOODNESS OF FIT TESTS, TRAP COORDINATES THAT ARE NOT INTEGERS  
 AND NON-RECTANGULAR TRAPPING GRIDS WILL CAUSE SPURIOUS RESULTS.

Fig. 6.7a. The grid definitions and capture matrix for an estimation of Richardson's ground squirrel density.

TEST FOR UNIFORM DENSITY. SEE THIS SECTION OF THE MONOGRAPH FOR DETAILS.  
 CHAPTER 6 EXAMPLE 7

## Evidence of nonuniform density in a horizontal (north-south) direction

CHI-SQUARE TEST OF UNIFORM DENSITY BY ROWS.

ROW	1	2	3	4	5	6	7	8	9	10
OBSERVED	12	13	14	14	15	24	20	25	27	27
EXPECTED	19.100	19.100	19.100	19.100	19.100	19.100	19.100	19.100	19.100	19.100
CHI-SQUARE	2.639	1.948	1.362	1.362	0.880	1.257	0.042	1.823	3.268	3.268

TOTAL CHI-SQUARE = 17.85 WITH 9 DEGREES OF FREEDOM. PROBABILITY OF LARGER VALUE = 0.0370

CHI-SQUARE TEST OF UNIFORM DENSITY BY COLUMNS.

COLUMN	1	2	3	4	5	6	7	8	9	10
OBSERVED	21	25	16	20	17	21	18	19	11	23
EXPECTED	19.100	19.100	19.100	19.100	19.100	19.100	19.100	19.100	19.100	19.100
CHI-SQUARE	0.189	1.823	0.503	0.042	0.231	0.189	0.063	0.001	3.435	0.796

TOTAL CHI-SQUARE = 7.27 WITH 9 DEGREES OF FREEDOM. PROBABILITY OF LARGER VALUE = 0.6088

CHI-SQUARE TEST OF UNIFORM DENSITY BY RINGS (OUTER RING IS NUMBER 1).

RING	1	2	3	4	5
OBSERVED	74	50	36	25	6
EXPECTED	68.760	53.480	38.200	22.920	7.640
CHI-SQUARE	0.399	0.226	0.127	0.189	0.352

TOTAL CHI-SQUARE = 1.29 WITH 4 DEGREES OF FREEDOM. PROBABILITY OF LARGER VALUE = 0.8625

## Number of animals captured in the eighth column of traps during the trapping period

Fig. 6.7b. Tests of three null hypotheses concerning uniformity of density over the grid.

POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 INNER GRID X=4-7 Y=4-7

OCCASION	J=	1	2	3	4	5
TOTAL CAUGHT	M(J)=	0	4	9	13	16
NEWLY CAUGHT	U(J)=	4	5	4	3	0

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 5				
1	17.67	2.872097	3.801	0.2837	0.3454	0.3454	0.3454	0.3454	0.3454
2	16.00	1.027285	3.340	0.1883	0.2500	0.5454	0.5454	0.5454	0.5454
3	16.00	0.5785123	2.198	0.1382	0.2500	0.4167	0.7000	0.7000	0.7000

POPULATION ESTIMATE IS 18 WITH STANDARD ERROR 2.8731

APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 12 TO 24

HISTOGRAM OF U(J)

FREQUENCY	4	5	4	3	0
5	.	.	.	.	.
4	.	.	.	.	.
3	.	.	.	.	.
2	.	.	.	.	.
1	.	.	.	.	.

Fig. 6.7c. Estimation of population size on the inner grid.

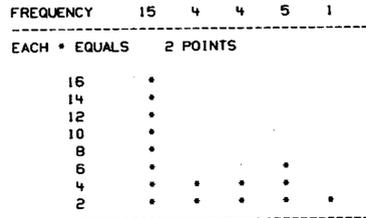
POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 MIDDLE INNER GRID X=3-8 Y=3-8

OCCASION	J=	1	2	3	4	5
TOTAL CAUGHT	M(J)=	0	15	19	23	29
NEWLY CAUGHT	U(J)=	15	4	4	5	1

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 5				
1	30.15	1.923854	4.969	0.1741	0.4412	0.4412	0.4412	0.4412	0.4412
2	32.44	5.616076	3.022	0.2207	0.4625	0.3129	0.3129	0.3129	0.3129
3	29.49	1.874384	2.727	0.0987	0.5086	0.2760	0.5410	0.5410	0.5410

POPULATION ESTIMATE IS  $N_2$  32 WITH STANDARD ERROR 5.6161  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 20 TO 44

HISTOGRAM OF U(J)



Average probability of capture of those members of the population not captured on the first occasion

Fig. 6.7d. Estimation of population size on the middle inner grid.

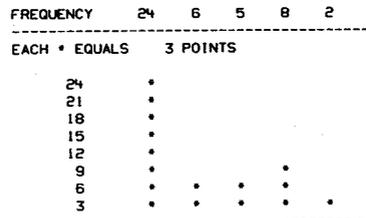
POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 MIDDLE OUTER GRID X=2-9 Y=2-9

OCCASION	J=	1	2	3	4	5
TOTAL CAUGHT	M(J)=	0	24	30	35	43
NEWLY CAUGHT	U(J)=	24	6	5	8	2

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 5				
1	47.41	2.633526	8.462	0.0374	0.4284	0.4284	0.4284	0.4284	0.4284
2	55.12	13.53864	4.022	0.1339	0.4354	0.2373	0.2373	0.2373	0.2373
3	48.71	5.969894	3.742	0.0531	0.4928	0.2429	0.3935	0.3935	0.3935

POPULATION ESTIMATE IS  $N_3$  55 WITH STANDARD ERROR 13.5386  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 28 TO 82

HISTOGRAM OF U(J)



Very wide confidence interval

Fig. 6.7e. Estimation of population size on the middle outer grid.

POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 TOTAL GRID X=1-10 Y=1-10

OCCASION	J=	1	2	3	4	5	
TOTAL CAUGHT	M(J)=	0	33	49	62	73	76
NEWLY CAUGHT	U(J)=	33	16	13	11	3	

K	N-HAT	SE(N)	CHI-SQ.	PROB.	ESTIMATED P-BAR(J), J=1, ..., 5				
1	82.73	4.642233	3.065	0.3817	0.3865	0.3865	0.3865	0.3865	0.3865
2	84.17	6.955550	2.713	0.2576	0.3921	0.3593	0.3593	0.3593	0.3593
3	79.21	3.955951	2.044	0.1528	0.4166	0.3462	0.5033	0.5033	0.5033

Good fit of the model  
to the observed data

POPULATION ESTIMATE IS  $N_4$  83 WITH STANDARD ERROR 4.6422  
 APPROXIMATE 95 PERCENT CONFIDENCE INTERVAL 73 TO 93

Relatively precise estimate

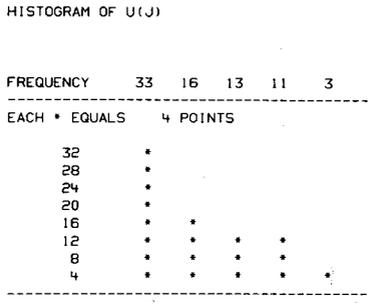


Fig. 6.7f. Estimation of population size on the total grid.

JOINT ESTIMATION OF DENSITY AND BOUNDARY STRIP WIDTH FROM CAPTURE DATA. SEE THIS SECTION OF THE MONOGRAPH FOR DETAILS.  
 CHAPTER 6 EXAMPLE 7

STARTING VALUES FOR DENSITY ESTIMATION--  
 NUMBER OF GRIDS 4  
 TRAP INTERVAL 10.00  
 UNITS CONVERSION 10000.00  
 INITIAL DENSITY ESTIMATE 75.3261  
 INITIAL STRIP WIDTH ESTIMATE 8.5581

Density will be in terms  
 of animals/hectare

GRID I	NAIVE DENSITY Y(I)	PERIMETER/AREA A(I)	PI/AREA B(I)	STARTING COVARIANCE MATRIX			
1	200.0000	0.1333333	0.3491E-02	0.102E+04			
2	128.0000	0.8000000E-01	0.1257E-02	353.	505.		
3	112.2449	0.5714286E-01	0.6411E-03	258.	368.	763.	
4	102.4691	0.4444444E-01	0.3879E-03	35.4	50.5	105.	32.8

RESULTS OF ITERATIONS  
 FUNCTION EVALUATIONS REQUIRED 65  
 ESTIMATED SIGNIFICANT DIGITS OF PARAMETER VALUES 9

Density estimates from each of  
 the subgrids assuming no edge effect

FITTED MODEL COMPARED TO THE DATA

GRID(I)	Y(I)	F(I)
1	200.000	198.641
2	128.000	137.594
3	112.245	114.520
4	102.469	102.503

A value close to 1 indicates a good  
 fit of the model to the data

MULTIPLE CORRELATION COEFFICIENT IS 0.99421

ESTIMATED DENSITY= 64.949 10.4286 = ITS STANDARD ERROR  
 ESTIMATED STRIP WIDTH= 11.796 4.3152 = ITS STANDARD ERROR  
 CORRELATION OF ESTIMATORS -.9580

TEST OF ESTIMATED STRIP WIDTH GREATER THAN ZERO.  
 Z-VALUE = 2.7335 PROBABILITY OF LARGER VALUE = 0.0031

Indication that addition of  
 a strip width is necessary  
 to compensate for edge effects

FINAL COVARIANCE MATRIX

1019.				
380.3	504.6			
289.2	383.8	763.4		
40.72	54.04	107.5	32.85	

Fig. 6.7g. Output from the density estimation algorithm.

# CHAPTER 7

## STUDY DESIGN

This chapter on study design is intended to assist the biologist in performing a capture-recapture study. Because this primer is oriented towards capture-recapture studies of closed populations, the suggestions made here do not always apply to other capture-recapture experiments, such as studies of open populations, or to experiments to estimate survival.

Before a capture-recapture experiment is conducted, the biologist must decide whether this approach is feasible. Other approaches, such as aerial or line transect surveys, may be more appropriate to estimate population size. If only relative differences in population are needed, indirect methods—pellet counts, catch-effort methods, or vegetation damage surveys—may suffice.

Once capture-recapture methods are decided on, the biologist must develop a design to ensure that (1) the assumptions can be tested, (2) the closure assumption is met, (3) the simplest appropriate model is selected from the eight possibilities, and (4) the number of animals captured is maximized (including recaptures). In this chapter, we begin by mentioning livetrapping versus removal methods, then discuss closure, proceed through ways of eliminating variation caused by time, behavior, and heterogeneity, and discuss sample size considerations such as grid size and number of traps.

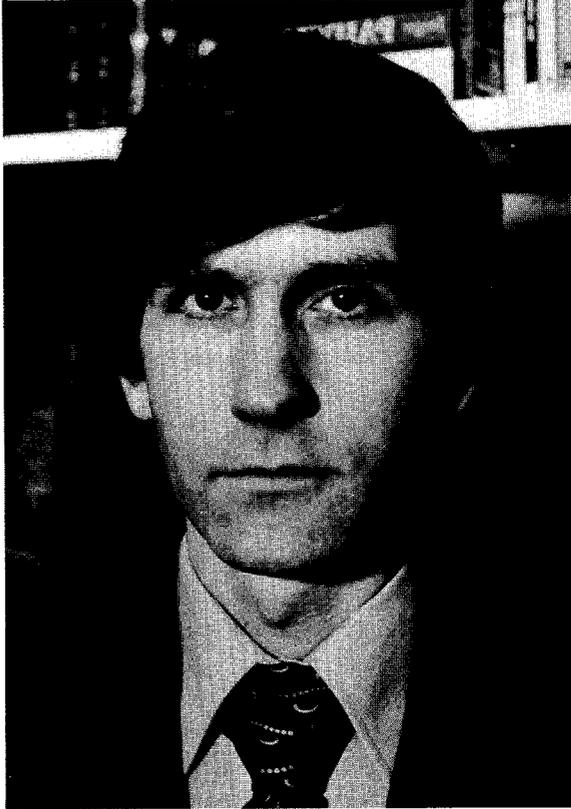
Much of this material is adapted freely from *Otis et al. (1978:74-81)* and *Arnason and Baniuk (1978:180-193)*. For additional discussion of considerations in grid trapping, the reader is referred to *Overton and Davis (1969)*, *M. H. Smith et al. (1969, 1971, 1975)*, *Tanaka (1970)*, and *Hansson (1974)*.

It is beyond the intended scope of this primer to provide guidance on the operational aspects of capture studies, although they are important. For example, if the method of marking (or trapping) is such that marks are lost, a basic assumption needed for meaningful results is violated. If the loss of marks is a significant problem, a double-tagging system can be used to estimate the loss of marks, and the number of unmarked animals in each occasion's sample can be corrected to reflect those that have lost their marks. Others who have worked with the species of interest or with a similar species should be consulted to determine an adequate marking technique. Reviews of marking techniques are given by *Tanner (1978)* and *Southwood (1978)*.

When making decisions about the field study, the biologist should consider the probable effects of the field methods on the subsequent data analysis. Will the field methods provide data that will meet the assumptions of the available data analysis methods? The consequences of the field methods must be evaluated before the data are taken, or the researcher may collect data that will be difficult or impossible to analyze and may generate estimates that will be misleading or absurdly biased. For additional discussion of field procedures, the reader is referred to *Davis (1956)*, *Southwood (1978)*, *Taber and Cowan (1969)*, *Begon (1979)*, and *Tanner (1978)*.

### Livetrapping Versus Removal Methods

As discussed in previous chapters, removal methods are a special case of livetrapping methods. That is, the removal estimators can be used on livetrapping data. The marked animals remain in the population, but they are "removed" by being marked. However, the following problems with removal studies suggest that livetrapping methods should be used to maintain design flexibility: (1) removal studies can fail because the estimate is impossible to compute (see Chapter 4), (2)  $\text{Var}(\hat{N})$  for removal studies is typically



A. Neil Arnason

Neil Arnason had a head start on the field of biology because both of his parents were entomologists. He studied mathematics and biology as an undergraduate and worked summers at the Statistical Research Service of Canada with E. C. Pielou and L. P. Lefkovich. His Ph.D. degree was taken at the University of Edinburgh under R. M. Cormack. The Statistics Department at Edinburgh had "captured" a large part of the market for people working on capture-recapture theory, and he worked there with G. M. Jolly and R. M. Cormack. His main work was related to the paper J. N. Darroch published in 1961 and focused on estimates of migration in animal populations.

Since leaving Edinburgh in 1970, he has taught modeling and simulation techniques in the Computer Science Department and quantitative ecology in the Zoology Department at the University of Manitoba. Arnason and his colleagues in Winnipeg have developed a comprehensive algorithm for the analysis of data using a variety of open-population methods. He sees the following developments in the future: (1) more powerful model-fitting techniques for capture-recapture data will allow the analyst to fit more specific models to his data; (2) more work will be done to find robust methods; (3) the analysis of capture data to estimate movement and to quantify habitat will be developed beyond its present very primitive state; and (4) new analytic and computational tools will be developed for application to these problems. (Recent photograph by F. Charron.)

greater than  $\text{Var}(\hat{N})$  for capture-recapture studies, and (3) the estimate obtained may be totally wrong because a sequence of  $p_i$  can be confounded with a constant  $p$ . Therefore, we recommend the use of livetrapping methods, if possible, because of the wider array of options available for the data analysis. Removal studies disrupt the population and, as a substantial proportion of the animals is removed, closure-violating immigration may occur.

During livetrapping studies, all possible precautions should be taken to prevent the deaths of animals in the traps, for example, by shading the traps in summer, checking traps more than once a day, shutting them in the afternoon, and avoiding periods of extreme cold. Similarly, it is assumed (implicitly) that the method of marking will not induce mortality. If 20% or greater mortality cannot be avoided, the data must be analyzed as a removal experiment.

## Closure

For the analysis methods presented here, the single most important assumption is closure. Testing for closure is very difficult, yet any violation of the assumption biases the tests and population estimators presented here. All closure tests assume  $M_0$  or  $M_t$  as the basic underlying model, and therefore, if the data are generated from other models (particularly from  $M_b$ ), the tests tend to reject the null hypothesis when in fact  $H_0$  is true—a Type I error. Because the tests are poor, we must ensure that the assumption of closure is met.

Methods to help ensure closure include timing the trapping to avoid suspected periods of migration or recruitment, or both, and keeping the experiment as short as possible. For example, when kill traps are

used, animals from the area around the grid will migrate onto the grid as local animals are removed (*Gentry et al. 1968; M. H. Smith et al. 1975:38*). The only good solution seems to be to keep the length of the study (in days) very short, so that it will end before significant immigration can occur. If a population must be studied during periods of change, open-population models should be used for data analysis (*Seber 1973; Pollock 1975; Arnason and Baniuk 1980*).

## Eliminating Variation Due to Time, Behavior, and Heterogeneity

If closure is satisfactorily achieved, the next most important considerations are twofold: (1) achieving a number of captures large enough to obtain reliable results, and (2) achieving a study for which the best model is the simplest possible one; for example, Model  $M_0$  rather than  $M_{tbh}$ , or Model  $M_h$  rather than  $M_{th}$ . In this section, we discuss methods of eliminating variation of capture probabilities due to time, behavioral response to first capture, and heterogeneity among individuals.

Of the three factors that affect capture probabilities, time is the most easily controlled by the biologist, but it is the most difficult to handle when present along with other sources of variation. The biologist can select the season of the year when studies are to be conducted, the length of the trapping period, and the time of day when trapping is to be done. In all the decisions, the objective is to reduce variation in capture probabilities over time. Among other things, this means that equal effort should be expended on each occasion, if the methods of this primer are to be used. For example, the number of traps should be the same throughout, trapping should be done at the same times of day, and if bait is used, the type and amount should be the same on all occasions. The study should be done when weather conditions are as constant as possible, because variable weather is likely to cause time variation in capture probabilities (*Getz 1961*). *Arnason and Baniuk (1978)* suggest that unequal effort on different trapping occasions may increase effectiveness per unit of effort with open-population models. However, the methods in program CAPTURE will work much more satisfactorily if equal effort results in little variation in capture probabilities due to time.

Behavioral response is common in small mammals, and it is doubtful whether much can be done to avoid it. Apparently for this reason, many biologists have studied and used strict removal methods for small-mammal population estimation. If there is any choice in capture methods for livetrapping, a method that will not result in a trap response should be used. One approach is to use different methods of capture on each occasion (*Overton and Davis 1969*). Although this procedure may reduce behavioral responses, it also may increase time variation. Because time variation combined with either behavioral response or heterogeneity results in the selection of models with no estimators, the possibility that a procedure may introduce time differences should be considered carefully. The key to reducing time variation is to catch the same number of animals on each occasion; that is,  $n_1 \cong n_2 \cong \dots \cong n_t$ .

One common source of heterogeneity is the lack of equal access to traps, if the traps are far apart relative to home range (*Eberhardt 1969*). We recommend the use of four traps per home range.

Other sources of heterogeneity are differences in activity or catchability related to measurable characteristics such as size, age, or sex. These characteristics probably affect an individual's position in the social hierarchy. Such sources can be removed by data stratification, if the characteristics are recorded and if the sample size permits. Data can be stratified to create subsets of the data that are most homogeneous with respect to capture probabilities. This involves nothing more complex than dividing the data into subsets on such variables as species, age, sex, and weight. Unfortunately, sample sizes are seldom large enough to allow stratification. However, if the data are sufficient, the device is valuable. The subsets are analyzed separately by the methods of program CAPTURE. The only additional testing needed might be tests for homogeneity among the strata or subsets (see *White 1975*).

Heterogeneity due to an unrecognized attribute, such as dominance, cannot be eliminated by stratification. The use of different methods of capture may increase the number of animals captured, but it will not eliminate the basic problem of heterogeneity of capture probabilities.

One possible method of eliminating heterogeneity, and perhaps trap response, is to locate traps randomly on each trapping occasion. Random placement would ensure that, for example, an individual's capture probability would not be consistently high or low because a trap was being placed very near or very far from its den. We are not sure how useful the technique is and would like to see further studies conducted to determine whether it results in significant reductions in heterogeneity. For logistical reasons, the randomization of trap locations on each occasion is probably not very practical, but it may be the only way to eliminate heterogeneity.

A valuable check on the livetrapping methods presented in this primer can be made by using a second method of estimating the proportion of marked animals in the population. A common example is killtrapping or hunting after the capture-recapture experiment has been completed. Then, using the number of marked animals in the population,  $M_{t+1}$ , and the ratio of marked to unmarked animals obtained from the killtrapping or hunting, a second and somewhat independent population estimate can be made with *Chapman's (1951)* version of the Lincoln estimator,

$$\hat{N} = \frac{(M_{t+1} + 1) (\text{all animals killtrapped} + 1)}{(\text{marked animals killtrapped} + 1)} - 1,$$

(*cf. Seber 1973:60*). Of course, traps are traps, and kill traps still present a strange object in the animal's environment. Another example, for small mammals marked by toe clipping, is to use tracks to obtain a ratio of marked to unmarked animals. Biologists have used tracks on smoked plastic (*Justice 1961*) or on dust-covered tiles to obtain an estimate of the proportion of animals with missing toes and, thus, a second population estimate to compare with the estimate obtained by livetrapping methods. The use of tracks has the distinct advantage of being a truly different method of sampling the population. Confidence in the population estimate is improved when both estimates give similar results.

## Sample Size

To obtain reliable estimates of population size, a sufficiently large sample must be taken. Here, "sample size" relates to the number of animals captured and recaptured, not to typical sample-size considerations, such as determining the number of plots to sample. For a livetrapping study, one must have both a large number of distinct animals captured and a sufficient number of recaptures, except for Models  $M_b$  and  $M_{bh}$ . The factors that control expected numbers of captures are (1) grid size, in terms of area covered and number of traps used; (2) capture probabilities; (3) number of trapping occasions; and (4) population size. We discuss these four factors in relation to the size of the experiment necessary to achieve precise population estimates.

The size of the grid is the first decision to be made. Grid size is a function of trap spacing, and of the number of rows ( $r$ ) and columns ( $c$ ) of traps. We suggest equal trap spacing and a systematic grid layout. At each grid station there will be one or more traps. If population densities are very high, we recommend two or more traps per station to avoid competition for traps. However, few studies have been made comparing two or more traps per station with a one-trap station, and further research will be required to determine whether the use of multiple traps per station increases the probability of capture for individual animals.

The objective of grid trapping over a short time period is to estimate the population size and, usually, the density at the grid site. Because of edge effect (discussed in Chapter 5), it is necessary to estimate the effective trapping area as well as  $N$ . In practice, this means we must be able to estimate the strip width  $W$ , which requires not only that the traps be placed in two dimensions rather than in a single line, but also that each trap can be associated with an X-Y coordinate. For practical reasons, this implies some sort of

regular grid layout, often a square or rectangle, with equal spacing between traps. This aspect of trap layout is not necessary if only  $N$  is to be estimated. If the trapping were done on a small island or in an enclosure, for example, knowledge of trap location in a coordinate system would not be needed. Finally, if the density estimate is to be meaningful, the grid should be placed in a homogeneous habitat to ensure nearly uniform density over the grid.

For a choice of  $r$  and  $c$  when the objective is density estimation, we suggest that both values be greater than or equal to 5; as a minimum we recommend  $r + c > 25$ . Examples are a square grid 13 by 13 or a rectangle 7 by 20. We note that much work in the literature relies on 16 by 16 grids (*Gentry et al. 1968; M. H. Smith et al. 1971*), and we suggest that grids should be at least that large for the density estimation method based on nested subgrids. We base this recommendation on the fact that a large number of captures is required in each subgrid to achieve a precise estimate; hence, the larger the size of the subgrids, the better the chance that numbers of captures will be large.

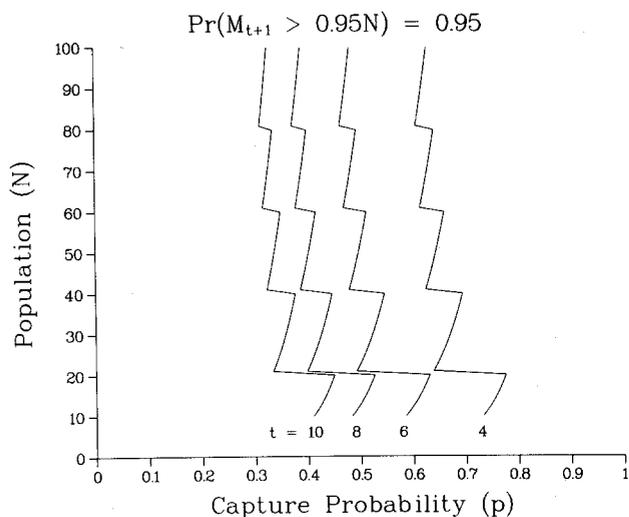
The next decision to be made is the spacing of traps. Most work with small mammals uses 15-m trap spacing or less (*Barbehenn 1974; M. H. Smith et al. 1971*). The rationale, when one is given, relates to the home range size. To illustrate, let  $s$  be the spacing between traps, and let  $2W$  be the average linear home range size. Home ranges are only rarely circular, but if we assume for design purposes that they are, then the radius here is  $W$ . We suggest at least four traps per home range, which implies  $s \leq (\sqrt{2})W$ . For best results we suggest  $s \leq W/2$ . Clearly, such a choice implies that some knowledge of home range sizes is needed before a good study can be designed. This requirement is not unreasonable—the biologist should have some behavioral knowledge of the species being studied, so that sampling decisions can be made intelligently. In fact, in any statistical sampling problem, a good study cannot be planned without some prior knowledge of the population parameters to be estimated.

As previously stated, the analyses for estimating population size presented in this primer require a sufficient number of captures to produce satisfactory results. Defining “sufficient numbers” is extremely complicated. Based on our experience with both real and simulated data, however, we have developed some crude guidelines. For instance, experiments in which  $M_{t+1}$  is on the order of 10 or 20 animals simply do not provide enough information for any procedure to perform well. The number of different animals captured must be several times larger than 10 or 20, depending heavily on the capture probabilities of the population members being studied. That is, if members’ average capture probabilities are as high as 0.40 or 0.50, program CAPTURE’s estimation and testing methods are useful for a population of about 50, but if the average capture probabilities are only 0.20, the population must be about 200.

In most studies, the number of recaptures must be relatively large before the experiment can produce useful results; again, the magnitude of the number relates to the magnitude of the capture probabilities. In general, the probabilities must be larger for smaller populations, but in no instance should  $N$  be less than 20 or average capture probabilities less than 0.30 (for  $N < 100$ ) when trapping occasions are few (say,  $t \leq 10$ ). These recommendations do not guarantee a satisfactory data analysis, but we have seen enough real and simulated data to say that if the data fail these criteria, the achievement of a precise or unbiased estimate is improbable.

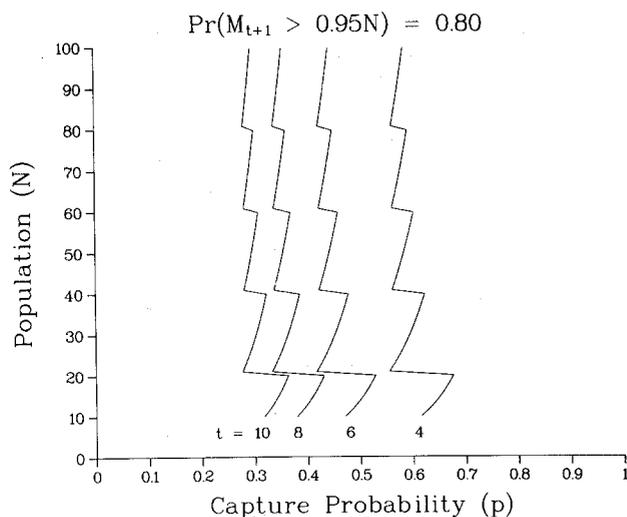
For small population sizes ( $N < 100$ ), we have developed two graphs (\*Figs. 7.1, 7.2) to help the biologist determine the number of trapping occasions necessary to catch more than 95% of the population. These figures were developed on the assumption that Model  $M_0$  applies, and they must be used very cautiously, because we do not believe that Model  $M_0$  is applicable to real data. In \*Fig. 7.1, the probability that  $M_{t+1}$  is greater than 95% of the population is equal to 0.95. In other words, for 95 of 100 experiments,  $M_{t+1}$  would exceed 0.95  $N$ . Four values of  $t$  are used. From \*Fig. 7.1, we conclude that more than 10 trapping occasions are necessary to catch 95% of the population if  $p$  is less than about 0.35. Because the assumption of closure is highly suspect for  $t > 10$ , we conclude that if  $N < 100$  and  $p < 0.35$ , no capture-recapture experiment will provide unbiased and precise estimates. Recall that \*Fig. 7.1 assumes Model  $M_0$ , which is an unrealistic model. If the data were actually from Model  $M_h$ , the results in \*Fig. 7.1 would be too conservative, with an even larger average  $p$  required to assume  $M_{t+1} > 0.95 N$ .

The plot in \*Fig. 7.2 is similar to the plot in \*Fig. 7.1, but the probability that  $M_{t+1} > 0.95 N$  is set at 0.8: that is, 80 of 100 experiments would provide  $M_{t+1}$  greater than 95% of  $N$ . By comparing the two figures, we see that not much has been gained. The curves are shifted to the left (smaller  $p$ ), but not drastically. Thus the conclusion still



\*Fig. 7.1. Plot of the number of trapping occasions required to catch 95% of the population in 95 of 100 experiments, where  $\Pr(M_{t+1} > 0.95 N) = 0.95$ .

\*Fig. 7.2. Plot of the number of trapping occasions required to catch 95% of the population in 80 of 100 experiments, where  $\Pr(M_{t+1} > 0.95 N) = 0.80$ .



remains: if  $N < 100$  and  $p < 0.3$ , a good capture-recapture study probably cannot be done. For  $N < 100$ , the probability of capture must be in the neighborhood of 0.5.

Estimation of density by nested subgrids requires even larger sample sizes; otherwise the data on the smaller subgrids will be too sparse for reliable results. We believe that reliable density estimates obtained with the subgrid approach require a grid at least 15 by 15 and, as a minimum, a total catch of 75 to 100 different animals.

We now consider an approach to determining a minimum grid area on the basis of these criteria. We start with the relation  $N = D A(W)$ , where  $A(W)$  is the effective trapping area. For a rectangular grid, the relation is

$$N = D [L_r L_c + 2(L_r + L_c) W + \pi W^2],$$

where  $L_r$  is the length of a row of traps [ $L_r = s \times (r - 1)$ ], and  $L_c$  is the length of a column of traps [ $L_c = s \times (c - 1)$ ]. Thus, the area covered by the grid is  $L_r L_c$ .

Without some knowledge of  $D$  and  $W$ , we cannot design a suitable study. We assume that  $D_0$  and  $W_0$  are the best guesses of the parameter values. To determine whether a grid study is at all feasible, we set  $N = 50$ , substitute  $D_0$  and  $W_0$  in the above equation, set  $L_r = L_c = L$  (a square grid), and solve for  $L$ .

$$L = \sqrt{0.8584(W_0)^2 + N/D_0} - 2W_0 .$$

Although the procedure is not difficult, we must be careful to use the same basic units for  $D$ ,  $L$ , and  $W$ . For example, we let  $W_0$  be 150 ft (45.7 m) and assume a density of three animals per acre (7.4 per ha). For compatibility of units, we put  $D_0$  in terms of square feet; then  $D_0 = (3/43\,560)$  animals/ft<sup>2</sup>. Solving for  $L$  gives

$$L = \sqrt{(0.8584) \times 150^2 + 50 \times 43\,560/3} - 300 = 563 \text{ ft} ,$$

which translates back into 7.3 acres (2.9 ha) as an absolute minimum grid size [ $7.3 = (563)^2/43\,560$ ].

This result is clearly conservative because not all animals will be caught. Improved planning requires that we establish grid size to ensure that a given number of animals  $M_{t+1}$  will be caught. But the expected number of different animals that will be caught depends on the true underlying capture probabilities, which are not known. The only practical approach is to make the best guess at the overall average probability of first capture,  $\bar{p}$ , applicable during the study and then to use the formula

$$M_{t+1} = N [1 - (1 - \bar{p})^t] .$$

Because the value of  $t$  (number of occasions) has been introduced, we can now compute several realistic values of  $t$ .

For example, we assume that  $\bar{p} = 0.10$ , set  $M_{t+1} = 50$ , and solve for  $N$  at several values of  $t$ . Given these values of  $N$ , we solve for the values of  $L$ , and obtain the following minimum grid sizes.

t	N	L	grid size in acres (ha)
4	145	1160	30.9 (12.5)
6	107	952	20.8 (8.4)
8	88	838	16.1 (6.5)

In practice, this example means that, if trapping were only for 4 days, we would need a 9 by 9 grid with traps spaced 150 ft (45.7 m) apart. For an 8-day trapping period, the same (expected) data could be obtained with a 7 by 7 grid with traps spaced 150 ft (45.7 m) apart.

This process can be reversed. Let us say a study is planned with a square grid of 15 traps, spaced 15 m apart ( $s = 15$  m). Then,  $L_r = L_c = 210$  m. Substitution in the basic equation gives

$$N = D \times [(210)^2 + 2(420)15 + \pi(15)^2] ,$$

$$= D \times 57\,407 \text{ (m}^2\text{)} ,$$

or

$$= D \times 5.74 \text{ (ha)} .$$

If the study is to last 7 days and the average capture probability is about  $\bar{p} = 0.15$ , then  $[(1 - 0.85^7) = (1 - 0.32) = 0.68]$ , or

$$M_g = N \times 0.68.$$

For density estimation, we must capture at least 230 animals because the inner 7 by 7 subgrid requires 50; so the density should be large enough that  $N \geq 338$  ( $338 = 230/0.68$ ).

$$D \geq 338/5.74 = 58.9 \text{ animals/ha.}$$

Thus, to get reliable results in this study, the true population density should equal or exceed 60 animals per hectare. If the biologist has good reason to believe that true density is only 10 or 20 animals per hectare, the study should be redesigned.

In addition to controlling the sampling effort through the trap spacing and the number of traps, the biologist also can select the number of trapping occasions. In theory, the more trapping times there are, the better will be the results, but this theory ignores the fact that the closure assumption becomes less realistic as more time passes. We recommend a minimum of 5 trapping occasions, but 7 to 10 are better, if closure is assured. The interval between occasions should be short, yet should allow marked individuals to recover from the stress of previous trapping. In practice, most trapping of small animals is done either once a day (morning) or twice a day (morning and evening). Trapping only in the mornings is far less likely to introduce time variation than trapping in the mornings and evenings, when a difference in capture probabilities between times is very likely to occur. If time variation is introduced and variation of behavior and heterogeneity also are present, the correct model will be  $M_{tbb}$ , for which no suitable estimator is available. Morning and evening trapping may, however, be aimed at different species. Then a workable design would be 5 (or 7) days of trapping in both morning and evening, with separate analyses of the morning and evening data. The possibility also exists of pooling the morning and evening data; pooling the two occasions may increase the probability of capture greatly and, hence, increase the sample size.

In removal studies, the absolute minimum is  $t = 3$  occasions (not 2, as is often suggested), because it is impossible to test for equal capture probabilities when  $t = 2$ . We recommend at least four removals when capture probability is less than 0.5, which is not uncommon (see Chapter 4, Removal Models).

Another valuable method of testing design adequacy before going to the field is to simulate the experiment on a computer with program CAPTURE. Approximate parameter values can be chosen, and the experiment can be replicated as many times as necessary. Among the criteria that can be observed are: the selection of the appropriate model, the bias of the selected estimators, and the achieved confidence level. The validity of the simulations to the field study will necessarily depend on the similarity of the selected parameters to the actual parameter values. A range of parameter values can be simulated, along with a variety of trapping efforts (the number of trapping occasions). Thus cost of the experiment (probably in manpower) can be related to the probable precision of the estimate to be achieved within the range of simulated parameter values. If only unreliable estimates result for the available manpower, the experiment should not be conducted. An excellent discussion of the use of the Monte Carlo simulation to design a study is given by *Arnason and Baniuk (1978:180-193)*.

### Unique Identification of Each Captured Animal

In capture-recapture studies, each captured animal is typically given a unique mark. Unique marking systems are preferable in studies where the animals are actually captured and, hence, where the technique is feasible. However, capture-recapture studies often can be performed on animals even when unique marking of each captured individual is not possible. Here, we discuss ways to fulfill the data requirements of all methods presented in this primer except density estimation, without a unique marking system.

The information required to use the model selection procedures discussed in Chapter 3 is that the  $X_{ij}$  matrix can be constructed. Each row of the  $X_{ij}$  matrix represents the capture history of an animal. No distinction is made

between two animals with the same capture history. Thus the data requirement for construction of the X matrix is that the capture history of each animal can be determined.

An example of the application of a non-unique marking system illustrates the possibilities. Suppose a biologist wants to estimate the number of moose (*Alces alces*) in a particular valley during the winter. Closure can be assumed because snow is deeper on the higher ground surrounding the area than in the valley. Moose are marked from a helicopter with a paint gun. On occasion 1, red paint is used, and 25 moose are given red spots on the rump. On occasion 2, yellow paint is used, and 20 moose are given yellow spots; 15 of these moose already had red spots. Thus after occasion 2, the X matrix would consist of 15 rows of (1,1), 10 rows of (1,0), and 5 rows of (0,1). On occasion 3, orange paint is used, and 22 moose are marked. Of these, 2 had no previous marks, 8 had only red marks, 2 had only yellow marks, and 10 had both yellow and red marks. The resulting X matrix is shown in \*Fig. 7.3. Notice that after occasion 1, the matrix consists of only the first 25 rows. After occasion 2, five rows and the second column are added. After occasion 3, two rows and the third column are added; column 3 reflects the

ROW	OCCASION		
	1	2	3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	0
12	1	1	0
13	1	1	0
14	1	1	0
15	1	1	0
16	1	0	1
17	1	0	1
18	1	0	1
19	1	0	1
20	1	0	1
21	1	0	1
22	1	0	1
23	1	0	1
24	1	0	0
25	1	0	0
26	0	1	1
27	0	1	1
28	0	1	0
29	0	1	0
30	0	1	0
31	0	0	1
32	0	0	1

\*Fig. 7.3. The X matrix resulting from the example of marking moose from the air, using a different mark for each occasion.

previous capture histories observed in the field. If the operation were to continue for more occasions, a row for each first capture and a column for each occasion would be added.

One potential problem of the marking scheme described above is that the person marking the animals may select animals already marked (or unmarked). The selection would result in Model  $M_b$ , because the marker has treated marked animals differently than unmarked animals. This problem is much more difficult if the animals are in a herd, because the smarter or faster animals may disappear before the marker gets a chance to paint them. An example is feral horse herds. As the marker in the helicopter starts singling out animals, the herd may break up, and some animals may escape being marked. On a later occasion, when the same herd is encountered by the helicopter, the same unmarked animals may escape being marked again. Thus, the technique would result in Model  $M_h$ , or in Model  $M_{bh}$  if the marker showed any selectivity as to which animals were to be marked. Although heterogeneity remains a problem, time variation can be eliminated by marking only on good weather days, and by capturing about the same number of animals each day—that is, by spending the same amount of effort on each occasion.

If a biologist cannot mark animals uniquely or cannot construct an X matrix from the data collected, the model selection procedure cannot be used. It may still be possible to run some of the tests and to calculate most of the estimators on less than the X matrix, but we strongly discourage such practices. Without the application of the model selection procedure to investigate what sort of variation of capture probabilities exists in the data, the biologist is likely to use an estimate from a model that does not fit the data.

## Recording Data

In recapture studies, the capture history of each animal caught must be known. Otherwise, substantial information may be lost, and it will be impossible to compute all tests for sources of variation. Although it may seem obvious, we insist that recording all data (animal numbers and trap locations) correctly is crucial.

If density is to be estimated on the basis of grid trapping data, trapping location also must be recorded. The minimum information that must be taken when a capture occurs includes the animal identification code and the trapping occasion. Usually, the species, sex, and age also are recorded for each animal, although for the analyses given here the only purpose is to enable subsequent partitioning of the data. We recommend analyzing the data separately by species, but often data are too few to be partitioned by sex and age.

For true removal studies, like electrofishing, there is no animal identification code. As suggested by *Raleigh and Short (1981)*, recording each individual by species and analyzing the data at least by major taxonomic groups are both very important in removal studies.

To know the trap location, traps must be numbered uniquely. Moreover, the biologist must know the relation between the trap number and its coordinate on some arbitrary rectangular X-Y coordinates. We strongly recommend use of this system, and we stress that density estimation with program CAPTURE requires that data be collected in this context. The northwest corner trap should be numbered (1,1), because (0,0) indicates no capture in the program CAPTURE code. Then the rows become the X-axis and the columns the Y-axis. For example, a 10 by 10 grid would look like Fig. 7.4. This system can be extended to cover any rectangular grid of rows and columns.

When traps are checked both morning and evening, it is necessary to record both the day and the time of capture. The data can be analyzed as two occasions per day, or they can be pooled as one occasion.

Standard field forms and standard conventions for trap numbering and animal identification facilitate data recording. One example of a standardized method is presented in *Brotzman and Giles (1966)*. Figure 7.5 shows a keypunch field form that is compatible with program CAPTURE. We recommend recording data directly on a keypunch form to avoid errors in transcribing the data from field notebooks. Keypunch forms can be duplicated for field use.

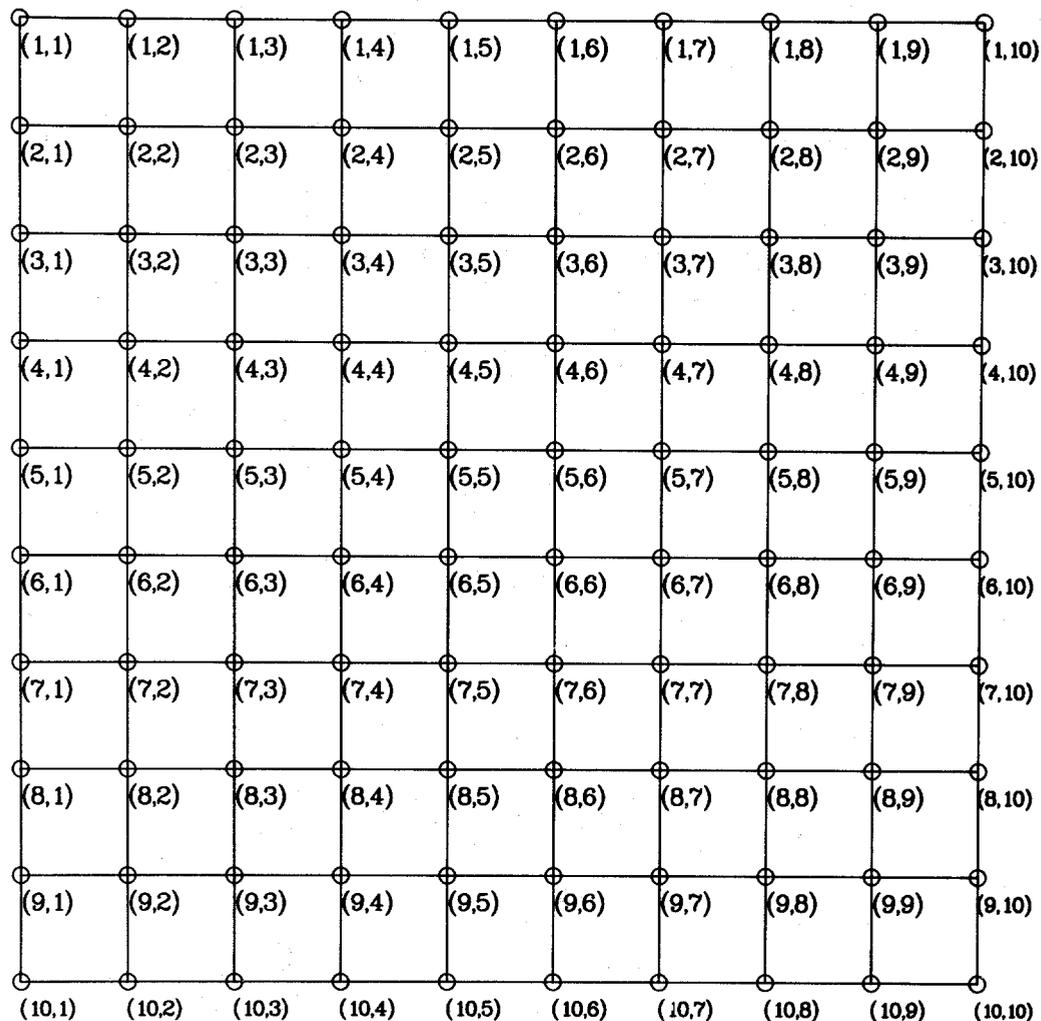


Fig. 7.4. The coordinates represent the trap numbering system used by program CAPTURE. The upper left (northwest) trap is labeled (1,1) not (0,0), because zero coordinates would indicate that the animal was not caught on that occasion.

## Data Anomalies

Various anomalies and unplanned events may occur during trapping. For example, (1) several animals may be found in one trap, (2) animals may be found dead in traps; (3) animals that have been released may be found further down the grid, trapped again on the same occasion; and (4) a trapped animal may escape when it is removed from the trap. We make the following suggestions regarding these happenings.

(1) More than one animal per trap presents no problem; each animal is recorded separately. This type of data does not invalidate the analyses presented here.

(2) A dead animal in a trap in a livetrapping study is a more serious problem. On the last trapping occasion, the dead animal does not matter; the capture is recorded as normal. Otherwise, the data analysis must be modified. Some of the methods described here (specifically, Models  $M_0$  and  $M_1$ ) can allow for known removals in a true livetrapping study. Because not all tests and estimators can be modified, we have not dealt here with such models. We recommend the following: if trap deaths are less than 5% of total captures, remove the dead-animal data from the total results, run the analyses, add the number of dead animals to  $\hat{N}$ , and then multiply the density estimate  $\hat{D}$  by  $(1 + \text{proportion dead})$ . If trap deaths are more than 20% of total captures, use the generalized removal method of analysis on first capture. For 5-20% trap deaths, the only "safe" analysis may be the removal method, although

ID	CAPTURE 1		CAPTURE 2		CAPTURE 3		CAPTURE 4		CAPTURE 5		CAPTURE 6		CAPTURE 7		CAPTURE 8		CAPTURE 9		CAPTURE 10		CAPTURE 11		CAPTURE 12			
	t	x	y	t	x	y	t	x	y	t	x	y	t	x	y	t	x	y	t	x	y	t	x	y		
1																										
2																										
3																										
4																										
5																										
6																										
7																										
8																										
9																										
10																										
11																										
12																										
13																										
14																										
15																										
16																										
17																										
18																										
19																										
20																										
21																										
22																										
23																										
24																										
25																										
26																										
27																										
28																										
29																										
30																										
31																										
32																										
33																										
34																										
35																										
36																										
37																										
38																										
39																										
40																										

Fig. 7.5. A computer form for recording capture histories for input to program CAPTURE. The TASK cards (Appendix A) to read this data form would be  
 TASK READ CAPTURES XY REDUCED OCCASIONS=12  
 FORMAT=(A3,5X,36 (F2.0)).

deletion of the dead animals may be preferable if the loss of information can be afforded. These modifications give  $\hat{N}$  and  $\hat{D}$  relative to the first day of the study. To make these estimates apply to the population remaining alive after the study, simply delete all trap deaths from the data set before analysis.

(3) When the same animal is caught more than once on a given occasion, the only added information provided is on movement. We recommend that both captures be recorded, but that only the results of the first capture be used for data analysis.

(4) When an animal escapes during handling before it is tagged, or before the mark is read, do nothing. In essence, this animal does not become part of the data.

### Simulation Example

To illustrate the use of Monte Carlo simulation for design of a capture-recapture experiment, we use the data from E. Larsen given in *Otis et al. (1978:42-43)*. The same data set is provided with program CAPTURE, so the reader can duplicate the following computer results.

Great Basin pocket mice (*Perognathus parvus*) were livetrapped in a desert community in Curlew Valley, Utah, during June 1977. The data are from 7 consecutive nights of livetrapping, with traps arranged in a 12 by 12 grid and spaced 15 m apart. The model selection procedure selected either Model  $M_{t_{bh}}$  or Model  $M_{b_{bh}}$  as appropriate for use in estimating  $N$ . The data on model selection and the results of using  $\hat{N}_{bh}$  to produce point and interval estimates of  $N$  are given in \*Fig. 7.6.

Note that the confidence interval for  $N$  in \*Fig. 7.6 is quite large. If we want to repeat the study, how should we modify it to obtain a confidence interval smaller than  $\pm 20\%$  of  $\hat{N}$ , assuming field conditions have not changed drastically?

From \*Fig. 7.6, we see that the estimate used there came from a model with two average capture probabilities: about 0.3 and 0.14. On the first trapping occasion, 23 animals were caught and  $\bar{p}_1 = 0.3$ , so the population must have consisted of a fair number of animals with large capture probabilities. If we assume that  $N = 77$ , the animals left uncaught had a mean capture probability of 0.15. Thus we might model the capture probabilities of the population as

Number of Animals	Capture Probability
20	0.10
20	0.20
37	0.65

We determined these numbers in a crude fashion by simulating various combinations until we obtained data and results similar to those of \*Fig. 7.6.

We first simulate the Larsen experiment from June 1977 to see whether the results from the hypothetical capture probabilities are consistent with the observed data. The input to program CAPTURE is as follows.

```
TASK SIMULATE POPULATION=77 OCCASIONS=7 REPLICATIONS=100
HETEROGENEITY=20,0.1,20,0.2,37,0.65
BEHAVIOR=77,1.4
```

The BEHAVIOR card guarantees that trap response is present, and thus that  $M_{b_{bh}}$  is the true underlying model for the simulated data. From the summary of the simulation results presented in \*Fig. 7.7, we note that Model  $M_{t_{bh}}$  was never selected, yet the model selection procedure output in \*Fig. 7.6 indicates that  $M_{t_{bh}}$  is just as likely to occur as  $M_{b_{bh}}$ . Another indication that the simulations are not realistic representations of the situation is the width of the confidence interval for the generalized removal estimator: the average confidence interval width in the simulations is about 11, whereas the confidence interval width for Larsen's data is 98. Other sets of parameters for Model  $M_{b_{bh}}$  likewise do not appear to provide a reasonable representation of the real data. The reader may try other parameter sets for Model  $M_{b_{bh}}$  to duplicate Larsen's data.

MODEL SELECTION PROCEDURE. SEE MODEL SELECTION SECTION OF THE MONOGRAPH FOR DETAILS.  
 DATA FROM E. LARSEN, PEROGNATHUS PARVUS, CURLEW VALLEY, UTAH.

OCCASION	J=	1	2	3	4	5	6	7
ANIMALS CAUGHT	N(J)=	23	22	17	17	22	25	34
TOTAL CAUGHT	M(J)=	0	23	32	35	40	46	55
NEWLY CAUGHT	U(J)=	23	9	3	5	6	6	3
FREQUENCIES	F(J)=	16	15	6	5	5	5	3

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

CHI-SQUARE VALUE = 61.688 DEGREES OF FREEDOM = 4 PROBABILITY OF LARGER VALUE = 0.00000

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

CHI-SQUARE VALUE = 7.844 DEGREES OF FREEDOM = 1 PROBABILITY OF LARGER VALUE = 0.00510

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

CHI-SQUARE VALUE = 15.596 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.01609

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

CHI-SQUARE VALUE = 18.236 DEGREES OF FREEDOM = 6 PROBABILITY OF LARGER VALUE = 0.00567

TEST OF MODEL M(H) BY FREQUENCY OF CAPTURE  
 (FREQUENCIES LESS THAN 21 ARE NOT CALCULATED.)

NUMBER OF CAPTURES	CHI-SQUARE	D.F.	PROBABILITY
1	5.000	6	0.54381
2	22.240	6	0.00110

MODEL SELECTION PROCEDURE. SEE MODEL SELECTION SECTION OF THE MONOGRAPH FOR DETAILS.  
 DATA FROM E. LARSEN, PEROGNATHUS PARVUS, CURLEW VALLEY, UTAH.

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

CHI-SQUARE VALUE = 17.856 DEGREES OF FREEDOM = 10 PROBABILITY OF LARGER VALUE = 0.05743

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

CHI-SQUARE VALUE = 9.966 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.07622

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

CHI-SQUARE VALUE = 7.891 DEGREES OF FREEDOM = 5 PROBABILITY OF LARGER VALUE = 0.16235

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

CHI-SQUARE VALUE = 60.390 DEGREES OF FREEDOM = 22 PROBABILITY OF LARGER VALUE = 0.00002

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

CHI-SQUARE VALUE = 17.988 DEGREES OF FREEDOM = 14 PROBABILITY OF LARGER VALUE = 0.20731

MODEL SELECTION CRITERIA. MODEL SELECTED HAS MAXIMUM VALUE.

MODEL	M(O)	M(H)	M(B)	M(BH)	M(T)	M(TH)	M(TB)	M(TBH)
CRITERIA	0.80	0.86	0.62	0.99	0.00	0.59	0.54	1.00

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

\*Fig. 7.6. Analysis from program CAPTURE of Larsen's data on pocket mice in Curlew Valley, Utah.

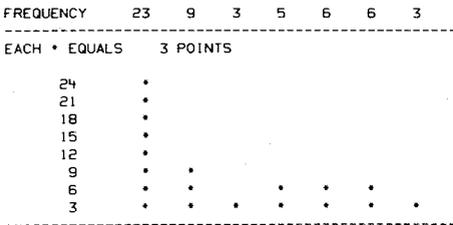
POPULATION ESTIMATION WITH VARIABLE PROBABILITY REMOVAL ESTIMATOR. SEE M(BH) OR REMOVAL MODELS OF THE MONOGRAPH FOR DETAILS.  
 DATA FROM E. LARSEN, PEROGNATHUS PARVUS, CURLEW VALLEY, UTAH.

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	0.0762	0.2802	0.2802	0.2802	0.2802	0.2802	0.2802	0.2802	0.2802
2	76.73	24.66909	3.486	0.4801	0.2998	0.1377	0.1377	0.1377	0.1377	0.1377	0.1377	0.1377
3	FAILURE CRITERION = -2 NO ESTIMATES FOR THIS STEP.											
4	67.44	18.18989	1.308	0.5201	0.3410	0.2025	0.0846	0.2067	0.2067	0.2067	0.2067	0.2067
5	58.65	5.896785	0.823	0.3643	0.3922	0.2525	0.1126	0.2114	0.3953	0.3953	0.3953	0.3953

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

HISTOGRAM OF U(J)



\*Fig. 7.6 (cont.)

Because of the contrast between the real and simulated data, we constructed a more complex set of capture probabilities from  $M_{tbh}$ . The input to program CAPTURE is as follows.

TASK SIMULATE POPULATION=77 OCCASIONS=7 REPLICATIONS=100  
 HETEROGENEITY=30,0.3,30,0.5,17,.90  
 BEHAVIOR=77,1.4  
 TIME=0.6,0.6,0.5,0.5,0.6,0.6,0.7

As seen in \*Fig. 7.8, the simulation output appears comparable to the results of Larsen. Two major differences are noted between the output in the two figures: in \*Fig. 7.8, the confidence interval width for  $M_{bh}$  is greatly expanded, and the average estimate for  $M_{bh}$  appears to be less biased. These two differences cause appreciably better coverage for the  $M_{bh}$  results—84% in \*Fig. 7.8 as opposed to 15% in \*Fig. 7.7.

What conclusions can we base on these simulations? We suggest that, if the study were to be redone, the variation in capture probabilities through time should be eliminated if any kind of precision is to be achieved. The estimates in \*Fig. 7.8 are accurate, but the precision is so poor that the investigator could have no confidence in the estimate. If a true Model  $M_{bh}$  data set were obtained, the estimate would be precise (although biased). However, the investigator, having access to the simulation results in *Otis et al. (1978:130)*, realizes that the generalized removal estimator typically is biased low. Even though a precise wrong answer is obtained, the investigator is aware of this fact and can correct for the bias on the basis of the simulation in *Otis et al. (1978)*, plus any additional simulations he may undertake. However, this approach should be documented carefully to enable other researchers to verify the results; in fact, the “doctoring” of an estimate should be avoided.

DATA SIMULATION PROCEDURE. SEE FIELD PROCEDURES SECTION OF THE MONOGRAPH FOR DETAILS.

RANDOM SEED IS 1234567

NUMBER OF REPLICATIONS 100

NUMBER OF OCCASIONS 7

NUMBER IN POPULATION 77

HETEROGENEITY STRUCTURE

NUMBER	VALUE
20	0.100
20	0.200
37	0.650

TIME VARIATION STRUCTURE

NUMBER	VALUE
1	1.000
2	1.000
3	1.000
4	1.000
5	1.000
6	1.000
7	1.000

BEHAVIOR STRUCTURE

NUMBER	VALUE
77	1.400

DATA SIMULATION PROCEDURE. SEE FIELD PROCEDURES SECTION OF THE MONOGRAPH FOR DETAILS.

MODEL SELECTION RESULTS

MODEL	TIMES SELECTED	PERCENT	FAILURES	AVERAGE	STANDARD ERROR	LCI	UCI
-----							
POINT ESTIMATES							
M(O)	0	0.000	0	0.00	0.0000	0.0	0.0
M(H)	23	23.000	0	77.00	2.7111	71.4	82.6
M(B)	31	31.000	0	63.55	0.5881	62.4	64.7
M(BH)	33	33.000	0	64.88	0.7843	63.3	66.4
M(T)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TH)	1	1.000	0	0.00	0.0000	0.0	0.0
M(TB)	9	9.000	0	0.00	0.0000	0.0	0.0
M(TBH)	3	3.000	0	0.00	0.0000	0.0	0.0

INTERVAL ESTIMATE LENGTHS  
 COVERAGE PERCENT

MODEL	COVERAGE	PERCENT	FAILURES	AVERAGE	STANDARD ERROR	LCI	UCI
M(O)	0	0.000	0	0.00	0.0000	0.0	0.0
M(H)	12	52.174	0	20.19	2.6614	14.7	25.7
M(B)	0	0.000	0	6.73	0.3946	6.0	7.5
M(BH)	5	15.152	0	10.82	1.8982	7.1	14.5
M(T)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TH)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TB)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TBH)	0	0.000	0	0.00	0.0000	0.0	0.0

TEST SIZE RESULTS

TEST	0.01	0.05	0.10
M(O) VS. M(H)	0.01	0.01	0.01
M(O) VS. M(B)	0.98	1.00	1.00
M(O) VS. M(T)	0.13	0.35	0.45
M(H) VS. NOT M(H)	0.44	0.58	0.70
M(B) VS. NOT M(B)	0.07	0.19	0.31
M(T) VS. NOT M(T)	0.99	0.99	0.99
M(H) VS. M(BH)	0.61	0.77	0.85
CLOSURE	0.30	0.57	0.73

\*Fig. 7.7. Program CAPTURE simulation output to duplicate Larsen's 1977 data. On this first attempt, only heterogeneity and behavioral response are included.

DATA SIMULATION PROCEDURE. SEE FIELD PROCEDURES SECTION OF THE MONOGRAPH FOR DETAILS.

RANDOM SEED IS 1234567

NUMBER OF REPLICATIONS 100

NUMBER OF OCCASIONS 7

NUMBER IN POPULATION 77

HETEROGENEITY STRUCTURE

NUMBER	VALUE
30	0.300
30	0.500
17	0.900

TIME VARIATION STRUCTURE

NUMBER	VALUE
1	0.600
2	0.600
3	0.500
4	0.500
5	0.600
6	0.600
7	0.700

BEHAVIOR STRUCTURE

NUMBER	VALUE
77	1.400

DATA SIMULATION PROCEDURE. SEE FIELD PROCEDURES SECTION OF THE MONOGRAPH FOR DETAILS.

MODEL SELECTION RESULTS

MODEL	TIMES SELECTED	PERCENT FAILURES	AVERAGE	STANDARD ERROR	LCI	UCI
M(O)	0	0.000	0	0.00	0.0000	0.0
M(H)	21	21.000	0	77.79	1.3300	75.0
M(B)	20	20.000	0	73.95	1.3128	71.2
M(BH)	43	43.000	0	80.93	3.9741	73.1
M(T)	0	0.000	0	0.00	0.0000	0.0
M(TH)	2	2.000	0	0.00	0.0000	0.0
M(TB)	1	1.000	0	0.00	0.0000	0.0
M(TBH)	13	13.000	0	0.00	0.0000	0.0

INTERVAL ESTIMATE LENGTHS

	COVERAGE		PERCENT	AVERAGE	STANDARD ERROR	LCI	UCI
M(O)	0	0.000	0	0.00	0.0000	0.0	0.0
M(H)	18	95.714	0	17.10	0.9393	15.1	19.1
M(B)	17	85.000	0	26.28	2.1869	21.7	30.9
M(BH)	36	83.721	0	83.49	40.7797	3.6	163.4
M(T)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TH)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TB)	0	0.000	0	0.00	0.0000	0.0	0.0
M(TBH)	0	0.000	0	0.00	0.0000	0.0	0.0

TEST SIZE RESULTS

TEST	0.01	0.05	0.10
M(O) VS. M(H)	0.87	0.95	0.96
M(O) VS. M(B)	0.65	0.83	0.88
M(O) VS. M(T)	0.36	0.64	0.75
M(H) VS. NOT M(H)	0.42	0.69	0.79
M(B) VS. NOT M(B)	0.04	0.20	0.27
M(T) VS. NOT M(T)	0.50	0.66	0.76
M(H) VS. M(BH)	0.17	0.37	0.53
CLOSURE	0.00	0.05	0.11

\*Fig. 7.8. The second attempt at simulating Larsen's data, by using Model M<sub>tbh</sub>.

## Check List

To summarize this chapter, we have prepared a brief check list of questions to be answered before a capture-recapture study is started.

1. Is a better approach available than the capture-recapture method? For example, line transect surveys?
2. Is a closed population being considered?
3. Is a reliable marking procedure available?
4. Will the data be fit by a model with an estimator? Stated differently, given that heterogeneity of capture probabilities may be assumed to exist, can time variation be eliminated to give Model  $M_{bh}$  or Model  $M_h$ ?
5. How precise an estimate is needed from the study?
6. Are the manpower and resources available to obtain the necessary precision?

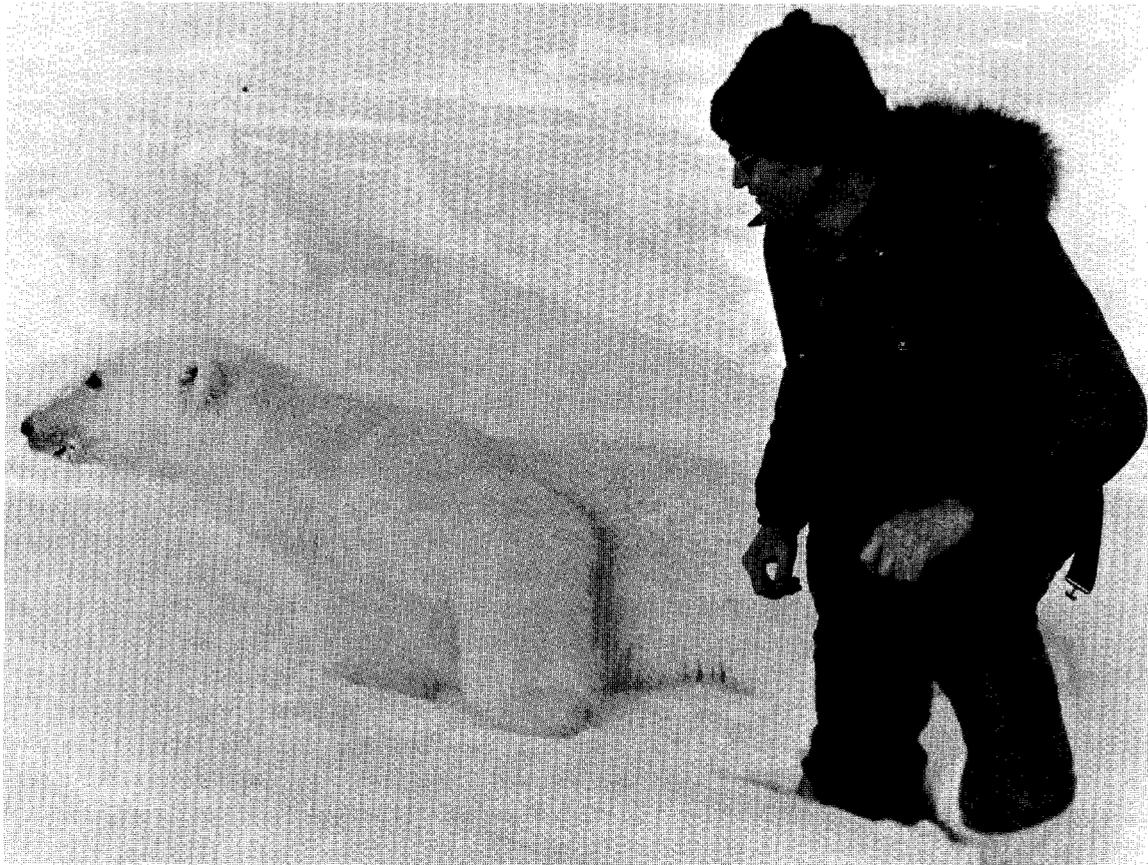
## Questions and Exercises

1. Compare the advantages and disadvantages of using 100 traps for 8 nights versus 200 traps for 4 nights to trap meadow voles (*Microtus pennsylvanicus*) in an old field. Assume that the grid size for both experiments is the same.
2. Assume that you know the home range of an animal is 10 000 m<sup>2</sup>. What trap spacing would you use?
3. What are the disadvantages if both time and behavior influence capture probabilities?
4. Comment on a capture-recapture experiment in which traps are checked morning and evening. What problems might you encounter? Would you pool the data?
5. What is the advantage of fencing the ends of a stream segment where you will conduct an electrofishing experiment?
6. The following input to program CAPTURE was used to simulate a capture-recapture experiment.

```
TASK READ CAPTURES POPULATION=50 OCCASIONS=4
HETEROGENEITY=20,0.05,10,0.10,20,0.15
TIME=1.0,1.5,1.0,0.5
BEHAVIOR=50,0.5
```

The data are obviously from Model  $M_{tbh}$ , yet the results from CAPTURE indicate that more than 90% of the time, Model  $M_o$  is selected. Why? (Hint: What is the average probability of capture on each occasion?)

7. You are doing a 5-night small-mammal capture-recapture study starting on Monday. On Tuesday night, the weather turns much colder with rain, but by Wednesday morning weather conditions have returned to normal. What problems would you expect when you analyze the data? What might you consider to rectify these problems, both during the remainder of the trapping sessions and later, during the analysis?



Tagging large animals like polar bears does not fit the methods in this primer because too few animals are recaptured, and the studies are generally long-term; that is, they are better modeled by open-population models. (Photograph courtesy of Jack Lentfer.)



Population estimates of medium-sized mammals such as this raccoon (*Procyon lotor*) may be made with capture-recapture methods, but often capture probabilities are small, and animals are not recaptured, making the methods in this primer inappropriate. (Photograph courtesy of Alan Sargeant and Raymond Greenwood.)



The use of radio collars on big game, such as elk (*Cervus elaphus*), does not generally provide reliable capture-recapture estimates unless recaptures are obtained through other methods, such as aerial observations of the collars (see Rice and Harder 1977). Survival rates can be obtained from biotelemetry data. (Photograph courtesy of Kenneth V. Bostick.)

# CHAPTER 8

## OPEN MODELS

### Basic Concepts

Capture-recapture estimation techniques are usually classified as belonging to one of two general categories—those appropriate for (demographically) closed populations and those appropriate for (demographically) “open” populations (*Arnason and Baniuk 1980; Seber 1981*). An open population is defined as one in which the processes of birth, death, and migration are allowed to operate. Just as there is more than one model for the closed-population case, there is a variety of open models for capture-recapture studies. However, the essential elements of open models are the types of parameters they involve. In the now classic Jolly-Seber model (see *Seber 1973:196-232*), the parameters are population size, survival rate, recruitment, and capture probability. In closed models, by way of comparison, one assumes that during the capture study, the animals have a survival rate of one and that the recruitment into the population is zero.

Field aspects of capture-recapture studies are not necessarily different when open models are used, except that the time period may be long relative to the population dynamics of the target species. For example, instead of trapping once a day for 6 days, one may trap only once a month or even once a year for studies on most mammals, birds, and fish. In many insect capture studies, open models are necessary even when capture occasions are on consecutive days.

The population size of open models varies with each capture occasion; thus, the one population size parameter  $N$  of the closed model becomes population sizes  $N_1, \dots, N_t$  for the  $t$  capture occasions of the open model. The survival rate parameters are  $S_1, \dots, S_{t-1}$ , where  $S_j$  represents an individual's survival rate, in the population exposed to trapping, between occasions  $j$  and  $j + 1$ . The survival rate applies to those animals alive in the trapped population just after the  $j^{\text{th}}$  trapping occasion. The recruitment parameters are  $B_1, \dots, B_{t-1}$ . Here,  $B_j$  is the total influx of new individuals into the population being trapped, between trapping occasions  $j$  and  $j + 1$ ; as such,  $B_j$  is not a rate. Finally, there are the capture probabilities  $p_1, \dots, p_t$ ; all the  $N_j$  animals in the population on the  $j^{\text{th}}$  capture occasion are assumed to have capture probability  $p_j$ . The Jolly-Seber model thus allows time variation in capture probabilities, but no variation due to behavioral response or heterogeneity. It is thus the open version of Model  $M_t$ .

Interrelationships exist among these parameters—primarily,  $N_{j+1} = N_j S_j + B_j$ , which simply indicates that the population size at occasion  $j + 1$  is the number of survivors ( $N_j S_j$ ) from the  $j^{\text{th}}$  occasion plus the recruits ( $B_j$ ) during this time period. Also, some parameters cannot be estimated for the first or last occasions—for example,  $N_1$ ,  $N_t$ , and  $S_{t-1}$ . Although both of these factors reduce the real (as opposed to apparent) number of parameters, open models do have more parameters than their closed counterparts. And the more realistic the model, the more parameters it has.

An additional parameter, totally derived from the survival rates and data, is  $M_j$ , the number of marked animals still alive just before the  $j^{\text{th}}$  capture occasion. This parameter is the key to estimation in the open models, but it is of no intrinsic biological interest.

Because the processes of population dynamics are always operating at least to some degree in natural populations, one might ask why so much attention has been given to capture-recapture data analysis methods based on models closed to these processes. The answer is familiar to statisticians and is becoming familiar to biologists as they gain experience in the construction and manipulation of mathematical models: as the model for a capture-recapture experiment is made increasingly realistic by the introduction of more parameters, such as time- and age-specific survival rates and birth rates, each individual parameter is estimated with less and less precision. Ultimately, the model has more parameters than data and becomes so general that no parameter estimates can be produced from it.

*Cormack (1979)* makes the valuable analogy that the data collected from an experiment contain only a fixed amount of information that can be used to estimate parameters of interest. As the number of parameters increases, this information is spread thinner and thinner among the parameters, and finally results in very little information about any individual parameter. Thus, because estimation techniques for closed populations are generally based on far fewer parameters than those for open populations the techniques can provide more precise estimates of population size  $N$  if the closure assumption is valid.

Of course, there are situations where open models must be used, but there are also situations where closed models should be used. When closed models are used, population size  $N$  is practically constant during the study and only one estimate of  $N$  is necessary. Use of an open model gives several  $(t - 2)$  separate estimates, which then must be combined into a single estimate. Before discussing the advantages, differences, and tradeoffs of closed versus open models, we present a brief synopsis of the state of the art of analytical techniques for capture data from open populations.

## State of the Art

Like the closed-population models discussed in Chapter 3, open models can vary widely in their generality, and as a result many methods and models have been introduced into the literature. (Note that open removal models are impossible.) Although there were earlier efforts to deal with the processes of birth, death, and migration in capture studies (*Fisher and Ford 1947; N. T. J. Bailey 1951; Leslie and Chitty 1951; Leslie 1952; Leslie et al. 1953*), the appropriate beginning point for our purposes is the appearance of the Jolly-Seber model, published independently by *Jolly (1965)* and *Seber (1965)*. The basic model allows birth, death, immigration, and permanent emigration to occur during the experiment.

The following points concerning the Jolly-Seber method must be made clear.

1. The estimated "survival" rate is really survival in the population at risk of capture. An animal fails to survive if it dies or emigrates. If  $D_j$  is the death rate during the time between the  $j^{\text{th}}$  and  $j + 1^{\text{th}}$  trapping occasions and  $E_j$  is the emigration rate during this time, the survival rate parameter is the product  $S_j = (1 - D_j)(1 - E_j)$ . If the emigration rate is large,  $\hat{S}_j$  will severely underestimate the true survival rate of the species under study. The true survival rate ( $= 1 - D_j$ ) can be estimated if one validly assumes that the emigration rate ( $E_j$ ) is zero.

Similarly, birth and immigration cannot be estimated separately. The influx or dilution parameter  $B_j$  is the sum of the number of recruits, due to reproduction in the population under study, and the number of immigrants, which are not produced by the population under study. In practice, a separate estimate of the number of immigrants is possible if these individuals are much older (bigger) than the locally generated recruits. However, the Jolly-Seber model applies to only one age class, which is typically adults, and no age effects on parameters are allowed. Thus, "recruitment" means entry into the adult segment of the population, and distinguishing whether recruits are from reproduction of the population being trapped or whether they are immigrants may be impossible. \*Figure 8.1 shows the fully open model and special cases that allow only influx or only outflux.

2. Parameters involving influx, outflux, and catchability are allowed to vary among sampling occasions, but not among animals. For example, survival rate may change from  $S_j$  between the  $j^{\text{th}}$  and the  $(j + 1)^{\text{th}}$  occasions to  $S_{j+1}$  between the  $(j + 1)^{\text{th}}$  and the  $(j + 2)^{\text{th}}$  occasions, but during each of those time periods, all members of the population are assumed to have the same probability of survival. Similarly, on the  $j^{\text{th}}$  sampling occasion, all  $N_j$  animals then in the population are assumed to have the same probability of capture  $p_j$ . (This capture probability structure corresponds to Model  $M_t$  of Chapter 3.)

3. As *Cormack (1979)* points out, the estimates of survival, birth, and probability of capture are optimal only if the parameters really do differ between sampling occasions. Thus, for example, if the biologist wants to assume that the survival rate remains constant between all sampling periods, the optimal estimate of this single survival rate parameter requires a substantially more complex estimation method.

The mathematical notation found in the publications by *Jolly (1965)*, *Pollock (1975)*, *Robson (1969)*, and *Seber (1965)* is somewhat intimidating. This being the case, it is worthwhile to note that the general problem is solved if the biologist can estimate two quantities for each sampling occasion  $j$ : the number of marked animals  $M_j$  alive in the population and the capture probability  $p_j$ . The problem is a little more complicated in the extensions proposed by *Robson (1969)* and *Pollock (1975)*, because at each sampling occasion  $M_j$  and  $p_j$  may differ over some subclasses of



George M. Jolly

George Jolly's interest in capture-recapture studies began while he was at the East Malling (Fruit) Research Station. There he encountered an entomologist using capture-recapture methods on a local population of an orchard predator. He began by studying the research papers by J. N. Darroch and by P. H. Leslie and his colleagues, and the work resulted in the well-known, general model for open populations published in 1965. That, of course, was before the computer era, when mathematically involved solutions had not yet acquired the respectability they enjoy today.

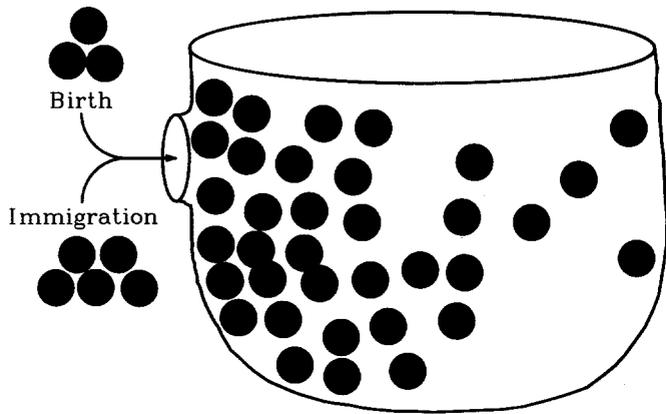
Jolly completed a mathematics degree at Aberdeen, Scotland, and studied statistics at Edinburgh University with a view to agricultural research. He then worked for 5 years at Rothamsted Experimental Station under Frank Yates. He has been with the Agricultural Research Council for many years.

His recent work has been in developing estimation methods allowing survival to remain constant over the study period. He believes that the full potential of capture-recapture methods is difficult to assess until further models, more complex but more realistic biologically, have been developed and studied. Although many estimators must be found numerically, Jolly feels it is helpful when formulas can be expressed in easily interpretable form. He believes it is unfortunate that some biologists cling to the out-dated and inefficient techniques still appearing in publications. (Photograph taken in mid-1960s.)

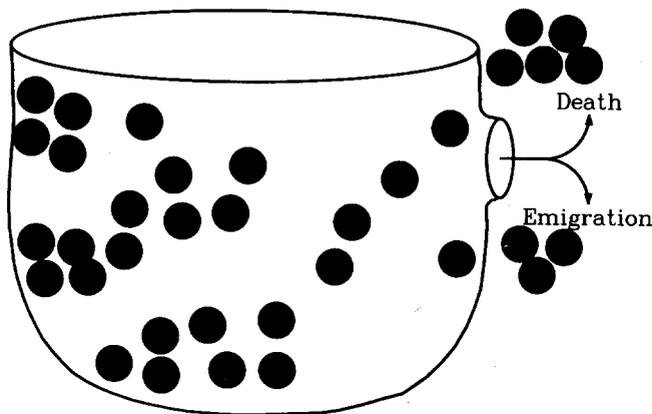
the population, such as those defined by previous capture history. Still, the essentials of the estimation problem for open models are related to the estimation of these two types of parameter sets: the number of surviving marked animals and the capture probabilities.

A subtle, but serious, problem arises with capture studies if simple correlation-regression analyses are applied to the estimated parameters in an attempt to infer anything about the underlying processes of population dynamics. The application is invalid primarily because both the estimators  $\hat{N}_2, \dots, \hat{N}_{t-1}$  and the estimators  $\hat{S}_1, \hat{S}_2, \dots, \hat{S}_{t-2}$  have very strong sampling correlations among themselves. The meaning of the sampling correlations is hard to make clear without presenting technical material beyond the intended level of this primer. However, the basic idea is that because all of the parameter estimates are computed from one common set of data, spurious relations—that is, relations that have nothing to do with the true dynamics of the population—are forced to appear among parameter estimates. As a simple analogy, if we have a set of data,  $x_1, \dots, x_n$ , define a new variable,  $y_i = 1/x_i$ , and then look at the correlation (relation) of  $x$  to  $y$ , we will find a strong, negative correlation. It is wrong to interpret this relation as implying anything about a physical or biological process underlying the “variables”  $x$  and  $y$ . This mistaken approach to investigating population processes is, unfortunately, encountered often in the literature pertaining to ecology. See, for example, *Eberhardt (1970)*; *Anderson and Burnham (1976:13-15, 31, 34-39)*; *Brownie et al. (1978:177-179)*.

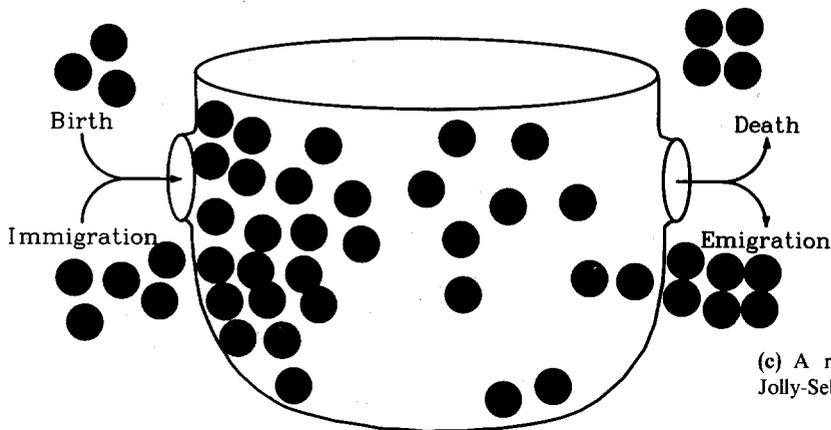
The interrelations of *estimated* parameters are spurious as regards any existing relations between the *true* parameters. As a result, we cannot validly explore population dynamics, such as density dependence of (true) survival rates, by the use of simple regression techniques that depend on estimates of survival rates and population sizes from capture-recapture studies. This is an important limitation of the use of



(a) A representation of the Jolly-Seber model with birth and immigration, but no losses: survival rate is assumed to be 1 and emigration rate zero.  $N = 35$  in this example.



(b) A representation of the Jolly-Seber model with death and emigration, but no gains: birth and immigration rates are assumed to be zero.  $N = 27$  in this example.



(c) A representation of the (demographically) fully open Jolly-Seber model.  $N = 31$  in this example.

\*Fig. 8.1. Open models allow the assumption of demographic closure to be relaxed. The concept of geographic closure is still required (something analogous to the sides of the glass container is needed), because without it, the population size parameter  $N$  has no well-defined meaning. Jolly (1965) and others have developed models for the special case represented by (a) and (b). However, birth and immigration are confounded in (a) as are death and emigration in (b). The rates of the two processes shown in each figure cannot be estimated separately. For example, the estimate of the "influx" parameter (a) is biologically meaningful only if one process, say, immigration, is zero. If immigration is zero, this parameter can be interpreted as an estimate of reproductive recruitment. When this influx parameter is estimated, it frequently is called a dilution rate. (c) represents the usual fully open population model for which several elegant methods of analysis have been developed (Jolly 1965; Seber 1965; Robson 1969; Pollock 1975). In the open models,  $\hat{N}_j$  is an estimator for the size of the population at the  $j^{\text{th}}$  sampling occasion. Of course,  $N_j$  changes through time as animals come and go.

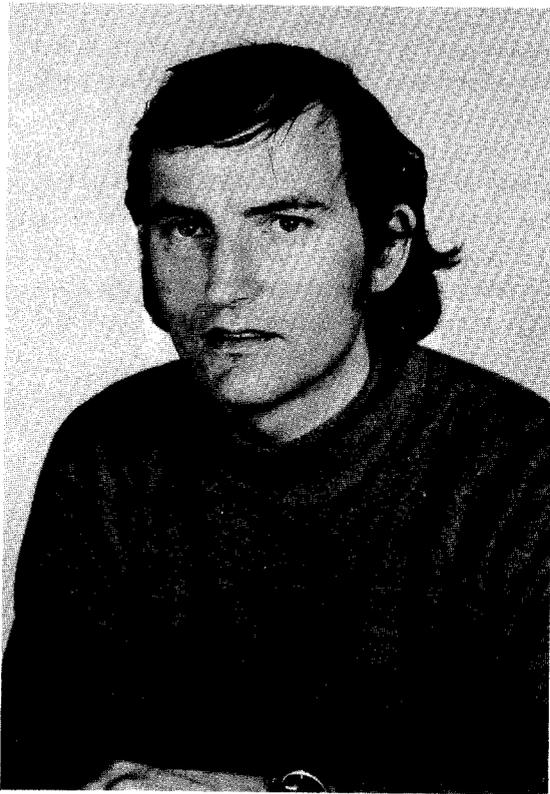
capture studies. As yet, the proper use of capture-recapture data to explore population processes has not been developed.

Clearly, the Jolly-Seber method can be too general (survival rates, in fact, may not vary over time), too specific (no age effects are allowed for), or not appropriate, depending upon the particular experimental situation. A more restrictive model in which survival rate is constant over time has been considered by *Jolly (1979)*, who provided ML equations for the parameters of interest. Direct solutions have not yet been developed. *Crosbie (1979)* considered a "modified" Jolly-Seber model, which accounts for the animals that enter and exit the population between two successive sampling occasions and thus have no chance of being marked. He then developed a sequence of models formed by making different assumptions, generally more restrictive than Jolly-Seber assumptions, concerning capture probabilities, survival rates, and birth distributions, and created an elaborate computer program that attempts to choose the most parsimonious model for the data at hand. In concept, this approach is very similar to that used by *Otis et al. (1978)* in their treatment of closed models. In addition, there are published models allowing different classes of animals in the population to have different probabilities of capture and survival. *Robson (1969)* has expanded the basic model so that the first capture is allowed to affect the animal's survival rate for either one or two subsequent time intervals. *Cormack (1972)* has investigated the case in which the first capture affects both the animal's probability of capture and its survival for the subsequent time period. The most general capture-recapture model for an open population is that formulated by *Pollock (1975)*. In his model, animals with different capture histories may have different survival and capture probabilities in subsequent sampling occasions, but he actually derives estimates only for the two cases in which the first capture affects survivability for  $\ell$  subsequent time periods and catchability for either  $\ell$  or  $\ell - 1$  subsequent periods. The fact that estimators for other forms of heterogeneous probabilities are theoretically possible, but computationally very difficult explains why Pollock's models, as well as those of Robson and Cormack, have not been put to much practical use. We hope that the availability of versatile, easy-to-use computer packages, like the one developed by *Arnason and Baniuk (1978, 1980)*, will facilitate practical implementation of these methods.

## Discussion

Now that we have developed a feel for what the open-model capture-recapture techniques have to offer, we can discuss when and how the two general classes (open and closed) of techniques compete with each other. First, however, we must deal with the common misconception that open-population models avoid the concept of geographic closure. Although the assumption of geographic closure is often approximate at best, both closed and open models suffer if it is violated. To illustrate this point, consider an experiment in which a grid of live traps has been placed in a large planting of sugar cane for the purpose of estimating the size of the cotton rat population using the field. Clearly, a closed-population model used for this experiment involves the parameter  $N$ , the absolute number of rats in the "population." A biologist may feel uncomfortable with this concept, because he knows that rats can come and go across irrigation ditches surrounding the cane field. Consider, however, that any open-population model, for example, the general Jolly-Seber model, postulates the existence of a population of  $N_j$  individuals present on the  $j^{\text{th}}$  sampling occasion. When a model contains a parameter representing the number of individuals in the population at a specific time, the idea of geographic closure is involved, because when an experimenter speaks in terms of a population of absolute size he implies the existence of a geographic area in which the population resides. Thus, the concept and assumption of geographic closure are required of both open and closed models of populations.

We have assumed that the objective of the capture-recapture experiment is to estimate population size. Often, this objective is only secondary, and the main purpose of the experiment is estimation of birth rates or death rates, or both (*Robson 1963; Cormack 1964*). In such instances, of course, closed-population models are not useful, and the investigator must use methods that allow these processes to occur.



Bryan F. J. Manly

Bryan Manly has made several contributions to the theory for open-population models. He was the first to develop estimators for age-dependent populations. He conducted simulation studies to explore the small-sample properties of the Jolly-Seber model. He has published some 15 papers on capture-recapture methods—many of them in an entomological setting.

Manly took a B.Sc. degree in mathematics from the City University in London and worked as an industrial statistician before going to the University of Salford. There, M. J. Parr introduced him to ecological problems and to some of the statistical questions arising in the analysis of capture-recapture data. They shared ideas and began computer simulation studies to see how theoretical results worked in practice. In 1970 Manly accepted a Lectureship in Statistics at the University of Papua and New Guinea. Since 1973 he has been in the Biometrics Unit at the University of Otago, New Zealand.

Manly's recent work in capture-recapture studies has been concerned with accounting for animal movement and parsimonious modeling of capture data. He believes these and other related problems deserve a great deal more work. (Recent photograph.)

Return now to the case where estimation of numbers is of primary concern and, therefore, where a choice must be made between the two general classes of estimation techniques. We can reduce the basic dilemma to one question that the investigator must ask himself: Which class of models best approximates the experimental situation—open models, which allow the processes of birth, death, and migration to operate, or closed models, which allow individuals to possess varying probabilities of capture, depending upon which of three sources of variation are present? This question must be answered because no practical estimation techniques that allow both types of assumptions are available. There is a good reason for this, one that we have alluded to before: capture-recapture models can be generalized only to a certain degree before they lose their ability to provide real information about the parameters of interest. Thus, the problem of choosing between the two classes of models may be with us indefinitely. Unfortunately, statistical studies of the models involved do not permit the conclusion that one or the other class of model will be robust to failure of the constraining assumptions each requires. For example, both *Carothers (1973a, b)* and *Gilbert (1973)* have documented the fact that Jolly-Seber estimates can exhibit significant negative bias when individuals show certain types of unequal catchability. Although the models of *Otis et al. (1978)* have not been examined for robustness to the presence of birth, death, and migration, it is reasonable to expect that the presence of these processes could effect significant biases in the population estimators.

Thus, we see that the experimenter must bring all his biological knowledge of a particular experimental situation to bear on the selection of proper estimation techniques. According to *Cormack (1968:456)*, "In all cases every iota of information, both biological and statistical, must be gathered to check and countercheck the unavoidable assumptions." Although each situation will be different at least to some degree, a few general principles may be of help. First, closed-population techniques are more likely to be appropriate for short-term studies. In this context, short-term is taken as relative to the mean life span of the species. For example, daily trapping of an area for a week might be appropriate for *Microtus* spp.,

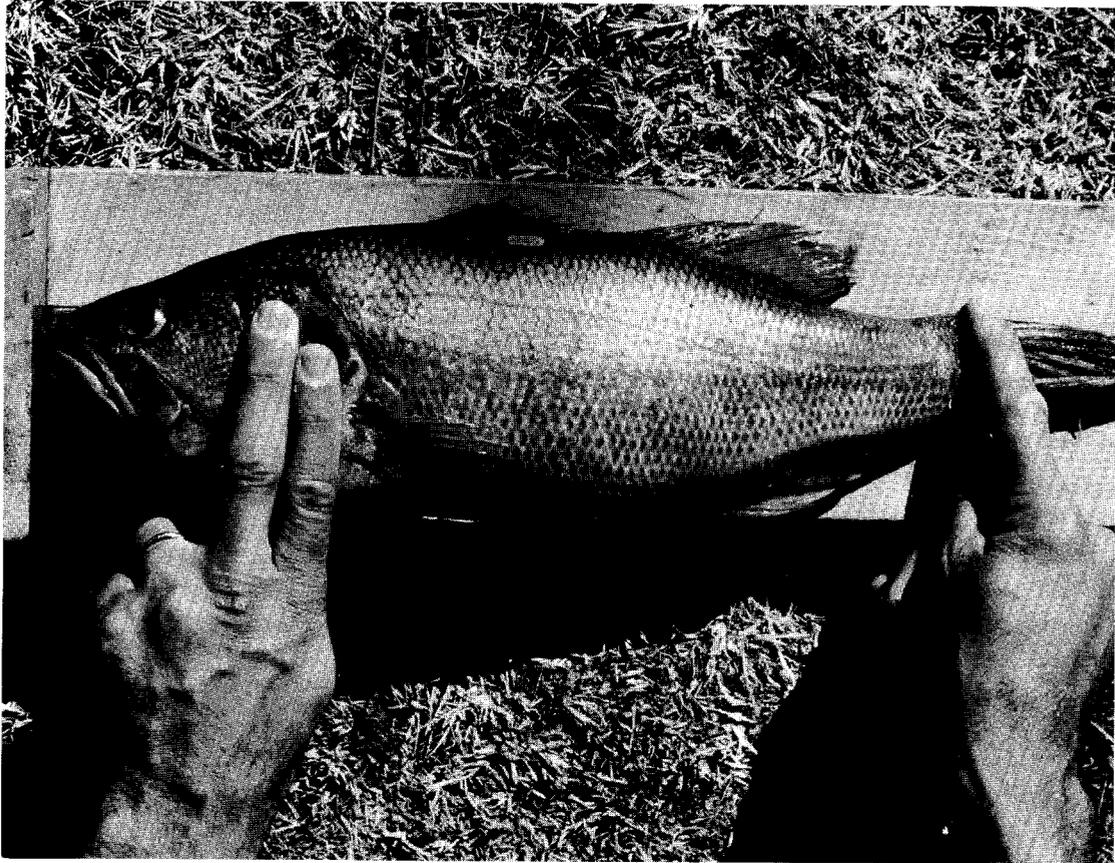
whereas weekly trapping for a month or two might be appropriate for *Lepus* spp. As the length of the trapping period increases, the probability increases that the processes of birth, death, and migration will affect the population significantly and, therefore, open models may become more appropriate. Further, there is the common sense notion that the size of the area trapped must be relatively large compared to the home range of the species under study if closed models that do not allow migration are to be used (see Chapter 5). For example, we would not assume that the effects of migration were negligible in an experiment in which coyotes were being trapped in a 1.6-km<sup>2</sup> grid of traps. Finally, we mention that closed-population models are likely to be most effective in producing good estimates if the trapping is not done during a period when young of the year appear in the population or during a time of significant dispersal.

### Summary

1. Open models allow the demographic closure assumption to be relaxed. The population size is allowed to vary among the sampling occasions, as  $N_1, N_2, \dots, N_t$ , because animals may enter the population, through birth or immigration, or leave the population, through death or emigration.
2. The most commonly used approach is the Jolly-Seber model, which is an extension of Model  $M_t$ . It allows estimates of population size, probability of capture, loss rate (deaths and emigration), and entry rate (births and immigration) for each sampling occasion  $j$  except, primarily, the first and last sampling occasions.
3. Heterogeneity has not been modeled for the open case, and, conceptually, no open removal model is possible. Behavior has been modeled for the open case only to a limited extent.
4. Geographic closure is a critical assumption for estimating population sizes  $N_j$ , but it can be ignored if only estimates of loss rates are of interest, (because these estimates are based only on the marked animals).
5. In most instances, open and closed models do not compete with each other: only one or the other is appropriate for satisfying the objectives of a given study.
6. Computer programs have been developed to allow a comprehensive analysis of open-population data (*Arnason and Baniuk 1980*).

### Questions and Exercises

1. A biologist is to study lake trout in Lake Superior and considers an open capture-recapture model in his work.
  - a. What problems might he face?
  - b. What parameters might be estimated?
  - c. What alternative methods should he consider?
2. You are interested in estimating the monthly survival rate of fish in a 1-ha pond. Would an open model be appropriate? Why?
3. Are the samples required for the open models relatively larger than those required for comparable closed models?
4. Using the Jolly-Seber model, a researcher obtains the following data.



A largemouth bass (*Micropterus salmoides*) has been tagged with a disk-dangler tag under the dorsal fin. (Photograph courtesy of Paul R. Turner.)

$j$	$\hat{N}_j(\hat{se}(\hat{N}_j))$	$\hat{S}_j(\hat{se}(\hat{S}_j))$	$\hat{B}_j(\hat{se}(\hat{B}_j))$
1	---	0.50 (0.09)	---
2	389 (42)	0.40 (0.06)	114 (70)
3	270 (51)	0.69 (0.08)	236 (106)
4	422 (59)	0.24 (0.05)	206 (72)
5	307 (67)	0.65 (0.04)	288 (106)
6	488 (96)	0.23 (0.07)	199 (90)
7	311 (60)	0.56 (0.15)	106 (130)
8	280 (100)	---	---
9	---	---	---

- Plot  $\hat{N}_j$  vs  $\hat{S}_j$ . Do these estimates provide support for the hypothesis that survival is low when the population is high (that the survival rate is density-dependent)?
  - What is the value of  $cv(\hat{N}_4)$ ?
  - What is the 95% confidence interval on  $S_5$ ?
  - Are the estimates of  $B_j$  related to the estimates of population size?
5. If a survival rate is assumed to be constant during a capture-recapture study, can this one survival rate be estimated with a Jolly-Seber model?

# CHAPTER 9

## THE FUTURE

Until recently, most biologists were isolated from quantitative methods and computing devices, but they now are beginning to be trained in basic statistics, deterministic models, simulation, and the use of digital computers. Scientists in the biological fields are progressing rapidly into sophisticated fields of analysis such as stochastic processes, multivariate analysis, and nonlinear and dynamic programming. We see this as a positive trend.

### Biological Developments

Major biological developments in the future probably will involve improvements in field techniques, to reduce the variability in capture probabilities. As our knowledge about the mechanisms that cause differences in capture probabilities increases, experiment design can be improved to eliminate the variability. Biotelemetry provides one means of studying the reasons why capture probabilities vary between animals, and the ways in which animals react to the placement of traps in their home ranges. For example, does the strangeness of a trap placed in the middle of an animal's home range tend to scare the animal away? Does its curiosity make the same animal investigate a trap placed in a strange area outside of its home range? Where in an animal's home range should a trap be placed to maximize the chance of catching that individual? Such questions will be answered as we learn more about the behavior of animals in relation to capture-recapture experiments.

Among the possibilities for eliminating some of the variability of capture probabilities are better ways to mark animals and later to differentiate marked from unmarked animals. One promising prospect involves methods in which marked and unmarked animals are distinguished without actually being recaptured. Possibilities include the use of electronic tags and an electric eye gate: one counter would tally all animals that pass through the gate and, in addition, the electronic tag would cause a second counter to tally only tagged animals. A number of animals would be tagged initially to introduce a marked segment into the population. The actual experiment to determine the proportion of marked animals would not increase the number of marks, but would only make counts of marked versus unmarked animals (*Rice and Harder 1977*). Because the animal would not be captured, behavioral response to trapping would be eliminated during the study's second stage. Moving the counting devices would help reduce heterogeneity of capture probabilities. Aerial surveys could be used to count animals such as mule deer, and an electronic device could be used to determine which animals carry the electronic tag. Such a scheme would eliminate the observer bias of spotting brightly colored neck collars before the animal is seen, or of failing to see a dull colored collar on a marked animal and thus counting it as unmarked.

A conceptually similar idea is to inject animals with a chemical, which is eliminated in the feces. Feces are collected, and the number of marked and unmarked feces are used to estimate the proportion of marked animals in the population. *Pelton and Marcum (1975)* used the method to estimate the number of black bears in the Great Smoky Mountains National Park, and *Davison (1980)* used it to estimate coyote populations in the Idaho National Environmental Research Park. In both studies, the radioactive materials used to tag the animals made detection in the feces relatively easy with radiation counters. An extension of the same idea could use stable chemical elements that are rare in the environment, but for which analytical techniques exist, to determine if the amount of material in the feces was above the

background for the element and thus from a marked animal. Fecal-tagging methods avoid the problem of behavioral response to trapping, unless the injected substance causes a change in defecation rates. However, differences in defecation rates between individuals or age and sex classes may cause significant heterogeneity in defecation rates and hence in the probability of finding the animals' feces.

The old adage about developing a "better mousetrap" is true. Improvements in all types of gear are needed, including increased overall reliability, the ability to catch several animals in a single trap, reliability in freezing temperatures, prevention of trap deaths or escape of captured animals, and reduced costs per unit. Often, the tags now being used fall off, or become unreadable, or cause mortality or infection. The use of microtaggants is one example of needed innovations (see *Johns 1979*).

Perhaps the least explored subject deals with alternative trapping designs. At present, only three general designs are widely used: (a) traps in a single line—a design totally without merit, even for use as an index; (b) traps in a rectangular or irregular grid; and (c) a grid with assessment lines. Other designs might be possible and could even facilitate the direct estimation of density.

Finally, there is a critical need for more studies like those conducted by *Mahon (1980)*, in which the true size of the animal population is known, and the capture-recapture or removal sampling is conducted to estimate  $N$ . This type of study enables a strict comparison of  $N$  versus  $\hat{N}$  and is most meaningful if it is replicated. Studies like those conducted by *Edwards and Eberhardt (1967)* and *Carothers (1973b)* are very informative, but they allow comparison and performance of estimators based on only one outcome of the stochastic sampling process.

We are extending a plea to authors and editors of manuscripts involving the use of capture-recapture methods to document any experiment features that might help others to design experiments with the same species. Most commonly, only the population estimate and its standard error are reported, with no reason given for using the particular estimation technique and no estimates of capture probabilities provided. An increase in the amount of such information would facilitate the construction of efficient and successful trapping designs. For example, when the results of several trapping experiments involving house mice and using the program CAPTURE to analyze the data are reported, other researchers planning similar work will want to know which models tended to be selected for estimation and the estimated probability of capture. In a sense, such information is as valuable to other experimentors as the actual population estimates.

## Statistical Developments

Statisticians have contributed to capture-recapture theory on two levels: first, in the development of new models and their associated methods of analysis, and second, in the evaluation of established methods (for example, investigations of the robustness and efficiency of estimation methods and evaluation of the power of tests of assumptions). More work needs to be done in both areas. Meaningful contributions on new, realistic models probably will be made by statisticians working with biologists, as opposed to mathematically trained statisticians with no background in biology, working alone. There is a large potential for improved capture-recapture studies if biologists and statisticians work together to design capture and removal studies and associated models (cf. comments of *Hartley 1980:2*).

We foresee the continuing development of both realistic new models and special cases of existing models. Development of special cases will be motivated by the real problems biologists face. Development of new, realistic models will be based on experience with the capture-recapture method and the insights to be gained by extensive analysis of existing data. In both regards, the joint efforts of biologists and statisticians are needed. For such cooperation to work well, the statistician needs to know something about biology, including the realities of field work, and the biologist needs to have a basic training in statistical methods, preferably including knowledge of and experience with computers. Given teams with experience in obtaining and analyzing capture-recapture data, we foresee the development both of models incorporating specific biological characteristics of target populations and of optimal statistical analyses.

When statisticians and biologists work alone, the statistician often builds elegant but unrealistic models, while the biologist, who knows the realism the models should have, cannot formulate efficient statistical analyses.

**Closed Models.** A difficulty plaguing both open and closed models, but better understood in closed models, is the poor confidence interval coverage on population size  $N$ . The difficulty derives from two problems: (1) the normal distribution is a poor approximation to the sampling distribution of  $\hat{N}$  for many models, even for moderate sample sizes; and (2) there is a very high sampling correlation between  $\hat{N}$  and  $\hat{se}(\hat{N})$ . The latter problem occurs because in most models  $\text{var}(\hat{N})$  generally equals  $N$  times some function of the capture probabilities and  $t$  (capture occasions), and therefore the estimated standard error of  $\hat{N}$  is proportional to  $\sqrt{\hat{N}}$ . This correlation causes  $\text{se}(\hat{N})$  to be overestimated when  $\hat{N}$  is larger than  $N$  (in this case coverage exceeds 95%) and to be underestimated when  $\hat{N}$  is less than  $N$ . Of course, in practice we do not know  $N$ , hence we do not know which situation prevails, nor its seriousness. Values of  $\hat{N}$  much smaller than  $N$  generally result in a 95% confidence interval that does not cover  $N$ , whereas values of  $\hat{N}$  much larger than  $N$  result in an interval that almost always covers  $N$ .

The problem of nonnormality may be solved by determining better approximations to the sampling distribution of  $\hat{N}$ , directly or through transformations that improve the normality assumption. For the jackknife estimator  $\hat{N}_h$  of Model  $M_h$ , the estimated variance typically underestimates the true sampling variance, although the normality assumption on  $\hat{N}_h$  seems acceptable. In this situation, a bootstrap method (*Effron 1979*) might be used to advantage to estimate the sampling variance of  $\hat{N}_h$ . In other models, extensive simulation may provide insights into the distribution of  $\hat{N}$ , so that more reliable confidence intervals can be constructed; the concept is related to the idea of using the bootstrap method based on a given set of data to investigate the distribution of  $\hat{N}$ .

Whether using the uniform minimum variance unbiased estimator (UMVUE) of  $\text{var}(\hat{N})$  will solve the correlation aspect of the coverage problem is not clear; however, developing these variance estimators is worthwhile. *Berg (1976)* has developed the UMVUE for  $\hat{N}$  of Model  $M_t$ . Similar work remains to be done for the removal estimators and Model  $M_o$ .

In general, asymptotic results for the estimators  $\hat{N}$  and  $\hat{se}(\hat{N})$  and for tests of assumptions should be avoided as much as possible. Such results can be especially poor for capture-recapture models. Alternative estimators to the ML approach have not been investigated widely for capture-recapture data. For example, minimum chi-square, minimax, and decision theoretic or information theoretic approaches may have merit. Calculations for all of them are highly computer oriented.

Bayesian statistical methods have found little favor in wildlife studies, partly because there has been almost no investigation of Bayesian methods for use with capture-recapture studies. We doubt they will find much favor, except in modified form. For example, specific types of probability distributions for the capture probabilities in the heterogeneity model may be appropriate. If the distributions could be deduced on the basis of the analysis of capture-recapture data, they might lead to improved estimators of population size under Model  $M_h$ . The same idea applies to the generalized removal model; if the heterogeneity could be modeled in some parametric form, improved removal estimation might result. Finally, the model selection rule of *Otis et al. (1978)* could be improved. As now formulated, it unrealistically assumes equal prior probabilities on all eight models. Unfortunately, we do not yet know on a species-by-species basis what better priors to use. An improved discriminant function rule should include the option of unequal prior distributions on the models in conjunction with specific information on which priors are appropriate for use in specific situations.

There are other ways to improve the model selection rule. One quick improvement would be to allow unequal covariance matrices in the discriminant function analysis by using quadratic discriminant functions. The quadratic equations could then replace the linear equations now available in program CAPTURE. There is also a problem with the coherence of the selection criteria with respect to Model  $M_t$ . Often, when the selection criterion is 1 for Model  $M_{th}$ ,  $M_{tb}$ , or  $M_{tth}$ , it is zero for Model  $M_t$ , but

substantially more than zero for Model  $M_o$ , indicating that the time function is a significant variable, along with others, but that Model  $M_o$  is preferred over Model  $M_t$  as an explanation of the data. We suspect that this incoherence relates to problems with the goodness of fit test for Model  $M_t$ , which requires a large amount of data and apparently has very low power. Conceptually, a selection rule that recognizes the relations among the models might help solve this "coherence" problem. In the extreme, a nonparametric discriminant function approach may be possible.

The power of the tests is a general problem; none appear to have very high power. Thus, the search for better tests of assumptions is a fruitful area for future research in capture-recapture methods. What approaches might be tried are not clear. Randomization methods are worth considering for these tests, as is the related bootstrap method of *Efron (1979)*.

There is a clear need for estimation procedures for Models  $M_{th}$ ,  $M_{tb}$ , and  $M_{tbb}$ . In each case, the parameters in these unrestricted models cannot be estimated individually. Thus, progress requires some way to reduce the number of parameters. In Model  $M_{tb}$ , the assumption that first-capture and subsequent recapture probabilities are proportionally related across the time effect would make  $N$  estimable. The reality of this assumption could be explored. A jackknife or bootstrap method might work for Model  $M_{th}$ . Alternatively, the introduction of moderately specific distributions on the capture probabilities for the heterogeneity component would reduce the number of parameters to a tractable number. We hold little hope for estimation under Model  $M_{tbb}$ .

Models intermediate to those in *Otis et al. (1978)*, such as models to allow recaptures to bear on the estimation of  $N$  in Models  $M_b$  and  $M_{bb}$ , also may be useful. To obtain such models, however, we must postulate a relation between first and subsequent capture probabilities.

The log-linear model approach (*Cormack 1981*) deserves attention; its properties for analysis of capture-recapture data should be determined. Whether it works well for real capture-recapture data is not yet known, and one good simulation study would go a long way toward answering this question. The study should generate the data by using models based on different capture probability structures. The log-linear theory has no obvious relation to any capture-recapture probability structure except for Models  $M_t$  and  $M_o$ . When the log-linear method of analysis selects a model with interaction terms, it is difficult to relate the model to heterogeneity or behavior (or both) or to some interaction among the sources of variation in capture probabilities. This difficulty will hamper biological understanding of capture results and hence hamper the design of future studies and the development of better methods of analysis. More research ought to lead to a better understanding of the results of log-linear analysis.

Log-linear models also may be improved by recognizing the time-ordering of captures. Because the time-ordering of the capture occasions is not taken into account in present log-linear models, a proper analysis of Models  $M_b$  and  $M_{bb}$ , the models requiring removal estimators, is especially difficult. True removal studies clearly do not fit into the log-linear framework. All these considerations lead us to believe that the log-linear approach may not solve the problem of developing a unified approach to analysis of capture data.

Relatively fewer areas of research are left for removal models, compared with capture-recapture models. All major research on mathematical aspects of the constant capture probability removal model has been done. There are ways to continue to investigate and modify this model (*Skalski and Robson 1979, Carle and Strube 1978*), but they are minor changes or refinements to the basic model and the ML estimator. However, investigations and improvements in the area of removal studies with heterogeneous capture probabilities are needed. The generalized removal estimator is the only existing treatment of this situation. Alternative approaches could model the degree of heterogeneity or could empirically model a relation between the conditional average capture probabilities on each occasion ( $\bar{p}_j$ ). For example, perhaps a model such as  $\bar{p}_j = a/(a + bj)$ ,  $\bar{p}_j = a/(a + b + j)$ , or  $\bar{p}_j = p + e^{-bj}$  would lead to more robust estimation for removal studies.

A final subject of concern is estimation of density. Improvements in the method presented in Chapter 5 and in *Otis et al. (1978:67-74)* are surely possible. The quite different approach of using assessment lines has not received enough attention to determine its validity and usefulness. A combination of statistical

modeling and empirical research on assessment lines seems appropriate. A method, based on trapping, for the direct estimation of density without first estimating  $N$ , then estimating the area, would be very attractive, if one could be found.

**Open Models.** This primer deals with robust analysis of capture data from closed populations. However, there is a tremendous need, and hence a challenge, to develop robust methods of estimating population size for the open models discussed briefly in Chapter 8. The number of possible open models is so large that a comprehensive approach would be very difficult. To simplify this discussion we will consider the Jolly-Seber model (*Jolly 1965; Seber 1965*) as a baseline from which to proceed with new methods.

For one age class and time-only variation, the Jolly-Seber model provides an adequate approach (Chapter 8). However, there is a need for both more restricted and more general models. More restricted models are obtained by reducing the number of parameters; the reduction tends to reduce the realism of the model, but may achieve an overall gain in effective estimation of the key parameters. In particular, the survival parameters can be reduced from  $t - 2$  to 1 by assuming a constant survival rate. Similarly, the recruitment parameters may be reduced by assuming a constant number of recruits per trapping period, assuming a relation between recruitment rate and population size, or otherwise modeling a relation into the recruitment parameters to reduce their numbers. Both approaches to developing more restricted Jolly-Seber models are being pursued.

The need for generalized open capture-recapture models to allow for other sources of variation in parameters is at least as great as the need for more restricted models. The generalized models will require more and/or different parameters to gain greater realism, hence robustness, with respect to the effects of behavior and heterogeneity. To date, there has been no incorporation of heterogeneity, such as in Model  $M_{11}$ , into open models. We believe the failing is very serious, because both the literature and our own experience with closed models strongly suggest that heterogeneity is always present to some degree. When it is a strong source of variation in capture probabilities, it will seriously bias estimation of population size. The only attempts to allow for heterogeneity that we are aware of have been attempts to develop age-stratified Jolly-Seber models, analagous to those of *Brownie et al. (1978)* for banding data. *Pollock (1981a)* has developed such age-stratified open models. Age, however, is only one potential source of heterogeneity; many sources are not identifiable and hence cannot be overcome by stratification.

Within the next decade we expect to see many developments of specialized and generalized Jolly-Seber models. The worth of these models must be evaluated by analytic and simulation methods and by application to real data. Application to real data, of special importance, will require prompt implementation of these methods in user-oriented computer programs. The state-of-the-art computer program for dealing with open models is POPAN-2 (*Arnason and Baniuk 1978, 1980*), which has many useful features. It implements all the Jolly-Seber models and many of Arnason's innovations, but it implements few of the special cases achieved by assuming restrictions on the survival, recruitment, or capture probability parameters. Nor has any of the work of *Pollock (1975)* been implemented yet. These comments are not a criticism of POPAN-2; rather, they are suggestions for future developments.

We recommend the extensive analysis, or re-analysis, of good capture-recapture data for open-model studies, using program POPAN-2. Publication of the results will help resolve the question about which of the existing models are useful with real data.

The problem of density estimation ( $D = N/A$ ) for open models also must be addressed. There is virtually nothing in the literature about this problem.

To summarize, a great deal of work remains to be done on open models. Research on most of the problems is ongoing, and we expect to see substantial progress in the next 5 years. The challenge will be to evaluate the methods quickly and to make them available for routine applied use. Certainly by the end of this decade there should be a complete synthesis of reliable, efficient methods for the analysis of

capture data assuming an open model. Although no such synthesis exists now and one would be premature in the next 5 years, it will be critical to further progress.

## Computer Programs

Computer programs have become essential in the analysis of capture-recapture and removal studies. The iterative nature of many estimators of population size under closure makes them nearly impossible to compute with a hand calculator, and the testing and model selection procedures are tedious. The notation and algebra for the estimators of open-population parameters are difficult, and recently developed models do not have closed-form estimators. In all cases, rounding errors, especially for the estimates of sampling variances, can be serious on a calculator. Now and in the future, a comprehensive analysis of any set of multiple capture data will require the use of sophisticated computer programs.

In the next decade, we foresee the development of a comprehensive, integrated series of computer algorithms, similar in concept to the SAS (Statistical Analysis System) program (*SAS 1979*). The program should be in FORTRAN and should be machine independent, at least for common makes of computers in the medium-to-large size class.

The program should include a data maintenance capability that is both sophisticated and easy to use, à la *Arnason and Baniuk (1978)*. The basic data must be accessible to any of the analysis, testing, model selection, or display routines. The system should operate in an interactive, remote-batch, or batch mode at the option of the user.

An extended graphics display routine will be advantageous. The routine should allow, for example, display of a three-dimensional histogram showing the number of animals caught per trap on the grid or subgrid, simple histograms of the data by age, sex, and weight, and figures showing estimates, standard errors, and confidence intervals for the results. A versatile routine to allow students or researchers to simulate data from fixed parameters is a powerful feature of such programs (*Arnason and Baniuk 1978*).

We expect the number of separate algorithms in the system to increase in a manner similar to the way additions and improvements have been made to the BMD series (*Dixon and Brown 1979*). Tables 9.1, 9.2, and 9.3 summarize some types of studies from which data are available, the analysis methods used, and the existing programs. Possible additions and improvements are discussed below.

Many other methods could be added, including the many catch-effort (*Pella and Tomlinson 1969*) and change-in-ratio methods (*Paulik and Robson 1969; Otis 1980*). Several ML estimators have regression-type approximations that might be added if they were more robust. Finally, many less general methods might find a place in a comprehensive computer system; estimators for instantaneous rates are an example.

A good computer system and its accompanying software are now mandatory, if biologists are to benefit from the statistical and theoretical advances made in the past decade. A general, flexible, easy-to-use system will be a great help in future research on biological populations. A prediction made in 1969 by *Overton and Davis (1969:404)* reinforces our argument.

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.

**TABLE 9.1.** Analysis methods for studies where data are available on only marked members at time j.

Method	Existing Program, Reference
Seber (1970); Robson and Youngs (1971)	ESTIMATE, Brownie et al. (1978)
Other models allowing age-independent parameters	ESTIMATE, Brownie et al. (1978)
Brownie and Robson (1976)	BROWNIE, Brownie et al. (1978)
Other models allowing age-dependent parameters	BROWNIE, Brownie et al. (1978)
Johnson (1974)	--- Johnson (1974)
Seber (1971); Cormack in Fordham (1970)	--- Fordham (1970)

**TABLE 9.2.** Analysis methods for studies where data are available on only unmarked members at time j.

Method	Existing Program, Reference
Zippin (1956, 1958); Moran (1952)	CAPTURE, Otis et al. (1978)
Generalized Removal Method; Otis et al. (1978)	CAPTURE, Otis et al. (1978)
Chapman and Robson (1960); Robson and Chapman (1961)	--- ---
Dupont (1976)	FISH, Dupont (1976)

**TABLE 9.3.** Analysis methods for studies where data are available on both marked and unmarked members at time j.

Method	Existing Program, Reference
<b>Closed Models</b>	
Darroch (1958, 1959); Burnham & Overton (1979); Pollock (1974)	CAPTURE, Otis et al. (1978)
Other estimators assuming closure	CAPTURE, Otis et al. (1978)
<b>Open models</b>	
Arnason (1972, 1973)	---
Cormack (1964)	---
Darroch (1961)	---
Fisher and Ford (1947)	POPAN-2, Arnason & Baniuk (1978, 1980)
Jolly (1965); Seber (1965)	POPAN-2, Arnason & Baniuk (1978)
Jolly (1979)	--- Bishop & Shepard (1973)
Manly and Parr (1968)	---
Pollock (1975); Robson (1969)	---
Pollock (1981a)	---
Seber (1962)	---

# REFERENCES

- Anderson, D. R. and K. P. Burnham, 1976, Population ecology of the mallard: VI, The effect of exploitation on survival, U.S. Fish and Wildl. Serv. Resour. Publ. 128, 66 pp.
- Andrezejewski, R. and W. Jezierski, 1966, Studies on the European hare: XI, Estimation of population density and attempt to plan the yearly take of hares, *Acta Theriol.* 11:433-448.
- Arnason, A. N., 1972, Parameter estimates from mark-recapture experiments on two populations subject to migration and death, *Res. Popul. Ecology* 13(2):97-113.
- Arnason, A. N., 1973, The estimation of population size, migration rates and survival in a stratified population, *Res. Popul. Ecology* 15(1):1-8.
- Arnason, A. N. and L. Baniuk, 1978, User's manual: POPAN-2, A data maintenance and analysis system for recapture data, Univ. Manitoba, Dept. of Computer Sci., Winnipeg, Canada, 269 pp.
- Arnason, A. N. and L. Baniuk, 1980, A computer system for mark-recapture analysis of open populations, *J. Wildl. Manage.* 44(2):325-332.
- Bailey, J. A., 1969, Trap response of wild cottontails, *J. Wildl. Manage.* 33(1):48-58.
- Bailey, N. T. J., 1951, On estimating the size of mobile populations from recapture data, *Biometrika* 38(3/4):293-306.
- Baker, J. J. W. and G. E. Allen, 1968, Hypothesis, prediction, and implication in biology, Addison-Wesley Publishing Company, Reading, Massachusetts. 143 pp.
- Barbehenn, K. R., 1974, Estimating density and home range size with removal grids: The rodents and shrews of Guam, *Acta Theriol.* 19:191-234.
- Begon, M., 1979, Investigating animal abundance: capture-recapture for biologists, E. Arnold, Ltd., London, 97 pp.
- Bell, G., 1974, Population estimates from recapture studies in which no recaptures have been made, *Nature* 248:616.
- Berg, S., 1976, A note on the UMVU estimate in a multiple-recapture census, *Scand. J. Stat.* 3:86-88.
- Beukema, J. J. and G. J. DeVos, 1974, Experimental tests of a basic assumption of the capture-recapture method in pond population of carp *Cyprinus carpio* L., *J. Fish Biol.* 6:317-329.
- Bishop, J. A. and P. M. Sheppard, 1973, An evaluation of two capture-recapture models using the technique of computer simulation, in M. S. Bartlett and R. W. Hiorns, eds., *The mathematical theory of the dynamics of biological populations*, Academic Press, London, pp. 235-252.

- Bliss, C., 1967, *Statistics in biology*, Vol. 1, McGraw-Hill, Inc., New York, 558 pp.
- Bohlin, T. and B. Sundstrom, 1977, Influence of unequal catchability on population estimates using the Lincoln index and the removal method applied to electro-fishing, *Oikos* 28:123-129.
- Brotzman, R. L. and R. H. Giles, Jr., 1966, Electronic data processing of capture-recapture and related ecological data, *J. Wildl. Manage.* 30:286-292.
- Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson, 1978, *Statistical inference from band recovery data: a handbook*, U.S. Fish and Wildl. Serv. Resour. Publ. 131, 212 pp.
- Brownie, C. and D. S. Robson, 1976, Models allowing for age-dependent survival rates for band return data, *Biometrics* 32(2):305-324.
- Burnham, K. P., 1972, *Estimation of population size in multiple capture-recapture studies when capture probabilities vary among animals*, Oregon State University, Corvallis, Oregon, Dissertation.
- Burnham, K. P., D. R. Anderson, and J. L. Laake, 1980, Estimation of density from line transect sampling of biological populations, *Wildl. Monogr.* 72:1-202.
- Burnham, K. P. and W. S. Overton, 1978, Estimation of the size of a closed population when capture probabilities vary among animals, *Biometrika* 65(3):625-633.
- Burnham, K. P. and W. S. Overton, 1979, Robust estimation of population size when capture probabilities vary among animals, *Ecology* 60(5):927-936.
- Carle, F. L., 1976, *An evaluation of the removal method for estimating benthic populations and diversity*, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, Thesis.
- Carle, F. L. and M. R. Strube, 1978, A new method for estimating population size from removal data, *Biometrics* 34(4):621-630.
- Carothers, A. D., 1973a, The effects of unequal catchability on Jolly-Seber estimates, *Biometrics* 29(1):79-100.
- Carothers, A. D., 1973b, Capture-recapture methods applied to a population with known parameters, *J. Anim. Ecol.* 42(1):125-146.
- Carothers, A. D., 1979, Quantifying unequal catchability and its effect on survival estimates in an actual population, *J. Anim. Ecol.* 48:863-869.
- Chapman, D. G., 1951, Some properties of the hypergeometric distribution with applications to zoological sample censuses, *University of California Publ. Stat.* 1(7):131-160.
- Chapman, D. G. and D. S. Robson, 1960, The analysis of a catch curve, *Biometrics* 16:354-368.
- Cochran, W. G., 1977, *Sampling techniques*, 3rd ed, John Wiley & Sons, Inc., New York, 428 pp.
- Cormack, R. M., 1964, Estimates of survival from the sightings of marked animals, *Biometrika* 51(3/4):429-438.

- Cormack, R. M., 1966, A test for equal catchability, *Biometrika* 22(2):330-342.
- Cormack, R. M., 1968, The statistics of capture-recapture methods, *Oceanogr. Mar. Bio. Ann. Rev.* 6:455-506.
- Cormack, R. M., 1972, The logic of capture-recapture estimates, *Biometrics* 28(2):337-343.
- Cormack, R. M., 1979, Models for capture-recapture, in R. M. Cormack, G. P. Patil, and D. S. Robson, eds., *Sampling biological populations*, International Co-operative Publishing House, Fairland, Maryland, pp. 217-255.
- Cormack, R. M., 1981, Log-linear models for capture-recapture experiments on open populations, in R. W. Hiorns and D. Cooke, eds., *The mathematical theory of the dynamics of biological populations*, Academic Press, London.
- Crosbie, S. F., 1979, The mathematical modelling of capture-mark-recapture experiments on animal populations, University of Otago, Dunedin, New Zealand, Thesis.
- Cross, D. G. and B. Stott, 1975, The effect of electric fishing on the subsequent capture of fish, *J. Fish Biol.* 7:349-357.
- Crowcroft, P. and J. N. R. Jeffers, 1961, Variability in the behavior of wild house mice (*Mus musculus* L.) towards live traps, *Proc. Zool. Soc. London* 137(4):573-582.
- Darroch, J. N., 1958, The multiple recapture census: I. Estimation of a closed population, *Biometrika* 45(3/4):343-359.
- Darroch, J. N., 1959, The multiple recapture census: II. Estimation when there is immigration or death, *Biometrika* 46(3/4):336-351.
- Darroch, J. N., 1961, The two-sample capture-recapture census when tagging and sampling are stratified, *Biometrika* 48(3/4):241-260.
- Davis, D. E., 1956, *Manual for analysis of rodent populations*, Edwards Brothers, Inc., Ann Arbor, Michigan, 82 pp.
- Davison, R. P., 1980, The effect of exploitation on some parameters of coyote populations, Utah State University, Logan, Utah, Dissertation.
- Dice, L. R., 1938, Some census methods for mammals, *J. Wildl. Manage.* 2(2):119-130.
- Dice, L. R., 1941, Methods for estimating populations of animals, *J. Wildl. Manage.* 5(4):398-407.
- Dixon, W. J. and M. B. Brown, eds., 1979, *BMDP Biomedical Computer Programs P-Series 1979*, University of California Press, Berkeley, California, 880 pp.

Dupont, W. D., 1976, A stochastic method for estimating animal abundance from catch-effort data, Johns Hopkins University, Baltimore, Maryland, Dissertation.

Eberhardt, L. L., 1969, Population estimates from recapture frequencies, *J. Wildl. Manage.* 33(1):28-39.

Eberhardt, L. L., 1970, Correlation, regression, and density dependence, *Ecology* 51(2):306-310.

Edwards, W. R. and L. L. Eberhardt, 1967, Estimating cottontail abundance from live trapping data, *J. Wildl. Manage.* 31(1):87-96.

Efron, B., 1979, Bootstrap methods: another look at the jackknife, *Ann. Stat.* 7(1):1-26.

Feller, W., 1950, An introduction to probability theory and its application, Vol. 1, John Wiley & Sons, Inc., New York, 509 pp.

Fisher, R. A. and E. B. Ford, 1947, The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* (L.), *Heredity* 1:143-174.

Fordham, R. A., 1970, Mortality and population change of dominican gulls in Wellington, New Zealand, with a statistical appendix by R. M. Cormack, *J. Anim. Ecol.* 39(1):13-27.

Friedman, R., 1974, Electrofishing for population sampling + a selected bibliography, U.S. Department of the Interior, Office of Library Services, Bibliography Series No. 31, Washington, D.C.

Geis, A. D., 1955, Trap response of the cottontail rabbit and its effect on censusing, *J. Wildl. Manage.* 19(4):466-472.

Gentry, J. B., F. B. Golley, and J. T. McGinnis, 1966, Effect of weather on captures of small mammals, *Am. Midl. Nat.* 75(2):526-530.

Gentry, J. B., F. B. Golley, and M. H. Smith, 1968, An evaluation of proposed International Biological Program census methods for estimating small mammal populations, *Acta Theriol.* 13:313-327.

Getz, L. L., 1961, Response of small mammals to live-traps and weather conditions, *Am. Midl. Nat.* 66(1):160-169.

Gilbert, R. O., 1973, Approximation of the bias in the Jolly-Seber capture-recapture model, *Biometrics* 29(3):501-526.

Gliwicz, J., 1970, Relation between trapability and age of individuals in a population of the bank vole, *Acta Theriol.* 15(2):15-32.

Goldstein, M. and I. F. Goldstein, 1978, How we know: an exploration of the scientific process, Plenum Publishing Corp., New York, 357 pp.

Gray, H. L. and W. R. Schucany, 1972, The generalized jackknife statistic, Marcel Dekker, Inc., New York, 308 pp.

Green, R. H., 1979, Sampling design and statistical methods for environmental biologists, John Wiley & Sons, Inc., New York, 257 pp.

Hansson, L., 1974, Influence area of trap stations as a function of number of small mammals exposed per trap, *Acta Theriol.* 19:19-25.

Hartley, H. O., 1980, Statistics as a science and as a profession, *J. Am. Stat. Assoc.* 75:1-7.

Hayne, D. W., 1949a, Two methods for estimating populations from trapping records, *J. Mammal.* 30(4):399-411.

Hayne, D. W., 1949b, Calculation of size of home range, *J. Mammal.* 30(1):1-18.

Hayne, D. W., 1978, Experimental designs and statistical analyses in small mammal population studies, in D. P. Snyder, ed., *Populations of small mammals under natural conditions*, Spec. Publ. Ser., Vol. 5, Pymatuning Lab. of Ecol., University of Pittsburgh, Pittsburgh, Pennsylvania, pp. 3-10.

Huber, J. J., 1962, Trap response of confined cottontail populations, *J. Wildl. Manage.* 26(2):177-185.

Hubert, G. F., Jr., G. L. Storm, R. L. Phillips, and R. D. Andrews, 1976, Ear tag loss in red foxes, *J. Wildl. Manage.* 40(1):164-167.

Hunter, J. R. and W. T. Wisby, 1964, Net avoidance behavior of carp and other species of fish, *J. Fish. Res. Board of Canada* 21:613-633.

Jensen, T. S., 1975, Trappability of various functional groups of the forest rodents *Clethrionomys glareolus* and *Apodemus flavicollis*, and its application in density estimations, *Oikos* 26(2):196-204.

Johns, B. E., 1979, Micotaggants—new wildlife marker, *U.S. Fish and Wildl. Serv. Research Info. Bull.* 8, 1 p.

Johnson, D. H., 1974, Estimating survival rates from banding of adult and juvenile birds, *J. Wildl. Manage.* 38(2):290-297.

Johnson, N. L. and S. Kotz, 1977, *Urn models and their application: an approach to modern discrete probability theory*, John Wiley & Sons, Inc., New York, 402 pp.

Jolly, G. M., 1965, Explicit estimates from capture-recapture data with both death and immigration—stochastic model, *Biometrika* 52(1/2):225-247.

Jolly, G. M., 1979, A unified approach to mark-recapture stochastic models, exemplified by a constant survival rate model, in R. M. Cormack, G. P. Patil, and D. R. Robson, eds., *Sampling biological populations*, International Co-operative Publishing House, Fairland, Maryland, pp. 227-282.

Justice, K. E., 1961, A new method for measuring home ranges of small mammals, *J. Mammal.* 42(4):462-470.

Kempthorne, O. and L. Folks, 1971, *Probability, statistics, and data analysis*, Iowa State University Press, Ames, Iowa, 555 pp.

Kendall, M. G., and W. R. Buckland, 1971, *A dictionary of statistical terms*, 3rd ed., Hafner Press, New York, 166 pp.

Larson, H. J., 1969, Introduction to probability theory and statistical inference, John Wiley & Sons, Inc., New York, 387 pp.

Leslie, P. H., 1952, The estimation of population parameters from data obtained by means of the capture-recapture method: II. The estimation of total numbers, *Biometrika* 38(3/4):368-388.

Leslie, P. H., 1958, Statistical appendix, *J. Anim. Ecol.* 27(1):84-86.

Leslie, P. H. and D. Chitty, 1951, The estimation of population parameters from data obtained by means of the capture-recapture method: I. The maximum likelihood equations for estimating the death rate, *Biometrika* 38(3/4):269-292.

Leslie, P. H., D. Chitty, and H. Chitty, 1953, The estimation of population parameters from data obtained by means of the capture-recapture method: II. An example of the practical application of the method, *Biometrika* 40(1/2):137-169.

Lindgren, B. W., 1968, Statistical theory, Macmillan & Co., London, 521 pp.

Mahon, R., 1980, Accuracy of catch-effort methods for estimating fish density and biomass in streams, *Environ. Biol. Fishes* 5(4):343-360.

Manly, B. F. J., 1970, A simulation study of animal population estimation using the capture-recapture method, *J. Appl. Ecol.* 7(1):13-39.

Manly, B. F. J. and M. J. Parr, 1968, A new method of estimating population size, survivorship, and birth rate from capture-recapture data, *Trans. Soc. Br. Entomol.* 18:81-89.

Medawar, P. B., 1969, Induction and intuition in scientific thought, Amer. Philos. Soc., Philadelphia, Pennsylvania, 62 pp.

Mendenhall, W., L. Ott, and R. L. Scheaffer, 1971, Elementary survey sampling, Wadsworth Publishing Co., Belmont, California, 247 pp.

Mendenhall, W., and R. L. Scheaffer, 1973, Mathematical statistics with applications, Duxbury Press, North Scituate, Massachusetts, 561 pp.

Montgomery, W. I., 1979, An examination of interspecific, sexual and individual biases affecting rodent captures in Longworth traps, *Acta Theriol.* 24(3):35-45.

Mood, A. M., F. A. Graybill, and D. C. Boes, 1974, Introduction to the theory of statistics, 3rd ed., McGraw-Hill Book Co., New York, 564 pp.

Moran, P. A. P., 1951, A mathematics theory of animal trapping, *Biometrika* 38(3/4):307-311.

Moran, P. A. P., 1952, The estimation of death-rates from capture-mark-recapture sampling, *Biometrika* 39(1/2):181-188.

Nichols, J. D., B. R. Noon, S. L. Stokes, and J. E. Hines, 1981, Remarks on the use of mark-recapture methodology in estimating avian population size, in J. M. Scott and C. J. Ralph, eds., Estimating numbers of birds, *Studies in Avian Biology* No. 6, The Cooper Ornithological Society, pp. 121-136.

- Nixon, C. N., W. R. Edwards, and L. L. Eberhardt, 1967, Estimating squirrel abundance from live trapping data, *J. Wildl. Manage.* 31(1):96-101.
- O'Farrell, M. J., D. W. Kaufman, and D. W. Lundohl, 1977, Use of live-trapping with the assessment line method for density estimation, *J. Mammal.* 58(4):575-582.
- Ostle, B., 1963, *Statistics in research*, Iowa State University Press, Ames, Iowa, 585 pp.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson, 1978, Statistical inference from capture data on closed animal populations, *Wildl. Monogr.* 62:1-135.
- Otis, D. L., 1980, An extension of the change in ratio method, *Biometrics* 36(1):141-147.
- Overton, W. S. and D. E. Davis, 1969, Estimating the number of animals in wildlife populations, in R. H. Giles, Jr., ed., *Wildlife management techniques*, 3rd ed, The Wildlife Society, Washington, D.C., pp. 403-455.
- Paloheimo, J. E., 1963, Estimation of catchabilities and population sizes of lobsters, *J. Fish. Res. Board Can.* 20(1):59-88.
- Paulik, G. J. and D. S. Robson, 1969, Statistical calculations for change-in-ratio estimators of population parameters, *J. Wildl. Manag.* 33(1):1-27.
- Pella, J. and P. Tomlinson, 1969, Generalized stock production model, *Inter-Amer. Trop. Tuna Comm. Bull.* 13(3):421-496.
- Pelton, M. R. and L. C. Marcum, 1975, The potential use of radioisotopes for determining densities of black bears and other carnivores, in R. L. Phillips and C. Jonkel, eds., *Proc. 1975 Predator Symp., Montana Forest and Conservation Exp. Sta., Missoula, Montana*, pp. 221-236.
- Pollock, K. H., 1974, The assumption of equal catchability of animals in tag-recapture experiments, Cornell University, Ithaca, New York, Ph.D. Thesis, 82 pp.
- Pollock, K. H., 1975, A K-sample tag-recapture model allowing for unequal survival and catchability, *Biometrika* 62(2):577-583.
- Pollock, K. H., 1981a, Capture-recapture models allowing for age-dependent survival and capture rates, *Biometrics* 37(3), pp. 521-529.
- Pollock, K. H., 1981b, Capture-recapture models: a review of current methods, assumptions and experimental design, in J. M. Scott and C. J. Ralph, eds., *Estimating numbers of birds, Studies in Avian Biology No. 6*, The Cooper Ornithological Society, pp. 426-435.
- Pollock, K. H., D. L. Solomon, and D. S. Robson, 1974, Tests for mortality and recruitment in a K-sample tag-recapture experiment, *Biometrics* 30(1):77-87.
- Popper, K. R., 1962, *Conjectures and refutations*, Basic Books, Inc., New York, 412 pp.
- Raleigh, R. F. and C. Short, 1981, Depletion sampling in stream ecosystems: assumptions and techniques, *Prog. Fish-Cult.* 43(3):115-120.
- Rice, W. R. and J. D. Harder, 1977, Application of multiple aerial sampling to a mark-recapture census of white-tailed deer, *J. Wildl. Manage.* 41(2):197-206.

Ricker, W. E., 1975, Computation and interpretation of biological statistics of fish populations, Information Canada Bulletin 191, Ottawa, Canada, 382 pp.

Robson, D. S., 1963, Maximum likelihood estimation of a sequence of annual survival rates from a capture-recapture series, *in* North Atlantic Fish Marking Symposium, I.C.H.A.F., Special Publication No. 4, pp. 330-335.

Robson, D. S., 1969, Mark-recapture methods of population estimation, *in* M. L. Johnson and H. Smith, Jr., eds., New developments in survey sampling, Wiley-Interscience, New York, pp. 120-140.

Robson, D. S. and D. G. Chapman, 1961, Catch curves and mortality rates, *Trans. Am. Fish. Soc.* 90(20):181-189.

Robson, D. S. and H. A. Regier, 1968, Estimation of population number and mortality rates, *in* W. E. Ricker, ed., Methods for assessment of fish production in fresh waters, IBP Handbook No. 3, Blackwell Scientific Publ., Oxford, pp. 124-158.

Robson, D. S. and W. D. Youngs, 1971, Statistical analysis of reported tag-recaptures in the harvest from an exploited population, Cornell University, Biometrics Unit report BU-369-M, 15 pp.

SAS Institute Inc., 1979, SAS User's Guide 1979 Edition, SAS (Statistical Analysis Systems) Institute Inc., Raleigh, North Carolina, 494 pp.

Schnabel, Z. E., 1938, The estimation of the total fish population of a lake, *Amer. Math. Mon.* 45(6):348-352.

Seber, G. A. F., 1962, The multi-sample single recapture census, *Biometrika* 49(3/4):339-350.

Seber, G. A. F., 1965, A note on the multiple-recapture census, *Biometrika* 52(1/2):249-259.

Seber, G. A. F., 1970, Estimating time-specific survival and reporting rates for adult birds from band returns, *Biometrika* 57(2):313-318.

Seber, G. A. F., 1971, Estimating age-specific survival rates from bird-band returns when the reporting rate is constant, *Biometrika* 58(3):491-497.

Seber, G. A. F., 1973, Estimation of animal abundance and related parameters, Griffin, London, 506 pp.

Seber, G. A. F., 1982, Estimation of animal abundance and related parameters, 2nd ed., 654 pp.

Seber, G. A. F. and R. Felton, 1981, Tag loss and the Petersen mark-recapture experiment, *Biometrika* 68(1):211-219.

Seber, G. A. F. and J. F. Whale, 1970, The removal method for two and three samples, *Biometrics* 26(3):393-400.

Skalski, J. R. and D. S. Robson, 1979, Tests of homogeneity and goodness-of-fit to a truncated geometric model for removal sampling, *in* R. M. Cormack, G. P. Patil and D. S. Robson, eds., Sampling biological populations, International Co-operative Publishing House, Fairland, Maryland, pp. 283-313.

Smith, H. D., C. D. Jorgansen, and H. D. Tolley, 1972, Estimation of small mammals using recapture methods: partitioning of estimator variables, *Acta Theriol.* 17:57-66.

Smith, M. H., J. B. Gentry, and F. B. Golley, 1969, A preliminary report on the examination of small mammal census methods, *in* K. Petruszewicz and L. Ruskowski, eds., *Energy flow through small mammal populations*, Polish Scientific Publ., Warsaw, Poland, pp. 25-29.

Smith, M. H., R. Blessing, J. G. Chelton, J. B. Gentry, F. B. Golley, and J. T. McGinnis, 1971, Determining density of small mammal populations using a grid and assessment lines, *Acta Theriol.* 16:105-125.

Smith, M. H., R. H. Gardner, J. B. Gentry, D. W. Kaufman, and M. J. O'Farrell, 1975, Density estimation of small animal populations, *in* F. B. Golley, K. Petruszewicz, and L. Ruskowski, eds., *Small mammals: their production and population dynamics*, Intern. Biol. Prog. 5., Cambridge University Press, London, pp. 25-53.

Southwood, T. R. E., 1978, *Ecological methods*, 2nd ed., Chapman and Hall, London, 524 pp.

Stickel, L. F., 1954, A comparison of certain methods of measuring ranges of small mammals, *J. Mammal.* 35(1):1-15.

Stuart, A., 1962, *Basic ideas of scientific sampling*, Charles Griffin and Co. Ltd., 99 pp.

Swift, D. M. and R. K. Steinhorst, 1976, A technique for estimating small mammal population densities using a grid and assessment lines, *Acta Theriol.* 21(4):471-480.

Swinebroad, J., 1964, Net-shyness and wood thrush populations, *Bird-Banding* 33(3):196-202.

Taber, R. D. and I. M. Cowan, 1969, Capturing and marking wild animals, *in* R. H. Giles, Jr., eds., *Wildlife management techniques*, 3rd ed., The Wildlife Society, Washington, D.C., pp. 277-317.

Tanaka, R., 1956, On differential response to live traps of marked and unmarked populations, *Annal. Zool. Jap.* 29:44-51.

Tanaka, R., 1970, A field study of the effect of prebaiting on censusing by the capture-recapture method in a vole population, *Res. Pop. Ecol.* 12(1):111-125.

Tanaka, R., 1972, Investigation into the edge effect by use of capture-recapture data in a vole population, *Res. Pop. Ecol.* 13(2):127-151.

Tanner, J. T., 1978, *Guide to the study of animal populations*, University of Tennessee, Knoxville, Tennessee, 186 pp.

White, E. G., 1975, Identifying population units that comply with capture-recapture assumptions in an open community of alpine grasshoppers, *Res. Pop. Ecol.* 16(2):153-187.

White, G. C., K. P. Burnham, D. L. Otis, and D. R. Anderson, 1978, *User's Manual for Program CAPTURE*, Utah State University Press, Logan, Utah, 40 pp.

Young, H., J. Neess, and J. T. Emlen, Jr., 1952, Heterogeneity of trap response in a population of house mice, *J. Wildl. Manage.* 16(2):169-180.

Zippin, C., 1956, An evaluation of the removal method of estimating animal populations, *Biometrics* 12(2):163-189.

Zippin, C., 1958, The removal method of population estimation, *J. Wildl. Manage.* 22(1):82-90.

# GLOSSARY

This glossary is intended for use with this primer and to help the reader interpret *Otis et al. (1978)*.

Accuracy	Freedom from error or defect; correctness, usually referring to numerical computations.
$\alpha$ (alpha)	Probability of a Type I error; that is, of rejecting the null hypothesis ( $H_0$ ) when it actually is true.
$\bar{y}$ or $\bar{x}$	The average value of the variable $y$ or $x$ taken over the sample. The bar indicates an average.
Behavioral response	Change in subsequent capture probability as a result of the first capture.
Bias (of an estimator)	The difference between the "expected" value of an estimator and the true value of the parameter being estimated. Bias is a measure of how much the average estimate and the true parameter value differ. Bias = $E(\hat{N}) - N$ , where $N$ is the parameter.
$\beta$ (beta)	Probability of a Type II error; that is, of not rejecting the null hypothesis ( $H_0$ ) when the alternative hypothesis ( $H_A$ ) is true.
Census	An exact count or total enumeration of the population. A true census of an animal population would require counting each member.
Confidence coefficient	The percentage of cases in repeated sampling in which the confidence interval covers the parameter of interest. For example, a 95% confidence coefficient means that for 100 experiments the 95% confidence interval would be expected to include the true value of the parameter 95 times.
CPUE	Catch per unit effort.
$c$	Recapture probability (a parameter to be estimated).
$\hat{c}$	Estimator of recapture probability.
$cv$	Coefficient of variation, usually seen as $cv(\hat{\theta})$ , or the coefficient of variation of the estimator $\theta$ . Defined as $se(\hat{\theta})/E(\hat{\theta})$ .
$\chi^2$	Chi-square, a commonly used test statistic.
Closure	The assumption that a population at risk of capture remains constant in size and composition over the period of investigation.
demographic	No changes due to population dynamics; that is, no births, deaths, emigration, or immigration during the investigation.
geographic	The assumption that the population at risk of capture occupies a distinct (but unknown) area on and around the trapping grid; that is, the assumption of something synonymous to the sides of an urn (or the assumption of boundary strip width $W$ not large relative to the space between traps).

CAPTURE	A FORTRAN computer program for selecting a model for the data and calculating population and density estimates ( <i>White et al. 1978</i> ).
Darroch estimator	The estimator of population size under Model $M_1$ .
$\frac{\partial}{\partial \theta} (\cdot)$	The partial derivative of some function $(\cdot)$ with respect to the parameter $\theta$ . This concept, taken from calculus, is useful in finding ML estimators.
D	Number of animals per area; that is, animal density (a parameter to be estimated).
df	Degrees of freedom, usually a function of the amount of data being used to test a hypothesis.
$\epsilon$	Denotes stochastic, or sampling, error in a statistical model.
Errors	
Type I	Rejection of a null hypothesis that is true.
Type II	Acceptance of a null hypothesis that is not true.
Estimate	The calculated value of an estimator, given a particular sample, designated by a caret or hat ( $\hat{\cdot}$ ) over the symbol for the parameter being estimated.
Estimator	A function of sample data that is used to estimate some parameter. An estimator is a random variable and is designated by a caret or hat ( $\hat{\cdot}$ ) over the symbol for the parameter.
$E(\hat{N})$	The expected value of the estimator of $N$ . If the experiment was repeated a very large number of times, the mean of all the $\hat{N}$ 's would be $E(\hat{N})$ . The symbol $E(\cdot)$ also can be used to express the expected value of a statistic, for example, $E(n)$ .
Eq.	Equation, for example, Eq. (5).
!	Factorial operator. For example, $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$ . In general $N! = N \times (N - 1) \times (N - 2) \times \dots \times 3 \times 2 \times 1$ . This odd-looking expression is frequently used as a "counting term" in probability models.
$f_i$	A capture frequency statistic; the number of individuals captured exactly $i$ times in $t$ days of trapping, $i = 1, 2, \dots, t$ . The symbol $f_0$ is used for the number of individuals never captured.
Generalized removal estimator	The estimator of population size under Model $M_{bh}$ .
Histogram	A graph that presents the frequencies with which the values of a variable fall into specified numerical categories.
$H_0$	The null hypothesis.
$H_A$	The alternative hypothesis.
Inductive inference	Generalization from a single experiment to the class of all similar experiments. Reasoning from the particular to the general.

$j$	Sampling occasion; $j = 1, 2, \dots, t$ .
Jackknife estimator	The estimator of population size under Model $M_h$ .
$L_r$	Length of a row in a grid of traps.
$L_c$	Length of a column in a grid of traps.
$\mathcal{L}(N, \underline{p}   \underline{X})$	Read as "the likelihood function of the parameters $N$ and $\underline{p}$ , given the data matrix $\underline{X}$ ." This particular example is for Model $M_o$ . The likelihood function is essential to the technique of deriving ML estimators.
$\ln(\cdot)$	The natural logarithm of a number or expression.
$\ln(\mathcal{L}(N, \underline{p}   \underline{X}))$	Read as "the logarithm of the likelihood function of the parameter $N$ and the vector of parameters $\underline{p}$ , given the data matrix $\underline{X}$ ." This is for Model $M_t$ because of the vector of capture probabilities, $\underline{p}$ .
ML	Maximum likelihood.
MLE	Maximum likelihood estimator.
Model $M_o$	Model in which capture probabilities are constant.
Model $M_t$	Model in which capture probabilities vary by time.
Model $M_b$	Model in which capture probabilities vary due to behavioral response.
Model $M_h$	Model in which capture probabilities vary by individual animal.
Model $M_{tb}$	Model in which capture probabilities vary due to time and behavioral response.
Model $M_{th}$	Model in which capture probabilities vary by time and individual animal.
Model $M_{bh}$	Model in which capture probabilities vary due to behavioral response and differences between individual animals.
Model $M_{tbh}$	Model in which capture probabilities vary due to time, behavioral response, and differences between individual animals.
Model $M_{Rj}$	A particular generalized removal model in which $j$ ( $j = 1, \dots, t - 2$ ) different average capture probabilities are used.
$M_{t+1}$	The number of different individuals caught during the trapping experiment (a statistic).
$M_j$	The number of marked animals in a population at the time of $j^{\text{th}}$ sample (a statistic).
$M$	The sum of the $M_j$ , $j = 1, \dots, t$ (a statistic).
$m_j$	The number of marked animals captured in the $j^{\text{th}}$ sample (a statistic).
$m$	The sum of the $m_j$ , $j = 1, \dots, t$ (a statistic).
$\mu$	Population mean.
$N(0,1)$	Shorthand for a normal distribution having a mean of 0 and variance of 1; called a "standard normal."
$N$	True population size; the number of animals (parameter to be estimated).

$\hat{N}$	Estimator of population size; estimator of the number of animals in the population. A particular value of the estimator (the estimate) is also designated as $\hat{N} =$ value.
$n_j$	The number of animals captured in the $j$ th sample, $j = 1, \dots, t$ (a statistic).
$n$ .	The total number of captures and recaptures during the study (a statistic).
Nonparametric	Refers to a statistical technique that does not depend on any distributional assumptions concerning the variables involved in the technique.
Null estimator	The estimator of population size under Model $M_0$ .
$p$	Capture probability (a parameter that we estimate to obtain an estimate of $N$ ).
$\hat{p}$	Estimator of capture probability.
$P(\cdot)$ or $\text{Pr}(\cdot)$	Probability of $(\cdot)$ .
$\bar{p}$	Average capture probability of animals in the population on a given occasion (parameter). The bar denotes an average or mean.
$\underline{p}$	An array or "vector" of several capture probabilities. For example, the row vector $(p_1, p_2, p_3, \dots, p_t)$ is often written $\underline{p}$ .
Parameter	A fixed quantity in a given population.
Parametric	Refers to a statistical technique that depends on at least one assumption that the variables involved in the technique have a specified distribution, such as chi-square or normal.
PRB	Percent relative bias, for example $100(E(\hat{N}) - N)/N$ is the PRB of $\hat{N}$ .
Precision	A property of an estimator related to the amount of variation among estimates from repeated samples.
$\pi(\cdot)$	Used as a short notation for a complicated probability expression; $\pi$ alone stands for 3.14159+.
$\prod_{i=1}^n (\cdot)$	Product operator for multiplication. Same idea as using $\sum_{i=1}^n$ for summation. For example, $10!$ can be written $\prod_{i=1}^{10} i$ .
POPAN-2	A FORTRAN computer program for the maintenance and analysis of open-population capture-recapture data. ( <i>Arnason and Baniuk 1978</i> ).
$\rho$	Rho, the correlation coefficient. This parameter has values between $-1$ and $1$ ; it measures the strength of the linear relationship between two variables.
Robustness (of an estimator to an assumption)	A robust estimator is one that is not sensitive to the breakdown of a particular assumption on which it is based. An estimator is more or less robust depending on the extent to which the validity of the assumption affects its performance.

$se(\hat{N})$	The standard error of the estimator of $N$ ; $se(\hat{N}) = \sqrt{\text{var}(\hat{N})}$ .
$\sigma^2$	Population variance.
Statistic	A function of the sample data.
$\sum_{i=1}^n (\cdot)$	Summation operator; a shorthand notation to indicate the addition of a number of terms. For example, $\sum_{j=1}^t n_j = n. = n_1 + n_2 + n_3 + \dots + n_t .$
Survey	A study of a portion of a population. In the context here, the portion is selected or "sampled" from the total population.
$t$	Number of capture occasions; also a common test statistic that has a Student's $t$ distribution.
Test statistic	A value, to be computed from the experimental data, that will determine the decision concerning a null hypothesis.
Type I and Type II errors	See error.
$u_j$	The number of new (unmarked) animals captured on the $j$ th sample (a statistic).
UMVUE	Uniform minimum variance unbiased estimator.
$\text{Var}(\cdot)$	Variance of the quantity in the parenthesis.
$\Sigma$	A matrix called the variance-covariance matrix of a vector of estimators. The variances are on the main diagonal and the covariances appear symmetrically throughout the rest of the matrix.
$\text{var}(\hat{N})$	The sampling variance of the estimator of $N$ ; $\text{var}(\hat{N}) = (se(\hat{N}))^2$ .
$\underline{X}$ or $[X_{ij}]$	The matrix of data composed of zeros (not captured) and ones (captured). All statistics used to calculate parameter estimates are obtained from the $X_{ij}$ matrix.
$W$	Strip width for nested grids in density estimation procedure.
$z$	A test statistic that is distributed $N(0,1)$ .
Zippin estimator	The estimator of population size under Model $M_b$ .