

1995

capture-recapture  
models: an overview

eskuratzea eta  
berreskuratzea  
bidezko laginketa

muestreo por  
captura y recaptura

KENNETH H. POLLOCK

33

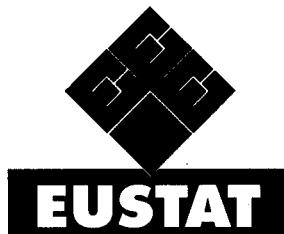
1995  
capture-recapture  
models: an overview

eskuratzea eta  
berreskuratzea  
bidezko laginketa

muestreo por  
captura y recaptura

KENNETH H. POLLOCK

33



*Lanketa / Elaboración:*  
**Euskal Estatistika-Erakundea /**  
Instituto Vasco de Estadística

*Argitalpena / Edición:*  
**Euskal Estatistika-Erakundea /**  
Instituto Vasco de Estadística  
C/Dato 14-16 - 01005 Vitoria-Gasteiz

**© Euskadiko K.A.ko Administratzioa**  
Administración de la C.A. de Euskadi

*Botaldia / Tirada*  
1.000 ejemplares  
XI-1995

*Inprimaketa eta koadernaketa /*  
*Impresión y encuadernación:*  
**ITXAROPENA, S.A.**  
Araba kalea, 45 - Zarautz (Gipuzkoa)

ISBN: 84-7749-221-2

*Lege-gordailua / Depósito legal: S.S. 1.170/95*

## AURKEZPENA

Estatistikako Mintegi Internazionalak sustatzean, hainbat xederekin bete nahi luke Euskal Estatistika-Erakundeak, hala nola:

- Unibertsitatearekiko eta, Estatistika Sailarekiko lankidetza bultzatu.
- Funtzionari, irakasle, ikasle eta estatistikaren alorrean interesaturik leudekeen guztien birziklapen profesionala erraztu.
- Estatistikako alorrean eta mundu-mailan irakasle prestu eta abangoardiako ikerlari diren pertsonaiak Euskadira ekarri, guzti horrek zuzeneko harremanei eta esperientziien ezagupenei dagokienez suposatzen duen ondorio positiboarekin.

Iharduketa osagarri bezala eta interesaturik leudekeen ahalik eta pertsona eta Erakunde gehienetara iristearren, Ikastaro hauetako txostenak argitaratzea erabaki da, beti ere txostenemaiaren jatorrizko hizkuntza errespetatuz, horrela gure Herrian gai honi buruzko ezagutza zabaltzen laguntzeko asmoarekin.

Vitoria-Gasteiz, 1995eko Azaroa  
FRANCISCO JOSE LLERA RAMO  
Zuzendari Orokorra

## PRESENTATION

In promoting the International Statistical Seminars, the Basque Statistical Institute wishes to achieve several aims:

- Encourage collaboration with the University, especially with the Statistics Departments.
- Facilitate the professional recycling of civil servants, university teachers, students and whoever else may be interested in the statistical field.
- Bring to the Basque Country illustrious professors and investigators in the vanguard of statistical subjects, on a worldwide level, with the subsequent positive effect of encouraging direct relationships and sharing knowledge of experiences.

As a complementary activity and in order to reach as many interested persons and institutions as possible, it has been decided to publish the papers of these Courses, always respecting the original language of the writer, to contribute in this way towards the growth of knowledge concerning this subject in our country.

Vitoria-Gasteiz, November 1995  
FRANCISCO JOSE LLERA RAMO  
General Director

## **PRESENTACION**

Al promover los Seminarios Internacionales de Estadística, el Instituto Vasco de Estadística pretende cubrir varios objetivos:

- Fomentar la colaboración con la Universidad y en especial con los Departamentos de Estadística.
- Facilitar el reciclaje profesional de funcionarios, profesores, alumnos y cuantos puedan estar interesados en el campo estadístico.
- Traer a Euskadi a ilustres profesores e investigadores de vanguardia en materia estadística, a nivel mundial, con el consiguiente efecto positivo en cuanto a la relación directa y conocimiento de experiencias.

Como actuación complementaria y para llegar al mayor número posible de personas e Instituciones interesadas, se ha decidido publicar las ponencias de estos Cursos, respetando en todo caso la lengua original del ponente, para contribuir así a acrecentar el conocimiento sobre esta materia en nuestro País.

Vitoria-Gasteiz, Noviembre de 1995  
**FRANCISCO JOSE LLERA RAMO**  
Director General

## AURKIBIDEA - CONTENTS - INDICE

<b>INDEX .....</b>	<b>7</b>
<b>SARRERA .....</b>	<b>13</b>
<b>1. INTRODUCTION .....</b>	<b>15</b>
<b>2. THE LINCOLN-PETERSEN MODEL .....</b>	<b>16</b>
2.1. Estimation of Population Size .....	16
2.2. Model Assumptions .....	17
2.3. Example .....	18
<b>3. CLOSED POPULATION MODELS .....</b>	<b>19</b>
3.1. Introduction .....	19
3.2. $M_0$ : The Equal Catchability Model .....	19
3.3. $M_h$ : The Heterogeneity Model .....	20
3.4. $M_b$ : Trap Response Model .....	20
3.5. $M_{bh}$ : The Heterogeneity and Trap Response Model .....	21
3.6. $M_t$ : The Time Model (Schnabel) .....	22
3.7. Other Time Dependent Models .....	23
3.8. Model Selection .....	23
3.9. Example .....	23
<b>4. THE JOLLY-SEBER MODEL .....</b>	<b>24</b>
4.1. Introduction .....	24
4.2. Assumptions and Notation .....	24
4.3. Parameter Estimation .....	25
4.4. Assumptions and Model Robustness .....	26
4.5. Example .....	29
4.6. Model Restrictions and Extensions .....	30
<b>5. OTHER OPEN MODELS .....</b>	<b>30</b>
5.1. Survival Modeling .....	30
5.2. Stage Structured Models .....	31
5.3. Comparative Capture-Recapture Studies .....	32

<b>6. COMBINATION CLOSED AND OPEN POPULATION MODELS .....</b>	<b>33</b>
6.1. Introduction .....	33
6.2. The Basic Design .....	33
6.3. Extensions .....	35
<b>7. DESIGN OF CAPTURE-RECAPTURE STUDIES .....</b>	<b>35</b>
7.1. Selection of an Estimation Method .....	35
7.2. Minimization of Assumption Violations .....	38
7.3. Sampling Intensity .....	41
<b>8. APPLICATIONS .....</b>	<b>41</b>
8.1. Animal Populations .....	41
8.2. Human Populations and Other Applications .....	42
<b>9. FRAME SIZE ESTIMATION .....</b>	<b>42</b>
9.1. Introduction .....	42
9.2. Estimation of Population Total .....	43
9.3. Use of Capture-Recapture Models in the Large Pelagic Survey .....	44
9.4. Other Issues .....	46
<b>10. COMPUTER PROGRAMS .....</b>	<b>48</b>
<b>11. CONCLUSIONS AND FUTURE RESEARCH .....</b>	<b>49</b>
<b>TABLES .....</b>	<b>51</b>
<b>REFERENCES .....</b>	<b>99</b>

ESTIMACIÓN DE LA POBLACIÓN DE MAMÍFEROS  
USANDO TRAMPAS Y CÁMARAS TRÍGUEO

ESTE LIBRO SE DEDICA A LOS ESTUDIOS DE POBLACIONES DE MAMÍFEROS EN EL MUNDO. SE TRABAJA CON UNA VARIADA GAMA DE TÉCNICAS, INCLUIDAS LAS TRAMPAS, LAS CÁMARAS TRÍGUEO, LOS SISTEMAS DE MONITOREO, LOS MODELOS DE POBLACIÓN Y LOS ESTUDIOS COMPARATIVOS.

ESTE LIBRO ES UN RESUMEN DE LOS PRINCIPALES CONCEPTOS Y TÉCNICAS UTILIZADAS EN ESTOS ESTUDIOS. SE DEDICA A LOS ESTUDIOS DE POBLACIONES DE MAMÍFEROS EN EL MUNDO. SE TRABAJA CON UNA VARIADA GAMA DE TÉCNICAS, INCLUIDAS LAS TRAMPAS, LAS CÁMARAS TRÍGUEO, LOS SISTEMAS DE MONITOREO, LOS MODELOS DE POBLACIÓN Y LOS ESTUDIOS COMPARATIVOS.

## ÍNDICE

<b>1. INTRODUCCIÓN .....</b>	<b>55</b>
<b>2. EL MODELO LINCOLN-PETERSEN .....</b>	<b>56</b>
2.1. Estimación del Tamaño de la Población .....	56
2.2. Suposiciones del Modelo .....	57
2.3. Ejemplo .....	58
<b>3. MODELOS DE POBLACION CERRADA .....</b>	<b>59</b>
3.1. Introducción .....	59
3.2. $M_o$ : El Modelo de Igualdad de Capturabilidad .....	60
3.3. $M_h$ : El Modelo de Heterogeneidad .....	60
3.4. $M_b$ : El Modelo de la Reacción ante las Trampas .....	61
3.5. $M_{bh}$ : El Modelo de Heterogeneidad y Reacción ante las Trampas .....	62
3.6. $M_t$ : El Modelo de Tiempo (Schnabel) .....	63
3.7. Otros Modelos Dependientes del Tiempo .....	64
3.8. Selección de Modelos .....	64
3.9. Ejemplo .....	64
<b>4. EL MODELO JOLLY-SEBER .....</b>	<b>65</b>
4.1. Introducción .....	65
4.2. Supuestos y Anotación .....	65
4.3. Estimación de Parámetros .....	66
4.4. Supuestos y Robustez de Modelos .....	68
4.5. Ejemplo .....	71
4.6. Restricciones y Extensiones de Modelos .....	72
<b>5. OTROS MODELOS ABIERTOS .....</b>	<b>72</b>
5.1. Modelado de Supervivencia .....	72
5.2. Modelos Estructurados por Etapas .....	73
5.3. Estudios Comparativos de Captura-Recaptura .....	74

<b>6. COMBINACIÓN DE MODELOS CERRADOS Y ABIERTOS DE POBLACIÓN .....</b>	<b>75</b>
<b>6.1. Introducción .....</b>	<b>75</b>
<b>6.2. El Diseño Básico .....</b>	<b>76</b>
<b>6.3. Extensiones .....</b>	<b>77</b>
<b>7. DISEÑO DE ESTUDIOS DE CAPTURA-RECAPTURA .....</b>	<b>78</b>
<b>7.1. Selección de un Método de Estimación .....</b>	<b>78</b>
<b>7.2. Minimización de las Violaciones de Suposiciones .....</b>	<b>81</b>
<b>7.3. Intensidad de Muestreo .....</b>	<b>84</b>
<b>8. APLICACIONES .....</b>	<b>85</b>
<b>8.1 .Poblaciones de Animales .....</b>	<b>85</b>
<b>8.2. Poblaciones Humanas y Otras Aplicaciones .....</b>	<b>86</b>
<b>9. ESTIMACIÓN DEL TAMAÑO DE ESTRUCTURAS .....</b>	<b>86</b>
<b>9.1. Introducción .....</b>	<b>86</b>
<b>9.2. Estimación de la Población Total .....</b>	<b>87</b>
<b>9.3. Uso de Modelos de Captura-Recaptura en la Encuesta Pelágica Grande .....</b>	<b>88</b>
<b>9.4. Otras Cuestiones .....</b>	<b>90</b>
<b>10. PROGRAMAS DE ORDENADOR .....</b>	<b>92</b>
<b>11. CONCLUSIONES Y FUTURAS INVESTIGACIONES .....</b>	<b>94</b>
<b>TABLAS .....</b>	<b>95</b>
<b>REFERENCIAS .....</b>	<b>99</b>

## **BIOGRAFI OHARRAK**

Kenneth Hugh Pollock Biometriean lizentziatu zen 1968an Sydney-ko Unibertsitatean, Hegoaldeko Gales Berrian (Australia). 1972an Biometriako masterra lortu zuen Cornell-eko Unibertsitatean, eta 1974an arlo horretako doktoretza eskuratu zuen unibertsitate berean.

Gaur egun, Estatistika, Biomatematika eta Zoologiako irakaslea da Ipar Carolinako Unibertsitate Publikoko Estatistika Departamentuan. Unibertsitate horren egoitza Ipar Carolinako Raleigh hirian dago.

Sari ugari jaso ditu "Journal of Wildlife Management" eta "Wildlife Monograph" aldizkarietan argitaratutako artikuluengatik, bai eta "Ecology" aldizkarian argitaratu dituenengatik ere. 1994an American Statistical Association-en sari berezi bat jaso zuen.

Argitalpen zientifiko askotan idatzi du, eta 80 artikulutik gora argitaratu ditu aldizkari zientifikoetan ekologi laginketako metodoei buruz.

## **BIOGRAPHICAL SKETCH**

Kenneth Hugh Pollock obtained his B.Sc. Agriculture (Biometry) at the University of Sydney, New South Wales, Australia, 1968. In 1972 obtained a M.S. in Biometry and in 1974 a Ph.D. in the same subject from the Cornell University, U.S.A.

At present, he is Professor of Statistics, Biomathematics and Zoology at the Department of Statistics at the North Carolina State University in Raleigh (North Carolina).

He has received numerous best paper awards for his articles in the magazines "Journal of Wildlife Management" and "Wildlife Monograph", as well as in the magazine "Ecology". In 1994, he received a distinguished achievement award from the American Statistical Association.

He has collaborated in numerous scientific publications and has published over 80 papers with the primary emphasis being on ecological sampling methods.

## NOTAS BIOGRAFICAS

**Kenneth Hugh Pollock** se licenció en Biometría por la Universidad de Sidney, New South Wales (Australia) en 1968. En 1972 obtuvo su Master en Biometría por la Universidad de Cornell y en 1974 se doctoró en esta materia en dicha universidad.

Actualmente, es profesor de Estadística, Biomatemáticas y Zoología en el Departamento de Estadística en la Universidad Pública de Carolina del Norte en Raleigh (Carolina del Norte).

Ha recibido numerosos galardones por sus artículos en las revistas "Journal of Wildlife Management" y "Wildlife Monograph" así como en la Revista "Ecology". En 1994 recibió un premio especial de la American Statistical Association.

Ha colaborado en numerosas publicaciones científicas y ha publicado más de 80 artículos en revistas científicas sobre métodos de muestreo ecológico.

Ha sido miembro de la junta editorial de la revista "Journal of Wildlife Management" y "Ecology".

Ha sido miembro de la junta editorial de la revista "Biometrika" y "Biometrics".

Ha sido miembro de la junta editorial de la revista "Biometrika" y "Biometrics".

Ha sido miembro de la junta editorial de la revista "Biometrika" y "Biometrics".

## NOTA BIOGRAFICA

Es un profesor de estadística y biomatemáticas en la Universidad Pública de Carolina del Norte en Raleigh (Carolina del Norte). Es autor de numerosas publicaciones científicas y ha recibido numerosos premios y distinciones.

Ha sido miembro de la junta editorial de la revista "Journal of Wildlife Management" y "Ecology".

Ha sido miembro de la junta editorial de la revista "Biometrika" y "Biometrics".

Ha sido miembro de la junta editorial de la revista "Biometrika" y "Biometrics".

## **ESKURATZEA ETA BERRESKURATZEA BIDEZKO LAGINKETA**

### **SARRERA**

Eskuratzea eta berreskuratzea metodoen lehen erabilpen ezaguña Laplace-k egin zuen (1786), zeinak Frantziako populazioaren tamaina neurtzeko erabili baitzituen (Seber 1982, 104 or.). Mende honetan azkar hedatu dira metodo horiek, eta arlo askotan aplikatu dira. Hasiera batean, biologi populazioei aplikatzen zitzazkien, baina gaur egun giza populazioetarako ere aplikazio asko daude (Sudman et al. 1988). Eskuratzea eta berreskuratzea ereduak laginketa-marko baten tamaina neurtzeko erabili izan dira (Pollock et al. 1990).

Eskuratzea eta berreskuratzea ereduak populazio bati buruzko demografi parametro nagusiak neurtzeko erabil daitezke. Eedu horiek populazioaren tamaina, "biziraupen"-indizea eta "jaiotzen" kopurua induzitzen dituzte. Normalean, ez da posible izaten benetako bizirau-pena emigraciotik bereiztea, edota benetako jaiotzak inmigraciotik, non eta informazio osagarrirrik biltzen ez den (Pollock et al. 1990).

Eskuratzea eta berreskuratzea saiakuntzetan, normalean, K aldiz lagintzen da aztergai den populazioa, nahiz eta K gehienetan  $>2$  izaten den. Aldi oro, markarik gabe atzematen diren animalia guztiak behin bakarrik markatzen da, aurretik markatutako animalien atzema-teak anotatzen dira, eta animalia guztiak berriro askatzen dira populazioaren artean. Zenbait kasutan, animaliak zauritu egin daitezke atzematerakoan, edo animaliak hil beharra izaten da: kasu horretan, ez dira askatuko eta "atzemate-galera" gisa erroldatuko dira. Aztelana amaitzean, suposatzen da saiakuntza egin duen ikerlariak erabili diren animalia guztien atzemateei buruzko historia osoa izango duela (Pollock et al. 1990).

Estatistikarien bi ataletan sailkatu ohi dituzte Eskuratzea eta berreskuratzea ereduak populazio irekietarako edo populazio itxietarako egiten diren kontuan izanik. Garrantzi handikoa da irakurleak ezagutzea zer esan nahi duen zehazki populazio irekiak eta itxiak. Populazio irekia diodanean, esan nahi dut populazioari egiten zaizkion gehiketak (jaiotzak eta/edo inmigratzaileak) eta populazioaren ezabaketa etengabeak (heriotzak eta/edo emigratzaileak) azterlana egiten den bitartean gertatzen direla. Populazio itxi batean, aldiz, ez dago gehiketa eta ezabaketarik, hau da, populazio itxi batek tamaina konstantea izaten du azterlana egiten den bitartean.

Estatistikariarentzat errazago izaten da populazio itxien eredu aztertzea. Garrantzi handiko erreferentziak dira Otis et al-en (1978) eta White et al.-en (1982) lanak. Populazio irekiko ereduak ere garrantzizkoak dira Pollock et al.-en (1990) eta Lebreton et al. (1992)-en lanetan ikus daitekeenez.

Horrekin zerikusia duen arlo bat populazio askotako biziraupen-indizeen alderaketa da; arlo horretan lan egin du Burnham et al.-ek (1987). Batzuetan posible izaten da eredu itxiak eta irekiak diseinu orokor batean biltzea (Pollock, 1982, Pollock et al. 1990, Kendall et al. 1995). Skalski eta Robson-ek (1992) kontuan izan zuten Eskuratzea eta berreskuratzea metodo soilak erabiltzea in situ-ko ekologi azterlanak eta ingurugiro-inpaktuko balioespenak diseinatu eta aztertzerakoan. Azken aldi honetan, asko garatu dira Eskuratzea eta berreskuratzea ereduak etapaka egituratutako populazioetarako (Hestbeck et al. 1991, Nichols et al. 1992, Brownie et al. 1993, Shwarz et al. 1993).

Ikuspegi orokor honetan, lehenik Lincoln-Petersen eredu aurkeztuko dut (2. atala); horrek eredu orokorragoetara eramango gaituelako: eredu itxieta (3. atala) eta irekietara (4. eta 5. atalak), hain zuzen. Ondoren, eredu itxiak eta irekiak diseinu bakarrean nola konbinatu daitezkeen azaltzen duen atal bat dago (6. atala), eta ondoren, diseinu-arazoei buruzko azalpen orokorragoa (7. atala). Animalientzako eta gizakientzako aplikazio-multzo zabala aurkeztuko dut (8. atala). Erakusten dut nola lor daitekeen laginketa-markoaren tamaina neurtea (9. atala), horrek estu lotzen baitu Eskuratzea eta berreskuratzeaa laginketako ohiko teoriarekin. Eskueran ditugun ordenadore-programa ugariak ere aipatuko ditut (10. atala) eta, azkenik, ondorio batzuk emango ditut eta aurrerantzean egin beharreko ikerketetarako iradokizun batzuk (11. atala).

Ez dut esan nahi txosten hau lan berria denik, baina aurreko lanak eguneratzen saiatu naiz (Pollock et al 1990, Pollock 1991), lan horiek baitira ikuspegi orokor honen oinarri. 8. ataleko gaia, batik bat, Pollock et al-en (1994) oinarritzen da.

## **1. INTRODUCTION**

The first known use of capture-recapture methods was by Laplace (1786) and who used it to estimate the population size of France (Seber 1982, p. 104). This century there has been a rapid proliferation of methods and applications in many fields. Originally the emphasis was on biological populations but now there are many applications to human populations (Sudman et al. 1988). Capture-recapture models have also been used to estimate the size of a sampling frame (Pollock et al. (1994)).

Capture-recapture models can be used to estimate fundamental demographic parameters for a population. These induce population size, "survival" rate and "birth" numbers. Usually it is not possible to separate true survival from emigration or true births from immigration unless additional information is collected (Pollock et al. 1990).

In capture-recapture experiments, one typically samples the population under study  $K$  times where  $K$  is usually  $>2$ . Each time, every unmarked animal caught is uniquely marked, previously marked animals have their captures recorded, and all animals are released back into the population. In some cases animals may be injured by capture or may be needed for sacrifice in which case they will not be released and are recorded as "losses on capture." At the end of the study, one assumes that the experimenter has the complete capture history of every animal handled (Pollock et al. 1990).

Statisticians have typically classified capture-recapture models into those suitable for closed or open populations. It is important that the reader be clear on the exact meaning of closed and open. By open I mean that additions (births and/or immigrants) to the population and permanent deletions (death and/or emigrants) from the population occur during the study. A closed population is one where permanent deletions or additions do not occur, i.e. a closed population had constant size during the study.

Closed population models are simpler for the statistician to study. Very important references are the monographs of Otis et al. (1978) and White et al. (1982). Open population models are also important with monographs by Pollock et al. (1990) and Lebreton et al. (1992). A related area is comparison of survival rates for different population which has been considered by Burnham et al. (1987). Sometimes it is possible to combine closed and open models into one overall design (Pollock, 1982, Pollock et al. 1990, Kendall et al. 1995). Skalski and Robson (1992) considered the use of simple capture-recapture methods in the design and analysis of ecological field studies and environmental impact assessments. Recently there has also been a strong development of capture-recapture models for stage structured populations (Hestbeck et al. 1991, Nichols et al. 1992, Brownie et al. 1993, Schwarz et al. 1993).

In this overview, I begin by presenting the Lincoln-Petersen model (Section 2) which leads into the more general closed models (Section 3) and open models (Sections 4 and 5). This is followed by a section on combining closed and open models in one design (Section 6) and then a more general discussion of design issues (Section 7). I present a variety of different applications on animal and humans (Section 8). I show how sampling frame size estimation can be accomplished (Section 9) which relates capture-recapture closely to traditional sampling theory. I refer to the plethora of computer programs available (Section 10) and then conclude with overall conclusions and suggestions for future research (Section 11).

I make no claim that this paper is new work although I have tried to update previous works (Pollock et al. 1990, Pollock 1991) which form the basis of this overview. The material in Section 8 is largely based on Pollock et al. (1994).

## 2. THE LINCOLN-PETERSEN MODEL

### 2.1 Estimation of Population Size

This simplest form of capture-recapture experiment has a long history going back to Laplace in 1786, who used it to estimate the human population size of France (Seber 1982:104). First uses in animal population ecology were by Petersen around the turn of the century (LeCren 1965) and Lincoln (1930). Our purpose here is to give a brief presentation that will set the stage for the more complex models that follow. A detailed discussion of the Lincoln-Petersen method is available in Seber (1982:59).

A sample of  $n_1$  animals is caught, marked, and released. Later a sample of  $n_2$  animals is captured, of which  $m_2$  have been marked. Intuitively one can derive an estimator of the population size ( $N$ ) based on the notion that the ratio of marked to total animals in the second sample should reflect the same ratio in the population so that

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

which gives the estimator

$$\hat{N} = \frac{n_1 n_2}{m_2}. \quad (2.1)$$

This estimator is sometimes called the Lincoln Index (Seber 1982:59), although this is a misuse of the word "index," which usually refers to a measure of relative abundance (Caughley 1977:12).

A modified version with less bias was originally given by Chapman (1951) as

$$\hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1 \quad (2.2)$$

with an approximately unbiased estimate of its variance given by

$$\text{var } \hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)(n_1 - m_2)(n_2 - m_2)}{(m_2 + 1)^2(m_2 + 2)} \quad (2.3)$$

from Seber (1970, 1982:60).

## 2.2 Model Assumptions

The Lincoln-Petersen method is based on a model with assumptions as follows:

1. The population is closed to additions (births or immigrants) and deletions (deaths or emigrants).
2. All animals are equally likely to be captured in each sample.
3. Marks are not lost and are not overlooked by the observer.

The closure assumption (1) can actually be weakened here. If additions occur, then they are always unmarked, and the Lincoln-Petersen estimator is valid for population size at the time of the second sample. If deletions occur randomly with respect to marked and unmarked animals, then the Lincoln-Petersen estimator is valid for population size at the time of the first sample. Later in this paper we consider completely open populations (additions and deletions permitted). The Lincoln-Petersen estimator in this simple form is no longer appropriate; however, a modified version of it is central to all the open population models studied.

The second assumption, which is often called the assumption of "equal catchability" of animals, is unlikely to be true in many wild populations. It is important to distinguish 2 very different types of alternatives.

1. Heterogeneity.- The probability of capture in any sample is a property of the animal and may vary over all the animals in the population. This variation in capture probability could be due to many factors such as age, sex, social status, or "territory" location in relation to trap position.
2. Trap response.- The probability of capture in any sample depends on the animal's prior history of capture, i.e., marked animals may have a lower (trap-shy) or a higher (trap-happy) capture probability than unmarked animals.

The type of response will depend on the trapping method used. For example, baiting traps is likely to lead to a trap-happy response where marked animals are more likely to be captured than unmarked animals. It should be emphasized that both heterogeneity and trap response could be present in a population.

If heterogeneity is operating, then animals with high capture probabilities will be more likely captured in the first sample and recaptured in the second. This result in  $m_2/n_2$  overestimating the true proportion marked in the population and hence in  $\hat{N}$  of eq. (2.1) being too small, i.e.,  $\hat{N}$  is a negatively biased estimator. (If capture probabilities are heterogeneous in each sample but independent from sample to sample, then no bias results.) If animals are trap-happy, then  $m_2/n_2$  will again be too large and  $\hat{N}$  too small, whereas a population with trap-shy

animals will have the opposite situation. To summarize, a trap response in capture probabilities can result in either a negative bias (trap-happy animals) or a positive bias (trap-shy animals) in the population size estimator. Although the validity of the Lincoln-Petersen estimator depends on there being no heterogeneity or trap response, it is possible for capture probabilities to vary between the 2 sampling times.

Assumption (3) is clearly important. If animals lose their tags,  $m_2/n_2$  will be too small,  $\hat{N}$  too large, and we are again in the situation of a positively biased estimator. If tag loss is likely to occur, an attempt should be made to estimate rate of loss so that a "correction" to the number of recaptures can be made. This is commonly done in fisheries studies where tag loss can be serious. One approach is to use a double-marking scheme (Caughley 1977:139, Seber 1982:94). A detailed treatment of Lincoln-Petersen method assumptions, effects of deviations from these assumptions, and tests for detecting such deviations is provided by Seber (1982:70-104).

### 2.3 Example

To conclude this section we present an example of use of the Lincoln-Petersen method taken from Skalski et al. (1983). In August 1974, a sample of 87 Nuttall's cottontail (*Sylvilagus nuttallii*) was live-trapped in central Oregon. The animals were marked with a picric acid dye on their tails and hind legs and released. On 5 September 1974, a drive count was conducted and 14 animals were counted of which 7 were marked.

This means that we have  $n_1 = 87$ ,  $n_2 = 14$ , and  $m_2 = 7$ . The Chapman estimate (eq. 2.2) is given by

$$\begin{aligned}\hat{N}_c &= \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1 \\ &= \frac{88 \times 15}{8} - 1 \\ &= 164\end{aligned}$$

with its approximate variance estimate (eq. 2.3)

$$\begin{aligned}\hat{\text{var}} \hat{N}_c &= \frac{(n_1 + 1)(n_2 + 1)(n_1 - m_2)(n_2 - m_2)}{(m_2 + 1)^2(m_2 + 2)} \\ &= \frac{88 \times 15 \times 80 \times 7}{8^2 \times 9} \\ &= 1,283.33.\end{aligned}$$

An approximate 95% confidence interval assuming normality for  $\hat{N}_c$  is given by

$$\begin{aligned}\hat{N}_c &\pm 1.96(\hat{\text{var}} \hat{N}_c)^{0.5} \\ &= 164 \pm 1.96 \times (1,283.33)^{0.5} \\ &= 164 \pm 70\end{aligned}$$

Hence we have a confidence interval ranging from 95 up to 234, which is rather wide. The precision of the experiment could have been increased if the drive count had sighted more animals. Notice that 2 different sampling methods (trapping and sighting) were used in this study. Therefore, there should be little concern with trap response or heterogeneity of the capture probabilities.

### 3. CLOSED POPULATION MODELS

#### 3.1 Introduction

A brief review of a series of 8 models for closed populations will be presented here. As noted in Chapter 1, closed population models assume that no births, deaths, or emigration in or out of population occur between sampling periods. For this reason, these models generally are used for experiments covering relatively short periods of time (e.g., trapping every day for 5 consecutive days). Capture histories for every animal caught are the data needed to obtain estimates under these models. These models were first considered as a set by Pollock (1974) and more fully developed by Otis et al. (1978) and White et al. (1982).

The models differ in the manner in which capture probability is modeled. Different models consider 3 sources of variation in capture probabilities: (1) heterogeneity (2) trap response, and (3) temporal variation or variation among sampling periods (e.g., capture probability for day  $i$  differs from that of day  $j$ ). A catalog of available models is present in Table 3.1. Otis et al. (1978) provided a detailed computer program, CAPTURE, for analysis of data using the models in Table 3.1 (also see White et al. 1978). As well as providing estimators for 5 of the models, CAPTURE contains a model selection procedure that aids the biologist in choosing the most appropriate model for a particular data set, but it does not perform well when capture probabilities are low and population sizes are small (Section 3.8). Recently Rextad and Burnham (1991) have adapted the program incorporating recent work. Estimators are now available for all models except  $M_{tbh}$ .

#### 3.2 $M_0$ : The Equal Catchability Model

Model  $M_0$  is unlikely to be realistic for natural animal populations because it assumes that every animal in the population has the same probability of capture ( $p$ ) for each sampling period in the entire study. It is mainly included for pedagogic reasons and to provide a basis for generalization.

Under this model the Maximum Likelihood Estimator (MLE) of  $N$ , the population size, can only be found iteratively. (This means it is not possible to give an explicit formula for the MLE.) Program CAPTURE has the capability of computing this estimator and its approximate standard error.

It should be emphasized that this estimator can be highly biased if there is unequal catchability. Heterogeneity causes a negative bias, whereas trap response can cause positive or negative bias depending on whether the animals are trap-shy or trap-happy, respectively. Otis

et al. (1978) found that the estimator is reasonably robust to changes in capture probabilities over time.

### 3.3 $M_h$ : The Heterogeneity Model

Model  $M_h$  allows heterogeneity but no trap response and assumes that each animal has its own unique capture probability ( $p_j, j = 1, \dots, N$ ), which remains constant over all the sampling times. The  $p_j$ 's are further assumed to be a random sample of size  $N$  from some probability distribution  $F(p)$ . This model was first considered by Burnham (1972) and later by Burnham and Overton (1978, 1979). The vector of capture frequencies, which is the numbers of animals captured, 1, 2, ...,  $k$  times, contains all the information for estimating  $N$  under this model.

Difficulties exist in finding a satisfactory estimator for this model. Burnham originally tried modeling capture probabilities as coming from a 2-parameter beta distribution (Johnson and Kotz 1970:37) and using Maximum Likelihood estimation, but found it unsatisfactory.

A naive estimator of  $N$  is simply the number of distinct animals seen in the whole experiment and is highly biased unless capture probabilities are very high or number of samples very large. Burnham applied the "jackknife" technique for bias reduction originally proposed by Quenouille (1956) to this estimator and came up with a series of estimators, which are given in Burnham and Overton (1978, 1979) and Otis et al. (1978:109). He also proposed an objective technique for choosing which estimator to use on a particular data set. This technique has been programmed into CAPTURE. Use of the jackknife estimator for point and interval estimation is not without difficulties (Otis et al. 1978:37). However, this jackknife estimation procedure is probably the most robust method so far proposed for Model  $M_h$  (Pollock and Otto 1983).

There has been a lot of recent interest in alternative estimators for this model. Chao (1988) came up with a moment estimator for this model. She found that when the mean capture probability is small, so that most animals are caught only once or twice, the estimator is usually less biased than the jackknife estimator. Another estimator was proposed by Lee and Chao (1994). Recently Norris and Pollock (1996) came up with a nonparametric maximum likelihood approach where  $N$  is estimated as well as the nonparametric density estimator of  $F$ , the distribution of capture probabilities.

Calculating population-size estimates separately by identifiable categories such as age and sex to reduce heterogeneity of capture probabilities is an alternative approach that sometimes works well. Unfortunately, however, it may not be feasible in small populations.

### 3.4 $M_b$ : The Trap Response Model

Model  $M_b$  allows trap response but not heterogeneity or temporal variation and makes the following assumptions:

1. Every unmarked animal in the population has the same probability of capture ( $p$ ) for all samples.
2. Every marked animal in the population has the same probability of recapture ( $c$ ) for all samples after it has been captured once.

Again, it is necessary to find the MLE for  $N$  numerically using program CAPTURE.

An important characteristic of this model is that animals do not contribute any information for population size estimation after first capture. Thus, this model is equivalent to the “removal” method (Zippin 1956, Seber 1982:309), except in this case an animal is not physically removed but is considered removed after initial marking.

Often a linear regression method has been used to estimate  $N$  in removal studies in the biological literature. This method has intuitive appeal and also is easy to compute. However, whenever computer programs (such as CAPTURE) are available, it is probably better to use the MLE. In practice we have found that there is usually little difference between the 2 estimators.

The regression method (Seber 1982:325) is presented briefly because of its intuitive appeal and because it will be helpful when we discuss  $M_{bh}$  in Section 3.5. It is based on

$$\begin{aligned} E(u_i | M_i) &= p(N - M_i) \\ &= pN - pM_i \end{aligned} \tag{3.1}$$

which can be described mathematically as follows. Given  $M_i$  (the cumulative number captured and marked before period  $i$ ) the expected or “average” catch of unmarked animals on day  $i$  ( $u_i$ ) is a linear function of the number of marked animals in the population. In the language of fisheries, where this technique has been applied most, we have a linear regression of catch ( $u_i$ ) versus cumulative catch ( $M_i$ ). The regression-based estimators of  $N$  and  $p$  are simple functions of the least squares estimators of the slope and intercept (Seber 1982:325).

### **3.5 $M_{bh}$ : The Heterogeneity and Trap Response Model**

Model  $M_{bh}$  allows for both heterogeneity and trap response and assumes that each animal has its own unique pair of potential capture probabilities ( $p_j, c_j; j = 1, \dots, N$ ) with  $p_j$  and  $c_j$  referring to whether the animal is unmarked or marked, respectively. These probabilities are assumed to remain constant over all sampling times.

This model was first considered by Pollock (1974) and was developed further by Otis et al. (1978:40), who give an estimation procedure called the “generalized removal method.” Here we describe their method intuitively by generalizing the regression method described in the previous section.

If heterogeneity (as well as trap response) is operating, there is no longer a linear relationship between catch ( $u_i$ ) and cumulative catch ( $M_i$ ), and thus eq. (3.1) is no longer valid. In fact we now have

$$E(u_i | M_i) = \bar{p}_i(N - M_i),$$

where  $\bar{p}_i$  is the average conditional probability of capture in the  $i$ th sample for those animals not previously captured. We would expect the  $\bar{p}_i$ 's to decrease gradually because the animals with higher first capture probabilities would tend to be caught earlier than those with lower first capture probabilities.

Otis et al. (1978) also suggest that the biggest differences between  $\bar{p}_i$ 's will be at the beginning of the experiment. This suggests that the regression should be more nearly linear if points corresponding to earlier sampling times are excluded. An objective method of doing this is to first test if all the  $\bar{p}_i$ 's are equal. If so, we use the removal method (Section 3.4), which is a regression using all sample points. If not, we test if  $\bar{p}_2 = \bar{p}_3 = \dots = \bar{p}_k$ , and if this hypothesis is not rejected, we use a regression based on all points except the first. We continue sequentially removing points until the remaining  $\bar{p}_i$ 's cannot be shown to differ.

Notice that although using fewer points in the regression gives us a more general model allowing heterogeneity, it also means we are estimating  $N$  based on less information, so that standard errors for  $N$  under this model tend to be larger than under the removal model. Also a substantial negative bias on estimators of  $N$  can still exist if the heterogeneity is severe and the number of sampling times small. The problem with heterogeneity is that some animals may be essentially "invisible" because of their very small capture probabilities, and no model can deal with this extreme situation.

Pollock and Otto (1983) considered some new estimators for Model  $M_{bh}$ . They conclude that a new jackknife-type estimator may be better than the generalized removal estimator described above. This estimator has been included in CAPTURE now so that biologists can use it as an alternative to the generalized removal estimator.

Recently there have been some new estimators for this model proposed by Chao and her co-workers based on the sample coverage method (Lee and Chao 1994). Norris and Pollock (1996) came up with a nonparametric maximum likelihood approach for this model and  $M_h$ . They estimate  $N$  as well as the nonparametric density estimator of  $F$ , the distribution of capture probabilities. They also develop a likelihood ratio test for comparing  $M_h$  versus  $M_{bh}$  based on bootstrapping.

### **3.6 $M_t$ : The Schnabel Model (Time Variation)**

Model  $M_t$  does not allow heterogeneity or trap response of the individual animal capture probabilities. The model instead assumes that every animal in the population has the same probability of capture at each sampling time,  $p_i$ ,  $i = 1, 2, \dots, k$ . This is the classic capture-recapture model for a closed population with a history going back to Schnabel (1938).

A detailed treatment of this model is given by Seber (1982:130). The MLE can be found numerically using program CAPTURE. This model, together with  $M_0$ , does not require detailed capture history. Batch marks are sufficient to allow estimation. The model  $M_t$  MLE, like that of Model  $M_0$ , can be highly biased if there is unequal catchability. Heterogeneity causes a negative bias, whereas trap response can cause positive or negative bias depending on whether the animals are trap-shy or trap-happy, respectively.

### **3.7 Other Time Dependent Models**

Model  $M_{tb}$  originally had no estimator of  $N$  provided but has been considered recently by Chao and co-workers (Lee and Chao 1994) and Rextad and Burnham (1991) also include an estimator in the most recent version of CAPTURE. Model  $M_{th}$  originally had no estimators of  $N$  provided but recently has been considered by Chao and co-workers (Chao et al. 1992, Lee and Chao 1994). Model  $M_{tth}$  is an important conceptual model but it has no estimator available (Table 3.1).

### **3.8 Model Selection**

In Otis et al. (1978) and in program CAPTURE, an important method of choosing 1 of the 8 possible models ( $M_o, M_p, M_{tb}, \dots, M_{tth}$ ) is presented. This model selection algorithm is based on goodness-of-fit tests and tests between models. The procedure often may need to be used, but with caution (Chapman 1980, Menkins and Anderson 1988), because the tests are not independent and often have very low power, especially for small populations (Menkins and Anderson 1988). If possible, biological information should be used to reduce the number of models considered in this study. For example, evidence may exist, based on behavior of the animal concerned, that trap response is unlikely with the trapping method used. In this case any models that allow trap response could be eliminated from consideration ( $M_b, M_{bh}, M_{tth}$ ). Menkins and Anderson (1988) suggest that, in situations where model selection is poor (i.e., low capture probabilities, small population), use of a Lincoln-Petersen estimate based on pooling of samples into an early versus late sample may be preferable to use of program CAPTURE.

### **3.9 Example**

We use meadow vole, *Microtus pennsylvanicus*, livetrapping data collected at Patuxent Wildlife Research Center, Laurel, Maryland, from June 1981 through January 1982. The grid contained a 10 ( 10 matrix of trapping stations spaced at 7.6-m intervals. A single modified Fitch live trap (Rose 1973) was placed at each station. Hay and dried grass were placed in the nest box section of the traps, and whole corn was used as bait. Traps were opened 1 evening, checked the following morning, and locked open until the evening when the procedure began again.

Here we use data from 5 consecutive trapping days in October 1981 to illustrate the output from program CAPTURE. Data include all animals of both sexes  $\geq 22$  g body weight. Output from the model selection procedure of CAPTURE for these data is presented in Table 3.2. In this case the heterogeneity model,  $M_h$ , is suggested.

Output corresponding to the model  $M_h$  estimates from program CAPTURE is presented in Table 3.3. The jackknife estimate of population size is 139 animals with an approximate 95% confidence interval of 117-161 animals. The average daily capture probability is 0.44, which explains the good precision of the estimate.

## 4. THE JOLLY-SEBER MODEL

### 4.1 Introduction

In many capture-recapture studies it is not possible to assume the population is closed to additions and permanent deletions. The basic open population model suitable for this situation is the Jolly-Seber model (Jolly 1965, Seber 1965), and it is crucial that the reader study this chapter carefully as a building block for the later ones. Cormack (1973) gave a brief intuitive description of this model and its estimators, and Seber (1982:196) gave the best detailed presentation. Comprehensive computer programs are available to help with analyses under this model (e.g., POPAN-3, Arnason and Schwarz 1987; SURGE, Lebreton et al. 1992; JOLLY, Pollock et al. 1990).

The Jolly-Seber model allows estimation of population size at each sampling time as well as survival rates and birth numbers between sampling times. However, as mentioned earlier, migration cannot be separated from the birth and death processes without additional information.

### 4.2 Assumptions and Notation

This model makes the following assumptions:

1. Every animal present in the population at the time of the  $i$ th sample ( $i = 1, 2, \dots, k$ ) has the same probability of capture ( $p_i$ ).
2. Every marked animal present in the population immediately after the  $i$ th sample has the same probability of survival ( $\phi_i$ ) until the  $(i + 1)$ th sampling time ( $i = 1, 2, \dots, k - 1$ ).
3. Marks are not lost or overlooked.
4. All samples are instantaneous and each release is made immediately after the sample.

Assumptions (1), (3), and (4) were required under the Lincoln-Petersen model. Because only marked animals are used to estimate survival rates, we do not need to assume that marked and unmarked animals have equal survival rates. However, in practice, biologists will want to use the survival estimates to refer to the whole population rather than just the marked component. The Jolly-Seber model allows for "loss or capture," i.e., some animals may not be returned to the population after capture. Implicit from the above assumptions is that all emigration from the population is permanent.

The following notation (Table 4.1) is used in this chapter and is based on Seber (1982:196). The statistics defined in Table 4.1 ( $m_i, u_i, n_i, R_i, r_i, z_i$ ) are summary statistics that we need to calculate from the complete capture-history information for each animal captured at least once in the study. Capture-history information is usually represented by series of zero's and one's that represent not captured or captured, respectively, in a particular sample. Let us illustrate by showing some typical capture histories for animals in 5-sample capture-

recapture study (Table 4.2). In Table 4.2, animal 1 was captured in the first sample, recaptured in the second sample, and never seen again. Animal 2 was captured in the first sample and never seen again. Animal 3 was captured in the first sample and then recaptured in the fourth sample. Animal 4 was first captured in sample 2 and then recaptured in samples 3, 4, and 5. We strongly advise this form of input rather than summary statistics. It allows more alternative analyses to be carried out and also allows more detailed goodness-of-fit tests to be run.

### 4.3 Parameter Estimation

Here an intuitive discussion of parameter estimation will be given. First imagine that  $M_i$ , the number of marked animals in the population just before the  $i$ th sample, is known for all values  $I = 2, \dots, k$  (there are no marked animals at the time of the first sample so that  $M_1 = 0$ ). Later we will discuss how to estimate the  $M_i$ , which are always unknown in an open population.

An intuitive estimator of  $N_i$ , the population size in sample  $i$ , is the Lincoln-Petersen estimator discussed in Chapter 2. If the model assumptions are valid, then the proportion of marked animals in the sample should equal that in the population,

$$\frac{m_i}{n_i} \cong \frac{M_i}{N_i}.$$

This yields the following estimator for  $N_i$ :

$$\hat{N}_i = \frac{n_i M_i}{m_i}, \quad (4.1)$$

where  $m_i$  and  $n_i$  are the marked and total numbers of animals captured in the  $i$ th sample, respectively.

The survival rate estimator is obtained by first considering the number of marked animals in the population immediately after sample  $i$ , which is given by  $(M_i - m_i + R_i)$ . The  $(M_i - m_i)$  in this expression represents the marked animals not caught in  $i$ , and released, with marks, back into the population. We then simply consider the number of animals in this group that are still alive in the population just before the  $(i + 1)$ th sample,  $M_{i+1}$ . A natural survival rate estimator is then the ratio of these 2 quantities:

$$\hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + R_i}. \quad (4.2)$$

An obvious estimator of the recruitment between  $i$  and  $(i + 1)$  is

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\phi}_i (\hat{N}_i - n_i + R_i). \quad (4.3)$$

This is simply the estimated difference between population size at sample ( $i + 1$ ) (which is  $N_{i+1}$ ) and the expected number of survivors from  $i$  to  $i + 1$  (which is  $\phi_i[N_i - n_i + R_i]$ ).

Capture probability,  $p_i$ , can be estimated as the proportion of marked or total (marked + unmarked) animals alive at  $i$  that are captured in  $i$ .

$$\hat{p}_i = \frac{m_i}{\hat{M}_i} = \frac{n_i}{N_i}. \quad (4.4)$$

To complete this intuitive outline, we need an estimator of the  $M_i$  because they are obviously unknown in an open population. This can be obtained by equating the 2 ratios

$$\frac{z_i}{M_i - m_i} \cong \frac{r_i}{R_i}$$

which are the future recovery rates of the 2 distinct groups of marked animals in the population at sampling period  $i$ : (1)  $(M_i - m_i)$  the marked animals not seen at  $i$  and (2)  $R_i$  are the animals seen at  $i$ , marked, and then released for possible recapture. Note that  $z_i$  and  $r_i$  are the members of  $(M_i = m_i)$  and  $R_i$ , respectively, which are captured again at least once. Therefore the estimator of  $M_i$  is given by

$$\hat{M}_i = m_i + \frac{R_i z_i}{r_i} \quad (4.5)$$

and is defined for  $i = 2, \dots, k - 1$  because we need animals seen before and after each  $i$ . It follows that  $N_i$  in eq. (4.1) is defined for  $i = 2, \dots, k - 1$ ;  $\hat{\phi}_i$  in eq. (4.2) for  $i = 1, \dots, k - 2$ ;  $\hat{B}_i$  in eq. (4.3) for  $i = 2, \dots, k - 2$ ; and  $\hat{p}_i$  in eq. (4.4) for  $i = 2, \dots, k - 1$ .

Two of the above estimators ( $\hat{\phi}_i, \hat{p}_i$ ) are maximum likelihood estimators and all 5 ( $\hat{\phi}_i, \hat{p}_i, N_i, \hat{B}_i, \hat{M}_i$ ) are intuitively reasonable. These estimators are biased, however, and the approximately unbiased estimators were recommended by Seber (1982:204; for  $(\hat{\phi}_i, \hat{N}_i, \hat{B}_i, \hat{M}_i)$ ) and Jolly (1982:304; for  $\hat{p}_i$ ). They are similar to the Chapman estimator used in Equation (2.2). Large sample variance estimators are also available but are not presented here (Pollock et al. 1990:Section 4.4). We also do not present goodness of fit tests based on sufficient statistics. See Pollock et al. (1990:Section 4.5).

#### 4.4 Assumptions and Model Robustness

The effects of heterogeneity of capture probability on Jolly-Seber estimators have been well studied. If certain individuals are more likely to be caught than others and if these different capture probabilities persist throughout the capture-recapture experiment, then the marked individuals in the population will tend to be those with the higher capture probabilities. Because of this difference in average capture probabilities between marked and unmarked animals, the proportion of marked animals in any sample ( $m_i/n_i$ ) will tend to over-estimate the true proportion of marked animals in the population ( $M_i/m_i$ ). This leads to a negative bias in  $\hat{N}_i$  (eq. 4.1). In addition, we have noted that the estimator  $\hat{M}_i$  (eq. 4.5) is derived intuitively.

vely by equating the proportions of 2 groups of marked animals in the population just after period  $i$  (those seen at  $i$ ,  $R_i$ ; and those not seen at  $i$ ,  $M_i - m_i$ ) that are captured at some subsequent period ( $>i$ ). Heterogeneity can result in a somewhat larger average capture probability for those animals caught at  $i$  ( $R_i$ ) than for marked animals not caught then ( $M_i - m_i$ ). This results in a relatively small negative bias in  $\hat{M}_i$  (Carothers 1973). Biased  $\hat{M}_i$  contribute not only to the bias of  $\hat{N}_i$ , but also result in biased  $\hat{\phi}_i$  (eq. 4.2).

Carothers (1973) and Gilbert (1973) used computer simulation and analytic expressions to approximate the bias in Jolly-Seber estimators resulting from the heterogeneous capture probabilities. These authors studied both  $\hat{N}_i$  and its bias-adjusted counterpart,  $\tilde{N}_i$  (eq. 4.7). Biases in both population size estimators resulting from heterogeneity tended to be negative and often were severe. Magnitude of the biases was found to depend largely on the average capture probability and the degree of variation in capture probability among individuals (e.g., as indicated by the coefficient of variation of the capture probability distribution, Carothers 1973). Heterogeneous capture probabilities are relatively unimportant when average capture probability is high, leading Gilbert (1973:524) to suggest that "an experimenter need not attempt to design an experiment so that all animals have the same probability of capture, but only that the distribution be shifted nearer one so that nearly all animals have probabilities of capture say  $\geq 0.50$ ." Of course this suggestion has limited practicality. Carothers (1973) also approximated the bias of  $\hat{\phi}_i$  and  $\hat{\phi}_i$  and found them to exhibit only small relative bias in response to heterogeneous capture probabilities. In an interesting example, Carothers (1979) found strong evidence of heterogeneous sighting probabilities in a capture-resighting study of fulmars (*Fulmarus glacialis*). Through computer simulation he was able to approximate the degree of heterogeneity reflected by his test statistic. He then approximated the relative bias in his survival rate estimate and found it to be negligible ( $<0.01$ ) for most years (Carothers 1979). This example illustrates well the robustness of Jolly-Seber survival rate estimator to heterogeneous capture probabilities. Finally, we note that the bias in survival rate estimators resulting from heterogeneity of  $p_i$  is generally negative, but can sometimes be positive (Nichols and Pollock 1983).

Temporary trap response in capture probability refers to the situation in which an animal exhibits a different capture probability for 1 or more periods after initial capture. Pollock (1975) developed a test of this phenomenon, as well as a set of models and appropriate estimators for this situation. Permanent trap response in capture probability refers to a situation in which the change in capture probability with initial capture persists for the remainder of the capture-recapture experiment. It is important to consider effects of permanent trap response on Jolly-Seber estimators because it is difficult to detect and because there are no models available that incorporate this form of trap response. The Jolly-Seber estimator,  $\hat{M}_i$ , depends only on marked animals (eq. 4.5) and is hence unaffected by permanent trap response. The survival rate estimator,  $\hat{\phi}_i$ , is based also on marked animals (see eq. 4.2) and remains unaffected by permanent trap response. The precision of the survival estimator is affected, however, and a trap-happy response increases precision whereas a trap-shy response decreases precision. The population size estimator (eq. 4.1), however, includes the proportion of marked animals in the sample,  $m_i/n_i$ , which is assumed to estimate the true proportion marked in the population,  $M/N_i$ . Under a trap-happy response, animals exhibit higher capture probabilities after initial capture and  $m_i/n_i$  will overestimate  $M/N_i$  resulting in a negative biased  $\hat{N}_i$  (Nichols et al. 1984a). Under a trap-shy response, animals exhibit lower capture probabilities after initial capture. This results in  $m_i/n_i$  underestimating the true  $M/N_i$ , thus producing a positive bias

in  $\hat{N}_i$ . Approximations of the magnitudes of the resulting biases are provided by Nichols et al. (1984a). Relative biases are most severe when the population exhibits substantial turnover (low survival and high recruitment) and also when the marked ratio in the population is small. Also we note that a trap-happy response increases precision of the population size estimate whereas a trap-shy response decreases precision.

If marking decreases the animal's survival rate, serious negative bias can occur to the survival rate estimators. Some methods of marking fish suffer severely from this problem (Ricker 1958). Population size estimators are still unbiased. Temporary reductions in survival rate associated with initial marking can be detected and then estimated using the models and associated tests and estimators of Robson (1969), Pollock (1975), and Brownie and Robson (1983).

Heterogeneity of survival probabilities sometimes can be detected using the goodness-of-fit test of Pollock et al. (1985). However, models (and associated estimators) that incorporate such heterogeneity are not available. Therefore, effects of heterogeneous survival on Jolly-Seber estimators are of interest. Cormack (1972:340) considered this problem intuitively and stated, "Since the parameter  $\phi_i$ , to be estimated is an average survival probability, variation of survival between individuals will not affect the Jolly-Seber estimates provided that the variation is not correlated with marking, either directly or indirectly through age or otherwise." We find that this is true only if an animal's survival probability in 1 period is independent of its survival probability in another period and if its survival probability is independent of its capture probability (Pollock and Raveling 1982). The nature of the bias on survival estimators has been studied by Pollock and Raveling (1982) and Nichols et al. (1982) in the band recovery model context, which has its own special problems. In the capture-recapture situation, no one has yet done a detailed study. A common case is likely to be heterogeneous survival rates, which are positively related for the same animal (i.e., if survival is high for an animal in 1 year, it is likely to be high also in the next year) but which are independent of capture probability. In this case there is a positive bias on survival rate estimators and also on population size estimators. It is not known how large these biases are likely to be in natural populations, but we suspect that they will be small.

Age-specific differences in survival probabilities can be thought of as a special kind of heterogeneity. Manly (1970) studied effects of age-specific differences in survival probability on Jolly-Seber estimators and examined several different patterns of age-specificity. For example, when survival probabilities are lower for young animals (a common situation in natural populations), resulting Jolly-Seber  $\hat{N}_i$  and  $\hat{\phi}_i$  were found to exhibit positive bias (Manly 1970:17). In situations where age of an animal can be ascertained at time of marking, the model of Pollock (1981b) can be used to test for age-specific differences in survival and capture probabilities and to estimate age-specific parameters when such differences are detected (also see Stokes 1984). However, age determination is not always possible (e.g., for many passerine birds banded during winter), and the results of Manly (1970) are of special interest in such situations. Manly (1970:18) concluded, "The results suggest that this method (Jolly-Seber) can reasonably be used whenever mortality rates are not strongly affected by age."

The assumption of no tag loss is also very important, because if animals lose their marks the number of recaptures will be too low. The results is underestimation of survival rates but no influence on population size estimates (Arnason and Mills 1981). Seber (1982:94-96) gives a method to estimate and adjust for tag loss using a double-marking scheme. It is

important to realize that tag loss will cause a decrease in precision of estimators even when it can be estimated and the estimators adjusted (Pollock 1981b).

An implied assumption is that any emigration from the study area is permanent. Balser (1981) derived a test for temporary emigration. Temporary emigration is likely to occur in a number of practical field situations, and our opinion is that the bias induced could be serious. We discuss this further in Section 6.

#### 4.5 Example

In Example 1 we analyzed data from a 2-year study on gray squirrels (*Sciurus carolinensis*) carried out in the mature oak woodland in Alice Holt Forest Research Station, Surrey, England, by A. Duboek. Squirrels were captured at approximately monthly intervals from November 1972 until September 1974. Multiple capture traps baited with grain were located throughout the area on the ground and disguised by leaves. The squirrels were uniquely marked using a toe clipping method.

Basic statistics required for Jolly-Seber estimates are given in Table 4.3, and the parameter estimates and approximate standard errors are given in Table 4.4. Notice that no estimates are presented for periods 12-14 because we believe they would be misleading because of the small numbers of captures (2, 1, and 4). Notice also that survival estimates are sometimes  $>1$  and have been recorded as 1. Similarly some birth number estimates are negative and have been recorded as zero.

This is a precise study, overall, with relatively small standard errors for most sampling periods. The precision results from the high capture and survival probabilities. Once a squirrel is marked it stays in the population a long time and provides information every time it is captured. Also the precision of estimates does vary considerably. This is mainly due to changes in capture probabilities. It should be emphasized, however, that even in studies where the capture probabilities are constant the precision will not be constant and the sampling periods towards the middle of the study will yield more precise estimates. The marked population gradually builds up and there are still plenty of sampling periods to get recaptures because the random variable  $r$  needs to be large for good precision.

In terms of validity of the estimates, there are many factors to consider. The biologist believed that migration was negligible and therefore interpreted survival and birth estimates as being representative of true survival and birth rates. There is likely to be heterogeneity of capture probabilities due to age, sex, and other factors and also a trap-happy response to capture probabilities. Both departures tend to cause underestimation of population size and, to a lesser extent, underestimation of survival rates.

Young animals should have been joining the catchable population in April and May. This shows up in high estimates of birth numbers in 1974 but not in 1973. The biologist had predicted this because 1973 was a very bad year for squirrel reproduction. After all the young animals joined the population in May 1974, there was a large drop in the survival estimate and the population size estimate. Unfortunately this occurred right at the end of the study, and there were so few captures in September 1974 that the estimates are very imprecise.

## **4.6 Model Restrictions and Extensions**

Sometimes it is possible to put restrictions on the parameters of the Jolly-Seber model and hence simplify the model and increase the precision of the remaining parameter estimates.

For some populations in which births and immigrations are known to be negligible, a model that allows only for losses (deaths and emigration) is reasonable. This model was originally developed by Darroch (1959). On rare occasions one may have the situation where there are no deaths or emigration for a population so that a model only allowing recruitment (births and immigration) is reasonable. This model was also first developed by Darroch (1959). (Seber 1982:217; Pollock et al. 1990:34).

Other important restricted models are those where capture and/or survival probabilities are constant over time. (Jolly 1982, Brownie et al. 1986, Pollock et al. 1990). Models with these restrictions can provide large gains in precision on the remaining parameter estimators. This approach has been extended by Lebreton et al. (1992) for survival estimation only and is discussed further in Section 5.1.

Sometimes it is necessary to generalize the Jolly-Seber model and some of the important generalizations are considered next in Section 5. One that is not is the temporary trap response models of Robson (1969), Pollock (1975) and Brownie and Robson (1983). Marking is allowed to have a short term effect on survival and capture probabilities. The estimators still have an intuitive form similar to the original Jolly-Seber estimators and contingency table tests of these short term effects are available. Precision of estimators is lower.

## **5. OTHER OPEN MODELS**

### **5.1 Survival Modeling**

Lebreton et al. (1992) have made a very important contribution to open population capture-recapture modeling. Due to space limitations we cannot do justice to the total content of their monograph but we attempt to summarize the important points.

In Jolly-Seber model it is possible to obtain survival and capture probability estimates directly from modeling the capture histories of the marked animals conditional on first captures. This goes back to the model of Cormack (1964). It is then possible using a generalization of the Lincoln-Petersen estimator to estimate population sizes using marked to unmarked ratios and further to combine all these estimators to obtain birth number estimators. Lebreton et al. (1992) emphasize the importance of estimating survival rates with capture rates as nuisance parameters and do not consider the population size or birth number estimation problem. The likelihoods they consider are conditional on first capture and are products of multinomial distributions. Estimation is straight forward using maximum likelihood methods.

Lebreton et al. (1992) emphasize that their "paper synthesizes, using a common framework, recent developments together with new ones, with an emphasis on flexibility in

modeling, model selection, and the analysis of multiple data sets. The effects on survival and capture rates of time, age, and categorical variables characterizing the individuals (e.g. sex) can be considered, as well as interactions between such effects. This "analysis of variance" philosophy emphasizes the structure of the survival and capture processes rather than the technical characteristics of any particular model. The flexible array of models encompassed in this synthesis uses a common notation. As a result of the great level of flexibility and relevance achieved, the focus is changed from fitting a particular model to model building and model selection."

They recommend the following general procedure of modeling "(1) start from a global model compatible with the biology of the species studied and with the design of the study, and assess its fit; (2) select a more parsimonious model using Akaike's Information Criterion to limit the number of formal tests; (3) test for the most important biological questions by comparing this model with neighboring ones using likelihood ratio tests; and (4) obtain maximum likelihood estimates of model parameters with estimates of precision." They emphasize the importance of computer software and suggest RELEASE (Burnham et al. (1987) for data summary and goodness-of-fit tests and SURGE for iterative model fitting and the computation of likelihood ratio tests.

Lebreton et al. (1992) develop their general approach carefully and in detail and make it easy to follow by illustrating the methodology with five comprehensive examples on a range of species.

Skalski et al. (1994) have developed a computer package SURPH which also does survival modeling. It is similar in philosophy to the work of Lebreton et al. (1992) but there are technical differences which will not be discussed here.

## 5.2 Stage Structured Models

Another extremely important and fast growing area of open capture-recapture modeling is for stage structured populations. These models which began with the work of Arnason (1972, 1973) are for biological populations where release and recapture data are available for each of a number of strata on every capture occasion. Strata may, for example, be geographic locations of physiological states. Movement of animals among strata occurs with unknown probabilities, and estimation of these unknown transition probabilities is the objective.

We emphasize that stratum membership is not known for any period where the animal is uncaptured. A good example would be weight class where an animal may or may not change weight class in the next period and the only way to know for sure is to capture the animal. However, it is possible to estimate the probability of transition from one weight class to another. We note that this is different and more complex than if the stratum where an age class where transfers to the next age class is automatic. The modeling of Lebreton et al. (1992) discussed in Section 5.1 can handle strata like age class where the transition is certain but cannot handle strata like weight class where the transition is stochastic.

There has been a lot of recent work in this area. Schwarz et al. (1993) present the Arnason-Schwarz model and maximum likelihood estimators for it. They illustrate using

some herring tag-return data. The model is basically the same as that of Arnason (1972, 1973) but the estimation procedure is better.

Hestbeck et al. (1991) considered a multiple capture-recapture data set on Canada geese banded in the three areas in the Atlantic flyway of North America. They described the biological importance of estimating the transition probabilities for movement between the three areas. They started with the Arnason-Schwarz model and then generalized to a model where the transition probabilities depend on location in the previous period (Markovian model). Nichols et al. (1992) ran similar models on a meadow vole data set where the strata are body mass classes.

Brownie et al. (1993) consider the Arnason-Schwarz model, the Markovian model and then generalize to non Markovian models where transition probabilities may depend on more than just the location at the last period (i.e. earlier ones also). James Nichols who was involved in the Hestbeck et al. (1991), Nichols et al. (1992) and Brownie et al. (1993) works has supervised James Hines in building a computer program for these models using program SURVIV (White 1983). They refer to their program as MSSURVIV.

There is likely to be a lot more work done on these models in the next few years. Nichols et al. (1992) emphasize that the robust design might be useful here. Having a small sample of radio tagged animals may also be very useful if the appropriate combined models could be developed.

### **5.3 Comparative Capture-Recapture Studies**

Burnham et al. (1987) present design and analysis methods for large class of comparative survival experiments based on release-capture of marked populations. The methods presented apply to any experiments involving treatment and control groups of marked animals. A primary application was to address fishery survival issues on the Columbia River in the Northwest of the U.S.A. Treatments involved fish passing through hydro electric turbines and controls involved fish who were added below the turbines on the river. Testing for differences between treatment and control groups was a primary focus.

Many other applications of this methodology now exist. For example, Deuel (1985) reported on a study that measured indirect mortality from ingested lead shot into northern pintails and Burnham et al. (1987) applied their methods to this data. Stromberg et al. (1988) used starlings as a surrogate native species to study survival of birds exposed to an organophosphate pesticide and applied these type of models.

Computer software is very important to the complex analyses here. RELEASE which is based on program SURVIV (White 1983) was developed by Gary White (Burnham et al. 1987) to handle this type of comparative design.

In this selection we have just been able to give a brief overview of material which Burnham et al. (1987) develop in book form. Their work is very important and deserves detailed study by the more dedicated reader of this article.

## 6. COMBINATION CLOSED AND OPEN POPULATION MODELS

### 6.1 Introduction

We believe that the design of capture-recapture experiments deserves much more attention from statisticians and biologists. Study design should be oriented around satisfaction of as many model assumptions as practically possible so that a simple and reasonably efficient model can be used for estimation.

At the basis of many capture-recapture sampling models is the assumption that all animals are equally likely to be caught in each sample (the equal catchability assumption), which was discussed in Section 2. This assumption is most unlikely to be realistic in natural animal populations, and either 1 or both of the heterogeneity and trap-response alternatives may be acting in a particular animal population.

The capture-recapture model most frequently used by biologists for open populations in long-term studies is the Jolly-Seber model, which was discussed in detail in Chapter 4. This model requires the equal catchability assumption and although some weakening of the assumption is possible, the complexity of open population models is likely to preclude general models that allow heterogeneity and/or permanent trap response. However, some closed models discussed in Section 3 do permit relaxation of the equal catchability assumption in a biologically realistic manner.

Pollock (1981a) asserted that statisticians have drawn a sharp distinction between closed and open population models, which is perhaps rather artificial. To quote (Pollock 1981a:435), "In practice a series of short-term studies may be carried out. One approach to analysis would be to analyze each short-term study using the closed population models...which allow unequal catchability. Then all the sampling periods in each short-term study could be pooled and survival estimators between these short-term studies could be estimated using the Jolly-Seber Model. This approach allows population size estimation under models allowing unequal catchability while survival estimation, which is not so affected by unequal catchability, is under the Jolly-Seber Model."

We now describe the design which was originally published by Pollock (1982) and was motivated by the desire to find a design for long-term studies that is robust to heterogeneity and/or trap response. We examine its robustness and consider some extensions. Pollock et al. (1990) provides a more detailed treatment.

### 6.2 The Basic Design

Consider the following presentation of a capture-recapture sampling experiment where we have  $k$  primary sampling periods (e.g., years), and within each one of these we have  $l$  consecutive days of trapping). It is not necessary that the number of secondary periods be equal in all the primary periods.

The design can be used to estimate the population size for each of the primary periods ( $N_1, N_2, \dots, N_k$ ) assuming that the population is constant over the secondary sampling periods within each primary sampling period. We also can estimate survival rates

$(\phi_1, \phi_2, \dots, \phi_{k-2})$  and numbers of new individuals entering the population between the primary sampling periods  $(B_1, B_2, \dots, B_{k-2})$ .

It would be possible to estimate all of these parameters except  $N_1$  and  $N_k$  using the standard Jolly-Seber model of Section 4. In order to use the Jolly-Seber model with the described design, all the secondary sampling periods within a primary sampling period would be "pooled." Here by "pooled" we mean that we are just interested in whether or not an animal is captured at least once during the primary sampling period.

When considering the use of the Jolly-Seber model, it is important to keep in mind requisite assumptions and model robustness. In Section 4 it was pointed out that heterogeneity and/or trap response will have a large effect on the population size estimators (eq. 4.1) because the sample ratio  $(m_i / n_i)$  will no longer accurately reflect the population ratio  $(M_i / N_i)$ . The marked population estimators (eq. 4.5) will not be so affected by unequal catchability because the 2 ratios of recoveries of marked animals will both tend to be influenced similarly. It follows that the survival rate estimators (eq. 4.2), which are simply ratios of the  $\hat{N}_i$ 's, also will be less affected by unequal catchability than the population size estimators. Cormack (1972) suggested this using an intuitive argument and Carothers (1973) documented it for heterogeneity of capture probabilities using analytic approximations and computer simulation. In the case of permanent trap response, survival rate estimates are not affected at all (Nichols et al. 1984a).

In the robust design we attempt to minimize the influence of unequal catchability on our estimators by exploiting our 2 levels of sampling. Survival rate estimators, which are not so influenced by unequal catchability, will be estimated exactly as under the Jolly-Seber procedure using pooled captures within a primary period. Recall that these survival estimators will be available only for  $i = 1, 2, \dots, k - 2$ .

We suggest that population size estimators for each primary sampling period  $(\hat{N}_1, \dots, \hat{N}_k)$  be obtained using closed population models allowing for unequal catchability (Section 3). These estimates use only the captures and recaptures within a primary sampling period. The easiest way for a biologist to obtain these captures is to use the program CAPTURE developed by Otis et al. (1978) (also White et al. 1982). It considers a range of different models allowing for heterogeneity and/or trap response and also gives an objective method of choosing the most appropriate model to use. Notice that an additional advantage of this approach is that population size estimators are available for all primary periods ( $i = 1, \dots, k$ ), whereas under the Jolly-Seber approach, estimators are only available for  $i = 2, \dots, k - 1$ . Once again, however, we suggest caution in the use of the model selection procedure from CAPTURE (Chapman 1980, Menkins and Anderson 1988). The tests are not independent and often have low power, especially for small populations (Menkins and Anderson 1988).

Finally, the birth rate of estimators can be obtained from eq. (4.3) as before, but now the population size estimators used are those described in the preceding paragraph. Notice that it is possible to estimate  $B_1, \dots, B_{k-2}$ , whereas under Jolly-Seber it is only possible to estimate  $B_2, \dots, B_{k-2}$ . If we had an estimator of  $\phi_{k-1}$  then  $B_{k-1}$  would also be estimable. In some experiments, it may be reasonable to estimate  $\phi_{k-1}$  by  $\phi_{k-2}$  or perhaps by an average of  $\hat{\phi}_1, \dots, \hat{\phi}_{k-2}$ .

### **6.3 Extensions**

Pollock and Mann (1983) and Pollock et al. (1990) pointed out that this design could also be useful in studies of age-stratified populations. Age specific survival rates can be obtained from capture-recapture data over the primary sampling periods while population size estimators can be obtained for each age class from the secondary sampling periods. They also show that it is now possible to separate the Jolly-Seber  $\hat{B}_i$  into its two components of production within the population and animals that have immigrated into the population.

William Kendall in his Ph.D. thesis and related publications (Kendall and Pollock 1992, Kendall et al. 1995) has studied the robust design in great detail. He considered ML estimation when heterogeneity was not present. He carried out some comprehensive simulation studies which showed the value of the robust design over regular open models like the Jolly-Seber model.

Recently work has been done on the study of temporary emigration using the robust design (Kendall et al. 1995). Temporary emigration is an assumption violation that is very difficult to deal with under the Jolly-Seber model. However, using the robust design it is possible to get unbiased estimates of population sizes and survival rates.

## **7. DESIGN OF CAPTURE-RECAPTURE STUDIES**

Although study design is the first task a biologist has to consider, we have left this topic until after all of the models have been presented. We believe that a scientist has to be aware of the different models and their assumptions before he or she can seriously design a capture-recapture study. In this chapter we first present some comments on selection of an estimation method for studying animal populations. We then narrow the focus to capture-recapture studies and consider questions of how to avoid or at least minimize assumption violations when designing a study. This is followed by a presentation on how comprehensive a study is required in terms of sampling intensity to achieve given levels of precision of parameter estimators.

### **7.1 Selection of an Estimation Method**

One of the most important decisions a scientist must make when beginning an investigation involves the selection of methods best suited to the attainment of study objectives. As we have discussed, open capture-recapture models can be used to estimate population size, survival rate, and number of new recruits to the population. These quantities also can be estimated using methods other than capture-recapture. Here we present a very brief general discussion of some other methods and of the strengths and weaknesses of capture-recapture relative to these methods.

A number of alternative methods exist for estimating size and density (population size per unit area) of free-ranging animal populations. Most of these alternatives involve counts or observations of animals. Some methods assume that all animals located within sam-

ple quadrats or transects are detected and seen with probability 1.0. In such instances, traditional sampling techniques (e.g., Cochran 1977) are used to "expand" these sample counts to estimate total population size over some area of interest (Eberhardt 1978, Seber 1982). In situations where animals are easily observed and detected with probability 1.0, such sample counts will almost certainly provide the best method of estimating population size.

In most field situations involving observations of animals, however, it is not realistic to assume a detection probability of 1.0. In these situations detection probability must be estimated and then used in conjunction with count data to estimate population size or density. A variety of methods are available for estimating detection probability. These include comparison of standard counts and intensive counts of sample areas, use of subpopulation of marked animals, mapping with multiple observers, formal line transect sampling, and multiple counts on the same areas (Seber 1982, Pollock and Kendall 1987, Buckland et al. 1993).

The method of estimating detection probability that has received the most detailed statistical development is line transect sampling (Buckland et al. 1993). In line transect sampling, the observer travels along the transect line counting animals. All animals located directly on the line are assumed to be detected with probability 1.0, and detection probability is assumed to decrease with distance away from the line. Perpendicular distances or paired sighting angles and distances are then used to estimate the detection function, which expresses detection probability as a function of distance from the transect line. The estimated detection function is then used with the total number of animals observed to estimate density. We believe the main limitations of line transect sampling models are (1) the assumption of all animals on the line being seen and (2) the practicality of recording distances. These limitations are particularly severe in aerial surveys.

When animals are readily observable, observation-based methods for estimating population size often will be preferable to capture-recapture methods. Field workers generally have recognized this and have tended to use observation-based methods for estimating population size for birds and large mammals. However, other vertebrates (small mammals, reptiles, amphibians, and fish) and most invertebrates are not so easily observed, and capture-recapture methods provide the most reasonable means of estimating population size for many such organisms.

Capture-recapture and band-recovery models are commonly used to estimate survival and mortality probabilities for animal populations. Band-recovery models (Brownie et al. 1985) represent special cases of general capture-recapture models in which all recaptures involve removal of the tagged animals from the population (e.g., removal usually occurs via fishing, hunting, trapping, or some other form of human exploitation). Generally, such data are not used to estimate population size or recruitment, although exceptions exist in special application areas (Brownie and Pollock 1985, Nichols et al. 1986). The relationship between band-recovery models and capture-recapture models is discussed by Brownie and Pollock (1985) and Brownie et al. (1985).

Despite their close conceptual relationship, band-recovery and capture-recapture models often are applied to very different kinds of field sampling situations, and hence, often yield estimates that carry different interpretations. Band recovery models often utilize data supplied by members of the general public involved in the exploitation process (e.g., hunters report bands of shot migratory birds; fishermen report tags of fish in the catch). Thus, the

sample of recovered animals is drawn from a large geographic area, which often covers the entire range of the studied population. Capture-recapture models, on the other hand, typically utilize samples provided by the efforts of a single investigator or team and thus often are restricted to local study areas. As a result, the complement of a survival rate estimate from a capture-recapture study ( $1 - \hat{\phi}$ ) generally includes both mortality and permanent emigration (failure of an animal to return to the area exposed to capture efforts), whereas the complement of survival estimates from many band-recovery studies (e.g., many studies of hunted migratory birds) includes only mortality. With regard to selecting an estimation method, if strict mortality rate is the parameter of interest and if permanent emigration is both likely to occur and difficult to estimate, then band-recovery models may provide the preferred means of estimation. Tag-recovery rates are low in many applications of band-recovery models, however, producing low precision in resulting survival rate estimates. Thus, in situations where permanent emigration does not present a problem or can be estimated, capture-recapture survival estimators often will be preferred to band-recovery estimators because of their greater precision.

In cases where both local recaptures and widespread exploitation recoveries occur for animals marked in a particular study, the distinction between band-recovery and capture-recapture survival estimates permits inference about permanent emigration (Anderson and Sterling 1974, Hepp et al. 1987). Specifically, the ratio of the capture-recapture survival estimate to the band recovery survival estimate should estimate the conditional probability that an animal alive at the time of recapture sampling efforts is in the local area exposed to these efforts. The complement of this conditional probability can be thought of as an estimate of the probability of permanent emigration.

Historically, band-recovery and capture-recapture models have been the 2 most important general methods for estimating survival and mortality rates in free-ranging animal populations. Recently, field researchers have begun to use another technique, based on radiotelemetry, to estimate mortality rates (Trent and Rongstad 1974, Heisey and Fuller 1985). Pollock et al. (1989a,b) point out the relationship between telemetry-based mortality estimation and the estimation of survival in medicine and engineering, and then exploit this relationship to provide general and flexible nonparametric models for estimation of animal mortality that are much more realistic than earlier models.

Radiotelemetry methods provide good estimates of survival probabilities if reasonable numbers of animals are followed, and we suspect that this method will be used with increasing frequency in the future. Telemetry offers advantages over capture-recapture because it often provides insight into sources of mortality (and sometimes even estimates of source-specific mortality risks) and information about emigration. The use of radiotelemetry is relatively expensive, however, in both equipment costs and the substantial field effort needed to obtain frequent locations of large numbers of animals. In addition, it is possible that there is sometimes a lack of independence between losses of contact (i.e., failure to relocate a transmitted animal) and animal status (alive or dead), and such a dependence could cause problems in estimation. In some cases the radio tag may also depress the animal's survival. Nevertheless, radiotelemetry offers a good means of estimating mortality in many situations and will become increasingly important in the future.

There are a variety of approaches available for estimating reproductive rate (we can define this as number of young produced per breeding age female) and number of new recruits entering a population. However, no single approach other than capture-recapture is

very general; instead methods are tailored to meet specific field situations. If alternative (i.e., not using capture-recapture models) methods are used to estimate both population size and survival rate, then number of new recruits can be estimated by subtraction as for capture-recapture  $\hat{B}_i$ . For exploited populations it is sometimes possible to estimate an age ratio (young per adult or adult female at a particular time of year), an indicator of reproductive rate, using harvest age ratios and age-specific band-recovery data (Martin et al. 1979). For easily observed animals, it is sometimes possible to estimate an age ratio using direct field observations.

Reproductive rate and numbers of young animals are frequently estimated by combining estimates of the various components of reproductive rate. For example, in many bird species the number of eggs in a typical clutch has been well estimated. There are also an increasing number of studies providing good estimates of nesting success (the probability that a newly initiated nest "succeeds" to produce at least 1 fledgling). When these estimates are combined with estimates of the proportion of adult females breeding and an estimate of post-fledgling survival, an estimator for reproductive rate of number of young can be developed. However, estimates of these latter quantities (proportion breeding and postfledgling survival) can be difficult to obtain.

When a biological question of interest involves a special component of reproductive rate, then it generally will be best to try to estimate that component directly, rather than make indirect inferences using capture-recapture  $\hat{B}_i$ . When estimates of total new recruits are needed, capture-recapture  $\hat{B}_i$  frequently may be the only available estimator. It should be remembered, however, that  $\hat{B}_i$  is estimated using estimates of both population size and survival rate, and hence has a large sampling variance. The standard capture-recapture  $\hat{B}_i$  also includes animals entering the population via both immigration and in situ reproduction. However, age-dependent models and the robust design (Section 6) permit separate estimation of these 2 components of recruitment in some situations.

## 7.2 Minimization of Assumption Violations

When capture-recapture is selected for use in a particular investigation, it is important to minimize violations of the model assumptions on which the estimators are based. We believe that the problem of assumption violations should be considered at 3 different stages of the capture-recapture study. First, possible violations should be considered when developing the field-sampling design for the study. Second, when the data are analyzed it is important to use general model goodness-of-fit tests, between-model tests, and tests of specific assumptions to aid in deciding which model is most appropriate for the data and to provide insight into whether specific assumptions of the selective models have been violated. Third, if it is concluded that specific assumptions have been violated, then it is important to use information from studies on effects of assumption violations when interpreting the estimates and using them to draw inferences. For example, if we conclude that heterogeneity of capture probabilities exists in our data, we still may be able to use resulting Jolly-Seber survival estimates, as they tend to exhibit relatively small bias under heterogeneity (Carothers 1973, 1979). On the other hand, if heterogeneity occurs and capture probabilities are not high (<0.5), then we would not be able to rely on our Jolly-Seber population size estimates although we would know that they likely underestimate true population size. In some cases, it is even possible to

"correct" capture-recapture estimators for specific violations (Carothers 1979, Arnason and Mills 1981). Tests of assumptions and effects of assumption violations on capture-recapture estimates have been discussed elsewhere. Here we briefly discuss aspects of study design that relate to model assumptions and their violation.

Heterogeneity of survival and capture probabilities is likely to exist in all capture-recapture studies to some degree. Although there is no study design that can completely eliminate such heterogeneity, there are certain ways to reduce it. When it is suspected that survival or capture probabilities are associated with some measurable characteristic of individual animals (e.g., age, sex, size), then such characteristics should be recorded so that the sample can be stratified for analysis. When animals cannot change strata during the course of the study (e.g., sex, also age and size in sufficiently short studies), the appropriate approach simply involves separate analyses for the different strata. When animals change strata during the course of the study, as with age, then special models must be developed to incorporate such heterogeneity.

Heterogeneity of capture probabilities can occur as a result of heterogeneous sampling intensities; therefore efforts should be made to ensure that all sections of the study area are sampled with equal intensity. When traps, mist nets, or other stationary capture devices are used, it is important that all animals in the sampled area are exposed to the devices. For example, even in trapping grids where traps are placed at equal intervals in a checkerboard pattern, it is important that the distance between traps be smaller than the average home range radius of the studied animal. If traps are spaced too far apart, then capture probability can vary substantially as a function of the location of an animal's home range. It is important to have at least 1 trap, and preferably more, within each animal's home range. If sufficient traps are not available for such intense sampling, then traps can be moved during the course of each sampling period to ensure adequate coverage.

In a 2-sample Lincoln-Petersen experiment (Section 2), a good method to eliminate bias due to heterogeneity is to use 2 different sampling methods. Unfortunately, open population models require at least 3 sampling periods and often have many more; therefore it is not feasible to use different methods in each sample.

As noted in the previous section, we believe that the robust design represents a very useful approach to dealing with heterogeneous capture probabilities. At this time, we believe that implementation of this design is the best available means of minimizing estimation problems associated with such heterogeneity.

Trap response in capture or survival probabilities is another problem that occurs commonly in capture-recapture studies. Evidence of trap response in capture probabilities has been found in a number of studies of vertebrates (Otis et al. 1978:29, Tanaka 1980:42-52, Nichols et al. 1981:123, Seber 1982:83-85). Trap-happy responses often are associated with the use of baited traps. Although such response could be eliminated by not baiting at all, this usually is a poor alternative because of the large losses of data (few animals are captured) that result. Prebaiting, the placement of traps that are baited and locked open in the study area prior to the initiation of trapping, is thought to reduce trap-happiness in some situations, but can lead to other problems (Tanaka 1980:52-55). When considering a trap-happy response, it is important to recall that survival rate estimators are not biased by trap response and that a trap-happy response results in higher precision to the survival rate estimator. Trap-shyness

also is a common response and can be minimized by using capture and handling techniques that exert minimal stress on animals. Minimization of capture and handling stress also helps protect against a trap response in survival rate, in which animals show low survival probabilities for the period immediately following initial capture.

Sometimes use of only resighting data (without capture) may be possible as, for example, with color-marked birds, and this method avoids the problem of trap response. Typically, studies using this design have obtained only survival rate estimators (Cormack 1964), but sometimes it is possible to get ratios of marked to unmarked animals at the same item and estimate population size as well. This design has good precision if it is possible to obtain high resighting probabilities. One possible disadvantage is that some biologists believe that color-marked animals may have lower survival rates than unmarked animals (Burley 1985).

Temporary (e.g., lasting 1 or perhaps 2 periods after marking) changes in capture and survival probability can be handled using the general models of Robson (1969), Pollock (1975), and Brownie and Robson (1983), and hence require no special design considerations. Permanent trap response in capture probability cannot be handled using available open models, but the robust design of Section 6 permits both testing for such response and estimating parameters of interest without bias when it does occur.

An important practical consideration when designing a study is to use a good marking method. The marking method needs to be such that it does not influence the animal's survival, but at the same time mark loss should be negligible. If a new marking method is being used, it would be advisable to carry out a small pilot study and investigate its practicality. If mark loss cannot be avoided completely, it should be estimated through techniques such as double marking (Section 2). Estimates of tag loss then can be used to correct open model estimates using methods such as those in Arnason and Mills (1981).

Most capture-recapture models cannot distinguish between births and immigration and between deaths and emigration. As noted in Section 6, the robust design permits separation of recruits arising from birth and immigration in some situations. In some cases migration may be ruled out on biological grounds by a careful choice of study area to take account of natural boundaries. This is the case for the study on the gray squirrel described in Section 4.5. The problem of temporary emigration should also be considered. This can be especially severe in small mammal studies and suggests the use of large grids, natural boundaries, and even "semipermeable" fences to minimize the problem. In cases where migration has the potential to cause large problems, it may be worthwhile to accompany the capture-recapture work with a radiotelemetry study designed to estimate the magnitude of both permanent and temporary emigration. The robust design may also be used if temporary emigration is suspected.

We believe that in some cases the biologist may be able to choose a population so that no births or immigration are occurring. In such cases the model that only allows losses (Section 4.6) can be used. It has the advantage of giving higher precision to population size and survival rate estimators. We believe that the model that only allows gains (births or immigrants) is less likely to be useful.

Finally we emphasize that simulation can be a useful tool at the design stage to evaluate assumption violations. POPAN-3 (Arnason and Schwarz 1987) has a flexible, powerful, simulation component as part of the program.

### 7.3 Sampling Intensity

In this section we consider the question of what sampling intensity is required to reach a specific level of precision for all parameter estimates. We shall use the quantity, coefficient of variation (CV), or proportional standard error as our measure of precision. We define

$$CV(\hat{\theta}) = \frac{SE(\hat{\theta})}{\theta},$$

which is just the standard error of the estimator of  $\theta$  divided by  $\theta$ , the parameter itself. Lowering CV indicates increasing precision for the study. A rough rule of thumb is that a study that provides a CV of 20% is reasonable. Many wildlife and fisheries studies will not achieve this level because of cost constraints. Of course the precision of our estimator will depend on the values of the various parameters in our model and on the model being considered.

Otis et al. (1978) considered sampling intensity in detail for the closed populations (Section 3). Pollock et al. (1990) considered in detail sampling intensities for the Jolly-Seber model and its simplifications and generalizations (Section 4). Burnham et al. (1987) considered the sampling intensity needed for good power for comparisons of survival rates. We believe that simulation will often be used routinely to check the feasibility of various designs with desirable levels of precisions in advance. The program POPAN-3 (Arnason and Schwarz 1987) can be used for this purpose for the Jolly-Seber model but other programs may need to be written to make simulation easier for other models.

## 8. APPLICATIONS

### 8.1 Animal Populations

A wide variety of biological applications beyond the usual mark and recapture studies on fish and wildlife populations have appeared in the literature. Standard mark and recapture methods have been applied to insect populations and Wileyto et al. (1994) have also developed what they term the Markov-recapture method for sampling insect populations. There are two types of traps a continuous marking trap and a continuous removal trap. A simple estimator related to the Lincoln-Petersen estimator results if the data is only collected once. Wileyto (1994) and Wileyto (1995) have extended the method to multiple markings and multiple time periods respectively.

The Lincoln-Petersen model has been used to estimate the number of animals missed in aerial surveys (Henny and Anderson 1979; Pollock and Kendall 1987). Buckland et al. (1993), Borchers (unpublished manuscript) and Alpizar (1994) have extended these methods to form a combined line transect and capture-recapture model.

A modified Lincoln-Petersen model has been applied to situations where feces (skats) are radioactively marked by injection of radio isotopes into the animal. This has been applied to black bears (Pelton and Marcom 1977), raccoons (Conner and Labisky 1987) and otters (Shirley et al. 1988).

Capture-recapture methods have been applied to the fossil record. Nichols and Pollock (1983) showed how to use the Jolly-Seber model to estimate number of species, speciation rates and extinction rates.

## **8.2 Human Populations and Other Applications**

Many authors have applied capture-recapture methods to estimating the census undercount (see Feinberg (1992) for a complete bibliography). Capture-recapture methods have been used to estimate the number of homeless people in a city (Cowan et al. 1986, Laska and Meisner (1993)). Greene (1983) has used the method to estimate demographic parameters on criminal populations. Witten (1974) and Witten et al. (1974) have used capture-recapture to estimate numbers of people with illnesses from hospitals and other lists (see also Feinberg (1972) and Bishop et al. (1975)). The sampling of elusive human populations using cluster sampling, network sampling, and capture-recapture sampling was discussed by Sudman et al. (1988). Authors including Nayak (1988, 1991) Goudie (1990) and Chao et al. (1993) have applied continuous sampling closed capture-recapture methods to the estimation of the number of errors in a computer program.

A closely related application is estimation of (establishment) frame size considered by Pollock et al. (1994) which we take up in the next section.

## **9. FRAME SIZE ESTIMATION**

### **9.1 Introduction**

In classical sampling theory it is assumed that a complete frame exists. There is, at least conceptually, a complete list of population units. It is then possible to draw a probability sample from the population. Estimators of population parameters such as the mean or total then have known properties and are easily studied theoretically or numerically. Cochran (1978) and other sampling theory books concentrate on this situation and give properties of estimators for common sampling designs such as simple random sampling, stratified random sampling, and multi-stage (or cluster) sampling.

In practice for many surveys a complete frame may not exist. For example in establishment or business surveys lists kept by professional associations or government agencies are often incomplete. A possible approach to solving this problem using capture-recapture methods has been proposed by Pollock et al. (1994), and that article forms the basic for this section.

There are several possible solutions to the problem of an incomplete list frame. The first and most obvious one is to establish a complete list and this is the approach taken by many national statistical agencies despite its expense. A second approach is to switch to an area frame approach and enumerate all the establishments in the sampled areas. This can be inefficient and expensive if the areas are large and the businesses sparse. Another widely used approach for sampling farms is to use the multi frame approach originally developed by

Hartley (1962, 1974) and now developed in the National Agricultural Statistics Service farm plus an area frame where all farms surveys in the U.S.A. (Kott and Vogel 1995). These surveys use an incomplete list frame of farms within a sample unit are enumerated. Thus the area frame is conceptually complete because within each sampled area all farms have to be enumerated and any area could actually be chosen probabilistically.

In the remainder of this section we consider another solution based on using several incomplete list frames and estimating frame size using capture-recapture models. Unlike the other solutions this approach is strongly model based.

## 9.2 Estimation of Population Total

The usual estimator of the population total for single random sampling without replacement is

$$\hat{Y} = N \bar{y} \quad (9.1)$$

where  $N$ , the frame size is known and  $\bar{y}$  is the sample mean. See, for example, Cochran (1978, p. 21). The variance of  $\hat{Y}$  is

$$\text{Var}(\hat{Y}) = N^2 \text{Var}(\bar{y}) \quad (9.2)$$

where

$$\text{Var}(\bar{y}) = \frac{S^2}{n} \left( \frac{N-n}{N} \right),$$

$S^2$  is the population variance and  $(N - n)/N$  is called the finite population correction factor.

The estimator (9.1) is unbiased.

If instead  $\hat{N}$  has to be estimated from a capture-recapture model we would have

$$\hat{Y} = \hat{N} \bar{y} \quad (9.3)$$

and now both  $\hat{N}$  and  $\bar{y}$  are random variables unlike in (9.1) where  $N$  was known exactly. The estimated variance of  $\hat{Y}$ , is

$$\begin{aligned} \text{Var}(\hat{Y}) &= (\hat{N})^2 \text{Var}(\bar{y}) + (\bar{y})^2 \text{Var}(\hat{N}) \\ &\quad - \text{Var}(\hat{N}) \text{Var}(\bar{y}). \end{aligned} \quad (9.4)$$

This variance will be larger than that in (9.2) because  $\hat{N}$  also has a variance due to its estimation. The estimator (9.3) will only be unbiased if both  $\hat{N}$  and  $\bar{y}$  are unbiased estimators.

It is possible to derive results like (9.3) and (9.4) for other parameters and other sampling designs.

### **9.3 Use of Capture-Recapture Models in the Large Pelagic Survey**

The large Pelagic survey is an angler survey conducted by the National Marine Fisheries Service in the U.S.A. using a combination telephone and access design. A sample of fishing boat owners on a list are telephoned to obtain fishing effort (*i.e.* number of fishing trips in a period) information. Catch per unit effort (*i.e.* catch per trip) information is obtained from a second sample of boat owners at access points at completion of their fishing trips. The information from the two surveys is combined to estimate total effort and total catch of important species such as Bluefin Tuna.

A serious problem with this survey is that the list of boat owners used in the telephone survey is very incomplete. Therefore, classical sampling theory which assumes a complete frame of known size ( $N$ ) is inadequate and has to be modified. The current method of estimating the size of the fishing boat list frame involves combining two lists, (a telephone list with a dockside list) and using the Lincoln-Petersen model.

#### **The Lincoln-Petersen Model**

##### **Estimation of Frame Size (N)**

Under the current method the "marked" boats ( $M$ ) are those on the master list which is primarily derived from previous telephone interviews. The recapture sample is carried out dockside at gas pumps and the total number of boats intercepted ( $n$ ) is checked to see which ones are "marked" ( $m$ ) (*i.e.*, on the original master list). The Chapman estimator (Section 2.1) can then be used to provide an estimator of the frame size ( $N$ ). Let us now consider the assumptions of this model and what effect violations might have on the bias of the estimator of  $N$ .

#### **Closure**

This assumption is likely to be violated. Fishing boats may be on the master list and then no longer take part in the fishery (losses). New fishing boats may join the fishery while it is in progress (gains). Ideally a separate estimate of frame size should be obtained for each two week time period. The advantage of using the Lincoln-Petersen closed model estimator is its simplicity and practicality. Biases in the estimator due to lack of closure could be either positive or negative.

Currently it is not known how the fishing fleet size is likely to change during the fishing season. A multiple capture-recapture sampling design would allow use of the Jolly-Seber model to estimate the fleet size during each period. Examination of these estimators and the survival rate and recruitment number estimators will enable us to evaluate the validity of the closure assumption.

## **Equal Catchability**

Violation of the assumption of equal catchability may be due to either inherent heterogeneity of capture probabilities between individuals or "trap response" where individuals that are marked have higher or lower capture probabilities than unmarked individuals. In either situation when the individuals on the lists are fishing boats we believe there is a potential for heterogeneity of capture probabilities among fishing boats. If heterogeneity is operating across both samples, individuals "caught" on the first list will tend to be those with high capture probabilities and therefore they will more likely to be "caught" again on the second list. This means that the proportion marked in the second sample (list ) will be too high and the estimator of  $N$  will be negatively biased. Note that this intuitive argument makes clear it is not heterogeneity per se which is the problem but the positive correlation of capture probabilities between the two samples. Another way of stating the equal catchability assumption is that capture probabilities in the two samples are independent. One method of attempting to achieve independence of the capture probabilities in the two samples is to use totally different sampling schemes for the two samples. This is why we recommended earlier that one sample list be based on the telephone interviews and the other on dockside interviews. However, we do suspect that there is still another heterogeneity and lack of independence in capture probabilities. We believe that fishing boats which take a very active part in the fishery are more likely to be on any lists gathered (telephone or dockside). This heterogeneity will cause a negative bias on the estimate of frame size but we have no idea of the degree of this negative bias. A more complete discussion of heterogeneity and independence of samples is given by Seber (1982, p. 86).

## **Marks Lost or Overlooked**

The situation here is a little confusing. At first one might think that in this application there is not a way that a mark could be lost or overlooked. However, this assumes that all boats have distinct names or if boats do have the same name there is additional information like captain's name which makes all individuals on the lists unique. If there is any problem with lack of uniqueness it may not be clear whether a marked boat has been captured or not. Another related point is that agents may make errors in the records which make it hard to match up a recapture with the original record. A standard operating procedure is being developed and documented to minimize these kinds of errors in the future.

## **Estimation of Total Effort and Total Catch**

Total Effort ( $E$ ) (i.e., the total number of fishing trips taken in a defined period ) is estimated by

$$\hat{E} = \hat{N} \bar{e}, \quad (9.5)$$

where  $\hat{N}$  is the frame size (Fleet Size) estimate and  $\bar{e}$  is the mean fishing effort (i.e., average number of fishing trips taken) obtained from the telephone sample. The evaluation of the pro-

perties of this estimator is more difficult than when  $N$  is known because both  $\hat{N}$  and  $\bar{e}$  are random variables. We suspect that  $\bar{e}$  is biased high because fishing boats that do not fish much are less likely to be on the list. Unfortunately we cannot say that  $\hat{N}$  will always be biased high or low. All three of the assumption violations could be important (closure, heterogeneity, and mark loss) and it is not clear what direction the overall bias on  $\hat{N}$  would take. The only possible approach is to use simulation with a variety of different scenarios for assumption violations. Using equation (9.4) the estimated variance of  $\hat{E}$  is given by

$$\begin{aligned}\hat{\text{Var}}(E) = & (\hat{N})^2 \hat{\text{Var}}(\bar{e}) + (\bar{e})^2 \hat{\text{Var}}(\hat{N}) - \\ & \hat{\text{Var}}(\bar{e}) \hat{\text{Var}}(\hat{N}).\end{aligned}\quad (9.6)$$

Total catch ( $C$ ) is estimated by  $\hat{C} = \hat{E}\bar{c}$  where  $\hat{E}$  is the estimated total fishing effort and  $\bar{c}$  is the average catch per unit effort calculated from the dockside interviews. Properties of this equation are likely to be subject to similar concerns as equation (9.5) and again simulation could be very useful.

## 9.4 Other Issues

### Methods of Dealing with Incomplete List Frames

#### (i) Complete the List Frame

The advantage is that the survey researcher has a complete frame and does not have to generalize results for an estimated frame size. The disadvantage is the cost and possible impracticality of completing the list frame.

#### (ii) Use an Area Frame

The advantage is that one only has to enumerate the establishments in the areas to be sampled. The disadvantage is possible inefficiency if businesses are sparse in each large area.

#### (iii) Using List and Area Frame (Multi-Frame Approach)

The advantages are obviously increased precision and having all establishments covered. The disadvantage could be expense and impracticality

#### (iv) Use of Capture-Recapture to Estimate List Frame Size

The advantage is having a practical method of lower expense than the first three approaches listed above. The disadvantages are potential bias if the assumptions of the capture-recapture method are violated and having to include variation due to frame size estimation in variance estimates of population total estimates.

## **Capture-Recapture Estimation of Frame Size**

In this section we consider model assumptions, precision of estimates, estimation of population totals and the special problems in more complex sampling designs when the capture-recapture approach to frame size estimation is used.

### **Model Assumptions**

#### **(i) Closure**

Can the frame size be considered constant so that the closed population models can be used? This will depend on whether the survey is just a snapshot at a single time point or whether a series of surveys over time are required. It will also depend on how quickly new ones arise. We suspect there will be the need for use of closed and open population models depending on the establishments being studied.

There is also the question of temporary emigration where establishments go out of the frame and then come back in again. This was considered a potential problem in the fishing boat example because boats could go inactive and then become active again. This may also be a problem in some other establishment surveys if establishments go in and out of business frequently and keep the same name when they come back into business.

#### **(ii) "Unequal Catchability" and Independence of Lists**

As we discussed earlier ideally the lists should be independent so that the estimates of frame size are unbiased. In practice it may not be easy to find two or more independent lists.

#### **(iii) Mark Loss-Unique Identification of Establishment**

Establishment names need to be unique and unmistakable or matches on different lists may be missed or mistaken. This is a problem in the fishing boat example in earlier years. We suspect this will not be such a problem in most establishment surveys.

### **Precision of Estimates**

The lists used need to be of sufficient size that the precision of the frame size ( $\hat{N}$ ) is adequate. Seber (1982, p. 96) discusses the Lincoln-Petersen estimate in detail and presents graphics of sample sizes required for various levels of precision. Pollock *et al.* (1990) presents sample size information for the open population models.

## **Estimation of Population Totals**

Once the estimate of frame size is obtained then that estimate will often be combined with a sample mean to estimate an estimate of the population total ( $\hat{Y} = \hat{N}\bar{y}$ ) The estimate of population total is subject to possible bias and additional variance because  $\hat{N}$  is estimated. The estimate may also be biased because  $\bar{y}$  is not based on a random sample of the complete frame.

## **More Complex Sampling Designs**

Here we have emphasized estimation of frame size in simple random sampling using the capture-recapture method. Further questions arise if more complex sampling designs are used. For example in stratified designs the question would arise of whether to estimate frame size in each stratum separately or to estimate the total frame size and then apportion it to the strata assuming equal probabilities of different strata on the incomplete lists. There is also the more complex question of how to estimate frame size in multi-stage designs. This is obviously an area that needs future research.

## **10. COMPUTER PROGRAMS**

There are now a very large number of computer programs that have been developed for use with various capture-recapture models. In each case a reference is provided.

### **Closed Populations**

For closed capture-recapture models the main program is CAPTURE which was originally developed by Otis et al. (1978). It has been substantially revised and updated by Rextad and Burnham (1991).

### **Jolly-Seber Models**

Pollock et al. (1990) provide JOLLY for the basic Jolly-Seber model. This program was written by James Hines and James Nichols. They also wrote an extension JOLLYAGE for populations with two age classes and sampling once per year.

Arnason and Schwarz (1987) have provided POPAN-3 which can be used for the Jolly-Seber type models.

## **Survival Modeling**

The first program developed here was SURVIV (White 1983) and it is still widely used because of its power and flexibility. It is not very "user friendly."

Recently SURGE (Lebreton et al. 1992) was developed. It allows complex survival models to be fitted easily and also allows for covariates.

SURPH (Skalski et al. 1995) has also been developed. It also allows complex survival modeling with covariates.

## **Stage-Structured Models**

A modified version of SURVIV has been written for use with stage structured populations. The program MSSURVIV is described in Brownie et al. (1993) but has its roots in the earlier papers of Hestbeck et al. (1991) and Nichols et al. (1992). The program was written by James Hines and James Nichols.

## **Comparative Capture-Recapture Models**

SURVIV has also been modified by White to consider comparative capture-recapture models. The program RELEASE is described in Burnham et al. (1987).

## **Generalized Linear Models (GLIM)**

Cormack (1989) and earlier works describes how to use the program GLIM (McCullagh and Nelder 1983) to compute closed and open capture-recapture model estimators.

## **11. CONCLUSIONS**

Recently there have been a lot of papers written using Bayesian methods for capture-recapture models. For example, Lincoln-Petersen models (Gazey and Staley 1986), closed population models (eg. Castledine 1981, Smith 1988, Rodrigues et al. 1988) and band return models (Vounatsou and Smith 1995) have been considered. Numerical problems are now being solved by the Gibbs sampling algorithm (Gelfand and Smith 1990) and numerical integration techniques. Bayesian methods are likely to become more widely used in the future.

Computational issues are very important and the use of resampling techniques like the bootstrap are likely to increase. Buckland and Garthwaite (1991) and Norris and Pollock (1995) have applied the bootstrap to closed capture-recapture models.

Lebreton et al. (1992) have taken a very strong model selection approach based on the Akaike Information Criteria (Akaike 1973). We believe model selection criteria will be studied further in the future.

We believe that in the future combination methods will receive a lot of study. For example, Alpizar (1994) and others have considered a combined capture-recapture and line transect method. Use of combined methods while complex may allow weakening of some of the assumptions of the individual methods. Burnham (1993) combined capture-recapture and band return models. Another combined method may involve combining capture-recapture models with radio-telemetry (Pollock et al. 1995).

Table 3.1. Capture-recapture models for closed populations that allow for unequal capture probabilities.

Model <sup>a</sup>	Source of variation in capture probability				Estimator available
	Heterogeneity	Trap Response	Time		
$M_o$					Yes
$M_h$	X <sup>b</sup>				Yes
$M_b$		X			Yes
$M_{bh}$	X	X			Yes
$M_t$			X		Yes
$M_{th}$	X		X		Yes
$M_{tb}$		X	X		Yes
$M_{tih}$	X	X	X		No

<sup>a</sup>This set of 8 models comes from Otis et al. (1978)

<sup>b</sup>X's denote the sources of variation in capture probability incorporated in the models.

Table 3.2. Model selection procedure from program CAPTURE for the meadow vole data collected by J. D. Nichols at Patuxent Wildlife Research Center, Laurel, Maryland, in October 1981.

Model	$M_o$	$M_h$	$M_b$	$M_{bh}$	$M_t$	$M_{th}$	$M_{tb}$	$M_{tih}$
Criteria	0.80	1.00 <sup>a</sup>	0.38	0.59	0.00	0.32	0.52	0.98 <sup>a</sup>

<sup>a</sup>The high values for  $M_h$  and  $M_{tih}$  suggest the appropriate model is probably  $M_h$  or  $M_{tih}$ . The suggested estimate is the jackknife corresponding to  $M_h$ , the simple model. There is no estimator for  $M_{tih}$ .

Table 3.3. Selected statistics and parameter estimates from program CAPTURE for the meadow vole data collected at Patuxent Wildlife Research Center, Laurel, Maryland, in October 1981 by J. D. Nichols. Model,  $M_h$  the heterogeneity model, is used.

Frequencies of capture <sup>a</sup>					
<i>i</i>	1	2	3	4	5
<i>F(i)</i>	29	15	15	16	27
Number of Animals Captured 102 <sup>b</sup>					
Average P-HAT = 0.44					
Interpolated Population Estimate is 139 With Standard Error 10.85.					
Approximate 95% Confidence Interval 117 to 161					

<sup>a</sup>These are the numbers of animals caught from 1 to 5 times.

<sup>b</sup>This is the number of distinct animals captured at least once.

Table 4.1. Notation for the Jolly-Seber model described in detail in Section 4.

#### PARAMETERS

- $M_i$  = the number of marked animals in the population at the time the  $i$ th sample is taken ( $i = 1, \dots, k; M_1 \equiv 0$ ).
- $N_i$  = the total number of animals in the population at the time the  $i$ th sample is taken ( $i = 1, \dots, k$ ).
- $B_i$  = the total number of new animals entering the population between the  $i$ th and  $(i + 1)$  the sample and still in the population at the time  $(i + 1)$  the sample is taken ( $i = 1, \dots, k - 1$ )
- $\phi_i$  = the survival probability for all animals between the  $i$ th and  $(i + 1)$  the sample ( $i = 1, \dots, k - 1$ )
- $p_i$  = the capture probability for all animals in the  $i$ th sample ( $i = 1, \dots, k$ )

#### STATISTICS

- $m_i$  = the number of marked animals captured in the  $i$ th sample ( $i = 1, \dots, k$ ).
- $u_i$  = the number of unmarked animals captured in the  $i$ th sample ( $i = 1, \dots, k$ ).
- $n_i$  =  $m_i + u_i$ , the total number of animals captured in the  $i$ th sample ( $i = 1, \dots, k$ ).
- $R_i$  = the number of  $n_i$  that are released after the  $i$ th sample ( $i = 1, \dots, k - 1$ ). This may not be all of the  $n_i$  due to losses on capture as discussed in the text.
- $r_i$  = the number of the  $R_i$  animals released at  $i$  that are captured again ( $i = 1, \dots, k - 1$ ).
- $z_i$  = the number of animals captured before  $i$ , not captured at  $i$ , and captured again later ( $i = 2, \dots, k - 1$ ).

Table 4.2. Hypothetical example to illustrate that capture-history information can be summarized by a series of zeros and one for each animal.

		Capture period				
Animal		1	2	3	4	5
1		1	1	0	0	0
2		1	0	0	0	0
3		1	0	0	1	0
4		0	1	1	1	1

Table 4.3. Capture-recapture statistics for a gray squirrel population at Alice Holt Forest Research Station, Surrey, England, November 1972-September 1974.

Period	Date	$n_i^a$	$m_i$	$R_i$	$r_i$	$z_i$
1	Nov 1972	46		46	43	
2	Dec 1972	46	42	46	44	1
3	Jan 1973	48	42	48	48	3
4	Feb 1973	46	42	46	45	9
5	Mar 1973	51	46	50	46	8
6	Apr 1973	37	37	37	35	17
7	May 1973	41	41	41	40	11
8	May-Jun 1973	42	39	42	37	12
9	Jun 1973	47	43	47	40	6
10	Jul 1973	31	26	31	26	20
11	Aug 1973	8	7	8	8	39
12	Sep 1973	2	2	2	2	45
13	Oct 1973	1	0	1	1	47
14	Nov 1973	4	3	4	3	45
15	Dec 1973	9	8	9	8	40
16	Jan 1974	19	17	18	17	31
17	Feb 1974	19	14	19	18	34
18	Mar 1974	27	20	27	24	32
19	Apr 1974	36	36	36	32	20
20	May 1974	45	34	44	33	18
21	Jul 1974	74	46	73	15	5
22	Aug 1974	22	20	22	2	0
23	Sep 1974	3	2	2		

<sup>a</sup>Notation is explained in Table 4.1.

Table 4.4. Jolly-Seber estimates and approximate standard errors<sup>a</sup> for a gray squirrel population at Alice Holt Forest Research Station, Surrey, England, November 1972-September 1974.

Period	Date	$\tilde{N}_i$ <sup>b</sup>	$\hat{SE}$	$\hat{\phi}_i$	$\hat{SE}$	$\tilde{B}_i$	$\hat{SE}$
1	Nov 1972			0.94	0.037		
2	Dec 1972	47.1	0.39	0.96	0.030	6.3	0.77
3	Jan 1973	51.3	0.70	1.00	0.004	4.5	1.27
4	Feb 1973	56.0	1.19	0.99	0.023	5.1	1.53
5	Mar 1973	60.5	1.51	0.94	0.041	0.0	1.06
6	Apr 1973	54.9	1.23	0.95	0.038	0.0	0.00
7	May 1973	52.3	0.60	1.00	0.030	3.9	1.22
8	May-Jun 1973	56.5	2.06	0.90	0.052	3.7	1.45
9	Jun 1973	54.6	1.57	0.92	0.067	8.7	3.30
10	Jul 1973	58.9	4.59	0.84	0.066	2.2	6.60
11	Aug 1973	51.8	5.99	1.00	0.000	0.0	5.99
12	Sep 1973						
13	Oct 1973						
14	Nov 1973						
15	Dec 1973	58.3	9.20	0.93	0.115	1.0	6.57
16	Jan 1974	55.3	4.30	0.98	0.068	12.1	8.25
17	Feb 1974	66.4	8.14	1.00	0.071	6.8	10.14
18	Mar 1974	74.5	7.91	0.93	0.067	0.0	6.28
19	Apr 1974	58.4	2.13	0.99	0.071	18.2	4.22
20	May 1974	76.0	6.12	1.00	0.168	33.9	8.86
21	Jul 1974	110.3	18.10	0.21	0.048	0.0	2.23
22	Aug 1974	21.9	0.00				
23	Sep 1974						

<sup>a</sup>  $\hat{SE}(\tilde{N}_i)$  and  $\hat{SE}(\tilde{B}_i)$  include only sampling variation or "error of estimation";  $\hat{SE}(\hat{\phi}_i)$  was obtained using the full variance estimator of Jolly (1965).

<sup>b</sup>Notation explained in Table 4.1.

## 1. INTRODUCCIÓN

El primer uso conocido de los métodos de captura-recaptura lo hizo Laplace (1786), que los empleó para estimar el tamaño de la población de Francia (Seber 1982, p.104). Durante este siglo ha habido una rápida proliferación de los métodos y de las aplicaciones en muchos campos. Al principio, se aplicaba sobre todo a las poblaciones biológicas, pero actualmente hay muchas aplicaciones para las poblaciones humanas (Sudman et al. 1988). Se han usado los modelos de captura-recaptura para estimar el tamaño de un marco de muestreo (Pollock et al. 1990).

Los modelos de captura-recaptura pueden usarse para estimar los parámetros demográficos fundamentales para una población. Éstos inducen tamaño de la población, índice de "supervivencia" y números de "nacimientos". Normalmente, no es posible separar la supervivencia verdadera de la emigración, o los nacimientos verdaderos de la inmigración al menos que se reúna información adicional (Pollock et al. 1990)

En los experimentos de captura-recaptura, normalmente, se muestrea  $K$  veces a la población bajo estudio, cuando  $K$  es generalmente  $>2$ . Cada vez, todo animal no marcado que se captura es marcado de forma única, se toma nota de las capturas de los animales marcados en ocasiones anteriores, y todos los animales son puestas en libertad de nuevo entre la población. En algunos casos, los animales pueden resultar heridos durante la captura, o pueden ser requeridos para sacrificiarlos, en cuyo caso no serán liberados y se registran como "pérdidas al capturar". Al final del estudio, se supone que el experimentador dispone de la historia completa de capturas de todos los animales manejados (Pollock et al. 1990).

Los estadísticos han clasificado típicamente los modelos de captura-recaptura como aptos para poblaciones abiertas o poblaciones cerradas. Es importante que el lector tenga claro el significado exacto de cerrada y abierta. Con abierta, quiero decir que las adiciones (nacimientos y/o inmigrantes) a la población y las deleciones permanentes (muertes y/o emigrantes) de la población ocurren durante el estudio.

Una población cerrada es una en que las adiciones y deleciones no ocurren, es decir, una población cerrada tiene un tamaño constante durante el estudio.

Resulta más sencillo para el estadístico el estudio de los modelos de las poblaciones cerradas. Unas referencias muy importantes son las monografías de Otis et al. (1978) y de White et al. (1982). Los modelos de población abierta son también importantes con las monografías de Pollock et al. (1990) y de Lebreton et al. (1992).

Un área relacionada es la comparación de los índices de supervivencia para poblaciones diferentes, que ha sido considerado por Burnham et al. (1987). A veces es posible combinar los modelos cerrados y abiertos en un diseño general (Pollock, 1982, Pollock et al. 1990,

Kendall et al. 1995). Skalski y Robson (1992) consideraron el uso de los métodos simples de captura-recaptura en el diseño y análisis de los estudios ecológicos *in situ* y de las evaluaciones de impacto ambiental. Últimamente, ha habido también un fuerte desarrollo de modelos de captura-recaptura para las poblaciones estructuradas por etapas (Hestbeck et al. 1991, Nichols et al. 1992, Brownie et al. 1993, Schwarz et al. 1993).

En esta vista general, empiezo por presentar el modelo Lincoln-Petersen (Sección 2) que conduce a los modelos más generales cerrados (Sección 3) y abiertos (Secciones 4 y 5). A continuación hay una sección sobre la combinación de los modelos cerrados y abiertos en un solo diseño (Sección 6) y, a continuación una discusión más general de las cuestiones de diseño (Sección 7). Presento una variedad de aplicaciones diferentes para los animales y los humanos (Sección 8). Demuestro como se puede conseguir la estimación del tamaño del marco de muestreo (Sección 9), que relaciona estrechamente la captura-recaptura con la teoría tradicional del muestreo. Hago referencia a la pléthora de programas de ordenador disponibles (Sección 10) y, a continuación, termino con conclusiones y sugerencias para investigación en el futuro (Sección 11).

No pretendo afirmar que esta ponencia sea un trabajo nuevo, aunque he intentado actualizar los trabajos anteriores (Pollock et al. 1990, Pollock 1991) que forman la base de esta visión general. La materia de la Sección 8 se basa principalmente en Pollock et al. (1994).

## 2. EL MODELO LINCOLN-PETERSEN

### 2.1 Estimación del Tamaño de la Población

Esta, la forma más sencilla de experimento de captura-recaptura, tiene una larga historia que se remonta a Laplace en 1786, que la usó para estimar el tamaño de la población humana de Francia (Seber 1982:104). Los primeros usos en la ecología de la población animal eran de Petersen, a mediados del siglo (LeCren 1965) y Lincoln(1930). Nuestro objetivo aquí es ofrecer una breve presentación que preparará el escenario para los modelos más complejos que vienen a continuación. Se dispone de una discusión detallada del método Lincoln-Petersen en Seber (1982:59).

Se captura, marca, y se pone en libertad una muestra de  $n_1$  animales. Más tarde se captura una muestra de  $n_2$  animales, de los cuales  $m_2$  han sido ya marcados. Por intuición, se puede derivar un estimador del tamaño de la población ( $N$ ) basado en la noción que la proporción de los animales marcados relativa al número total de animales de la segunda muestra, debe reflejar la misma proporción en la población tal que

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

que da el estimador

$$\hat{N} = \frac{n_1 n_2}{m_2}. \quad (2.1)$$

Este estimador se llama a veces el Índice de Lincoln (Seber 1982:59), aunque esto es un uso incorrecto de la palabra "índice", que normalmente se refiere a una medición de abundancia relativa (Caughley 1977:12).

Una versión modificada con menos sesgo fue dada originalmente por Chapman (1951) como

$$\text{var } \hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)(n_1 - m_2)(n_2 - m_2)}{(m_2 + 1)^2(m_2 + 2)} \quad (2.2)$$

con una estimación aproximadamente insesgada de su varianza dada por

$$\hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1 \quad (2.3)$$

de Seber (1970, 1982:60)

## 2.2 Suposiciones del Modelo

El método Lincoln-Petersen se basa en un modelo con las siguientes suposiciones:

1. La población está cerrada a las adiciones (nacimientos o inmigrantes) y a las delegaciones (muertes o emigrantes).
2. Existen posibilidades iguales de captura para todos los animales de cada muestra.
3. Las marcas no se pierden ni eluden la atención del observador.

La suposición de la población cerrada (1) puede, de hecho, debilitarse aquí. Si ocurren adiciones, entonces quedan siempre sin marcar, y el estimador Lincoln-Petersen es válido para el tamaño de la población a la hora de la segunda muestra. Si las delegaciones ocurren al azar con respecto a los animales marcados y no marcados, entonces el estimador Lincoln-Petersen es válido para el tamaño de la población en el momento de la primera muestra. Más tarde en esta ponencia, consideraremos las poblaciones completamente abiertas (adiciones y delegaciones permitidas). El estimador Lincoln-Petersen ya no es apropiado en esta forma simple; sin embargo, una versión modificada de ello es esencial en todos los modelos de población abierta estudiadas.

Es improbable que la segunda suposición, que a menudo se llama la suposición de "igualdad de capturabilidad" de los animales, sea verídica en muchas poblaciones salvajes. Es importante distinguir 2 tipos muy diferentes de alternativo.

1. La heterogeneidad—La probabilidad de captura en cualquier muestra es una propiedad del animal y puede variar entre todos los animales de la población. Esta variación en la probabilidad de captura podría deberse a muchos factores como, por ejemplo, edad, sexo, rango social, o situación de "territorio" con relación a la posición de la trampa.

2. Reacción a la trampa—La probabilidad de la captura en cualquier muestra depende de la historia anterior de capturas del animal, es decir, los animales marcados pueden tener una probabilidad de captura más baja (asustadizo con las trampas), o mayor (propenso a captura) que los animales no marcados. El tipo de reacción dependerá del método de captura empleado. Por ejemplo, es probable que las trampas cebadas conduzcan a una reacción de predisposición a captura, en que los animales marcados resultan más propensos a la captura que los no marcados. Se debe subrayar que tanto la heterogeneidad como la reacción a la trampa podrían estar presentes en una población.

Caso de estar presente la heterogeneidad, es más probable que los animales con predisposición alta de ser capturados resulten capturados en la primera muestra y recapturados en la segunda. El resultado de esto es la sobreestimación en  $m_2/m_3$  de la verdadera proporción marcada en la población, y de ahí  $\hat{N}$  de ecuación (2.1) resulta demasiado pequeño, es decir,  $\hat{N}$  es un estimador negativamente sesgado. ( Si las probabilidades de captura son heterogéneas en cada muestra pero independientes de una muestra a otra, entonces no ocurre ningún sesgo.) Si los animales son propensos a ser capturados,  $m_2/m_3$  será de nuevo demasiado grande y  $\hat{N}$  será demasiado pequeño, mientras una población de animales reacios a las trampas causarán la situación opuesta.

Resumiendo, una reacción a la trampa en probabilidades de captura puede resultar en un sesgo negativo (animales propensos a las trampas) o en un sesgo positivo (animales reacios a las trampas), en el estimador del tamaño de la población. Aunque la validez del estimador de Lincoln-Petersen depende de la ausencia de heterogeneidad y reacción a las trampas, es posible que las probabilidades de captura varíen entre los dos momentos de muestreo.

La suposición (3) es claramente importante. Si los animales pierden sus marcas,  $m_2/n_2$  será demasiado pequeño,  $\hat{N}$  será demasiado grande, y estamos de nuevo en la situación de un estimador de sesgo positivo. Si es probable que ocurra la perdida de marcas, se debe intentar estimar el índice de perdidas para que pueda realizarse una "corrección" al número de recapturas. Esta es una práctica común en los estudios de pesca, donde la pérdida de etiquetas puede ser importante. Un método consiste en el uso de un sistema de doble marcado (Caughley 1977:139, Seber 1982:94). Seber (1982:70-104) presenta un tratamiento detallado de suposiciones para el método Lincoln-Petersen, de los efectos de las desviaciones de estas suposiciones, y ensayos para la detección de tales desviaciones.

### 2.3 Ejemplo

Para terminar esta sección, presentamos un ejemplo del uso del método Lincoln-Petersen, tomado de Skalski et al. (1983). En agosto 1974, una muestra de 87 conejos de rabo blanco de Nuttal (*Sylvilagus nuttalii*) fue atrapado en vivo en Oregon central. Los animales fueron marcados en los rabos y en las patas traseras con un tinte de ácido pícrico, y puestos en libertad. El día 5 Septiembre 1974, se llevó a cabo un una batida de recuento y se contaron 14 animales de los cuales 7 estaban marcados.

Esto quiere decir que tenemos  $n_1 = 87$ ,  $n_2 = 14$ , y  $m_2 = 7$ . La estimación de Chapman (ec. 2.2) se obtiene con

$$\begin{aligned}\hat{N}_c &= \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1 \\ &= \frac{88 \times 15}{8} - 1 \\ &= 164\end{aligned}$$

con su estimación de varianza aproximada (ec. 2.3)

$$\begin{aligned}\hat{\text{var}} N_c &= \frac{(n_1 + 1)(n_2 + 1)(n_1 - m_2)(n_2 - m_2)}{(m_2 + 1)^2(m_2 + 2)} \\ &= \frac{88 \times 15 \times 80 \times 7}{8^2 \times 9} \\ &= 1,283.33.\end{aligned}$$

Suponiendo la normalidad para  $\hat{N}_c$ , se da un intervalo de confianza de aproximadamente 95% con

$$\begin{aligned}\hat{N}_c &\pm 1.96(\hat{\text{var}} N_c)^{0.5} \\ &= 164 \pm 1.96 \times (1,283.33)^{0.5} \\ &= 164 \pm 70.\end{aligned}$$

Por lo tanto, tenemos un intervalo de confianza desde 95 hasta 234, que es bastante amplio. Se podría haber aumentado la precisión del experimento si la batida de recuento hubiera visto más animales. Tengan en cuenta que se usaron 2 métodos diferentes de muestreo (captura y observación) en este estudio. De ahí, debe haber poca preocupación por la reacción a la trampa o la heterogeneidad de las probabilidades de captura.

### 3. MODELOS DE POBLACIÓN CERRADA

#### 3.1 Introducción

Se presentará aquí un análisis breve de 8 modelos para poblaciones cerradas. Como se indicó en Capítulo 1, los modelos de población cerrada suponen que no tendrán lugar nacimientos, muertes, emigración o inmigración en la población entre los períodos de muestreo. Por esta razón, se suele usar éstos modelos en experimentos que cubren períodos de tiempo relativamente cortos (por ejemplo, capturando todos los días durante 5 días consecutivos). Las historias de captura de todos los animales atrapados son los datos necesarios para obtener las estimaciones bajo éstos modelos. Estos modelos fueron considerados como conjunto primero por Pollock (1974) y desarrollados de forma más completa por Otis et al. (1978) y White et al. (1982).

Los modelos se diferencian por la manera en que se modela la probabilidad de captura. Los diferentes modelos consideran 3 fuentes de variación en las probabilidades de cap-

tura: (1) Heterogeneidad (2) Reacción a la trampa, y (3) Variación temporal o variación entre los períodos de muestreo (por ejemplo, la probabilidad de captura para día  $i$  es distinta a la de día  $j$ ). La Tabla 3.1 muestra un catálogo de los modelos disponibles. Otis et al. (1978) aportaron un programa detallado de ordenador, CAPTURE, para el análisis de datos usando los modelos de Tabla 3.1 (ver también White et al. 1978). Además de ofrecer estimadores para 5 de los modelos, CAPTURE contiene un procedimiento para la selección de modelos, que ayuda al biólogo a elegir el modelo más apropiado para un grupo concreto de datos, pero este no se porta bien cuando las probabilidades de captura son bajas y los tamaños de población son pequeños (sección 3.8). Recientemente, Rextad y Burnham (1991) han adaptado el programa, incorporando trabajos recientes. Ya hay estimadores disponibles para todos los modelos excepto  $M_{th}$ .

### **3.2 $M_o$ : El Modelo de Igualdad de Capturabilidad**

Es improbable que el modelo  $M_o$  resulte realista para las poblaciones naturales de animales puesto que supone que todos los animales de la población tienen la misma probabilidad de captura ( $p$ ) para cada periodo de muestreo durante todo el estudio. Este modelo se incluye principalmente por razones pedagógicas y para dar una base para la generalización.

Con este modelo, El Estimador de Probabilidad Máxima (MLE) de  $N$ , el tamaño de la población, puede encontrarse solo de forma iterativa. ( Esto quiere decir que no es posible dar una fórmula explícita para el MLE). El programa CAPTURE tiene la capacidad de computar este estimador y su error típico aproximado.

Se debe subrayar que este estimador puede ser muy sesgado si existe desigualdad de capturabilidad. La heterogeneidad causa un sesgo negativo, mientras la reacción a la trampa puede causar un sesgo positivo o negativo, según si los animales son asustadizos ante las trampas o propensos a captura, respectivamente. Otis et al. (1978) encontraron bastante robusto el estimador ante cambios en las probabilidades de captura con el tiempo.

### **3.3 $M_h$ : El Modelo de Heterogeneidad**

El modelo  $M_h$  permite la heterogeneidad pero no la reacción ante las trampas, y supone que cada animal tiene su propia y única probabilidad de captura ( $p_j$ ,  $j = 1, \dots, N$ ), que permanece constante durante todos los períodos de muestreo. Se supone también que los  $p_j$  son una muestra aleatoria de tamaño  $N$  de alguna distribución de probabilidad  $F(p)$ . Este modelo fue considerado por primera vez por Burnham (1972) y más tarde por Burnham y Overton (1978, 1979). El vector de las frecuencias de captura, que son los números de animales capturados, 1, 2, ...,  $k$  veces, contiene toda la información para la estimación de  $N$  con este modelo.

Existen dificultades para encontrar un estimador satisfactorio para este modelo. Al principio, Burnham intentó modelar las probabilidades de captura como provenientes de una distribución beta de 2-parámetros (Johnson y Kotz 1970:37) y usando estimación de Probabilidad Máxima, pero lo encontró poco satisfactorio.

Un estimador ingenuo de  $N$  es sencillamente el número de animales distintos observado durante todo el experimento, y es altamente sesgado al menos que las probabilidades de captura sean muy altas, o el número de muestras muy grande. Burnham aplicó a este estimador la técnica de la “navaja” para la reducción del sesgo, propuesta inicialmente por Quenouille (1956), y sacó una serie de estimadores, que se dan en Burnham y Overton (1978, 1979) y Otis et al. (1978:109). Él propuso también una técnica objetiva para elegir el estimador a usar con un grupo concreto de datos. Esta técnica ha sido programado en CAPTURE. El uso del estimador de la “navaja” para la estimación de punto y de intervalo no carece de dificultades (Otis et al. 1978:37). No obstante, este procedimiento de estimación de la navaja es probablemente el método más robusto propuesto hasta ahora para el Modelo  $M_b$  (Pollock y Otto 1983).

Últimamente, ha habido mucho interés en los estimadores alternativos para este modelo. Chao (1988) propuso un estimador de momentos para este modelo. Ella encontró que cuando la probabilidad de captura media es pequeña, de manera que la mayoría de los animales son capturados solo una o dos veces, el estimador es normalmente menos sesgado que el estimador de la navaja. Lee y Chao (1994) propusieron otro estimador. Recientemente, Norris y Pollock (1996) propusieron un método de probabilidad máxima no paramétrica, donde  $N$  es estimado además del estimador de densidad no paramétrica de  $F$ , la distribución de probabilidades de captura.

El cálculo de las estimaciones de tamaño de población por separado en categorías identificables, como la edad y el sexo, para reducir la heterogeneidad de las probabilidades de captura, es un método alternativo que a veces funciona bien. Sin embargo, y por desgracia, puede que no sea factible en las poblaciones pequeñas.

### **3.4 $M_b$ : El Modelo de la Reacción ante las Trampas**

El modelo  $M_b$  permite la reacción ante las trampas, pero no la heterogeneidad ni la variación temporal, y hace las siguientes suposiciones:

1. Todo animal no marcado de la población tiene la misma probabilidad de captura ( $p$ ) para todas las muestras.
2. Todo animal marcado de la población tiene la misma probabilidad de recaptura ( $c$ ) para todas las muestras después de haber sido capturado una vez.

De nuevo, es necesario encontrar la MLE para  $N$  numéricamente usando el programa CAPTURE.

Una característica importante de este modelo es el hecho de que los animales no contribuyen ninguna información para la estimación del tamaño de la población después de la primera captura. Por lo tanto, este modelo es equivalente al método de “eliminación” (Zippin 1956, Seber 1982:309), salvo que en este caso no se quita físicamente el animal, sino que se le considera eliminado después del marcado inicial.

Se ha usado a menudo un método de regresión lineal para estimar  $N$  en los estudios de eliminación en la literatura biológica. Este método tiene atractivo intuitivo y es también

fácil de computar. Sin embargo, siempre que se disponga de programas de ordenador (como CAPTURE), es probablemente mejor usar la MLE. En la práctica, hemos encontrado que existe normalmente poca diferencia entre los dos estimadores.

El método de regresión (Seber 1982:325) se presenta brevemente a causa de su atractivo intuitivo y porque resultará útil cuando hablamos de  $M_{hi}$  en la Sección 3.5. Se basa en:

$$\begin{aligned} E(u_i | M_i) &= p(N - M_i) \\ &= pN - pM_i \end{aligned} \quad (3.1)$$

que puede describirse matemáticamente de la siguiente manera. Dado  $M_i$  (el número acumulativo capturado y marcado antes del periodo  $i$ ), el número esperado o “medio” de animales no marcados capturados en el día  $i$  ( $u_i$ ) es una función lineal del número de animales marcados en la población. En el lenguaje de la pesca, donde más se ha aplicado esta técnica, tenemos una regresión lineal de capturas ( $u$ ) contra la captura cumulativa ( $M$ ). Los estimadores de  $N$  y  $p$  basados en la regresión son simples funciones de los estimadores de los cuadrados mínimos de la pendiente e intersección (Seber 1982:325).

### 3.5 $M_{hi}$ : El Modelo de Heterogeneidad y Reacción ante las Trampas

El modelo  $M_{hi}$  tiene en cuenta la heterogeneidad y la reacción ante las trampas, y supone que cada animal tiene sus dos propios y únicos probabilidades potenciales de captura ( $p_j$ ,  $c_j$ ,  $j = 1, \dots, N$ ) donde  $p_j$  y  $c_j$  se refieren a si el animal está marcado o no marcado, respectivamente. Se supone que estas probabilidades permanecen constantes durante todos los períodos de muestreo.

Este modelo fue considerado por primera vez por Pollock (1974) y fue desarrollado más por Otis et al. (1978:40), que dan un procedimiento de estimación llamado el “método de eliminación generalizada.” Aquí, describimos su método de manera intuitiva, mediante la generalización del método de regresión descrito en la sección anterior.

Si está funcionando la heterogeneidad (además de la reacción ante las trampas), ya no hay relación lineal entre la captura ( $u_i$ ) y la captura cumulativa ( $M_i$ ), y por lo tanto ecuación (3.1) ya no es válida. De hecho, ahora tenemos

$$E(u_i | M_i) = \bar{p}_i(N - M_i),$$

donde  $\bar{p}_i$  es la probabilidad condicional media de captura en la  $i$ -isima muestra para aquellos animales no capturados anteriormente. Supondríamos que los  $\bar{p}_i$ 's disminuyesen de forma gradual, pues los animales con las probabilidades más altas de primera captura tenderían a resultar capturados antes de aquellos con probabilidades de primera captura más bajas.

Otis et al. (1978) también sugieren que las diferencias más importantes entre los  $\bar{p}_i$ 's ocurrirán al principio del experimento. Esto sugiere que la regresión se acercaría más a ser lineal si los puntos correspondientes a los primeros períodos de muestreo se excluyesen. Un

método objetivo de hacer esto consiste en comprobar si todos los  $\bar{p}_i$ 's son iguales. Si éste es el caso, usamos el método de la eliminación (Sección 3.4), que es una regresión usando todos los puntos de muestreo. Si no es el caso, comprobamos si  $\bar{p}_2 = \bar{p}_3 = \dots = \bar{p}_k$  y si no se rechaza esta hipótesis, usamos una regresión basada en todos los puntos excepto el primero. Seguimos eliminando secuencialmente puntos hasta que no se pueda demostrar diferencias entre los  $\bar{p}_i$ 's restantes.

Tenga en cuenta que, aunque el uso de menos puntos en la regresión nos da un modelo más general, permitiendo la heterogeneidad, significa también que estamos estimando  $N$  en base de menos información, con el resultado que los errores normales para  $N$  con este modelo tienden a ser mayores que los del modelo de eliminación. También puede existir todavía un sesgo negativo substancial en los estimadores de  $N$  si la heterogeneidad es severa y el número de períodos de muestreo es pequeño. El problema de la heterogeneidad es que algunos animales pueden ser esencialmente "invisibles" debido a las probabilidades muy pequeñas de su captura, y ningún modelo puede ocuparse de esta situación extrema.

Pollock y Otto (1983) consideraron algunos estimadores nuevos para el modelo  $M_{bh}$ . Concluyen que un nuevo estimador tipo "navaja" puede resultar mejor que el estimador de eliminación generalizada anteriormente descrita. Este estimador ha sido incluido ya en CAPTURE, así que los biólogos pueden usarlo como alternativo al estimador de eliminación generalizada.

Últimamente ha habido algún estimador nuevo para este modelo, propuestos por Chao y sus colegas y basados en el método de cobertura de muestras (Lee y Chao 1994). Norris y Pollock (1996) propusieron un método de probabilidad máxima no paramétrica, para este modelo y para  $M_h$ . Estiman  $N$ , además del estimador de densidad no paramétrica de  $F$ , las probabilidades de distribución de capturas. También desarrollan un ensayo de relación de probabilidades para comparar  $M_h$  contra  $M_{bh}$  basado en la autoelevación.

### 3.6 $M$ : El Modelo Schnabel (Variación de Tiempo)

El modelo  $M$  no permite ni la heterogeneidad ni la reacción ante las trampas de las probabilidades de captura de los animales individuales. El modelo supone que todos los animales de la población tienen la misma probabilidad de captura en cada tiempo de muestreo,  $p_i$ ,  $i = 1, 2, \dots, k$ . Este es el modelo clásico de captura-recaptura para una población cerrada, con una historia que remonta hasta Schnabel (1938).

Seber (1982:130) ofrece un tratamiento detallado de este modelo. La MLE puede encontrarse numéricamente usando el programa CAPTURE. Este modelo, junto con  $M_o$ , no requiere la historia detallada de capturas. Las marcas de lote son suficientes para permitir la estimación. La MLE del modelo  $M$ , como la del modelo  $M_o$ , pueden resultar altamente sesgados si existe una capturabilidad desigual. La heterogeneidad causa un sesgo negativo, mientras la reacción ante las trampas puede causar sesgo positivo o negativo, según si los animales son asustadizos ante las trampas o propensos a la captura, respectivamente.

### **3.7 Otros Modelos Dependientes del Tiempo**

El modelo inicial de  $M_{tb}$  no estaba provisto de estimador de  $N$ , pero ha sido considerado recientemente por Chao y colegas (Lee y Chao 1994), y Rextad y Burnham (1991) incluyen también un estimador en la última versión de CAPTURE. El modelo inicial de  $M_{tb}$  no disponía de estimadores de  $N$ , pero ha sido considerado recientemente por Chao y colegas (Chao et al. 1992, Lee y Chao 1994). El modelo  $M_{tb}$  es un modelo conceptual importante, pero no dispone de estimador (Tabla 3.1).

### **3.8 Selección de Modelos**

Otis et al. (1978) y el programa CAPTURE presentan un método importante para la selección de 1 de los 8 modelos posibles ( $M_o$ ,  $M_t$ ,  $M_{tb}$ , ,  $M_{tth}$ ). Este algoritmo para la selección del modelo se basa en ensayos de bondad del ajuste y en ensayos entre modelos. Puede que sea necesario el uso frecuente del procedimiento, pero con precaución (Chapman 1980, Menkins y Anderson 1988), pues los ensayos no son independientes y a menudo tienen potencia baja, sobre todo para las poblaciones pequeñas (Menkins y Anderson 1988). De ser posible, se debe usar información biológica para reducir el número de modelos considerados en este estudio. Por ejemplo, puede existir evidencia, basada en el comportamiento del animal en cuestión, que la reacción ante las trampas es improbable con el método de captura empleado. En este caso, todos los modelos que permiten la reacción ante las trampas pueden eliminarse ( $M_b$ ,  $M_{bh}$ ,  $M_{thb}$ ). Menkins y Anderson (1988) sugieren que, en las situaciones en que la selección de modelos es pobre (es decir, probabilidades bajas de captura, población pequeña), el uso de una estimación de Lincoln-Petersen basada en aunar las muestras en una muestra de pronto contra tarde, puede resultar preferible al uso del programa CAPTURE.

### **3.9 Ejemplo**

Usamos los datos de la captura viva del campañol, *Microtus pennsylvanicus*, recogidos en el Centro de Investigación de la Fauna de Patuxent, de Laurel, Maryland, entre Junio 1981 y Enero 1982. La cuadrícula contenía una matriz de 10 x 10 estaciones de captura separadas a intervalos de 7,6-m. Una sola trampa para capturar vivo del tipo Fitch, modificado, (Rose 1973) fue colocado en cada estación. Se colocó heno y hierba seca en la sección cajonado de las trampas, y se usó maíz integral como cebo. Las trampas fueron abiertas por la tarde, comprobadas la mañana siguiente, y cerradas en posición abierta hasta la tarde, cuando se volvía a empezar el procedimiento.

Usamos aquí los datos correspondientes a 5 días consecutivos de capturas en Octubre 1981 para demostrar el rendimiento del programa CAPTURE. Los datos incluyen a todos los animales de ambos sexos  $\geq 22$  g de peso corporal. El rendimiento del procedimiento para la selección de modelos de CAPTURE para estos datos aparece en la Tabla 3.2. En este caso, se sugiere el modelo de heterogeneidad,  $M_t$ .

El rendimiento correspondiente a las estimaciones para el modelo  $M_t$  del programa CAPTURE se presenta en la Tabla 3.3. La estimación "navaja" del tamaño de la población es de 139 animales con un intervalo aproximado de confianza del 95% de 117-161 animales. La probabilidad de captura media diaria es 0.44, que explica la buena precisión de la estimación.

## 4. EL MODELO JOLLY-SEBER

### 4.1 Introducción

En muchos estudios de captura-recaptura, no es posible suponer que la población quede cerrada a las adiciones y a las posibles eliminaciones. El modelo básico para población abierta apta para esta situación es el modelo Jolly-Seber (Jolly 1965, Seber 1965), y es preciso que el lector estudie este capítulo cuidadosamente, como primera piedra para las posteriores. Cormack (1973) dio una breve descripción intuitiva de este modelo y sus estimadores, y Seber (1982:196) dio la mejor presentación detallada.

Existen programas amplias de ordenador para ayudar con los análisis para este modelo (por ejemplo, POPAN-3, Arnason y Schwarz 1987; SERGE, Lebreton et al. 1992; JOLLY, Pollock et al. 1990).

El modelo Jolly-Seber permite la estimación del tamaño de la población en cada periodo de muestreo, además de los índices de supervivencia y los números de nacimientos entre los períodos de muestreo. Sin embargo, como se ha mencionado anteriormente, no se puede separar la migración de los procesos de nacimiento y muerte sin información adicional.

### 4.2 Supuestos y Anotación

Este modelo hace los siguientes supuestos:

1. Todo animal presente en la población en el momento de la  $i$ -ésima muestra ( $i = 1, 2, \dots, k$ ) tiene la misma probabilidad de captura ( $p_i$ ).
2. Todo animal marcado presente en la población inmediatamente después de la  $i$ -ésima muestra tiene la misma probabilidad de supervivencia ( $\phi$ ) hasta el  $(i+1)$  periodo de muestreo ( $i = 1, 2, \dots, k-1$ ).
3. No se pierden las marcas, ni se dejan de notar por descuido.
4. Todas las muestras son instantáneas y cada liberación se lleva a cabo inmediatamente después de la muestra.

Las suposiciones (1), (2), y (4) eran requisitos del modelo Lincoln-Petersen. Como solamente los animales marcados se usan para estimar los índices de supervivencia, no es necesario suponer que los animales marcados y los no marcados tienen índices de supervivencia iguales. En la práctica, sin embargo, los biólogos querrán usar las estimaciones de supervivencia para referirse a la población entera, y no solamente al componente marcado. El modelo

Jolly-Seber tiene en cuenta la "pérdida al capturar", es decir, es posible que algunos animales no se devuelvan a la población después de su captura. Es implícita en estas suposiciones que toda emigración de la población es permanente.

La siguiente Anotación (Tabla 4.1) se usa en este capítulo y está basada en Seber (1982:196). Las estadísticas definidas en la Tabla 4.1 ( $m_i$ ,  $u_i$ ,  $n_i$ ,  $R_i$ ,  $r_i$ ,  $z_i$ ) son estadísticas sumarias que necesitamos calcular a partir de la información completa de la historia de capturas de cada animal capturado al menos una vez durante el estudio. La información de la historia de captura se representa normalmente mediante una serie de ceros y unos que representan no capturado o capturado, respectivamente, en una muestra concreta. Pongamos como ejemplo unas historias típicas de captura para animales en un estudio de captura-recaptura de 5 muestras (Tabla 4.2). En la Tabla 4.2, animal 1 fue capturado en la primera muestra, recapturado en la segunda muestra, y no se le volvió a ver nunca. Animal 2 fue capturado en la primera muestra y no volvió a aparecer. Animal 3 fue capturado en la primera muestra y luego recapturado en la cuarta muestra. Animal 4 fue capturado por primera vez en la muestra 2 y luego recapturado en las muestras 3, 4, y 5. Consideramos muy aconsejable esta clase de entrada en vez de las estadísticas sumarias. Permite que se lleven a cabo más análisis alternativos, y permite, también, que se lleven a cabo ensayos de bondad del ajuste más detallados.

### 4.3 Estimación de Parámetros

Aquí se dará una discusión intuitiva de la estimación de parámetros. Primero, imagínese que  $M_i$ , el número de animales marcados en la población justo antes de la muestra  $i$ , es conocido por todos los valores  $I = 2, \dots, k$  (no hay animales marcados en el momento de la primera muestra, así que  $M_1 = 0$ ). Más tarde, hablaremos de como estimar los  $M_i$ , que siempre son desconocidos en una población abierta.

Un estimador intuitivo de  $N_i$ , el tamaño de la población en muestra  $i$ , es el estimador de Lincoln-Petersen tratado en el Capítulo 2. Si las suposiciones del modelo son válidas, entonces la proporción de animales marcados de la muestra debe igualar la de la población,

$$\frac{m_i}{n_i} \approx \frac{M_i}{N_i}.$$

Esto da el siguiente estimador para  $N_i$ .

$$\hat{N}_i = \frac{n_i M_i}{m_i}, \quad (4.1)$$

donde  $m_i$  y  $n_i$  son los números marcados y los números totales de animales capturados en la muestra  $i$ , respectivamente.

El estimador del índice de supervivencia se obtiene al considerar primero el número de animales marcados en la población inmediatamente después de la muestra  $i$ , que se da mediante  $(M_i - m_i + R_i)$ . El  $(M_i - m_i)$  de esta expresión representa a los animales marcados no capturados en  $i$ , y liberados, con marcas, de nuevo a la población. A continuación, considera-

mos el número de animales en este grupo que siguen vivos en la población justo antes de la muestra ( $i+1$ ),  $M_{i+1}$ . Entonces, un estimador del índice de supervivencia natural es la relación de estas dos cantidades:

$$\hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + R_i}. \quad (4.2)$$

Un estimador claro del reclutamiento entre  $i$  y  $(i+1)$  es

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\phi}_i(\hat{N}_i - n_i + R_i). \quad (4.3)$$

Este es simplemente la diferencia estimada entre el tamaño de la población en la muestra ( $i+1$ ) (que es  $\hat{N}_{i+1}$ ) y el número esperado de supervivientes de  $i$  a  $i+1$  (que es  $\hat{\phi}_i(\hat{N}_i - n_i + R_i)$ ).

La probabilidad de captura,  $p_i$ , puede estimarse como la proporción de animales marcados, o total de animales (marcados + no marcados) vivos en  $i$  que son capturados en  $i$ .

$$\hat{p}_i = \frac{m_i}{\hat{M}_i} = \frac{n_i}{\hat{N}_i}. \quad (4.4)$$

Para completar este esquema intuitivo, necesitamos un estimador de los  $M_i$ , puesto que están claramente desconocidos en una población abierta. Esto puede obtenerse poniendo en ecuación las dos razones

$$\frac{z_i}{M_i - m_i} \cong \frac{r_i}{R_i}$$

que son los futuros índices de recuperación de los dos grupos distintos de animales marcados de la población en el periodo de muestreo  $i$ : (1)  $(M_i - m_i)$ , los animales marcados no observados en  $i$ , y (2)  $R_i$ , son los animales observados en  $i$ , marcados, y entonces liberados para su posible recaptura. Nótese que  $z_i$  y  $r_i$  son los miembros de  $(M_i - m_i)$  y  $R_i$ , respectivamente, que son capturados de nuevo al menos una vez. Por lo tanto, el estimador de  $M_i$  se da con

$$\hat{M}_i = m_i + \frac{R_i z_i}{r_i}. \quad (4.5)$$

y se define para  $i = 2, \dots, k-1$ , porque necesitamos animales observados antes y después de cada  $i$ . Se deduce que  $N_i$  en ecuación (4.1) se define para  $i = 2, \dots, k-1$ ;  $\hat{\phi}_i$  en ecuación (4.2) para  $i = 1, \dots, k-2$ ;  $\hat{B}_i$  en ecuación (4.3) para  $i = 2, \dots, k-2$ ; y  $\hat{p}_i$  en ecuación (4.4) para  $i = 2, \dots, k-1$ .

Dos de estos estimadores ( $\hat{\phi}_i, \hat{p}_i$ ) son estimadores de máxima probabilidad y los cinco ( $\hat{\phi}_i, \hat{p}_i, N_i, \hat{B}_i, \hat{M}_i$ ) son intuitivamente razonables. Sin embargo, estos son estimadores sesgados,

y los estimadores aproximadamente no sesgados fueron recomendados por Seber (1982:204; para  $(\hat{\phi}_i, \hat{N}_i, \hat{B}_i, \hat{M}_i)$ ) y Jolly (1982:304; para  $\hat{p}_i$ ). Son parecidos al estimador de Chapman usado en la ecuación 2.2. Se dispone también de estimadores de varianza de muestras grandes, pero no se presentan aquí (Pollock et al. 1990: Sección 4.4). Tampoco presentamos ensayos de bondad del ajuste basados en suficientes estadísticas. Ver Pollock et al. (1990:Sección 4.5).

#### 4.4 Supuestos y Robustez de Modelos

Los efectos de la heterogeneidad de la probabilidad de captura sobre los estimadores de Jolly-Seber han sido bien estudiados. Si ciertos individuos tienen más probabilidad de ser capturados que otros, y si estas probabilidades distintas de captura persisten durante todo el experimento de captura-recaptura, entonces los individuos marcados de la población tendrán tendencia a ser los de las probabilidades más altas de captura.

Debido a esta diferencia en las probabilidades medias de captura entre los animales marcados y los no marcados, la proporción de animales marcados de cualquier muestra ( $m/n$ ) tenderá a sobreestimar la verdadera proporción de animales marcados de la población ( $M/m$ ). Esto conduce a un sesgo negativo en  $\hat{N}_i$  (ecuación 4.1). Además, hemos notado que el estimador  $\hat{M}_i$  (ecuación 4.5) se derive intuitivamente poniendo en ecuación las proporciones de dos grupos de animales marcados de la población justo después del periodo  $i$  (los observados en  $i$ ,  $R_i$ , y los no observados en  $i$ ,  $M_i - m_i$ ) que son capturados en algún periodo posterior ( $>i$ ). La heterogeneidad puede resultar en una probabilidad de captura media algo mayor para aquellos animales capturados en  $i$  ( $R_i$ ) que para los animales marcados no capturados entonces ( $M_i - m_i$ ). Esto resultó en un sesgo negativo relativamente pequeño en  $\hat{M}_i$  (Carothers 1973).  $\hat{M}_i$  sesgados contribuyen no solo al sesgo de  $\hat{N}_i$ , sino también resultan en  $\hat{\phi}_i$  sesgado (ecuación 4.2).

Carothers (1973) y Gilbert (1973) usaron la simulación por ordenador y las expresiones analíticas para aproximar el sesgo de los estimadores de Jolly-Seber resultantes de las probabilidades de captura heterogénea. Estos autores estudiaron  $\hat{N}_i$  y su pareja con sesgo ajustado  $\tilde{N}_i$  (ecuación 4.7). Los sesgos de los dos estimadores de tamaño de población resultantes de la heterogeneidad tendían a ser negativos y eran a menudo muy severos. Se encontró que la magnitud de los sesgos dependía, en gran medida, de la probabilidad media de captura y del grado de variación en la probabilidad de captura entre individuos (por ejemplo, como indica el coeficiente de variación de la distribución de la probabilidad de captura, Carothers 1973). Las probabilidades de capturas heterogéneas tienen poca importancia en términos relativos cuando la probabilidad de captura es alta, tal que Gilbert (1973:524) sugirió que "no hace falta que un experimentador intente diseñar un experimento para que todos los animales tengan la misma probabilidad de captura, sino solo para que la distribución sea trasladada más cerca de uno de manera que casi todos los animales tengan probabilidades de captura de, digamos,  $\geq 0.50$ ."

Naturalmente, esta sugerencia es de factibilidad limitada. Carothers (1973) también approximó los sesgos de  $\hat{\phi}_i$  y  $\tilde{\phi}_i$  y encontró que mostraban solamente un sesgo relativo pequeño como respuesta a las probabilidades de captura heterogénea. En un ejemplo interesante, Carothers (1979) encontró gran evidencia de probabilidades de observación heterogénea en un estudio de captura-reobservación de fulmares (*Fulmarus glacialis*). Mediante la simulación por ordenador, logró aproximar el grado de heterogeneidad reflejado por su estadística de ensayo. A continuación, approximó el sesgo relativo de su estimación del índice de superviven-

cía, y encontró que era insignificante (0.01) para la mayoría de los años (Carothers 1979). Este ejemplo demuestra bien la robustez del estimador del índice de supervivencia de Jolly-Seber para las probabilidades de captura heterogénea. Finalmente, notamos que el sesgo en los estimadores del índice de supervivencia resultantes de la heterogeneidad de  $p_i$ , es normalmente negativo, pero puede, a veces, resultar positivo (Nichols y Pollock 1983).

La reacción temporal ante las trampas en la probabilidad de captura se refiere a la situación en que un animal muestra una probabilidad de captura diferente para 1 o más períodos después de la captura inicial. Pollock (1975) desarrolló un ensayo para este fenómeno, además de un grupo de modelos y estimadores apropiados para esta situación. La reacción permanente ante las trampas en la probabilidad de captura se refiere a una situación en que el cambio en la probabilidad de captura con la captura inicial persiste durante el resto del experimento de captura-recaptura. Es importante considerar los efectos de la reacción permanente ante las trampas en los estimadores Jolly-Seber, porque es difícil de detectar y porque no se dispone de modelos que incorporan este tipo de reacción ante las trampas. El estimador Jolly-Seber,  $\hat{M}_i$ , depende solamente de los animales marcados (ecuación 4.5) y, por lo tanto, no le afecta la reacción permanente ante las trampas. El estimador del índice de supervivencia,  $\hat{\phi}_i$ , se basa también en los animales marcados (ver ecuación 4.2) y no le afecta la reacción permanente ante las trampas.

La exactitud del estimador queda afectada, sin embargo, y una reacción de propensión a la captura aumenta la exactitud, mientras una reacción asustadiza ante las trampas reduce la exactitud. El estimador del tamaño de la población (ecuación 4.1), sin embargo, incluye la proporción de animales marcados en la muestra,  $m/n$ , que se supone estimará la verdadera proporción marcada de la población,  $M/N$ . Con una reacción de propensión a la captura, los animales muestran probabilidades de captura más altas después de la captura inicial y  $m/n$  sobreestimará  $M/N$ , resultando en  $\hat{N}_i$  de sesgo negativo (Nichols et al. 1984a). Con una reacción asustadiza ante las trampas, los animales muestran probabilidades de captura más bajas después de la captura inicial. El resultado es la subestimación por  $m/n$  del verdadero  $M/N$ , produciendo así un sesgo positivo en  $\hat{N}_i$ . Nichols et al. (1984a) ofrecen aproximaciones de las magnitudes de los sesgos resultantes. Los sesgos relativos son más severos cuando la población muestra una rotación importante (supervivencia baja y reclutamiento alto) y también cuando la proporción marcada de la población es pequeña. Notamos también que una reacción de propensión a la captura aumenta la exactitud de la estimación del tamaño de la población, mientras una reacción asustadiza ante las trampas reduce la exactitud.

Si el marcado reduce el índice de supervivencia del animal, puede resultar en un sesgo negativo grave de los estimadores del índice de supervivencia. Algunos métodos de marcar peces padecen mucho de este problema (Ricker 1958). Los estimadores de tamaño de la población siguen sin sesgo. Las reducciones temporales del índice de supervivencia asociadas con el marcado inicial pueden detectarse y luego estimarse usando los modelos y los ensayos y estimadores asociados de Robson (1969), Pollock (1975), y Brownie y Robson (1983).

Las probabilidades de la heterogeneidad de supervivencia puede detectarse, a veces, usando el ensayo de la bondad del ajuste de Pollock et al. (1985). Sin embargo, no se dispone de modelos (y estimadores asociados) que incorporan tal heterogeneidad. Por lo tanto, los efectos de la supervivencia heterogénea en los estimadores Jolly-Seber son de interés.

Cormack (1972:340) consideró este problema de manera intuitiva y declaró: "puesto que el parámetro a estimar es una probabilidad de supervivencia media, la variación de supervivencia entre individuos no afectará a las estimaciones Jolly-Seber, siempre que la variación no se correlacione con el marcado, directamente o indirectamente por edad u otra consideración." Encontramos que esto es cierto solo si la probabilidad de supervivencia de un animal durante un periodo es independiente de su probabilidad de supervivencia en otro periodo, y si su probabilidad de supervivencia es independiente de su probabilidad de captura (Pollock y Ravelling (1982). La naturaleza del sesgo en los estimadores de supervivencia ha sido estudiado por Pollock y Ravelling (1982) y Nichols et al. (1982) en el contexto del modelo de la recuperación de bandas, que tiene sus propios problemas especiales. En la situación captura-recaptura, nadie ha llevado a cabo un estudio detallado todavía. Es probable que un caso común sea los índices de supervivencia heterogénea, que son relacionados positivamente para el mismo animal (es decir, si la supervivencia es alta para un animal durante un año, es probable que lo sea también en el año siguiente), pero que son independientes de la probabilidad de captura. En este caso, hay un sesgo positivo en los estimadores del índice de supervivencia y también en los estimadores del tamaño de la población. No se sabe los tamaños probables de estos sesgos en las poblaciones naturales, pero sospechamos que serán pequeños.

Las diferencias por edad en las probabilidades de supervivencia pueden considerarse como un tipo especial de heterogeneidad. Manly (1970) estudió los efectos de las diferencias por edad en la probabilidad de supervivencia en los estimadores Jolly-Seber, y examinó a varios modelos distintos de especificación de edad. Por ejemplo, cuando las probabilidades de supervivencia son más bajas para los animales jóvenes (una situación corriente en las poblaciones naturales), se ha encontrado que los resultantes  $\hat{N}$ , y  $\hat{\phi}$ , de Jolly-Seber muestran un sesgo positivo (Manly 1970:17). En las situaciones en que se puede comprobar la edad del animal en el momento del marcado, el modelo de Pollock (1981b) puede usarse para ensayar para diferencias por edad en las probabilidades de supervivencia y captura, y para estimar los parámetros por edad cuando se detectan tales diferencias (ver también Stokes 1984). Sin embargo, no siempre es posible determinar la edad (por ejemplo, para muchos aves transeúntes bandados durante el invierno), y los resultados de Manly (1970) son de interés especial en tales situaciones. Manly (1970) sacó la conclusión: "Los resultados sugieren que este método (Jolly-Seber) puede emplearse razonablemente siempre que los índices de mortalidad no resulten muy afectados por la edad."

La suposición de la nula pérdida de marcas es también muy importante, pues si los animales pierden sus marcas, el número de recapturas será demasiado bajo. El resultado es la subestimación de los índices de supervivencia, pero no influye en las estimaciones de los tamaños de población (Arnason y Mills 1981). Seber (1982:94-96) ofrece un método para estimar, y compensar, la pérdida de marcas usando un sistema de doble marcado. Es importante tener en cuenta que la pérdida de marcas causará una disminución en la precisión de los estimadores, incluso cuando se puede estimar la pérdida y ajustar los estimadores (Pollock 1981b).

Es una suposición implícita que cualquier emigración del estudio es permanente. Balser (1981) derivó un ensayo para la emigración temporal. Es probable que ocurra la emigración temporal en varias situaciones prácticas sobre el terreno, y es nuestra opinión que el sesgo inducido podría resultar serio. Esto se discute más en la sección 6.

#### 4.5 Ejemplo

En el ejemplo 1 analizamos los datos de un estudio de 2 años sobre las ardillas grises (*Sciurus carolinensis*), llevado a cabo por A. Duboek en la zona boscosa de robles maduros de la Estación de Investigación del Bosque Alice Holt, en Surrey, Inglaterra. Las ardillas fueron capturados a intervalos aproximados de un mes desde Noviembre 1972 hasta Septiembre 1974. Se colocaron trampas para capturas múltiples, cebadas con grano, en el suelo por toda la zona, y disfrazadas con hojas. Las ardillas fueron marcados utilizando un método único de corte de dedos.

En la Tabla 4.3, se dan las estadísticas básicas para las estimaciones Jolly-Seber, y en la Tabla 4.4 se dan las estimaciones de parámetros y los errores típicos aproximados. Nótese que no se presentan estimaciones para los períodos 12-14, porque opinamos que resultarían engañosas debido a los números pequeños de capturas (2, 1, y 4). Nótese también que, a veces, las estimaciones de supervivencia son de  $>1$  y han sido anotadas como 1. De la misma manera, algunas estimaciones del número de nacimientos son negativas y han sido anotadas como cero.

Este es un estudio exacto, en general, con errores típicos relativamente pequeños para la mayor parte de los períodos de muestreo. La exactitud es el resultado de las probabilidades altas de captura y supervivencia. Una vez marcado una ardilla, esta permanece en la población durante un periodo largo de tiempo, y aporta información cada vez que resulte capturada. Es cierto que la precisión de las estimaciones varía considerablemente. Esto se debe, principalmente, a los cambios en las probabilidades de captura. Se debe subrayar, sin embargo, que incluso en los estudios en que las probabilidades de captura son constantes, la precisión no será constante y los períodos de muestreo hacia la mitad del estudio rendirán estimaciones más exactas. La población marcada aumenta poco a poco, y quedan todavía muchos períodos de muestreo para obtener recapturas porque el variable aleatorio  $r_i$  necesita ser grande para una buena precisión.

En términos de la validez de las estimaciones, hay muchos factores a considerar. El biólogo creía que la migración era insignificante y, por lo tanto, interpretó las estimaciones de supervivencia y nacimiento como representativas de los verdaderos índices de supervivencia y nacimiento. Es probable que existen probabilidades de heterogeneidad de captura debido a edad, sexo, y otros factores, además de una reacción propensa a la captura en las probabilidades de captura. Las dos desviaciones tienden a ocasionar la subestimación del tamaño de la población y, en menor medida, la subestimación de los índices de supervivencia.

Los animales jóvenes deberían haber estado uniéndose a la población capturable en Abril y Mayo. Esto se percibe en estimaciones altas de los números de nacimientos en 1974, pero no en 1973. El biólogo había previsto esto, pues 1973 fue un año muy malo para la reproducción de las ardillas. Una vez que todos los animales jóvenes se habían unido a la población en Mayo de 1974, hubo una gran bajada en la estimación de supervivencia y en la estimación del tamaño de la población. Desgraciadamente, esto ocurrió justo al final del estudio, y hubo tan pocas capturas en Septiembre 1974 que las estimaciones son muy inexactas.

#### **4.6 Restricciones y Extensiones de los Modelos**

A veces, es posible poner restricciones en los parámetros del modelo Jolly-Seber y, así, simplificar el modelo y aumentar la exactitud de las estimaciones de parámetros restantes.

Para algunas poblaciones en que se sabe que los nacimientos y las inmigraciones son insignificantes, un modelo que solo tiene en cuenta las pérdidas (muertes y emigración) es razonable. Este modelo fue desarrollado inicialmente por Darroch (1959). En escasas ocasiones, puede existir la situación en que no existen muertes ni emigración para una población, por lo que un modelo que solo contempla el reclutamiento (nacimientos y inmigración) es razonable. Este modelo fue desarrollado inicialmente por Darroch (1959). (Seber 1982:217; Pollock et al. 1990:34).

Otros modelos restringidos importantes son los en que las probabilidades de captura y/o supervivencia son constantes en el tiempo. (Jolly 1982, Brownie et al. 1986, Pollock et al. 1990) Los modelos con estas restricciones pueden aportar grandes aumentos en exactitud sobre el resto los estimadores de parámetros. Este método ha sido ampliado por Lebreton et al. (1992) solamente para la estimación de supervivencia, y se habla más de él en la sección 5.1.

A veces, es necesario generalizar al modelo Jolly-Seber, y consideramos algunas de las generalizaciones importantes a continuación en la sección 5. Una de las que no, trata de los modelos de reacción temporal ante las trampas de Robson (1969), Pollock (1975), y Brownie y Robson (1983). Se considera que el marcado tiene un efecto a corto plazo sobre las probabilidades de supervivencia y captura. Los estimadores siguen teniendo una forma intuitiva parecida a los estimadores originales de Jolly-Seber, y disponemos de ensayos de tablas de contingencia para estos efectos a corto plazo. La precisión de los estimadores es menor.

### **5. OTROS MODELOS ABIERTOS**

#### **5.1 Modelado de Supervivencia**

Lebreton et al. (1992) han hecho una contribución muy importante al modelado de captura-recaptura para poblaciones abiertas. Debido a las limitaciones de espacio, no podemos tratar adecuadamente el contenido total de su monografía, pero intentamos resumir los puntos más importantes.

En el modelo Jolly-Seber es posible obtener estimaciones de probabilidad de supervivencia y captura directamente al modelar las historias de captura de los animales marcados en función de las primeras capturas. Esto se remonta al modelo de Cormack (1964). Entonces, es posible, usando una generalización del estimador de Lincoln-Petersen, estimar los tamaños de población usando marcados en razón con no marcados, y combinar más a todos estos estimadores para obtener estimadores de número de nacimientos. Lebreton et al. (1992) subrayan la importancia de estimar los índices de supervivencia usando los índices de captura como parámetros de molestia, y no consideran el problema del tamaño de la población ni el de la estimación del número de nacimientos. Las probabilidades que ellos consideran son depen-

dientes de las primeras capturas y son productos de distribuciones multinominales. La estimación es sencilla usando los métodos de probabilidad máxima.

Lebreton et al. (1992) subrayan que su "trabajo sintetiza, usando un marco común, los desarrollos recientes junto con otros nuevos, con énfasis especial en la flexibilidad en el modelado, en la selección de modelos, y en el análisis de series de datos múltiples. Los efectos sobre los índices de supervivencia y captura de tiempo, edad, y variables categóricos que caracterizan los individuos (por ejemplo, el sexo) pueden tenerse en cuenta, además de las interacciones entre tales efectos."

Esta filosofía del "análisis de varianza" subraya la estructura de los procesos de supervivencia y captura, más que las características técnicas de algún modelo concreto. La serie flexible de modelos abarcada en este síntesis usa una Anotación común. Como resultado del gran nivel de flexibilidad y pertinencia alcanzado, el foco cambia de encajar con un modelo determinado a la construcción y selección de modelos."

Recomiendan el siguiente procedimiento general de modelado: "(1) empezar a partir de un modelo global compatible con la biología de la especie estudiada y con el diseño del estudio, y evaluar su ajuste; (2) seleccionar un modelo más parsimonioso, usando Los Criterios de Información de Akaike para limitar el número de ensayos formales; (3) Ensayar para las cuestiones biológicas más importantes comparando este modelo con modelos vecinos usando ensayos de razón de probabilidad; y (4) Obtener estimaciones de probabilidad máxima de parámetros de modelo con estimaciones de precisión." Hacen hincapié en la importancia de software para ordenador y sugieren RELEASE (Burnham et al. (1987) para resumen de datos y ensayos de bondad del ajuste, y SURGE para el ajuste de modelo iterativo y la computación de los ensayos de razón de probabilidad.

Lebreton et al. (1992) desarrollan su método general cuidadosamente y con detalle, y hacen que sea fácil de seguir, ilustrando la metodología con cinco ejemplos amplios sobre una gama de especies.

Skalski et al. (1994) han desarrollado un paquete para ordenador SURPH que también hace modelado de supervivencia. Su filosofía se parece al del trabajo de Lebreton et al. (1992) pero hay diferencias técnicas que no se discutirán aquí.

## 5.2 Modelos Estructurados por Etapas

Otra esfera de modelado de captura-recaptura abierta, de gran importancia y crecimiento rápido, es para las poblaciones estructuradas por etapas.

Estos modelos, que empezaron con el trabajo de Arnason (1972, 1973) son para poblaciones biológicas donde los datos de liberación y recaptura son disponibles para cada uno de varios estratos en cada ocasión de captura. Los estratos pueden ser, por ejemplo, situaciones geográficas de estados fisiológicos. El movimiento de animales entre estratos ocurre con probabilidades desconocidas, y la estimación de estas probabilidades desconocidas de transición es el objetivo.

Subrayamos que la pertenencia a un estrato no se conoce para cualquier periodo en que el animal no este capturado. Un buen ejemplo sería la categoría de peso, donde un animal puede o no cambiar de categoría de peso en el siguiente periodo y la única manera segura de averiguarlo es capturar al animal. Sin embargo, es posible estimar la probabilidad de transición de una categoría de peso a otra. Notamos que esto es diferente y más complejo que si el estrato fuera una categoría de edad, cuando la transición a la siguiente categoría de edad es automática. El modelado de Lebreton et al. (1992) tratado en la Sección 5.1 puede manejar estratos como la categoría de edad, en que la transición es seguro, pero no puede manejar estratos como categorías de peso, en que la transición es estocástico.

Ha habido mucho trabajo reciente en esta esfera. Schwarz et al. (1993) presentan el modelo Arnason-Schwarz, y los estimadores de probabilidad máxima para ello. Ilustran con el uso de datos de marcado-regreso de arenques. El modelo es fundamentalmente igual que el de Arenson (1972,1973) pero el procedimiento de estimación es mejor.

Hestbeck et al. (1991) consideraron una serie de datos múltiples de captura-recaptura sobre ocas Canadienses bandadas en las tres zonas del paso de aves Atlántico de América del Norte. Describen la importancia biológica de estimar las probabilidades de transición para el movimiento entre las tres zonas. Empezaron con el modelo Arnason-Schwarz, y entonces generalizaron hacia un modelo en que las probabilidades de transición dependen de la posición en el periodo anterior (modelo Markoviano).

Nichols et al. condujeron modelos parecidos con una serie de datos sobre campañoles donde los estratos son categorías de masa corporal.

Brownie et al. (1993) consideran el modelo Arnason-Schwarz, el modelo Markoviano, y entonces generalizan hasta los modelos no Markovianos en que las probabilidades de transición dependen de más que la simple posición en el periodo anterior (es decir, de las anteriores también). James Nichols, que participó en los trabajos Hestbeck et al. (1991), Nichols et al. (1992) y Brownie et al. (1993), ha supervisado a James Hines en la construcción de un programa de ordenador para estos modelos, usando el programa SURVIV (White 1983). Se refieren a su programa como MSSURVIV.

Probablemente, se hará mucho más trabajo en estos modelos durante los próximos años. Nichols et al. (1992) sostienen que el diseño robusto podría ser útil aquí. El disponer de una pequeña muestra de animales marcados con radio podría también resultar útil si se pudiera desarrollar los modelos combinados apropiados.

### 5.3 Estudios Comparativos de Captura-Recaptura

Burnham et al. (1987) presentan métodos de diseño y análisis para categoría grande de experimentos de supervivencia comparativa, basándose en la liberación-captura de poblaciones marcadas. Los métodos presentados son aplicables a cualquier experimento relacionado con el tratamiento y control de grupos de animales marcados. Una aplicación primaria se dirigió a cuestiones de supervivencia de peces en el Río Columbia en el Noroeste de Los Estados Unidos de América. Los tratamientos se ocupaban de los peces que pasaban a través de las turbinas hidroeléctricas y los controles trataban de los peces que fueron añadidos río

abajo de las turbinas. Los ensayos en busca de diferencias entre los grupos de tratamiento y los de control fue objetivo primario.

Existen ya muchas aplicaciones más de esta metodología. Por ejemplo, Deuel (1985) informó sobre un estudio que midió la mortalidad indirecta de ánades del norte, causado por plomo ingerido como resultado de disparos, y Burnham et al. (1987) aplicaron sus métodos a éstos datos.

Stromberg et al. (1988) usaron estorninos como sustituto de especie nativo para estudiar la supervivencia de los aves expuestos a un pesticida organofosfático, y aplicaron modelos de este tipo.

El software de ordenador es muy importante aquí en los análisis complejos. RELE-ASE, que se basó en el programa SURVIV (White 1983) fue desarrollado por Gary White (Burnham et al. 1987) para ocuparse de este tipo de diseño comparativo.

En esta selección, solamente hemos podido dar un breve resumen del material que Burnham et al. (1987) desarrollan en forma de libro. Su trabajo es muy importante, y merece el estudio detallado por parte del lector más dedicado de este artículo.

## 6. COMBINACION DE MODELOS CERRADOS Y ABIERTOS DE POBLACION

### 6.1 Introducción

Creemos que el diseño de los experimentos de captura-recaptura merece mucho más atención de los estadísticos y biólogos. El diseño de los estudios debe orientarse alrededor de la satisfacción de tantas suposiciones de modelo como sea prácticamente posible, para que un modelo sencillo y razonablemente eficiente pueda usarse para la estimación.

En la base de muchos modelos de muestreo de captura-recaptura está la suposición que existe la misma probabilidad de captura en cada muestra para todos los animales (la suposición de igualdad de capturabilidad), que discutimos en la sección 2. Es muy poco probable que esta suposición sea realista en las poblaciones naturales de animales, y puede estar actuando o 1 o ambas de las alternativas de heterogeneidad y reacción ante las trampas en una población concreta de animales.

El modelo de captura-recaptura más frecuentemente usado por los biólogos para las poblaciones abiertas en los estudios a largo plazo, es el modelo Jolly-Seber, que se comentó de forma detallada en el capítulo 4.

Este modelo requiere la suposición de igualdad de capturabilidad, y aunque algún debilitamiento de la suposición es posible, es probable que la complejidad de los modelos de población abierta excluyan a los modelos generales que permiten la heterogeneidad y/o la reacción permanente ante las trampas. Sin embargo, algunos de los modelos cerrados tratados en la Sección 3 permiten el relajamiento de la suposición de igualdad de capturabilidad, de una manera biológicamente realista.

Pollock (1981a) afirmó que los estadísticos han hecho una distinción marcada entre los modelos cerrados y abiertos de población, que es, quizás, bastante artificial. Para citar a

(Pollock 1981a:435), "En la práctica, se puede llevar a cabo una serie de estudios a corto plazo. Un método de análisis consistiría en analizar cada estudio de corto plazo, usando los modelos de población cerrada ... que permiten igualdad de capturabilidad. Luego, se podría reunir todos los períodos de muestreo de cada estudio de corto plazo, y podrían estimarse los estimadores de supervivencia entre estos estudios de corto plazo usando el Modelo Jolly-Seber. Este procedimiento permite la estimación de tamaño de población con modelos que permiten desigualdad de capturabilidad, mientras la estimación de supervivencia, que no resulta tan afectada por la desigualdad de capturabilidad, está bajo el Modelo Jolly-Seber."

Ahora describimos el diseño que Pollock (1982) publicó inicialmente, y que fue motivado por el deseo de encontrar un diseño para los estudios a largo plazo, que es robusto a la heterogeneidad y/o la reacción ante las trampas. Examinamos su robustez y consideramos algunas extensiones. Pollock et al. (1990) ofrecen un tratamiento más detallado.

## 6.2 El Diseño Básico

Considere la siguiente presentación de un experimento de muestreo de captura-recaptura, donde tenemos  $k$  períodos primarios de muestreo (por ejemplo, años), y dentro de cada uno de éstos tenemos  $l$  días consecutivos de apresamiento).

No es necesario que el número de períodos secundarios sea igual en todos los períodos primarios.

El diseño puede usarse para estimar el tamaño de la población para cada uno de los períodos primarios ( $N_1, N_2, \dots, N_k$ ) suponiendo que la población es constante a lo largo de los períodos secundarios de muestreo dentro de cada período primario de muestreo. Podemos estimar también los índices de supervivencia ( $\phi_1, \phi_2, \dots, \phi_{k-2}$ ) y los números de individuos nuevos entrando en la población entre los períodos primarios de muestreo ( $B_1, B_2, \dots, B_{k-2}$ ).

Sería posible estimar todos estos parámetros excepto  $N_1$  y  $N_k$  usando el modelo standard de Jolly-Seber de la Sección 4. Para usar el modelo Jolly-Seber con el diseño descrito, todos los períodos secundarios de muestreo dentro de un período primario serían "reunidos." Con "reunidos" en este caso, queremos decir que solamente nos interesa saber si se captura o no un animal al menos una vez durante el período primario de muestreo.

Al usar el modelo Jolly-Seber, es importante tener en cuenta las suposiciones requeridas y la robustez del modelo. En la Sección 4, se señaló que la heterogeneidad y/o la reacción ante las trampas tendrán un gran efecto sobre los estimadores del tamaño de la población (ecuación 4.1) porque la razón de muestras ( $m_i / n_i$ ) ya no reflejará con exactitud la razón de población ( $M_i / N_i$ ). Los estimadores de la población marcada (ecuación 4.5) no resultarán tan afectados por la desigualdad de capturabilidad porque las dos razones de recuperaciones de animales marcados tenderán a sufrir influencias parecidas. De ahí se deduce que los estimadores del índice de supervivencia (ecuación 4.2), que son simplemente razones de los  $M_i$ , resultarán también menos afectados por la desigualdad de capturabilidad que los estimadores de tamaño de la población. Cormack (1972) sugirió esto usando un argumento intuitivo, y Carothers (1973) lo documentó para las probabilidades de heterogeneidad de captura usando aproximaciones analíticas y simulación por ordenador. En el caso de reacción permanente

ante las trampas, las estimaciones de los índices de supervivencia no están afectados en absoluto. (Nichols et al. 1984a).

En el diseño robusto, intentamos minimizar la influencia de la desigualdad de capturabilidad sobre nuestros estimadores mediante la explotación de nuestros dos niveles de muestreo. Los estimadores del índice de supervivencia, que no se influyen tanto por la desigualdad de capturabilidad, se estimarán exactamente como se hace con el procedimiento Jolley-Seber, usando capturas reunidas dentro de un periodo primario. Recuerde que se dispondrán de todos estos estimadores solamente para  $i = 1, 2, \dots, k, -2$ .

Sugerimos que los estimadores de tamaño de población para cada periodo primario de muestreo ( $\hat{N}_1, \dots, \hat{N}_k$ ) se obtengan usando modelos de población cerrada que tengan en cuenta la desigualdad de capturabilidad (Sección 3). Estas estimaciones solamente usan las capturas y recapturas dentro de un periodo primario de muestreo. La manera más fácil para que un biólogo obtenga estas capturas consiste en el uso del programa CAPTURE desarrollado por Otis et al. (1978) (también White et al. 1982). Este programa considera una gama de modelos diferentes teniendo en cuenta la heterogeneidad y/o reacción ante las trampas y ofrece, también, un método objetivo para seleccionar el modelo a usar más apropiado. Nótese que una ventaja adicional de este procedimiento consiste en la disponibilidad de estimadores de tamaño de población para todos los periodos primarios ( $i = 1, \dots, k$ ), mientras con el procedimiento de Jolly-Seber, solo se dispone de estimadores para  $i = 2, \dots, k-1$ . De nuevo, sin embargo, sugerimos prudencia en el uso del procedimiento para la selección de modelos de CAPTURE (Chapman 1980, Menkins y Anderson 1988). Los ensayos no son independientes y tienen potencia baja, sobre todo para las poblaciones pequeñas (Menkins y Anderson 1988).

Finalmente, los estimadores de los índices de nacimiento pueden obtenerse de la ecuación (4.3) como antes, pero ya los estimadores de tamaño de población usados son los descritos en el párrafo anterior. Nótese que es posible estimar  $B_1, \dots, B_{k-2}$ , mientras con Jolly-Seber solo se puede estimar  $B_1, \dots, B_{k-2}$ .

Si tuviésemos un estimador de  $\phi_{k-1}$ , entonces  $B_{k-1}$  sería estimable. En algunos experimentos, puede ser razonable estimar  $\phi_{k-1}$  por  $\phi_{k-2}$  o quizás por un promedio de  $\phi_1, \dots, \phi_{k-2}$ .

### 6.3 Extensiones

Pollock y Mann (1983) y Pollock et al. (1990) señalaron que este diseño podría resultar útil también en los estudios de poblaciones estratificadas por edad. Los índices de supervivencia por edad pueden obtenerse de los datos de captura-recaptura a lo largo de los periodos primarios de muestreo, mientras los estimadores de tamaño de población pueden obtenerse para cada categoría de edad, de los periodos secundarios de muestreo. Demuestran también que ya es posible separar el  $\hat{B}$  de Jolly-Seber en sus dos componentes de producción dentro de la población, y animales que han inmigrado a la población.

William Kendall, en su tesis de Ph.D. y en publicaciones relacionadas (Kendall y Pollock 1992, Kendall et al. 1995), ha estudiado el diseño robusto con gran detalle. Él consideró la estimación ML cuando no está presente la heterogeneidad. Llevó a cabo algunos estu-

dios amplios de simulación que demostraron el valor del diseño robusto sobre modelos abiertos regulares, como el modelo Jolly-Seber.

Recientemente, se ha trabajado en el estudio de la emigración temporal usando el diseño robusto (Kendall et al. 1995). La emigración temporal es una violación de suposición que resulta muy difícil de tratar con el modelo Jolly-Seber. Sin embargo, usando el diseño robusto, es posible conseguir estimaciones sin sesgo de tamaños de población y índices de supervivencia.

## 7. DISEÑO DE ESTUDIOS DE CAPTURA-RECAPTURA

Aunque el diseño del estudio es la primera tarea que el biólogo tiene que considerar, hemos dejado este tema hasta que se hayan presentado todos los modelos. Creemos que el científico debe ser consciente de los diferentes modelos y sus suposiciones antes de poder diseñar con seriedad un estudio de captura-recaptura.

En este capítulo, presentamos primero algunos comentarios sobre la selección de un método de estimación para el estudio de las poblaciones de animales. A continuación, ajustamos para enfocar en los estudios de captura-recaptura y consideramos cuestiones de cómo evitar, o al menos minimizar, las violaciones de suposiciones cuando diseñamos un estudio. A esto le sigue una presentación sobre la amplitud que requiere un estudio, en términos de la intensidad del muestreo, para lograr los niveles dados de precisión de los estimadores de parámetros.

### 7.1 Selección de un Método de Estimación

Una de las decisiones más importantes que un científico tiene que tomar al empezar una investigación implica la selección de los métodos más apropiados para la consecución de los objetivos del estudio. Como ya hemos comentado, los modelos abiertos de captura-recaptura pueden usarse para estimar el tamaño de población, el índice de supervivencia, y el número de reclutas nuevas en la población. Estas cantidades pueden estimarse usando métodos que no sean los de captura-recaptura. Presentamos aquí una breve discusión general de algunos de los otros métodos, y de las fuerzas y debilidades de captura-recaptura en comparación con éstos métodos.

Existen varios métodos alternativos para estimar el tamaño y la densidad (tamaño de población por unidad de área) de poblaciones de animales de libre distribución. Algunos métodos suponen que todos los animales situados dentro de los cuadros de muestra se detectan y observan con una probabilidad 1.0. En tales casos, se usan las técnicas tradicionales de muestreo (por ejemplo, Cochran 1977) para "ampliar" estas cuentas para estimar el tamaño total de la población en un área de interés (Eberhardt 1978, Seber 1982). En la situaciones donde se observan y detectan con facilidad a los animales con una probabilidad de 1.0, es casi seguro que tales cuentas de muestra ofrecerán el mejor método de estimar el tamaño de la población.

En la mayor parte de las situaciones que implican observaciones de los animales sobre el terreno, sin embargo, no es realista suponer una probabilidad de detección de 1.0. En estas situaciones hay que estimar la probabilidad de detección y luego usarla conjuntamente con los datos de cuentas para estimar el tamaño o la densidad de la población. Existen una variedad de métodos para estimar la probabilidad de detección. Estos incluyen la comparación de cuentas estándar y cuentas intensivas de áreas de muestra, el uso de subpoblación de animales marcados, la aplicación con múltiples observadores, muestreo formal por líneas transversales, y cuentas múltiples sobre los mismos áreas (Seber 1982, Pollock y Kendall 1987, Buckland et al. 1993).

El método de estimar la probabilidad de detección que ha recibido el desarrollo estadístico más detallado, es el muestreo por líneas transversales (Buckland et al. 1993). En el muestreo por líneas transversales, el observador viaja a lo largo de la línea transversal, contando los animales. Se supone que todos los animales localizados directamente en la línea son detectados con una probabilidad de 1.0, y se supone que la probabilidad de detección disminuye de acuerdo con la distancia de la línea. Se usan, a continuación, las distancias perpendiculares o los ángulos y distancias emparejados de observación, para estimar la función de detección, que expresa la probabilidad de detección como función de la distancia desde la línea transversal. Entonces, se usa la función de detección estimada, con el número total de animales observados, para estimar la densidad. Creemos que las limitaciones principales de los modelos de muestreo por líneas transversales son (1) la suposición que todos los animales en la línea serán observados, y (2) la factibilidad de registrar las distancias. Estas limitaciones son especialmente severas en los estudios aéreos.

Cuando los animales son fácilmente observables, los métodos para la estimación del tamaño de la población basados en la observación resultarán, a menudo, preferibles a los métodos de captura-recaptura. Los trabajadores sobre el terreno normalmente han reconocido esto, y han tendido a usar métodos basados en la observación para la estimación del tamaño de la población para aves y mamíferos grandes. Sin embargo, no resulta tan fácil observar a otros vertebrados (mamíferos pequeños, reptiles, anfibios, y peces), ni a los invertebrados, y los métodos de captura-recaptura ofrecen el medio más razonable de estimar el tamaño de la población para muchos de estos organismos.

Los modelos de captura-recaptura y de recuperación de bandas se usan corrientemente para estimar las probabilidades de supervivencia y de mortalidad para las poblaciones de animales. Los modelos de recuperación de banda (Brownie et al. 1985) representan casos especiales de modelos generales de captura-recaptura, en que todas las recapturas implican la eliminación de la población de los animales etiquetados (por ejemplo, la eliminación normalmente ocurre mediante la pesca, la caza, la captura en trampas, u otra forma de explotación humana). Generalmente, estos datos no se usan para estimar el tamaño o el reclutamiento de la población, aunque existen excepciones en los áreas de aplicaciones especiales (Brownie y Pollock 1985, Nichols et al. 1986). La relación entre los modelos de recuperación de bandas y los de captura-recaptura se discute en Brownie y Pollock (1985) y Brownie et al. (1985).

A pesar de la relación conceptual estrecha entre sí, los modelos de recuperación de bandas y los de captura-recaptura, se aplican a menudo en tipos muy diferentes de situaciones de muestreo sobre el terreno, y, por lo tanto, producen con frecuencia estimaciones que llevan interpretaciones diferentes. Los modelos de recuperación de bandas usan frecuentemente los datos aportados por miembros del público general implicados en el proceso de la explotación

(por ejemplo, los cazadores registran las bandas de los aves migratorios abatidos; los pescadores registran las etiquetas de los peces capturados). De esta manera, la muestra de animales recuperados se recoge de un área geográfico grande, que a menudo cubre el área total de distribución de la población estudiada. Los modelos de captura-recaptura, por otra parte, utilizan típicamente las muestras aportadas por los esfuerzos de un solo investigador o equipo y, por lo tanto, quedan frecuentemente restringidos a áreas locales de estudio. Como resultado, el complemento de una estimación del índice de supervivencia de un estudio de captura-recaptura ( $1 - \hat{\phi}_i$ ) normalmente incluye la mortalidad y la emigración (el no regreso de un animal al área expuesto a los esfuerzos de captura), mientras el complemento de las estimaciones de supervivencia de muchos estudios de recuperación de bandas (por ejemplo, muchos estudios de los aves migratorios cazados) incluye solo la mortalidad. En cuanto a la selección de un método de estimación, si el parámetro de interés es el índice estricto de mortalidad, y si resulta probable que ocurra la emigración permanente y que esta sea difícil de estimar, entonces, los modelos de recuperación de bandas pueden ofrecer el medio preferido de estimación. Los índices de recuperación de etiquetas son bajos en muchas aplicaciones de los modelos de recuperación de bandas, produciendo una baja precisión en las estimaciones de los índices de supervivencia resultantes. Por lo tanto, en las situaciones en que la emigración no constituye un problema, o puede estimarse, se preferirán a menudo los estimadores de supervivencia de captura-recaptura a los estimadores de recuperación de bandas, debido a su mayor precisión.

En los casos en que las recapturas locales y las recuperaciones amplias por explotación ocurren para los animales marcados en un estudio concreto, la distinción entre las estimaciones de supervivencia de la recuperación de bandas y de la captura-recaptura, permite inferencia sobre la emigración permanente (Anderson y Sterling 1974, Hepp et al. 1987). Concretamente, la razón de la estimación de supervivencia de captura-recaptura con la estimación de la supervivencia de la recuperación de bandas debe estimar la probabilidad condicional de que un animal vivo en el momento de los esfuerzos de muestreo de recaptura está en el área local expuesto a éstos esfuerzos. El complemento de esta probabilidad condicional puede considerarse como una estimación de la probabilidad de emigración permanente.

Históricamente, los modelos de recuperación de bandas y de captura-recaptura han sido los dos métodos generales más importantes para la estimación de los índices de supervivencia y mortalidad en las poblaciones de animales de libre distribución. Recientemente, los investigadores sobre el terreno han empezado a usar otra técnica, basada en la radiotelemetría, para estimar los índices de mortalidad (Trent y Rongstad 1974, Heisy y Fuller 1985). Pollock et al. (1989a,b) señalan la relación entre la estimación de la mortalidad basada en la telemetría y la estimación de la supervivencia en la medicina y la ingeniería, y a continuación, explotan esta relación para aportar modelos no paramétricos generales y flexibles para la estimación de la mortalidad de los animales, que son mucho más realistas que los modelos anteriores.

Los métodos de radiotelemetría dan buenas estimaciones de las probabilidades de supervivencia si se sigue a un número razonable de animales, y opinamos que se usará este método con una frecuencia cada vez mayor en el futuro. La telemetría ofrece ventajas sobre la captura-recaptura porque a menudo da ideas sobre las fuentes de la mortalidad (y a veces incluso aporta estimaciones de los riesgos de mortalidad de origen específico) e información sobre la emigración. El uso de la radiotelemetría es relativamente costoso, sin embargo, tanto en el coste del equipo como en el importante esfuerzo sobre el terreno necesario para obtener las posiciones frecuentes de grandes números de animales. Además, es posible que exista a

veces una carencia de independencia entre las pérdidas de contacto (es decir, el no lograr volver a localizar a un animal portador de transmisor) y el estado del animal (muerto o vivo), y semejante dependencia podría causar problemas de estimación. En ciertos casos, el transmisor puede también reducir la supervivencia del animal. No obstante, la radiotelemetría ofrece un buen medio para estimar la mortalidad en muchas situaciones y tendrá cada vez más importancia en el futuro.

Hay una variedad de procedimientos disponibles para estimar el índice de reproducción (podemos definir esto como el número de crías producido por hembra en edad de criar) y el número de nuevas reclutas uniéndose a una población. Sin embargo, solo el procedimiento de captura-recaptura es muy general; en lugar de este, se diseñan métodos a medida para las situaciones específicas sobre el terreno. De usarse métodos alternativos (es decir, no usando modelos de captura-recaptura) para estimar el tamaño de la población y el índice de supervivencia, entonces puede estimarse el número de reclutas nuevos mediante la resta al igual que para  $\hat{B}_i$ , de captura-recaptura. Para las poblaciones explotadas, es posible a veces estimar una razón de edad (crías por adulto o hembra adulta en una época concreta del año), un indicador del índice de reproducción, usando relaciones cosecha edad y datos de recuperación de bandas de edad específica (Martin et al. 1979).

Para los animales que resultan fáciles de observar, es posible, a veces, estimar una relación de edades usando las observaciones directas sobre el terreno.

El índice de reproducción y los números de animales jóvenes se estiman frecuentemente combinando las estimaciones de los distintos componentes del índice de reproducción. Por ejemplo, para muchas especies de aves, el número de huevos de una puesta típica ha sido bien estimado. También, hay un número cada vez mayor de estudios que aportan buenas estimaciones de éxito de nidificación (la probabilidad de que un nido recién establecido tenga el "éxito" de producir al menos 1 ave volantón). Cuando se combinan estas estimaciones con las de la proporción de hembras adultas criando, y una estimación de la supervivencia pos-volantón, se puede desarrollar un estimador para el índice de reproducción del número de crías. Sin embargo, las estimaciones de estas últimas cantidades (proporción criando y supervivencia pos-volantón) pueden resultar difíciles de obtener.

Cuando una cuestión biológica de interés implica un componente especial de índice de reproducción, generalmente será mejor intentar estimar este componente directamente, y no hacer inferencias indirectas usando  $\hat{B}_i$ , captura-recaptura. Cuando se necesitan estimaciones del total de reclutas nuevas,  $\hat{B}_i$ , captura-recaptura puede frecuentemente ser el único estimador disponible. Se debe recordar, sin embargo, que  $\hat{B}_i$  se estima usando estimaciones del tamaño de la población y del índice de supervivencia, y por lo tanto, tiene una varianza de muestreo grande. El  $\hat{B}_i$  de captura-recaptura estándar incluye también a los animales que entran en la población vía la inmigración y la reproducción en situ. Sin embargo, los modelos dependientes de la edad y el diseño robusto (Sección 6) permiten la estimación separada de estos dos componentes del reclutamiento, en ciertas situaciones.

## 7.2 Minimización de las Violaciones de Suposiciones

Cuando se selecciona la captura-recaptura para usarla en una investigación concreta, es importante minimizar las violaciones de las suposiciones del modelo en que se basan los

estimadores. Nosotros creemos que el problema de las violaciones de las suposiciones debería considerarse en 3 etapas diferentes del estudio de captura-recaptura. Primero, se debe considerar las posibles violaciones al desarrollar el diseño para el muestreo directo del estudio. Segundo, cuando se analizan los datos, es importante usar ensayos de bondad de ajuste de modelo general, ensayos de entre modelos, y ensayos de suposiciones específicas para ayudar a decidir cual modelo es el más apropiado para los datos y para comprobar si las suposiciones específicas de los modelos selectivos han sido violadas. Tercero, si se concluye que se han violado suposiciones específicas, entonces es importante usar información de estudios sobre los efectos de las violaciones de las suposiciones, al interpretar las estimaciones y usarlas para sacar inferencias. Por ejemplo, si sacamos la conclusión que existe heterogeneidad de las probabilidades de captura en nuestros datos, puede que todavía sea posible el uso de las estimaciones de supervivencia de Jolly-Seber resultantes, pues estas tienden a mostrar un sesgo relativamente pequeño bajo la heterogeneidad (Carothers 1973, 1979). Por otra parte, si la heterogeneidad ocurre y las probabilidades de captura no son altas ( $<0.5$ ), entonces no podríamos fiarnos de nuestras estimaciones del tamaño de la población de Jolly-Seber, aunque se sabría que probablemente subestiman el tamaño verdadero de la población. En algunos casos, es incluso posible "corregir" los estimadores de captura-recaptura para violaciones específicas (Carothers 1979, Arnason y Mills 1981). Los ensayos sobre las suposiciones y los efectos de las violaciones de las suposiciones en las estimaciones de captura-recaptura han sido discutidos en otro lugar. Aquí, y de manera breve, discutimos aspectos del diseño de estudios, que están relacionados con las suposiciones de modelos y su violación.

Es probable que exista la heterogeneidad de probabilidades de captura y supervivencia hasta cierto punto en todos los estudios de captura-recaptura. Aunque no hay diseño de estudio que pueda eliminar completamente tal heterogeneidad, hay ciertas maneras de reducirla. Cuando se sospecha que las probabilidades de supervivencia o captura están asociadas con alguna característica de los animales individuales (por ejemplo, edad, sexo, tamaño), entonces, estas características deberían registrarse para poder estratificar la muestra para su análisis. Cuando los animales no pueden cambiar de estrato durante el curso del estudio (por ejemplo, el sexo, y también la edad y el tamaño en los estudios suficientemente cortos), el procedimiento apropiado simplemente implica análisis separados para los estratos distintos. Cuando los animales cambian de estrato durante el curso del estudio, como en el caso de la edad, entonces hay que desarrollar modelos especiales para incorporar tal heterogeneidad.

La heterogeneidad de la probabilidades de captura puede ocurrir como resultado de intensidades de muestreo heterogéneas; por lo tanto, se debe hacer lo posible para asegurar que todas las secciones del área del estudio se muestrean con intensidad igual. Cuando se usan trampas, redes finas, u otros aparatos de captura estacionarios, es importante que todos los animales del área muestreado sean expuestos a los aparatos. Por ejemplo, incluso en las cuadrículas de captura donde se colocan las trampas a intervalos iguales en forma de tablero de damas, es importante que la distancia entre trampas sea inferior al radio medio de alcance desde casa del animal a estudiar. Si se colocan las trampas demasiado lejos entre sí, entonces la probabilidad de captura puede variar substancialmente como función de la posición del radio de alcance desde casa del animal. Es importante tener al menos una trampa, y preferiblemente más, dentro de este radio de alcance de cada animal. Si no se disponen de bastantes trampas para un muestreo tan intenso, entonces pueden trasladarse las trampas en el curso de cada periodo de muestreo para asegurar una cobertura adecuada.

En un experimento de 2 muestras de Lincoln-Petersen (Sección 2), un buen método para eliminar sesgo causado por heterogeneidad consiste en el uso de 2 métodos diferentes de muestreo. Desgraciadamente, los modelos para poblaciones abiertas requieren al menos 3 períodos de muestreo, y a menudo tienen muchos más; por lo tanto, no es factible usar métodos diferentes en cada muestra.

Como se notó en la sección anterior, creemos que el diseño robusto representa un procedimiento muy útil para ocuparse de las probabilidades de capturas heterogéneas. En este momento, creemos que la implementación de este diseño es el mejor medio disponible para minimizar los problemas de estimación asociados con tal heterogeneidad.

La reacción ante las trampas en las probabilidades de captura o de supervivencia es otro problema que ocurre corrientemente en los estudios de captura-recaptura. Evidencia de reacción ante las trampas en las probabilidades de captura ha sido encontrado en varios estudios de vertebrados (Otis et al. 1978:29, Tanaka 1980:42-52, Nichols et al. 1981:123, Seber 1982:83-85). Las reacciones propensas a la captura se asocian a menudo con el uso de trampas cebadas. Aunque se podría eliminar esta reacción al no poner ningún cebo, esta es, normalmente, una alternativa pobre debido a las grandes pérdidas de datos (se capturan pocos animales) que resultan. Se cree que el pre-cebar, la colocación de trampas cebadas y cerradas en posición abierta dentro de la zona de estudio, antes del comienzo de capturas, reduce la propensión a la captura en ciertas situaciones, pero puede conducir a otros problemas (Tanaka 1980:52-55). Al considerar una reacción de propensión a la captura, es importante recordar que los estimadores del índice de supervivencia no resultan sesgados por la reacción ante las trampas y que una reacción propensa a la captura resulta en una mayor precisión para el estimador del índice de supervivencia. El asustarse ante las trampas es también una reacción común y puede minimizarse usando técnicas de captura y manejo que causan un mínimo de tensión nerviosa a los animales. La minimización de la tensión nerviosa en la captura y manipulación ayuda también a proteger contra una reacción ante las trampas en el índice de supervivencia, en que los animales muestran probabilidades de supervivencia bajas para el periodo inmediatamente después de la captura inicial.

A veces, el uso exclusivo de datos de reobservación (sin captura) puede resultar posible, como por ejemplo en el caso de aves marcadas con colores, y este método evita el problema de la reacción ante las trampas. Típicamente, los estudios que usan este diseño solamente han obtenido estimadores del índice de supervivencia (Cormack 1964), pero a veces es posible conseguir razones de animales marcados con animales no marcados en el mismo punto y estimar el tamaño de la población también. Este diseño tiene buena precisión si es posible obtener altas probabilidades de reobservación.

Una desventaja posible es la opinión de algunos biólogos que los animales marcados con colores pueden tener índices de supervivencia más bajos que los animales no marcados (Burley 1985).

Los cambios temporales (por ejemplo, con una duración de 1 o quizás 2 períodos después del marcado) en la probabilidad de captura y supervivencia pueden tratarse usando los modelos generales de Robson (1969), Pollock (1975), y Brownie y Robson (1983), y, por lo tanto, no requieren ninguna consideración especial de diseño.

No se puede tratar la reacción permanente ante las trampas en la probabilidad de captura mediante el uso de los modelos abiertos disponibles, pero el diseño robusto de la Sección

6 permite ensayar para tal reacción además de la estimación de parámetros de interés sin sesgo cuando ocurre.

Una consideración práctica importante al diseñar un estudio es la de usar un buen método de marcado. El método de marcado debe ser tal que no influye en la supervivencia del animal, pero, al mismo tiempo, la perdida de marcas debe ser insignificante. Si se está usando un método de marcado nuevo, sería aconsejable llevar a cabo un pequeño estudio piloto para investigar su factibilidad. Si no se puede evitar completamente la pérdida de marcas, esta debe estimarse mediante técnicas como la del doble marcado (Sección 2). Las estimaciones de la pérdida de etiquetas pueden entonces usarse para corregir las estimaciones de modelos abiertos, usando métodos como los de Arnason y Mills (1981).

La mayor parte de los modelos de captura-recaptura no pueden distinguir entre nacimientos e inmigración, ni entre muertes y emigración. Como se notó en la Sección 6, el diseño robusto permite la separación de reclutas procedentes del nacimiento y de la inmigración, en algunas situaciones. En algunos casos, la migración puede excluirse por razones biológicas mediante la elección cuidadosa del área de estudio para tener en cuenta las fronteras naturales. Este es el caso para el estudio sobre la ardilla gris descrito en la Sección 4.5. Debe considerarse, también, el problema de la emigración temporal. Esto puede ser especialmente severo en los estudios de pequeños mamíferos, y sugiere el uso de cuadrículas grandes, fronteras naturales, y incluso vallas "semipermeables", para minimizar el problema. En los casos en que la migración tiene el potencial de causar grandes problemas, puede merecer la pena acompañar el trabajo de captura-recaptura con un estudio de radiotelemetría diseñado para estimar la magnitud de la emigración permanente y la emigración temporal. También, puede usarse el diseño robusto si se sospecha la presencia de emigración temporal.

Creemos que, en algunos casos, el biólogo pueda tener la oportunidad de elegir una población donde no ocurren nacimientos ni inmigración. En tales casos, el modelo que solo permite pérdidas (Sección 4.6) puede ser usado. Tiene la ventaja de dar mayor precisión a los estimadores de tamaño de población y del índice de supervivencia. Creemos que es menos probable que sea útil el modelo que solamente permite ganancias (nacimientos o inmigrantes).

Finalmente, subrayamos que la simulación puede ser una herramienta útil, en la fase de diseño, para evaluar las violaciones de las suposiciones. POPAN-3 (Arnason y Schwarz 1987) tiene un componente de simulación potente y flexible, como parte del programa.

### 7.3 Intensidad de Muestreo

En esta sección consideramos la cuestión de que grado de intensidad de muestreo se requiere para alcanzar un nivel específico de precisión para todas las estimaciones de parámetros. Usaremos la cantidad, coeficiente de variación (CV), o error estándar proporcional como nuestra medición de precisión. Definimos

$$CV(\hat{\theta}) = \frac{SE(\hat{\theta})}{\theta},$$

que es simplemente el error estándar del estimador dividido entre , el mismo parámetro. La reducción de la CV indica un aumento en la precisión para el estudio. Una aproximación general es que un estudio que aporta una CV de 20% es razonable. Muchos estudios de fauna y pesca no alcanzarán este nivel debido a las limitaciones de presupuesto. Naturalmente, la precisión de nuestro estimador dependerá de los valores de los distintos parámetros de nuestro modelo, y del modelo bajo consideración.

Otis et al. (1978) consideraron de forma detallada la intensidad de muestreo para las poblaciones cerradas (Sección 3). Pollock et al. (1990) consideraron de forma detallada las intensidades de muestreo para el modelo Jolly-Seber, y sus simplificaciones y generalizaciones (Sección 4). Burnham et al. (1987) consideraron la intensidad de muestreo necesaria para buena potencia para las comparaciones de los índices de supervivencia. Creemos que la simulación será usada frecuentemente, de forma rutinaria, para comprobar de antemano la factibilidad de varios diseños con niveles deseables de precisiones. El programa POPAN-3 (Arnason y Schwarz 1987) puede usarse con este fin para el modelo Jolly-Seber, pero puede que sea necesario escribir otros programas para hacer más fácil la simulación para otros modelos.

## 8. APLICACIONES

### 8.1 Poblaciones de animales

Ha aparecido en la literatura una amplia variedad de aplicaciones biológicas, más allá de los estudios de marcado y recaptura sobre las poblaciones de peces y fauna. Los métodos estándar de marcado y recaptura han sido aplicados a las poblaciones de insectos, y Wileyto et al. (1994) han desarrollado lo que ellos llaman el método Markov-recaptura para el muestreo de las poblaciones de insectos. Hay dos tipos de trampa: una trampa de marcado continuo y una trampa de traslado continuo. Un estimador sencillo, relacionado con el estimador Lincoln-Petersen, resulta si se recogen los datos solamente una vez. Wileyto (1994) y Wileyto (1995) han extendido el método a marcados múltiples y períodos múltiples de tiempo, respectivamente.

Se ha usado el modelo Lincoln-Petersen para estimar el número de animales no observados en los reconocimientos aéreos (Henny y Anderson 1979; Pollock y Kendall 1987). Buckland et al. (1993), Borchers (manuscrito sin publicar), y Alpizar (1994) han extendido estos métodos para formar un modelo combinado de líneas transversales y captura-recaptura.

Un modelo modificado de Lincoln-Petersen ha sido aplicado a las situaciones en que los excrementos (heces) se marcan por radioactividad mediante la inyección de radioisótopos en el animal. Este ha sido aplicado a los osos negros (Pelton y Marcom 1977), mapaches (Conner y Labisky 1987) y nutrias (Shirley et al. 1988).

Los métodos de captura-recaptura han sido aplicados para el registro de fósiles. Nichols y Pollock (1983) demostraron como usar el modelo Jolly-Seber para estimar el número de especies, índices de evolución de las especies, y índices de extinción.

## **8.2 Poblaciones Humanas y Otras Aplicaciones**

Muchos autores han aplicado los métodos de captura-recaptura a la estimación de la parte del censo no contada (ver Feinberg (1992) para bibliografía completa). Se han empleado los métodos de captura-recaptura para estimar el número de personas sin hogar de una ciudad (Cowan et al. 1986, Laska y Meisner (1993)). Greene (1983) ha usado el método para estimar los parámetros demográficos sobre las poblaciones criminales. Witter (1974) y Witter et al. (1974) han usado Captura-recaptura para estimar los números de personas con enfermedades, de los hospitales y de otras listas (ver también Feinberg (1972) y Bishop et al. (1975)). El muestreo de las poblaciones humanas elusivas usando muestreo por universos, muestreo por redes, y muestreo por captura-recaptura, fue discutido por Sudman et al. (1988). Autores como Nayak (1988, 1991) Goudie (1990) y Chao et al. (1993) han aplicado métodos cerrados de captura-recaptura de muestreo continuo a la estimación del número de errores en un programa de ordenador.

Una aplicación muy relacionada es la estimación de tamaño de estructura (establecimiento), considerada por Pollock et al. (1994) y que consideramos en la próxima sección.

## **9. ESTIMACIÓN DEL TAMAÑO DE ESTRUCTURAS**

### **9.1 Introducción**

En la teoría clásica del muestreo, se supone que existe una estructura completa. Existe, al menos de forma conceptual, una lista completa de unidades de población. Entonces, es posible sacar una muestra de probabilidad de la población. Los estimadores de los parámetros de la población, como la media o el total, tienen, entonces, propiedades conocidas y resultan fáciles de estudiar teóricamente o numéricamente. Cochran (1978) y otros libros sobre la teoría del muestreo, se concentran en esta situación y dan propiedades de estimadores para diseños comunes de muestreo como el muestreo aleatorio sencillo, el muestreo aleatorio estratificado, y el muestreo multi-etapa (o por universos).

En la práctica, para muchos estudios, puede que no exista una estructura completa. Por ejemplo, en las encuestas sobre establecimientos o negocios, las listas mantenidas por asociaciones profesionales o agencias gubernamentales son, frecuentemente, incompletas. Un procedimiento posible para resolver este problema usando los métodos de captura-recaptura, ha sido propuesto por Pollock et al. (1994) y este artículo forma la base para esta sección.

Hay varias soluciones posibles para el problema de una estructura de listas incompletas. La primera, y la más obvia consiste en establecer una lista completa, y este es el procedimiento seguido por muchas agencias nacionales de estadística, a pesar de su carestía. El segundo procedimiento consiste en cambiar a un método de estructuras por áreas y enumerar todos los establecimientos en los áreas muestreados.

Esto puede resultar ineficiente y costoso si las áreas son grandes y los negocios escasos. Otro procedimiento ampliamente empleado para el muestreo de granjas consiste en el uso del método multiestructural desarrollado inicialmente por Hartley (1962, 1974) y ahora desarrollado en la granja del Servicio Nacional de Estadísticas Agrarias, más una estructura de

áreas donde todas las encuestas sobre granjas se hacen en los E.E.U.U. (Kott y Vogel 1995). Estas encuestas que usan una estructura de listas incompletas de granjas dentro de una unidad de muestra están enumeradas. Por lo tanto, una estructura de áreas es conceptualmente completa, porque dentro de cada área muestreado, todas las granjas tienen que estar enumeradas y, de hecho, se podría elegir probabilísticamente a cualquier área. En lo que queda de esta sección, consideramos otra solución basada en el uso de varios estructuras de listas incompletas, y en la estimación del tamaño de la estructura usando modelos de captura-recaptura. Al contrario de las demás soluciones, este procedimiento se basa fuertemente en modelos.

## 9.2 Estimación de la Población Total

El estimador normal de la población total para un solo muestreo aleatorio sin reposición es

$$\hat{Y} = N \bar{y} \quad (9.1)$$

cuando se conoce  $N$ , el tamaño de la estructura, y  $\bar{y}$  es la media de la muestra. Ver, por ejemplo, Cochran (1978, p.21). La varianza de  $\hat{Y}$  es

$$\text{Var}(\hat{Y}) = N^2 \text{Var}(\bar{y}) \quad (9.2)$$

cuando

$$\text{Var}(\bar{y}) = \frac{s^2}{n} \left( \frac{N-n}{N} \right),$$

$s^2$  es la varianza de población y  $(N-n)/N$  se llama el factor de corrección de la población finita. El estimador (9.1) no tiene sesgo.

Si, en lugar de esto, tuviéramos que estimar  $\hat{N}$  a partir de un modelo de captura-recaptura, tendríamos

$$\hat{Y} = \hat{N} \bar{y} \quad (9.3)$$

y ahora tanto  $\hat{N}$  como  $\bar{y}$  son variables aleatorias, al contrario de (9.1) donde  $N$  se sabía exactamente. La varianza estimada de  $\hat{Y}$ , es

$$\begin{aligned} \text{Var}(\hat{Y}) &= (\hat{N})^2 \text{Var}(\bar{y}) + (\bar{y})^2 \text{Var}(\hat{N}) \\ &\quad - \text{Var}(\hat{N}) \text{Var}(\bar{y}). \end{aligned} \quad (9.4)$$

Esta varianza será mayor que la de (9.2) porque  $\hat{N}$  tiene, también, una varianza debida a su estimación. El estimador (9.3) solamente resultará sin sesgo si  $\hat{N}$  y  $\bar{y}$  son estimadores sin sesgo.

Es posible derivar resultados como (9.3) y (9.4) para otros parámetros y otros diseños de muestreo.

### **9.3 Uso de Modelos de Captura-Recaptura en la Encuesta Pelágica Grande**

La encuesta Pelágica grande es una encuesta de pescadores llevada a cabo por el Servicio Nacional de Pesca Marina de los E.E.U.U., usando una combinación de teléfono y diseño de acceso. Se telefona a una muestra de propietarios de barcos de pesca, de una lista, para obtener información sobre el esfuerzo de pesca (es decir, el número de viajes de pesca durante un periodo). Se obtiene información sobre las capturas por unidad de esfuerzo (es decir, capturas por viaje) de una segunda muestra de propietarios de barcos, en puntos de acceso al final de sus viajes de pesca. Se combina la información de las dos encuestas para estimar el esfuerzo total y la captura total de especies importantes, como el Atún de Aleta Azul.

Un problema de esta encuesta es que la lista de armadores usada en la encuesta telefónica es muy incompleta. Por lo tanto, la teoría clásica del muestreo, que supone una estructura completa de tamaño conocido  $N$ , no es adecuada y hay que modificarla. El método actual de estimar el tamaño de la estructura de la lista de barcos de pesca implica la combinación de dos listas, (una lista telefónica con una lista de muelles de puerto) y el uso del modelo Lincoln-Petersen.

#### **El Modelo Lincoln-Petersen**

##### **Estimación del Tamaño de la Estructura ( $N$ )**

Con el método actual, los barcos "marcados" ( $M$ ) son los que están en la lista maestra, que se derive principalmente de las entrevistas telefónicas anteriores. La muestra de recaptura se lleva a cabo en los surtidores de carburantes de los muelles, y se comprueba el número total de barcos interceptados ( $n$ ), con el fin de averiguar cuales están "marcados" ( $m$ ) (es decir, en la lista maestra inicial). A continuación, se puede usar el estimador Chapman (Sección 2.1) para dar un estimador del tamaño de la estructura( $N$ ). Consideremos, ahora, las suposiciones de este modelo y el efecto que las violaciones podrían tener sobre el sesgo del estimador de  $N$ .

#### **Cierre**

Es probable que esta suposición resulte violada. Puede haber barcos de pesca en la lista maestra que ya no participan en la pesca (pérdidas). Barcos de pesca nuevos pueden unirse a la pesca estando esta ya en curso (ganancias). Lo ideal sería la obtención de una estimación de tamaño de estructura, por separado, para cada periodo de tiempo de dos semanas. La ventaja de usar el modelo cerrado de Lincoln-Petersen está en su simplicidad y en su factibilidad. Los sesgos del estimador, debidos a la falta de clausura, podrían ser positivos o negativos.

Actualmente, no se sabe que probabilidad hay de cambio del tamaño de la flota pesquera, durante la temporada de pesca. Un diseño de muestreo de captura-recaptura múltiple

permitiría el uso del modelo Jolly-Seber, para estimar el tamaño de la flota durante cada periodo. El estudio de éstos estimadores y de los estimadores del índice de supervivencia y del número de reclutas, nos permitirá evaluar la validez de la suposición de cierre.

### Igualdad de Capturabilidad

La violación de la suposición de igualdad de capturabilidad puede deberse, o bien a probabilidad de la heterogeneidad inherente de captura entre individuos, o a la "reacción ante las trampas", cuando los individuos que están marcados tienen probabilidades de captura mayores o menores que los individuos no marcados. En cualquier de las dos situaciones, cuando los individuos en las listas son barcos de pesca, creemos que existe una potencial para las probabilidades de heterogeneidad de captura entre los barcos de pesca. Si la heterogeneidad está operando a través de las dos muestras, los individuos "cogidos" en la primera lista tenderán a ser los de probabilidades altas de captura y, por lo tanto, es más probable que sean "cogidos" de nuevo en la segunda lista. Esto significa que la proporción marcada en la segunda muestra (lista) será excesivamente alta, y el estimador de  $N$  tendrá un sesgo negativo. Nótese que este argumento intuitivo deja clara que el problema no es la heterogeneidad en sí, sino la correlación positiva de probabilidades de captura entre las dos muestras. La suposición que las probabilidades de captura en las dos muestras son independientes, es otra manera de exponer la suposición de igualdad de capturabilidad. Un método para intentar lograr la independencia de las probabilidades de captura de las dos muestras, consiste en el uso de sistemas de muestreo totalmente diferentes para las dos muestras. Por esta razón, recomendamos antes, que una lista de muestra se basara en las entrevistas telefónicas, y la otra en las entrevistas en los muelles. Sin embargo, sospechamos que existe otra heterogeneidad y falta de independencia, en las probabilidades de captura. Creemos que los barcos de pesca que toman una parte muy activa en la pesca, tienen una probabilidad mayor de aparecer en cualquier lista elaborada (telefónica, o de muelles). Esta heterogeneidad causará un sesgo negativo sobre la estimación del tamaño de la estructura, pero no tenemos ni idea del grado de este sesgo negativo. Seber (1982, p. 86) ofrece una discusión más completa de la heterogeneidad e independencia de las muestras.

### Marcas Perdidas o No Observadas

La situación aquí es algo confusa. Al principio, se podría pensar que, en esta aplicación, no ha manera de poder perder, o no ver, una marca. Sin embargo, esto supone que todos los barcos tienen nombres distintos, o que, cuando los barcos tienen el mismo nombre, existe información adicional, como el nombre del capitán, que hace que hace que todos los individuos de la lista sean únicos. Si existe algún problema de falta de unicidad, puede que no quede claro si un barco marcado ha sido capturado o no.

Otro punto relacionado es que los agentes pueden equivocarse en el registro de datos, creando así dificultades para que una recaptura corresponda con los datos iniciales. Se está desarrollando y documentando un procedimiento estándar para reducir al mínimo estos errores en el futuro.

## **Estimación del Esfuerzo Total y de la Captura Total**

El Esfuerzo Total ( $E$ ) (es decir, el número total de viajes de pesca realizados durante un periodo definido) se estima con

$$\hat{E} = \hat{N} \bar{e}, \quad (9.5)$$

cuando  $\hat{N}$  es la estimación del tamaño de la estructura (Tamaño de la Flota) y  $\bar{e}$  es la media del esfuerzo de pesca (es decir, el número medio de viajes de pesca realizados) obtenida de la muestra telefónica. La evaluación de las propiedades de este estimador es más difícil que cuando  $N$  es conocido, porque tanto  $\hat{N}$  como  $\bar{e}$  son variables aleatorias. Sospechamos que  $e$  tiene sesgo alto porque los barcos de pesca que no pescan mucho tienen menos probabilidades de aparecer en la lista. Por desgracia, no podemos decir que  $\hat{N}$  siempre tendrá sesgo alto o bajo. Las tres violaciones de suposiciones podrían tener importancia (cierre, heterogeneidad, y pérdida de marcas), y no está claro que dirección tomaría el sesgo general sobre  $\hat{N}$ . El único procedimiento posible es el de usar la simulación con una variedad de escenarios diferentes para las violaciones de suposiciones. Usando la ecuación (9.4), la varianza estimada de  $\hat{E}$  se da con

$$\begin{aligned} \hat{\text{Var}}(E) = & (\hat{N})^2 \hat{\text{Var}}(\bar{e}) + (\bar{e})^2 \hat{\text{Var}}(\hat{N}) - \\ & \hat{\text{Var}}(\bar{e}) \hat{\text{Var}}(\hat{N}). \end{aligned} \quad (9.6)$$

La captura total ( $C$ ) se estima con  $\hat{C} = \hat{E} \bar{c}$ , cuando  $\hat{E}$  es el esfuerzo total de pesca estimado, y  $\bar{c}$  es la captura media por esfuerzo de unidad, calculado a partir de las entrevistas en los muelles. Es probable que las propiedades de esta ecuación estén sujetas a las mismas cuestiones que la ecuación (9.5), y, de nuevo, la simulación podría resultar de gran utilidad.

## **9.4 Otras Cuestiones**

### **Métodos para Ocuparse de las Estructuras de Listas Incompletas**

#### **(i) Completar la Estructura de la Lista**

La ventaja es que el investigador de la encuesta tiene una estructura completa, y no tiene que generalizar los resultados para un tamaño estimado de estructura. La desventaja es el coste, y la posible impracticabilidad de completar la estructura de la lista.

#### **(ii) Uso de una Estructura de Áreas**

La ventaja es que solamente hay que enumerar los establecimientos de las áreas a muestrear. La desventaja es la posible ineficacia, en el caso de haber escasez de negocios en cada área grande.

### **(iii) Uso de Lista y Estructura de Áreas (Procedimiento Multi-Estructura)**

Está claro que las ventajas son el aumento de precisión, y el tener cubiertos a todos los establecimientos. La desventaja podría ser la carestía y la impracticabilidad.

### **(iv) Uso de Captura-Recaptura para Estimar el Tamaño de Estructura de Lista**

La ventaja consiste en tener un método práctico, de menor gasto que los tres procedimientos listados arriba. Las desventajas son el posible sesgo si se violan las suposiciones del método de captura-recaptura, y el tener que incluir la variación debida a la estimación del tamaño de estructura, en las estimaciones de varianza de las estimaciones de la población total.

#### **Estimación de Tamaño de Estructura con Captura-Recaptura**

En esta sección consideramos las suposiciones de modelos, la precisión de las estimaciones, la estimación de los totales de población, y los problemas especiales de los diseños de muestreos más complejos cuando se usa el procedimiento de captura-recaptura para la estimación del tamaño de estructura.

#### **Suposiciones de Modelos**

##### **(i) Cierre**

¿Puede considerarse constante el tamaño de la estructura, para que puedan usarse los modelos cerrados de población? Esto dependerá de si la encuesta es simplemente un instante tomado en un punto único de tiempo, o si se requieren una serie de encuestas a través del tiempo. Dependerá, también, de la rapidez con que surgen las nuevas. Sospechamos que habrá necesidad de usar modelos de población cerrada y abierta según los establecimientos a estudiar.

Tenemos, también, la cuestión de emigración temporal cuando los establecimientos salen de la estructura y entonces vuelven a entrar. Esto fue considerado como un posible problema en el ejemplo de los barcos de pesca, porque los barcos podrían quedar inactivos, y luego volver a la actividad de nuevo. Esto puede ser problema, también, en otras encuestas sobre establecimientos, si algunos establecimientos se dan de alta y de baja en los negocios con frecuencia, y si guardan el mismo nombre cuando vuelven a darse de alta.

##### **(ii) “Desigualdad de Capturabilidad” y la Independencia de las Listas**

Como discutimos anteriormente, lo ideal sería que las listas fuesen independientes para que las estimaciones de tamaño de estructura quedan sin sesgo. En la práctica, puede que no sea fácil encontrar dos o más listas independientes.

### **(iii) Pérdida de Marca - Identificación Única de Establecimiento**

Los nombres de los establecimientos deben ser únicos e inconfundibles, para que su identificación en listas distintas no se eche a perder ni a equivocar. Esto era un problema del ejemplo de los barcos de pesca en años anteriores. Sospechamos que no vaya a resultar problemático en la mayoría de las encuestas sobre los establecimientos.

### **Precisión de las Estimaciones**

Las listas usadas deben tener tamaño suficiente tal que la precisión del tamaño de la estructura ( $\hat{N}$ ) sea adecuada. Seber (1982, p.96) trata de forma detallada la estimación de Lincoln-Petersen, y presenta gráficos de los tamaños de muestra requeridos para varios niveles de precisión. Pollock et al. (1990) presenta información sobre tamaños de muestra para los modelos de población abierta.

### **Estimación de los Totales de Población**

Una vez obtenida la estimación del tamaño de la estructura, esta estimación se combinará, a menudo, con una media de muestra, para estimar una estimación del total de población ( $\hat{Y} = \hat{N} \bar{y}$ ). La estimación del total de población queda sujeta a un posible sesgo y a varianza adicional, porque  $\hat{N}$  es estimado. La estimación puede, también, tener sesgo porque  $\bar{y}$  no se basa en una muestra aleatoria de la estructura completa.

### **Diseños más complejos de Muestreo**

Hemos subrayado aquí la estimación de tamaño de estructura en el muestreo aleatorio sencillo, usando el método captura-recaptura. Surgen cuestiones adicionales si se usan diseños más complejos de muestreo. Por ejemplo, en los diseños estratificados, surgiría la cuestión de si se debe estimar el tamaño de la estructura de cada estrato por separado, o estimar el tamaño total de estructura y, a continuación, repartirlo entre los estratos, suponiendo igualdad de probabilidades de los diferentes estratos en las listas incompletas. Existe, también, la cuestión más compleja de como estimar el tamaño de estructura en los diseños multi-etapa. Está claro que esta es una área que requerirá investigación en el futuro.

## **10. PROGRAMAS DE ORDENADOR**

Existen ya un gran número de programas de ordenador que han sido desarrollados para usar con varios modelos de captura-recaptura. En cada caso, se da una referencia.

## **Poblaciones Cerradas**

Para los modelos cerrados de captura-recaptura, el programa principal es CAPTURE, que fue desarrollado inicialmente por Otis et al. (1978). Ha sido revisado y actualizado considerablemente por Rextad y Burnham (1991).

## **Modelos de Jolly-Seber**

Pollock et al. (1990) ofrecen JOLLY para el modelo básico de Jolly-Seber. James Hines y James Nichols escribieron este programa. También, escribieron una extensión, JOLLYAGE, para las poblaciones con dos clases de edad y muestreo una vez al año.

Arnason y Schwarz (1987) han aportado POPAN-3, que puede usarse para los modelos del tipo Jolly-Seber.

## **Modelado de Supervivencia**

El primer programa desarrollado aquí fue SURVIV (White 1983), que se sigue usando extensivamente, debido a su potencia y flexibilidad. No es "amigable al usuario".

SURGE (Lebreton et al. 1992) fue desarrollado recientemente. Permite la adaptación con facilidad de modelos complejos de supervivencia, y tiene en cuenta, también, a los covariantes.

SURPH (Skalski et al. 1995) también ha sido desarrollado. Permite, también, el modelado complejo de supervivencia con covariantes.

## **Modelos Estructurados por Etapas**

Se ha escrito una versión de SURVIV para usar con las poblaciones estructuradas por etapas. Brownie et al. (1993) describen al programa MSSURVIV, pero este tiene sus raíces en los trabajos anteriores de Hestbeck et al. (1991) y Nichols et al. (1992). James Hines y James Nichols escribieron el programa.

## **Modelos Comparativos de Captura-Recaptura**

SURVIV ha sido modificado por White para considerar los modelos comparativos de captura-recaptura. El programa RELEASE se describe en Burnham et al. (1987).

## **Modelos Lineales Generalizados (GLIM)**

Cormack (1989) y trabajos anteriores describen como usar el programa GLIM (McCullagh y Nelder 1983) para computar los estimadores de los modelos de captura-recaptura abiertos y cerrados.

## **11. CONCLUSIONES**

Recientemente, se han escrito muchos trabajos usando los métodos Bayesianos para los modelos de captura-recaptura. Por ejemplo, los modelos Lincoln-Petersen (Gazey y Staley 1986), modelos de población cerrada (por ejemplo, Castledine 1981, Smith 1988, Rodrigues et al. 1988) y los modelos de recuperación de banda (Vounatsou y Smith 1995) han sido considerados. Ya se están solucionando los problemas numéricos con el algoritmo de muestreo de Gibbs (Gelfand y Smith 1990) y las técnicas de integración numérica. Es probable que los métodos Bayesianos se usen más extensivamente en el futuro.

Las cuestiones de la computación son muy importantes y es probable que haya un aumento en el uso de técnicas de muestreo repetido, como la de la autoelevación. Buckland y Garthwaite (1991) y Norris y Pollock (1995) han aplicado la autoelevación a modelos cerrados de captura-recaptura.

Libreton et al. (1992) han elegido un método muy fuerte de selección de modelos, basado en los Criterios de Información de Akaike (Akaike 1973). Creemos que los criterios para la selección de modelos se estudiarán más en el futuro.

Creemos que, en el futuro, los métodos de combinaciones recibirán mucho estudio. Por ejemplo, Alpizar (1994) y otros han considerado un método combinado de captura-recaptura y líneas transversales. El uso de los métodos combinados, aunque complicado, puede permitir la debilitación de algunas de las suposiciones de los métodos individuales. Burnham (1993) combinó modelos de captura-recaptura y modelos de recuperación de bandas. Otro método combinado puede implicar la combinación de los modelos de captura-recaptura con la radiotelemetría (Pollock et al. (1995).

Tabla 3.1 Modelos de captura-recaptura para poblaciones cerradas, que tienen en cuenta las probabilidades de desigualdad de captura.

Modelo <sup>a</sup>	Fuente de variación de la probabilidad de captura			Estimador disponible
	Heterogeneidad	Reacción ante Trampas	Tiempo	
$M_o$				Si
$M_h$	X <sup>b</sup>			Si
$M_b$		X		Si
$M_{bh}$	X	X		Si
$M_t$			X	Si
$M_{th}$	X		X	Si
$M_{tb}$		X	X	Si
$M_{tgh}$	X	X	X	Si

<sup>a</sup> Este grupo de 8 modelos es de Otis et al. (1978)

Los <sup>b</sup>X's denotan las fuentes de variación de probabilidad de captura incorporadas en los modelos.

Tabla 3.2. Procedimiento para la selección de modelos del programa CAPTURE, para los datos sobre el campañol, recogidos por J. D. Nichols en el Centro para la Investigación de la Fauna de Patuxent, en Laurel, Maryland, en Octubre 1981.

Modelo	$M_o$	$M_h$	$M_b$	$M_{bh}$	$M_t$	$M_{th}$	$M_{tb}$	$M_{tgh}$
Criterio	0.80	1.00 <sup>a</sup>	0.38	0.59	0.00	0.32	0.52	0.98 <sup>a</sup>

<sup>a</sup> Los valores altos para  $M_h$  y  $M_{tgh}$  sugieren que el modelo apropiado es, probablemente,  $M_h$  o  $M_{tgh}$ . La estimación sugerida es la "navaja" correspondiente a  $M_h$ , el modelo simple. No hay estimador para  $M_{bh}$ .

Tabla 3.3. Estadísticas seleccionadas y estimaciones de parámetros del programa CAPTURE, para los datos sobre el campañol recogidos por J. D. Nichols en el Centro Patuxent para la Investigación de la Fauna, Laurel, Maryland, en Octubre 1981. Se utiliza el modelo Mh, el modelo de la heterogeneidad.

Frecuencias de captura<sup>a</sup>

<i>i</i>	1	2	3	4	5
<i>F(i)</i>	29	15	15	16	27

Número de Animales Capturado 102

Media de P-HAT = 0.44

Estimación Interpolada de Población es 139 Con error típico 10.85.

Intervalo de Confianza 95% Aproximadamente, 117 a 161

<sup>a</sup> Estos son los números de animales capturados de 1 a 5 veces

<sup>b</sup> Este es el número de animales distintos capturados al menos una vez.

Tabla 4.1. Anotación para el modelo Jolly-Seber descrito detalladamente en Sección 4.

**PARÁMETROS**

- $M_i$  = número de animales marcados en la población en el momento de tomar la  $i$ -esima muestra ( $i = 1, \dots, k, M_1 = 0$ ).
- $N_i$  = el número total de animales en la población en el momento de tomar la  $i$ -esima muestra ( $i = 1, \dots, k$ ).
- $B_i$  = el número total de animales nuevos entrando en la población entre las muestras nº  $i$  y nº  $(i + 1)$ , ( $i = 1, \dots, k - 1$ )
- $\phi_i$  = la probabilidad de supervivencia para todos los animales entre las muestras nº  $i$  y nº  $(i + 1)$ , ( $i = 1, \dots, k - 1$ )
- $P_i$  = la probabilidad de captura para todos los animales en la muestra nº  $i$ , ( $i = 1, \dots, k$ )

**ESTADÍSTICAS**

- $m_i$  = el número de animales marcados capturados en la muestra nº  $i$ , ( $i = 1, \dots, k$ ).
- $u_i$  = el número de animales no marcados capturados en la muestra nº  $i$ , ( $i = 1, \dots, k$ ).
- $n_i$  =  $m_i + u_i$ , el número total de animales capturados en la muestra nº  $i$ , ( $i = 1, \dots, k$ ).
- $R_i$  = el número de ni puestos en libertad después de la muestra nº  $i$ , ( $i = 1, \dots, k - 1$ ).  
Es posible que esto no sea todo el  $n_i$ , debido a pérdidas al capturar, como se comentó en el texto.
- $r_i$  = el número de los animales  $R_i$  puestos en libertad en  $i$ , que se vuelven a capturar ( $i = 1, \dots, k - 1$ )
- $z_i$  = el número de animales capturados antes de  $i$ , no capturados en  $i$ , y capturados de nuevo más tarde ( $i = 2, \dots, k - 1$ ).

Tabla 4.2 Ejemplo hipotético para ilustrar que la información de la historia de capturas puede resumirse en una serie de ceros y uno para cada animal.

Periodo de captura

Animal	1	2	3	4	5
1	1	1	0	0	0
2	1	0	0	0	0
3	1	0	0	1	0
4	0	1	1	1	1

Tabla 4.3. Estadísticas de captura-recaptura para una población de ardillas grises en la Estación de Investigación de Bosque de Alice Holt, Surrey, Inglaterra, Noviembre 1972-Septiembre 1974

Período	Fecha	$n_t^a$	$m_t$	$R_t$	$r_t$	$z_t$
1	Nov. 1972	46	46	43		
2	Dic. 1972	46	42	46	44	1
3	Ene. 1973	48	42	48	48	3
4	Feb. 1973	46	42	46	45	9
5	Mar. 1973	51	46	50	46	8
6	Abr. 1973	37	37	37	35	17
7	May. 1973	41	41	41	40	11
8	May.-Jun. 1973	42	39	42	37	12
9	Jun. 1973	47	43	47	40	6
10	Jul. 1973	31	26	31	26	20
11	Agos. 1973	8	7	8	8	39
12	Sept. 1973	2	2	2	2	45
13	Oct. 1973	1	0	1	1	47
14	Nov. 1973	4	3	4	3	45
15	Dic. 1973	9	8	9	8	40
16	Ene. 1974	19	17	18	17	31
17	Feb. 1974	19	14	19	18	34
18	Mar. 1974	27	20	27	24	32
19	Abr. 1974	36	36	36	32	20
20	May. 1974	45	34	44	33	18
21	Jul. 1974	74	46	73	15	5
22	Agos. 1974	22	20	22	2	0
23	Sept. 1974	3	2	2		

\*Anotación explicada en Tabla 4.1.

**Tabla 4.4. Estimaciones de Jolly-Seber y errores<sup>a</sup> típicos aproximados para una población de ardillas grises en la Estación de Investigación de Bosque Alice Holt, Surrey, Inglaterra, Noviembre 1972-Septiembre 1974**

Periodo	Fecha	$\tilde{N}_i^b$	$\hat{SE}$	$\hat{\phi}_i$	$\hat{SE}$	$\bar{B}_i$	$\hat{SE}$
1	Nov 1972	0.94	0.037				
2	Dic 1972	47.1	0.39	0.96	0.030	6.3	0.77
3	Ene 1973	51.3	0.70	1.00	0.004	4.5	1.27
4	Feb 1973	56.0	1.19	0.99	0.023	5.1	1.53
5	Mar 1973	60.5	1.51	0.94	0.041	0.0	1.06
6	Abr 1973	54.9	1.23	0.95	0.038	0.0	0.00
7	May 1973	52.3	0.60	1.00	0.030	3.9	1.22
8	May-Jun 1973	56.5	2.06	0.90	0.052	3.7	1.45
9	Jun 1973	54.6	1.57	0.92	0.067	8.7	3.30
10	Jul 1973	58.9	4.59	0.84	0.066	2.2	6.60
11	Ago 1973	51.8	5.99	1.00	0.000	0.0	5.99
12	Sep 1973						
13	Oct 1973						
14	Nov 1973						
15	Dic 1973	58.3	9.20	0.93	0.115	1.0	6.57
16	Ene 1974	55.3	4.30	0.98	0.068	12.1	8.25
17	Feb 1974	66.4	8.14	1.00	0.071	6.8	10.14
18	Mar 1974	74.5	7.91	0.93	0.067	0.0	6.28
19	Abr 1974	58.4	2.13	0.99	0.071	18.2	4.22
20	May 1974	76.0	6.12	1.00	0.168	33.9	8.86
21	Jul 1974	110.3	18.10	0.21	0.048	0.0	2.23
22	Ago 1974	21.9	0.00				
23	Sep 1974						

<sup>a</sup>  $\hat{SE}(\tilde{N}_i)$  y  $\hat{SE}(\bar{B}_i)$  incluyen solamente variación de muestreo o “error de estimación”;  $\hat{SE}(\hat{\phi}_i)$  se obtuvo usando el estimador de varianza total de Jolly (1965).

<sup>b</sup>Anotación explicada en Tabla 4.1.

## REFERENCES

- Akiake, H. 1973. Information theory as an extension of the maximum likelihood principal. Pages 267-281. In B. N. Petrov and F. Csaki editors. Second International Symposium on Information Theory, Academiai Kaido, Budapest, Hungary.
- Alpizar, R. O. 1994. A combination line transect and capture-recapture sampling model for multiple observers in aerial surveys. Masters Thesis, Biomathematics Program, Department of Statistics, North Carolina State University.
- Anderson, D. R., and Sterling, R. T.. 1974. Population dynamics of molting pintail drakes banded in south-central Saskatchewan. *Journal of Wildlife Management* 38, 266-274.
- Arnason, A. N. 1972. Parameter estimation for mark-recapture experiments on two populations subject to immigration and death. *Researches in Population Ecology* 13, 97-113.
1973. The estimation of population size, migration rates, and survival in a stratified population. *Researches in Population Ecology* 15, 1-8.
- , and Mills, K. H. 1981. Bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. *Canadian Journal of Fisheries and Aquatic Sciences* 38, 1077-1095.
- , and Schwarz, C. J. 1987. POPAN-3. Extended analysis and testing features for POPAN-2. Charles Babbage Research Center, St. Norbert, Manitoba, Canada, 83pp.
- Balser, J. P. 1981. Confidence interval estimation and tests for temporary outmigration in tag-recapture studies. Ph.D. Thesis, Cornell University, Ithaca, N.Y. 205pp.
- Bishop, Y.M.M., Fienberg, S. E., and Holland, P. W. 1995. Discrete Multivariate Analysis: Theory and Practice. The MIT Press, Cambridge, Mass.
- Brownie, C., Anderson, D. R., Burnham, K. P., and Robson, D. S. 1985. Statistical inference from band recovery data: a handbook. Second ed. U.S. Fish and Wildlife Service Resource Publication 156, 305pp.
- , Hines, J. E., and Nichols, J. D. 1986. Constant-parameter capture-recapture models. *Biometrics* 42, 561-574.

- , Hines, J. E., Nichols, J. D., Pollock, K. H., and Hestbeck, J. B. 1993. Capture-recapture studies for multiple strata including non-Markovian transitions. *Biometrics* 49, 1173-1187.
- , and Pollock, K. H. 1985. Analysis of multiple capture-recapture data using band-recovery methods. *Biometrics* 41, 411-420.
- , and Robson, D. S. 1983. Estimation of time-specific survival rates from tag-resighting samples: a generalization of the Jolly-Seber model. *Biometrics* 39, 437-453.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., and Laake, J. L. 1993. *Distance Sampling: Estimating Abundance of Biological Populations*. Chapman and Hall, London.
- , and Garthwaite, P. H. 1991. Quantifying precision of mark-recapture estimates using the bootstrap and related methods. *Biometrics* 47, 255-268.
- Burley, N. 1985. Leg-band color and mortality patterns in captive breeding populations of zebra finches. *Auk* 102, 647-651.
- Burnham, K. P. 1972. Estimation of population size in multiple capture-recapture studies when capture probabilities vary among animals. Ph.D. Thesis, Oregon State University, Corvallis. 168pp.
- . 1993. A theory for combined analysis of ring recovery and recapture data. Pages 199-213, in J. D. Lebreton and P. M. North editors. *Marked Individuals in the Study of Bird Population*. Birkhauser. Verlag, Basel, Switzerland.
- , Anderson, D. R., White, G. C., Brownie, C., and Pollock, K. H. 1987. *Design and Analysis Methods for Fish Survival Experiments Based on Release-Recapture*. American Fisheries Society Monographs 5, Bethesda, MD.
- , and Overton, W. S. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65, 625-633.
- , and . 1979. Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60, 927-936.
- Carothers, A. D. 1973. The effects of unequal catchability on Jolly-Seber estimates. *Biometrics* 29, 79-100.
- . 1979. Quantifying unequal catchability and its effect on survival estimates in an actual population. *Journal of Animal Ecology* 48, 863-869.
- Castledine, B. J. 1981. A Bayesian analysis of multiple-recapture sampling for a closed population. *Biometrika* 67, 197-210.
- Caughley, G. 1977. *Analysis of Vertebrate Populations*. John Wiley and Sons, New York, N.Y. 234pp.
- Chao, A. 1988. Estimating animal abundance with capture frequency data. *Journal of Wildlife Management* 52, 295-300.

- , Ma, M.-C., and Yang, M. C. K. 1993. Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* 80, 193-201.
- , Lee, S.-M., and Jeng, S.-L. 1992. Estimating population size for capture recapture data when capture probabilities vary by time and individual animals. *Biometrics* 48, 201-216.
- Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. *University of California Publ. Stat.* 1, 131-160.
- .1980. Book review of statistical inference from capture data on closed animal populations. *Biometrics* 36, 362.
- Cochran, W. G. 1977. *Sampling Techniques*. Third ed. John Wiley and Sons, New York, N.Y. 428pp.
- Conner, M. C., and Labisky, R. F. 1985. Evaluation of radio isotope tagging for estimating abundance of raccoon populations. *Journal of Wildlife Management* 49, 326-332.
- Cormack, R. M. 1964. Estimates of survival from the sighting of marked animals. *Biometrika* 51, 429-438.
- .1972. The logic of capture-recapture estimates. *Biometrics* 28, 337-343.
- .1973. Commonsense estimates of capture-recapture studies. Pages 225-234 in M. S. Bartlett and R. W. Hiorns, eds. *The Mathematical Theory of the Dynamics of Biological Populations*. Academic Press, New York, N.Y.
- .1989. Loglinear models for capture-recapture. *Biometrics* 45, 395-413.
- Cowan, C. D., Breakey, W. R., and Fischer, P. J. 1986. The methodology of counting the homeless, in *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 170-175.
- Darroch, J. N. 1959. The multiple recapture census: II. Estimation when there is immigration or death. *Biometrika* 46, 336-351.
- Deuel, B. 1985. Experimental lead dosing of northern pintails in California. *California Fish and Game* 71, 125-128.
- Eberhardt, L. L. 1978. Appraising variability in population studies. *Journal of Wildlife Management* 42, 207-238.
- Fienberg, S. 1972. The multiple recapture census for closed populations and Incomplete 2k contingency tables. *Biometrika* 59, 591-603.
- .1992. Bibliography to capture-recapture modelling with application to census undercount adjustment. *Survey Methodology* 18, 143-154.
- Gazey, W. J. and Staley, M. J. 1986. Population estimation for mark-recapture experiments using a sequential Bayes algorithm. *Ecology* 67, 941-951.
- Gelfand, A. E., and Smith, A. F. M. 1990. Sampling based approaches to calculating marginal densities. *Journal of American Statistical Association* 85, 398-409.

- Gilbert, R. O. 1973. Approximations of the bias in the Jolly-Seber capture-recapture model. *Biometrics* 29, 501-526.
- Goudie, I. B. J. 1990. A likelihood-based stopping rule for recapture debugging. *Biometrika* 77, 203-206.
- Greene, M. A. 1983. Estimating the size of the criminal population using an open population approach, in *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 8-13.
- Hartley, H. O. 1962. Multiple frame surveys. *Proceedings of the Social Statistics Section, American Statistical Association*, 203-206.
- . 1974. Multiple frame methodology and selected applications. *Sankhya* 36, 99-118.
- Heisey, D. M., and Fuller, T. K. 1985. Evaluation of survival and cause-specific mortality rates using telemetry data. *Journal of Wildlife Management* 49, 668-674.
- Henny, C. J., and Anderson, D. 1979. Osprey distribution, abundance, and status in western North America III. The Baja California and Gulf of California population. *Bulletin Southern California Academy of Science* 78, 89-106.
- Hepp, G. R., Hoppe, R. T., and Kennamer, R. A. 1987. Population parameters and philopatry of breeding female wood ducks. *Journal of Wildlife Management* 51, 401-404.
- Hestbeck, J. G., Nichols, J. D., and Malecki, R. A. 1991. Estimates of movement and site fidelity using mark-resight data of wintering Canada Geese. *Ecology* 72, 523-533.
- Johnson, N. L. and Kotz, S. 1970. *Continuous Univariate Distributions--Vol. 2*. Houghton Mifflin, Boston, Mass. 306pp.
- Jolly, G. M. 1965. Explicit estimates for capture-recapture data with both death and immigration-stochastic model. *Biometrika* 52, 225-247.
- . 1982. Mark-recapture models with parameters constant in time. *Biometrics* 38, 301-321.
- Kendall, W. L., Nichols, J. D., and Hines, J. E. 1995. Estimating temporary emigration and breeding proportions using capture-recapture data with Pollock's robust design. *Ecology (in review)*
- , and Pollock, K. H. 1992. The robust design in capture-recapture studies: a review and evaluation by Monte Carlo simulation, in *Wildlife 2001: Populations*, edited by Dale R. McCullough and Reginal H. Barrett, Elsevier Applied Science, New York, 31-43.
- , Pollock, K. H., and Brownie, C. 1995. A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics* 51, 293-308.
- Kott, P. S., and Vogel, F. A. 1995. Multiple-Frame Surveys, in *Survey Methods for Businesses, Farms and Institutions* (B.G. Cox, ed.), John Wiley and Sons, New York, pp. 185-203.

- Laska, E. M., and Meisner, M. 1993. A plant-capture method for estimating the size of a population from a single sample. *Biometrics* 49, 209-220.
- Lebreton, J.-D., Burnham, K. P., Clobert, J., and Anderson, D. R. 1992. Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* 62, 67-118.
- LeCren, E. D. 1965. A note on the history of mark-recapture population estimates. *Journal of Animal Ecology* 34, 453-454.
- Lee, S.-M., and Chao, A. 1994. Estimating population size via sample coverage for closed capture-recapture models. *Biometrics* 50, 88-97.
- Lincoln, F. C. 1930. Calculating waterfowl abundance on the basis of banding returns. U.S. Department of Agric. Circ. 118, 4pp.
- McCullough, P., and Nelder, J. A. 1983. Generalised Linear Models. Monograph on Statistics and Probability 37, Chapman and Hall, London.
- Manly, B. F. J. 1970. A simulation study of animal population estimation using the capture-recapture method. *Journal of Applied Ecology* 7, 13-39.
- Martin, F. W., Pospahala, R. S., and Nichols, J. D. 1979. Assessment and population management of North America migratory birds. Pages 187-239 in J. Cairns, G. P. Patil, and W. E. Waters, eds. Environmental biomonitoring, assessment, predicting, and management--certain case studies and related quantitative issues. Statistical ecology series. Vol. 11. International Co-operative Publ. House, Burtonville, Md.
- Menkins, G. E., Jr., and Anderson, S. H. 1988. Estimation of small-mammal population size. *Ecology* 69, 1952-1959.
- Nayak, T. K. 1988. Estimating population size by recapture sampling. *Biometrika* 75, 113-120.
- . 1991. Estimating the number of component processes of a superimposed process. *Biometrika* 78, 75-81.
- Nichols, J. D., Hines, J. E., and Pollock, K. H. 1984a. Effects of permanent trap response in capture probability on Jolly-Seber capture-recapture model estimates. *Journal of Wildlife Management* 48, 289-294.
- , Morris, R. W., Brownie, C., and Pollock, K. H. 1986. Sources of variation in extinction rates, turnover and diversity of marine invertebrate families during the Paleozoic. *Paleobiology* 12, 421-432.
- , Noon, B. R., Stokes, S. L., and Hines, J. E., 1981. Remarks on the use of mark-recapture methodology in estimating avian population size. Pages 121-136 in C. J. Ralph and M. J. Scott eds. Estimating the numbers of terrestrial birds. Proc. international symposium at Asilomar, Calif., 26-31 October 1980. Studies in Avian Biology 6, Cooper Ornithological Society.
- , and Pollock, K. H. 1983. Estimating taxonomic diversity, extinction rates and speciation rates from fossil data using capture-recapture models. *Paleobiology* 9, 150-163.

- , and. 1983. Estimation methodology in contemporary small mammal capture-recapture studies. *Journal of Mammalogy* 64, 253-260.
- , Sauer, J. R., Pollock, K. H., and Hestbeck, J. B. 1992. Estimating transition probabilities for stage based population projection matrices using capture-recapture data. *Ecology* 73, 306-312.
- , Stokes, S. L., Hines, J. E., and Conroy, M. J. 1982. Additional comments on the assumption of homogeneous survival rates in modern bird banding estimation models. *Journal of Wildlife Management* 46, 953-962.
- Norris, J. L., and Pollock, K. H. 1996. Nonparametric maximum likelihood estimators for population size under two closed capture-recapture models with heterogeneity. *Biometrics* (To appear).
- Norris, J. L., and Pollock, K. H. 1995. Including model uncertainty in estimating variances in multiple capture studies. Unpublished manuscript in review.
- Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monograph* 62, 135pp.
- Pelton, M. R., and Marcum, L. C. 1977. Potential use of radioisotopes for determining densities of black bear and other carnivores, in *Proceedings 1975 Predator Symposium*, eds. R. L. Phillips and C. Jonkel, Montana Forestry and Conservation Experiment Station, University of Montana, Missoula, pp. 221-236.
- Pollock, K. H. 1974. The assumption of equal catchability of animals in tag-recapture experiments. Ph.D. Thesis, Cornell University, Ithaca, N.Y. 82pp.
- .1975. A K-sample tag-recapture model allowing for unequal survival and catchability. *Biometrika* 62, 577-583.
- .1981a. Capture-recapture models: a review of current methods, assumptions and experimental design. Pages 426-435 in C. J. Ralph and J. M. Scott, eds. *Estimating the numbers of terrestrial birds. Proc. international symposium at Asilomar, Calif., 26-31 October 1980. Studies in Avian Biology* 6, Cooper Ornithology Society.
- .1981b. Capture-recapture models allowing for age-dependent survival and capture rates. *Biometrics* 37, 521-529.
- .1982. A capture-recapture design robust to unequal probability of capture. *Journal of Wildlife Management* 46, 752-757.
- .1991. Modelling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present, and future. *Journal of the American Statistical Association* 86, 225-238.
- , Bunck, C. M., Winterstein, S. R., and Chen, C. L. 1995. A capture-recapture survival analysis model for radio-tagged animals. *Journal of Applied Statistics* (To appear).
- , Hines, J. E., and Nichols, J. D. 1985. Goodness-of-fit tests for open capture-recapture models. *Biometrics* 41, 399-410.

- , and Kendall, W. L. 1987. Visibility bias in aerial surveys: a review of estimation procedures. *Journal of Wildlife Management* 51, 502-510.
- , and Mann, R. H. K. 1983. The use of an age-dependent mark-recapture model in fisheries research. *Canadian Journal of Fisheries and Aquatic Sciences* 40, 1449-1455.
- , Nichols, J. D., Brownie, C., and Hines, J. E. 1990. Statistical Inference for Capture-Recapture Experiments. *Wildlife Society Monographs* 107, 1-97.
- , and Otto, M. C. 1983. Robust estimation of population size in closed animal populations from capture-recapture experiments. *Biometrics* 39, 1035-1049.
- , and Raveling, D. G. 1982. Assumptions of modern band-recovery models with emphasis on heterogeneous survival rates. *Journal of Wildlife Management* 46, 88-98.
- , Turner, S. C., and Brown, C. A. 1994. Use of capture-recapture techniques to estimate population size and population totals when a complete frame is unavailable. *Survey Methodology* 20, 117-124.
- , Winterstein, S. R., and Conroy, M. J. 1989a. Estimation and analysis of survival distributions for radio-tagged animals. *Biometrics* 45, 99-109.
- , Bunck, C. M., and Curtis, P. D. 1989b. Survival analysis in telemetry studies: the staggered entry design. *Journal of Wildlife Management* 53, 7-15.
- Quenouille, M. 1956. Notes on bias in estimation. *Biometrika* 43, 343-360.
- Rextad, E., and Burnham, K. P. 1991. User's Guide for Interactive Program CAPTURE. Ft. Collins, Colorado: Colorado Cooperative Fish and Wildlife Unit.
- Ricker, W. E. 1958. Handbook of computations for biological statistics of fish populations. *Fisheries Res. Board Can. Bull.* 119. 300pp.
- Robson, D. S. 1969. Mark-recapture methods of population estimation. Pages 120-140 in N. L. Johnson and H. Smith, Jr., eds. *New developments in survey sampling*. John Wiley and Sons, New York, N.Y.
- Rodrigues, J., Bolfarine, H., and Leite, J. G. 1988. A Bayesian analysis in closed animal populations from capture-recapture experiments with trap response. *Communications in Statistics Simulation* 17, 407-430.
- Rose, R. K. 1973. A small mammal live trap. *Transactions of the Kansas Academy of Science* 76, 14-17.
- Schnabel, Z. E. 1938. The estimation of the total fish population of a lake. *American Math. Monthly* 45, 348-352.
- Schwarz, C. J. Schweigert, J. F., and Arnason, A. N. 1993. Estimating migration rates using tag-recovery data. *Biometrics* 49, 177-193.

- Seber, G. A. F. 1965. A note on the multiple-recapture census. *Biometrika* 52, 249-259.
- .1970. The effects of trap-response on tag-recapture estimates. *Biometrics* 26, 13-22.
- .1982. The estimation of animal abundance and related parameters. Second ed. MacMillan, New York, N.Y. 654pp.
- Shirley, M. G., Linscomb, R. G., Kinler, N. W., Knaus, R. R. and Wright, V. L. 1988. Population estimates of river otters in a Louisiana coastal marshland. *Journal of Wildlife Management* 52, 512-515.
- Skalski, J. R., and coauthors. 1994. Statistical Survival Analysis of fish and Wildlife Tagging Studies SURPH.1: Center for Quantitative Sciences, School of Fisheries, University of Washington, Seattle, Washington.
- , and Robson, D. S. 1992. Techniques for Wildlife Investigations: Design and Analysis of Capture Data. Academic Press Inc., San Diego.
- , and Simmons, M. A. 1983. Comparative census procedures using single mark-recapture methods. *Ecology* 64, 752-760.
- Smith, P. J. 1988. Bayesian methods for multiple capture-recapture surveys. *Biometrics* 44, 1177-1189.
- Stokes, S. L. 1984. The Jolly-Seber method applied to age-stratified populations. *Journal of Wildlife Management* 48, 1053-1059.
- Stromberg, K. L., Grue, C. E., Nichols, J. D., Hepp, G. R., Hines, J. E., and Brown, H. C. 1988. Postfledgling survival of European starlings exposed as nestlings to an organophosphorus insecticide. *Ecology* 69, 590-601.
- Sudman, S., Sirken, M. G., and Cowan, C. D. 1988. Sampling rare and elusive populations. *Science* 240, 991-995.
- Tanaka, R. 1980. Controversial problems in advanced research on estimating population densities of small rodents. *Res. Popul. Ecol., Suppl.* 2. 67pp.
- Trent, T. T., and Rongstad, C. J. 1974. Home range and survival of cottontail rabbits of southwestern Wisconsin. *Journal of Wildlife Management* 38, 459-472.
- Vounastou, P., and Smith, A. F. M. 1995. Bayesian analysis of ring-recovery data via Markov chain Monte Carlo simulation. *Biometrics* 51, 687-708.
- White, G. C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. *Journal of Wildlife Management* 47, 716-728.
- , Anderson, D. R., Burnham, K. H., and Otis, D. L. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, LA 8787-NERP, Los Alamos, N.M. 235pp.
- Wileyto, E. P. 1994. Improving Markov-recapture population estimation with multiple marking. *Environmental Entomology* 23, 1129-1137.

- .1995. Expanding Markov-recapture models to incorporate multiple observations. *Environmental Entomology* 24, 000-000.
- , Ewens, W. J., Mullen, M. A. 1994. Markov-recapture population estimates: a tool for improving interpretation of trapping experiments. *Ecology* 75, 1109-1117.
- Wittes, J. T. 1974. Applications of a multinomial capture-recapture model to epidemiological data. *Journal of the American Statistical Association* 69, 93-97.
- , Colton, T., and Sidel, V. W. 1974. Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *Journal of Chronic Diseases* 27, 25-36.
- Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. *Biometrics* 12, 163-189.

## **NAZIOARTEKO ESTATISTIKA MINTEGIA SEMINARIO INTERNACIONAL DE ESTADÍSTICA**

3

### **ORAIN ARTE ARGITARATUTAKO GAIAK TEXTOS PUBLICADOS HASTA LA FECHA**

- 1.- LINEAL STATISTICAL INFERENCE  
(Inglés, Euskera, Español)  
**C.R.RAO**  
1983
- 2.- MUESTREO Y APLICACIONES  
(Español, Euskera)  
**E. CANSADO**  
1983
- 3.- STATISTICAL EDUCATION  
(Inglés, Euskera, Español)  
**V. BARNETT**  
1983
- 4.- ANALYSE DES DONNÉES  
(Francés, Euskera, Español)  
**P. CLAPIER**  
1983
- 5.- DESIGN OF EXPERIMENTS  
(Inglés, Euskera, Español)  
**D.J.FINNEY**  
1984
- 6.- ASPECTOS DE TEORÍAS Y APLICACIONES EN EL MUESTREO  
(Español, Euskera)  
**F.AZORIN POCH**  
1984

- 7.- CURSO BASICO INTENSIVO DE MUESTREO  
(Español, Euskera)  
J.L.SANCHEZ-CRESPO  
1985
- 8.- ANALYSE DES SERIES CHRONOLOGIQUES: LES INDICES STATISTIQUES  
(Francés, Euskera, Español)  
J. FOURASTIE  
1985
- 10.- METHODOLOGY AND TREATMENT FOR NON-RESPONSE  
(Inglés, Euskera, Español)  
R.PLATEK  
1986
- 11.- STATISTICAL OPERATION BY SAMPLING  
(Inglés, Euskera, Español)  
L.KISH  
1986
- 12.- ANALISIS DE SERIES TEMPORALES: ALGUNAS TECNICAS DE PREDIC-  
CION  
(Español, Euskera)  
I.GALLASTEGI  
1986
- 13.- BASES DE DATOS  
(Español, Euskera)  
F.SALTOR  
1987
- 14.- METODOS ESTADISTICOS PARA LA INVESTIGACION EPIDEMIOLOGICA  
(Español)  
L.C.SILVA  
1987
- 15.- SAMPLING AND NON-SAMPLING ERRORS IN SURVEYS  
(Inglés)  
A.MARTON  
1988
- 16.- LES ENQUÊTES TELEPHONIQUES  
(Francés)  
V.SALVY  
1988

- 17.- GENERALIZED LINEAR MODELS IN EPIDEMIOLOGY  
(Inglés)  
J.C.DUFFY  
1989
- 18.- NEW TECHNOLOGIES IN COMPUTER ASSISTED SURVEY PROCESSING  
(Inglés)  
J.G.BETHLEHEM and W.J.KELLER  
1989
- 19.- EVALUATION OF QUESTIONNAIRE DESIGN EFFECTS  
(Inglés)  
G.NATHAN  
1990
- 20.- PROCEDIMIENTO DE DEPURACION DE DATOS ESTADISTICOS  
(Inglés)  
I.VILLAN CRIADO y M.S.BRAVO CABRIA  
1990
- 21.- THE X11 ARIMA/88 SEASONAL ADJUSTMENT METHOD  
(Inglés)  
E.BEE DAGUM  
1990
- 22.- RENTA Y DISTRIBUCION DE LA RIQUEZA, DESIGUALDAD Y POBREZA:  
TEORIA, MODELOS Y APLICACIONES  
(Español)  
C.DAGUM  
1991
- 23.- LA CONTABILIDAD NACIONAL COMO MARCO DE LAS ESTIMACIONES DE  
VARIABLES ECONOMICAS  
(Español)  
V.ANTON VALERO  
1991
- 24.- MACRO-EDITING. METHODS FOR TAIONALIZING THE EDITING OF QUAN-  
TITATIVE DATA  
(Inglés)  
L.GRANQUIST  
1991
- 25.- METHODOLOGICAL ISSUE IN FAMILY EXPENDITURE SURVEYS  
(Inglés, Euskera, Español)  
M.KANTOROWITZ  
1992

- 26.- QUALITY CONTROL IN STATISTICS FROM ADMINISTRATIVE REGISTERS AND RECORDS  
(Inglés, Euskera, Español)  
HANS PETTERSSON  
1992
- 27.- ANALYSE DES DONNÉES ET CLASSIFICATION AUTOMATIQUE NUMÉRIQUE ET SYMBOLIQUE  
(Francés)  
E.DIDAY  
1992
- 28.- METHODOLOGIE DE LA RECHERCHE EXPERIMENTALE  
(Francés, Euskera, Español)  
ROGER PHAN TAN LUU  
1993
- 29.- STRUCTURAL EQUATION MODELING WITH LISREL  
(Español)  
KARL G. JÖRESKOG  
1993
- 30.- METODOS ESTADISTICO-ECONOMETRICOS PARA EL ANALISIS DE LA COYUNTURA ECONOMICA  
(Español)  
ANTONI ESPASA  
1994
- 31.- ESTADISTICAS DE EMPLEO Y PARO EN LA UNION EUROPEA  
(Español)  
BERNARD GRAIS, ALOIS VAN BASTELAER, ANDRÉ PERSENAIRE  
1994
- 32.- EL SISTEMA EUROPEO DE ESTADISTICAS INTEGRADAS DE PROTECCION SOCIAL (SEEPROS)  
(Inglés, Euskera, Español)  
ALFONSO BARRADA, MERCEDES ALCALDE, COR N.GORTER  
1995
- 33.- CAPTURE-RECAPTURE MODELS: AN OVERVIEW.  
(Inglés, Euskera, Español)  
KENNETH H. POLLOCK  
1995
- 34.- RESIDENTS REGISTRATION SYSTEMS AND STATISTICS  
(Inglés, Euskera, Español)  
PEKKA MYRSKYLÄ, WOLFGANG MOHR  
1995