"ON THE CAPTURE-RECAPTURE METHODS OF ESTIMATING POPULATION SIZE"

BY

NYANGOMA, STEPHEN ODUNDO

DEPARTMENT OF MATHEMATICS

UNIVERSITY OF NAIROBI

This dissertation is submitted in partial fulfillment of the degree of Master of Science in Mathematics in the University of Nairobi.

UNIVERSITY OF NAIROBI.

JULY, 1991.

DECLARATION

This dissertation is my original work and has not been presented for a degree in any University.

Signature Sodundo.

NYANGOMA STEPHEN ODUNDO.

This dissertation has been submitted for examination with my approval as a University Supervisor.

Signature Odhiambo 24-4-92

PROF. J. W. ODHIAMBO.

ACKNOWLEDGEMENT

This dissertation would not have been written without the assistance of many.

First and foremost, I am grateful to my sponsor the University of Nairobi for granting me a Scholarship that enabled me to pursue the postgraduate course.

I would like to express my sincere thanks to my supervisor Prof. J. W. Odhiambo, without whose guidance and close supervision, I may not have completed this work. Prof. Odhiambo has been of utmost assistance in all the stages of preparation of this dissertation and indeed the whole course. I am also grateful to my lecturers Dr. Manene and Prof. Ottieno, whose lectures in various fields provided me with background knowledge and inspiration to embark on this study. I am also very grateful to Prof. G. F. Seber for his advise and assistance during the preparation of this dissertation.

I owe the former Director of the Kenya Marine and Fisheries Research Institute (K. M. F. R. I.), Mr. Alela a word of thanks for releasing me to pursue the postgraduate course. I am also grateful to my colleague at the K.M.F.R.I., Mr. Ogunja for all the assistance he offered me at the end of the course. I wish to express my sincere gratitude to my colleagues in the course Mr. Bodo and Miss. Matu, who have been a source of great encouragement and company during the entire course.

I hereby wish to express deep appreciation to my parents Mr. and Mrs. Nyang'oma for their active support and encouragement during the entire education. Finally, I want to say thank you to all those whose names do not appear here but who in one way or the another helped me attain my goal.

CONTENTS.

SUMMARY OF CONTENTS.

This dissertation is an attempt to study the **capture-recapture** models estimation of closed population size. Where possible, the assumptions underlying various models have been discussed in some detail. Departures from various assumptions have also been discussed.

Chapter one, outlines the basic principles underlying the **capture-recapture** method of estimating population size. Some statistical methods used **capture-recapture** studies are also presented.

In Chapter Two, a very extensive literature review of the most work done on both open and closed populations is presented. A statement of the problem is also outlined in this Chapter.

In Chapter Three, we discuss various models based on Single-Mark release experimental set-up. The various assumptions underlying these models have also been discussed in some detail. The properties of the Petersen estimate have also been discussed. Inverse sapling scheme and the models based on it are also discussed.

In Chapter four, we discuss the models based on the Schnabel sapling scheme. Properties of the estimates derived from these models have also been discussed. Some Regression models have been presented. The testing of the underlying assumptions have been presented. We also discuss the models based on constant probability of capture. Inverse Multiple sampling census together with the models based on it have also been discussed. The models based on the Multi-Single recapture census have also been presented. Conclusions are in section 4.4 .

Lastly we give an appendix and references.

CHAPTER ONE

BASIC PRINCIPLES UNDERLYING THE CAPTURE-RECAPTURE METHOD

1.1 THE CAPTURE-RECAPTURE METHOD

The estimation of the total population size of animal populations is of great importance in a variety of biological problems. These problems may relate to population growth, ecological adaptation, genetic constitution, natural selection and evolution and so on. Obvious practical consequences are the maintenance of human food supplies and control of insect pests. For human communities,procedures employing fixed sampling units are available, but for mobile populations, other methods must be used. Techniques for estimating total population size of organisms which are mobile and wary of man are still in relatively primitive stage of development and while indices of abundance may be available in a variety of forms, the assessment of total size with any degree of precision generally requires considerable ingenuity and effort. Thus the number of fish of a given species present in a lake is not an easily accessible parameter. By a judicious selection of times and places to set nets each year, however, the fishery biologist may be able to monitor change in relative abundance with little difficulty.

Among the techniques which have been developed for estimating the total population size, the capture-recapture method is the most widely used. In its simplest and most commonly applied form, the capture-recapture experiment is a two sample experiment in which the members of the first sample are marked in some recognizable manner and returned to the population. The proportion of marked individuals appearing in the second sample is then regarded as an estimate of the proportion marked in the population. Since the number of marked individuals in the population is known, this reasoning leads directly to an estimate of the total population size. Thus if n_1 individuals are marked and released in the first sample and m_2 marked individuals are subsequently recaptured in a sample of size n_2, then the population size is estimated by $\hat{N} = n_1 n_2/m_2$ on the assumption that m_2/n_2 estimates n_1/N.

In more extensive investigations the sampling and marking continues intermitently over a period of time, the unmarked individuals captured on each occasion being marked before being returned with the others to the population. Distinct "batch marks" are sometimes used in such studies; that is on each sampling occasion all unmarked individuals are given an identical batch-mark but recognizably different mark is used for each successive batch. More commonly in the multiple sample experiment the "mark" consists of a numbered tag which is attached to the individual and thereafter uniquely identifies it. In some multiple sample experiments, both the marked and unmarked individuals are distinctively marked any time they are caught and then released to the population. Each time an individual is captured, a record is marked for it or on it to show the recapture history. We shall here consider statistical aspects of k-sample capture-recapture experiment.

The k-sample capture-recapture experiment

Sample size on any one occasion is usually limited by the amount of capture gear and manpower which can be brought into operation at any one time. Under such economic restrictions the only feasible means of increasing the precision of the experiment may be steadily increased by marking all unmarked individuals captured on each sampling occasion and returning the entire sample to the population.

The general field experiment is similar for all capture-recapture studies. At the beginning of the study, a sample of size n_1 is taken from the population. On each occasion the catch is considered as a random sample of individuals from the population. That is each individual in the population has an equal chance of being captured on any given occasion. Each time an individual is caught a record is marked for it or on it to show the occasion of capture. The individual is then returned into the population.

After allowing time for the marked and unmarked animals to mix, a second sample of size n_2 is taken. The second sample normally contains both marked and unmarked animals. In some methods the unmarked animals are marked and all captured animals are released back into the population, where as some methods demand that, both the marked and unmarked be marked distinctively and then released into the population. This procedure proceeds for k periods where $k \geq 2$.

During the course of this sampling experiment the population itself may undergo changes through such processes as mortality, emigration and immigration; and conceivably, the risks associated with these processes may vary with the previous capture history of an individual. In particular, the untagged portion of the population may be subject to different rates of mortality, emigration and immigration from the tagged portion. The early models assumed deterministic changes, constant over different periods, where as the later models (stochastic) considers variable changes. Since stochastic models describing capture-recapture studies have recently been shown not only to be less complicated to analyse than their corresponding deterministic models but also to provide more valid results, these methods should totally supersede their deterministic counterperts. However, if it is assumed that no changes occur over the sampling period, then the population is considered closed.

ASSUMPTIONS:

Before any mathematical formulation of capture-recapture models can be done, there are certain basic assumptions that have to be made about the population under study. These assumptions vary from one model to the other. We now give some of the assumptions which are made for most of the capture-recapture models; these are:

(i) animals do not loose their marks;

(ii) sampled animals are classified correctly as marked and unmarked according to when they were recaptured;

(iii) sampling is random with respect to mark status so that, either

(a) every animal has the same chance of recapture, or

(b) if there exist strata within the population such that, by size, behaviour or any other variation, different strata have different chances of recapture, then the marked animals belong to these strata in exactly the same proportions as the occurrence of the strata in the whole population.

The following assumption holds for strictly closed population models.

(iv) either (a) the population is really closed, or

(b) there is neither recruitment nor immigration (both of which affect unmarked animals only), and death and emigration affect marked and unmarked animals equally, or

(c) knowledge is available from other sources which permit an allowance to be made for migration, birth, and death prior to the analysis of the data.

We shall give more assumptions later as we study each model.

NOTATION AND TERMINOLOGY:

NOTATION

In this desertation we shall adopt the international notations used by F.A.O. for fishery research and especially the "mnemonic" notation. For example, N and n denote the number of individuals in the population and sample respectively; M and m refer to the number of marked (or tagged members) of the population and sample respectively s represents the number of samples and so on. Each chapter will be self contained as far as the notation is concerned.

Some statistical symbols are required: $E[y]$, $\sigma[y]$ $v[y]$ (= $\sigma^2[y]$, $c[y]$ (= $\sigma[y]/E[y]$) will represent the mean, standard deviation, variance and coefficient of variation respectively, of the random variable y, where as $cov[x,y]$ will denote the covariance of the random variables x and y, and $E[x|y]$, $v[x|y]$

are the mean and variance of x conditional on fixed y. The symbols z_α and $t_k[\alpha]$ will represent the 100α percent upper tail for the values of standard normal distribution and the t-distribution with k-degrees of freedom, respectively.

Occasionally the symbols $O[N]$ and $o[N]$ will be used; if g is a function of N, then $g(N) = O[N]$ if there exists an integer N_1 and a positive number A such that, for $N > N_1$, $d(|g(N)/N| < A$; $g(N) = o[N]$ if $\lim_{N \to \infty} \{g(N)/N\} = 0$. Roughly speaking, $O[N]$ means "of the same order of magnitude as N when N is large", while $o[N]$ means "of smaller order of magnitude than N when N is large".

If $\underline{x}$ has a multivariate normal distribution with mean vector $\underline{\theta}$ and variance covariance (dispersion) matrix Σ, we shall write $\underline{x} \sim N(\underline{\theta}, \Sigma)$.

All logarithms written $\log x$ will be to base e unless otherwise.

TERMINOLOGY

The size of an animal population in a given area will be determined by the process of immigration (or movement into an area), emigration (or movement out of the area), total mortality, and recruitment.

Total mortality: In dealing with exploited populations, we shall usually subdivide total mortality into mortality due to exploitation and natural mortality, i.e. mortality due to natural processes such as predation, disease, climatic conditions: Contrary to some authors, emigration is not included here under "mortality". We shall also distinguish between mortality rate and Instantenous mortality rate as follows:

Let ϕ_t be the probability that an animal survives for the period of time $[0,t]$, then if N_o animals are alive at time zero we would expect $N_t = N_o\phi_t$ to be alive at time t.

The proportion ϕ_t, sometimes expressed as a percentage, is called the survival rate over period t, and $1 - \phi_t$ is called the mortality rate over period t. If, however, the mortality rate may be regarded as a Poisson process with parameter μ, that is the probability that an individual dies in the time interval $(t, t+\delta t)$ is $\mu\delta t + o(\delta t)$, then

$$\phi_t = e^{-\mu t}$$

$$\frac{dN_t}{dt} = -\mu N_t,$$

and the parameter μ is called the instantanous mortality rate.

Mean Life Expectancy: Let Y be the time at which a member of N_o dies.

Then,

$$\begin{aligned} F[y] &= \Pr[Y \leq y] \\ &= 1 - \Pr[Y > y] \\ &= 1 - \Pr[\text{animal survives until time } y \end{aligned}$$

$$= 1 - \exp(-\mu y)$$

and Y has the probability density function

$$f(y) = F'(y) = \mu e^{-\mu y} \quad , \quad (y \geq 0).$$

Therefore the <u>mean life expectancy</u> is

$$E[Y] = \int_0^\infty \mu y e^{-\mu y}\, dy \;.$$

$$= 1/\mu$$

$$= - 1/\log \phi_1 .$$

<u>Recruitment</u>: By recruitment, we shall refer to those animals born into the population or, where applicable, those animals which grow into the catchable part of the population. In fishery research, recruitment some times denote those fish which grow into the class of legally catchable fish. Thus we do not treat immigration as a component of recruitment.

<u>Open and closed populations</u>: A population which remains unchanged during the period of investigation (i.e. the effects of migration, mortality and recruitment are negligible) is called a closed population. If a population is changing due to one or more of the above processes operating, then the population is said to be open.

1.3 SOME STATISTICAL METHODS USED IN CAPTURE-RECAPTURE STUDIES:

MAXIMUM LIKELIHOOD ESTIMATION:

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n, which are independently and identically distributed (i.i.d) as $f(x_i, \theta)$ then

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

is the likelihood function. Then, under regularity conditions, $\hat{\theta}$, the maximum-likelihood estimate of θ, is a solution of the equation

$$\frac{\partial \log L(\theta)}{\partial \theta} = 0.$$

and as $n \to \infty$, $\hat{\theta}$ is asymptotically distributed as $N(\theta, \sigma_{\hat{\theta}}^2)$, where

$$\sigma_{\theta}^2 = -E\left\{\left[\frac{\sigma^2 \log L(\theta)}{\partial \theta^2}\right]\right\}^{-1}$$

Replacing θ by $\hat{\theta}$ leads to the estimate $\hat{\sigma}_{\theta}^2$, say of σ_{θ}^2, and an approximate large-sample $100(1-\alpha)$ percent confidence interval for θ is given by

$$\hat{\theta} \pm z_{\alpha/2} \sigma_{\theta}$$

Coefficient of variation: The coefficient of variation $\hat{\theta}$ is asymptotically given by

$$C(\theta) = \sigma_{\theta}/\theta$$

which can be estimated by

$$\hat{C} = \hat{\sigma}_\theta/\hat{\theta}$$

Here $\hat{C}$ is related to the width of the interval $\hat{\theta} \pm z_{\alpha/2}\sigma$ and is therefore a useful measure of the "accuracy" of $\hat{\theta}$.

Bias Consider an estimate $\hat{\theta}$ of θ, suppose that

$$E[\hat{\theta}] = \theta + b_\theta .$$

Then, the $b_\theta = E[\hat{\theta}] - \theta$ is called the bias of $\hat{\theta}$, as an estimate of θ. The quantity b_θ/θ is called the bias of θ.

Several Parameters: Let $x_1, x_2, \ldots, x_n$ be a random sample from $f(x, \underline{\theta})$, where $\underline{\theta}$ is now a vector of parameters $\theta_1, \theta_2, \ldots, \theta_r$. Then if f satisfies the regularity conditions, then, $\underline{\hat{\theta}}$, the vector of the maximum likelihood estimates, is a solution of the r equations

$$\frac{\partial \log L(\underline{\theta})}{\partial \theta_i} = 0 \quad (i = 1,2,\ldots,r).$$

And is asymptotically distributed as a multivariate normal distribution $N(\underline{\theta}, \Sigma_\theta)$, where Σ_θ is an $r x r$ matrix with i,jth element

$$- E\left[\frac{\partial^2 \log L(\theta)}{\partial \theta_i \, \partial \theta_j}\right]^{-1}$$

The matrix Σ_θ is sometimes called the information matrix.

Moment Estimates:

The maximum-likelihood theory is applicable to more general situations than those stated above. For example, the x_i's may have different distributions or x_i's may not be independent but have a joint multinomial distribution. In this case the number of random variables equals the number of unknown parameters, then the maximum-likelihood estimates can usually be obtained by equating each random variable to its expected value and solving the resulting equations for the unknown parameters. This method is called moment estimation and the estimates are called moment estimates.

Estimating a Mean

Let $x_i (i = 1,2,\ldots,n)$ be n independent random variables with known variances σ_i^2 and common mean θ. For the class of estimates of the form

$$\bar{x}_w = \left(\sum_{i=1}^{n} w_i x_i\right)/(\Sigma w_i) ,$$

it is readily shown that $\bar{x}_w$ has minimum variance when w_i is proportional to $1/\sigma_i^2$. In particular if $w_i \sigma_i^2 = a$ say then

$$v[\bar{x}_w] = a/(\Sigma w_i).$$

It can be shown that

$$\hat{v}(\bar{x}_w) = \frac{\Sigma w_i (x_i - \bar{x}_w)^2}{(n-1)\ \Sigma w_i}$$

where $n = \sum_{i=1}^{n} w_i$, is an unbiased estimate of this minimum variance.

Unweighted Mean: If the variances σ_i^2 are unknown, then we can simply use the sample mean

$$\bar{x} = \Sigma x_i/n$$

as our estimate of θ. In this case, it transpires that

$$\hat{v}[\bar{x}] = \frac{\Sigma(x_i-x)^2}{n(n-1)}$$

is an unbiased etimate of $v[\bar{x}]$.

A similar estimate of $v[\bar{x}]$ can also be obtained when x_i's are not independent but correlated. Suppose that

$$\operatorname{cov}(x_i,x_j) = \begin{cases} \sigma_{ij} & j = i+1 \\ 0 & j > i+1 \end{cases}$$

so that the (unknown) non zero covariances are σ_{12}, σ_{23}, ..., $\sigma_{n-1;n}$.

Then

$$v[\bar{x}] = \frac{1}{n^2}\ \{\sum_{i=1}^{n} \sigma_i^2 + 2 \sum_{i=1}^{n-1} \sigma_{i,i+1}\}$$

$$= \frac{1}{n^2}\ \{A + 2B\} \text{ say,}$$

and the problem reduces to finding the estimates of A and B.

Let

$$S_1^2 = \sum_{i=1}^{n} (x_i-\bar{x})^2$$

$$S_2^2 = \sum_{i=1}^{n} (x_{i+1}-x_i)^2$$

where $x_{n+1} = x_1$, then

$$E[S_1^2] = \frac{(n-1)}{n} A - \frac{2B}{n}$$

$$E[S_2^2] = 2A - 2B .$$

and using moment estimation, unbiased estimates of A and B are

$$\hat{A} = \frac{n S_1^2 - S_2^2}{(n-3)}$$

$$\hat{B} = \hat{A} - \frac{1}{2} S_2^2 .$$

Therefore an unbiased estimate of $v[\bar{x}]$ is given by

$$\hat{v}[\bar{x}] = \frac{\hat{A} + 2\hat{B}}{n^2} = \frac{3S_1^2 - S_2^2}{n(n-3)}$$

we note in passing that $\hat{A} > 0$ since

$$nS_1^2 = \frac{1}{2} \sum_i \sum_j (x_i - x_j)^2 > S_2^2 .$$

If x_i's actually have different means θ_i, then

$$E[v[\bar{x}]] = v[\bar{x}] + (3c_1 - c_2)/[n(n-3)]$$

where, $c_1 = \sum_{i=1}^{n} (\theta_i - \hat{\theta})^2$ and $c_2 = \sum_{i=1}^{n} (\theta_{i+1} - \theta_i)^2$.

THE DELTA METHOD:

A useful method used repeatedly in this dissertation for finding approximate means, variances and covariances is demonstrated by the following examples.

Mean Let x_i be a random variable with mean $\theta_i (i=1,2,\ldots,n)$ and suppose we wish to find the mean of some function $g(x_1, x_2, \ldots, x_n)$ $(= g(\underline{x})$, say). Then using the first few terms of Taylor expansion about $\underline{\theta}$, we have

$$g(\underline{x}) \cong g(\underline{\theta}) + \sum_{i=1}^{n} (x_i - \theta_i) \frac{\partial g}{\partial x_i} + \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \theta_i)(x_j - \theta_j) \frac{\partial^2 g}{\partial x_i, \partial}$$

where all partial derivatives are evaluated at $x = \theta$. Therefore taking expected values,

$$E[g(\underline{x})] = g(\underline{\theta}) + b$$

where

$$b = \sum_i \sum_j \frac{1}{2} \operatorname{cov}[x_i, x_j] \frac{\partial^2 g}{\partial x_i \partial x_j}$$

$$= \sum_{i=1}^{n} \frac{1}{2} \ v[x_i] \frac{\partial^2 g}{\partial x_i^2} + \sum_i \sum_j \operatorname{cov}[x_i, x_j] \frac{\partial^2 g}{\partial x_i \partial x_j}$$

<u>Variance</u>: If we ignore the bias b and neglect quadratic terms in the above Taylor expansion, then

$$v[g(x)] \cong E[\{g(x) - g(\theta)\}^2]$$

$$\cong \sum_{i=1}^{n} v[x_i] {\frac{\partial g}{\partial x_i}}^{2} + 2 \sum_i \sum_j \operatorname{cov}[x_i, x_j] \frac{\partial g}{\partial x_i} \frac{\partial g}{\partial x_j} .$$

<u>AN EXACT FORMULA</u>. If x and y are independent random variables then we have the exact relation (Goodman [1960])

$$v[x,y] = (E[x])^2 v[y] + [E(y)]^2 v[x] + v[x] \ v[y].$$

<u>CONDITIONAL VARIANCES.</u>

Let x and y be a pair of random variables. Then, we have

$$E[x] = E_y\{E[x|y]\}$$

and

$$v[x] = \underset{y}{E}\{v[x|y)\} + v_y\{E[x|y]\}.$$

where $\underset{y}{E}$, and so on, denotes taking the expected value with respect to the distrubiton of y. We note that if $E[x|y]$ does not depend on y, then the second term of the last expression is zero. By the delta method

$$\begin{aligned} v[x] &= \underset{y}{E}\{v[x|y]\}. \\ &= \underset{y}{E}\{g(y)\}\ ,\quad \text{say} \\ &= g(\theta) \\ &\cong \{v[x|y]\}_{y=\theta} \end{aligned}$$

where $\theta = E[y]$.

REGRESSION MODELS:

Weighted Linear Regression

Consider the regression line

$$Y_i = \beta_o + \beta x_i + e_i \tag{1.1}$$

where x_i's are constants, the e_i are random variables independently distributed as $N(0,\sigma^2/w_i)$, the weights w_i are known, β_o, β and σ^2 are unknown parameters. Thus

$$\hat{\beta} = \Sigma w_i(Y_i-\bar{Y})(x_i-\bar{x})/\Sigma w_i(x_i-\bar{x})^2$$

and $\hat{\beta}_o = Y - \hat{\beta}\bar{x}$

where $\bar{Y} = \Sigma w_i Y_i/\Sigma w_i$ and $\bar{x} = \Sigma w_i x_i/\Sigma w_i$

Also,

$$v[\hat{\beta}] = \sigma^2/\Sigma w_i(x_i-\bar{x})^2$$

and an unbiased estimate of σ^2 is

$$\hat{\sigma}^2 = \Sigma w_i(Y_i-\bar{Y}-\hat{\beta}(x_i-\bar{x}))^2/(n-2)$$

A $100(1-\alpha)$ percent confidence interval for β can be obtained in the usual manner from the t-distribution, namely

$$\hat{\beta} \pm t_{n-2}[\alpha/2](\hat{\sigma}^2/\Sigma w_i(x_i-\bar{x})^2)^{\frac{1}{2}}$$

There are two cases:

CASE 1. When $\beta_o = 0$ in the regression model (1.1), the least-squares estimate of β is now

$$\tilde{\beta} = \Sigma w_i \; Y_i x_i \; / \; \Sigma w_i x_i^2$$

The corresponding confidence interval for β is

$$\tilde{\beta} \pm t_{n-1}[\sigma/2](\sigma^2 \; / \; \Sigma w_i x_i^2)^{\frac{1}{2}}$$

where

$$(n-1)\;\tilde{\sigma}^2 = \Sigma w_i(Y_i-\tilde{\beta}x_i)^2$$

$$= \Sigma w_i Y_i^2 - (\Sigma \; w_i Y_i x_i)^2/\Sigma w_i x_i^2$$

CASE 2. When $\beta = 0$ in the regression model (1.1), the least-squares estimate of β_o is now

$$\tilde{\beta}_o = \Sigma w_i y_i \ / \ \Sigma w_i = Y$$

with confidence interval

$$\tilde{\beta}_o \pm t_{n-1}[\alpha/2](\tilde{\sigma}_o^2 \ / \ \Sigma w_i)^{\frac{1}{2}}$$

where

$$(n-1)\ \tilde{\sigma}_o^2 = \Sigma w_i (Y_i - \tilde{\beta}_o)^2$$

$$= \Sigma w_i y_i^2 - (\Sigma\ w_i y_i)^2 \ / \ (\Sigma\ w_i) \ .$$

Weighted multiple linear regression

A generalization of model (1.1) is the multiple regression model

$$\underline{y} = X\underline{\beta} + \underline{e}$$

where $\underline{e}$ has a multivariate normal distribution $N(\underline{0}\ ,\ \sigma^2 B)$, X is a known $n x r$ matrix of rank r, B is a known nxn positive definite matrix, and $\underline{\beta}$ and σ^2 are unknown parameters. The weighted least-squares estimates of β obtained by minimizing

$$(\underline{y} - X\underline{\beta})'\ (\underline{y} - X\underline{\beta})$$

with respect to $\underline{\beta}$, is

$$\hat{\underline{\beta}} = (X'B^{-1}X)^{-1}X'B^{-1}\underline{y} \quad .$$

The variance matrix of this estimate is

$$v[\hat{\underline{\beta}}] = \sigma^2(X'B^{-1}X)^{-1}$$

and σ^2 is estimated by

$$\hat{\sigma}^2 = (\underline{y}'B^{-1}\underline{y} - \underline{y}'B^{-1}X\hat{\underline{\beta}})\ /\ (n-r)$$

The confidence interval for β can be obtained in the usual manner.

GOODNESS-OF-FIT TESTS

Binomial distribution

Let $x_1, x_2, \ldots, x_n$ be a random sample from the binomial distribution

$$f(x) = \binom{N}{x} p^x q^{N-x}, \quad (q=1-p, \quad x=1,2,\ldots,N).$$

suppose x_i takes a value x with frequency $f_x (\Sigma_x f_x = n)$ then it is readily shown that the maximum-likelihood estimate of p is

$$\hat{p} = \bar{x}/N = \sum_{x=o}^{n} x f_x / nN$$

and the expected frequencies E_x, are given by

$$E_x = n \binom{N}{x} \hat{p}^x \hat{q}^{N-x} \quad , \quad x = 1,2,\ldots,N \, .$$

Since the joint distribution of the random variables f_x is multinomial with $N+1$ categories, the goodness-of-fit statistic for testing the appropriateness of the binomial model is

$$T_1 = \sum_{x=o}^{N} (f_x - E_x)^2 / E_x$$

where f_x is as defined above. T_1 is approximately distributed as χ^2_{N-1} when n is large.

An alternative test statistic can be obtained by putting the data in the form of a contigency table, namely

x_1	x_2	...	x_n	Σx_i
$N-x_1$	$N-x_2$	...	$N-x_n$	$nN-\Sigma x_i$
N	N	...	N	nN

and carrying out a test for homogeneity. The test statistic is then

$$T_2 = \sum_{i=1}^{n} (x_i - Np)^2 / N \hat{p} \hat{q}$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 / \{\bar{x}(1-\bar{x}/N)\} .$$

This is the so-called Binomial Dispersion Test Statistic and it is asymptotically distributed as χ^2_{n-1}.

We note that $T_2/(n-1)$ is effectively based on comparing observed variance estimate

$$\Sigma (x_i - \bar{x})^2/(n-1) \quad \text{with} \quad \hat{N} \hat{p} \hat{q}$$

an estimate of the expected variance under a binomial model.

Poisson distribution

Let $x_1, x_2, \ldots, x_n$ be a random sample from a Poisson distribution

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \qquad x = 0,1,2,\ldots$$

Then the maximum-likelihood estimate of λ is $\hat{\lambda} = \bar{x}$ and the expected frequences are

$$E_x = \frac{n\, e^{-\hat{\lambda}} \hat{\lambda}^x}{x!}$$

usually the expected frequencies are pooled for $x \geq X$ so as to ensure that

$$n - \sum_{x=0}^{X-1} E_x \quad (= E_{x+} \text{ say })$$

is greater than about 5 (though a value as small as 1 can usually be tolerated if $X \geq 4$) and the goodness-of-fit statistic

$$T_1 = \sum_{x=0}^{X-1} \{(f_x - E_x)^2/E_x\} + \frac{(f_{x+} - E_{x+})^2}{E_{x+}}$$

is then approximately distributed as χ^2_{X-1} .

Alternatively, we can use the <u>Poisson Dispersion Test</u> (Purey and Mead (1979)). The test statistic of which is given by

$$T_2 = \sum_{i=0}^{n} (x_i - \bar{x})^2/\bar{x} = \sum_{x} f_x (x-\bar{x})^2/\bar{x},$$

is asymptotically distributed as χ^2_{n-1}. Since the mean of a Poisson variable equals its variance, $T_2/(n-1)$ can be regarded as a statistic for comparing the observed variance estimate with the estimate, $\bar{x}$, of the expected variance under a Poisson model.

In general, T_2 will provide a more sensitive test than T_1, though when the underlying distribution is not Poisson a comparison of the f_x and E_x may give some idea as to the form of depature from Poisson. Also T_2 can be used for quite small values of n, ($n > 20$, or if $\bar{x} > 1$, $n > 6$ Kathirgumatamby (1953)), while T_1 requires a much larger sample size n in order to ensure that $E_x \geq 5$ for several values of x.

Multinomial Distribution with N unknown

Let $y_1, y_2, \ldots, y_k$, have a multinomial distribution

$$f(y_1, y_2, \ldots, y_k) = \frac{N!}{(\prod_{i=1}^{k} y_i!)(N-r)!} (\prod_{i=1}^{k} p_i^{y_i}) p_{k+1}^{N-r}$$

where $r = \sum_{i=1}^{k} y_i$, $p_{k+1} = 1 - \sum_{i=1}^{k} p_i$

we wish to test the hypothesis H_o that $p_i = p_i(\underline{\theta})$ $(i = 1,2,\ldots,a)$ where $p_i(\underline{\theta})$ is a function of t unknown parameters $\theta_1, \theta_2, \ldots, \theta_t$. When N is known, we can test it using the standard multinomial goodness-of-fit statistic

$$T_1 = \sum_{i=1}^{k} (y_i - N\tilde{p}_i)^2 / N\breve{p}_i + (N-r-N\breve{p}_{k+1})^2 / N\breve{p}_{k+1}$$

where

$$\tilde{p}_i = p_i(\tilde{\underline{\theta}}), \quad \tilde{p}_{k+1} = 1 - \sum_{i=1}^{k} \tilde{p}_i,$$

$\tilde{\underline{\theta}}$ is the M.L.E. of $\underline{\theta}$ for the multinomial model above.

When N is unknown, then we can work with the conditional multinomial distribution.

$$f_2(y_1,y_2,\ldots,y_k|r) = \frac{r!}{\prod_{i=1}^{k} y_i!} \prod_{i=1}^{k} \left[\frac{p_i}{1 - p_{k+1}}\right]^{y_i}$$

and use

$$T_2 = \sum_{i=1}^{k} \{y_i - rp_i/(1-\hat{p}_{k+1})\}^2/r\,\hat{p}_i(1-\hat{p}_{k+1})$$

$$= \sum_{i=1}^{k} (y_i - \hat{N}\,\hat{p}_i)^2 / \hat{N}\,\hat{p}_i \quad ,$$

where

$\hat{N} = r/(1-\hat{p}_{k+1})$, $\hat{p}_i = p_i(\hat{\underline{\theta}})$ and $\hat{\underline{\theta}}$ is the M.L.E. of $\underline{\theta}$ for f_2.

It can be shown that when H_o is true, T_2 is asymptotically distributed as χ^2_{k-t-1} as $N \to \infty$.

By solving the equations $\partial \log f_1/\partial\theta_j = 0$ $j = 1,2,\ldots,r$ and $\nabla \log f_1 = 0$ (∇ denotes backward difference with respect to N), we find that when N is unknown, $\hat{\underline{\theta}}$ and $\hat{N}$ are close to the maximum-likelihood estimates of $\tilde{\theta}$ and $\tilde{N}$ for the model f_1.

SOME CONDITIONAL DISTRIBUTIONS

Poisson Variables

If x_1 and x_2 are independent Poisson random variables with means θ_1 and θ_2 respectively, then it is readily shown that the distribution of x_1 conditional on $y = x_1+x_2$ is binomial, namely

$$f(x_1,|y) = \binom{y}{x_1} p^{x_1} q^{x_2}$$

where $p = \theta_1/(\theta_1+\theta_2)$.

Conversely, if x_1 and y are a pair of random variables such that the conditional distribution of x_1 given y is binomial with parameters y and p, then y is a Poisson with mean λp.

Multinomial Variables

Let $x_1,x_2,\ldots,x_k$ have a multinomial distribution

$$f(x_1,x_2,\ldots,x_k) = \frac{n!}{\prod_{i=1}^{k} x_i!} \prod_{i=1}^{k} p_i^{x_i}$$

then the joint marginal distribution of x_1 and x_2 is

$$f_1(x_1,x_2) = \frac{n!}{x_1!x_2!(n-x_1-x_2)!} p_1^{x_1} p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}$$

If $y = x_1+x_2$ then y has probability function,

$$f(y) = \binom{n}{y} (p_1+p_2)^y (1-p_1-p_2)^{n-y}$$

and the condition probability function of x given $y = x_1 + x_2$ is

$$\binom{y}{x_1} \left[\frac{p_1}{p_1+p_2}\right]^{x_1} \left[\frac{P_2}{P_1+p_2}\right]^{x_2}.$$

CHAPTER TWO

LITERATURE REVIEW AND STATEMENT OF THE PROBLEM

2.1 LITERATURE REVIEW

We shall start by giving the history of a closed population (single marking). The structure is extremly simple. A closed population of unknown size N is under study, n_1 individuals of which are marked and released. From this population, a sample of size n_2 is taken at a single instant over a period of time, m_2 of which are found to be marked.

The first recorded use of this technique is due to Laplace (1786). He estimated the population of France by recording the number m_2, of births in some parishes of known population n_2, whose names were recorded amongst the n_1 names in the birth registrations for the whole country. Petersen (1886) first suggested the use of records of the proportion of marked individuals in the study of fish population. When we spread the labelled fish over the whole fishing ground, we may with some reason suppose that, proportionally, "as many of unlabelled fish which are living there will be caught as those that are labelled." Then intuitively the proportion of marked individuals should be the same in the sample as in the population, that is $m_2/n_2 = n_1/N$. This leads to the Petersen estimate $\hat{N} = n_1 n_2/m_2$. The first recorded use of Petersen's idea is due to Dahl(1919). Lincoln (1930) used bird returns from sportsmen to estimate the size of the wild fowl population of North America. Jackson (1933) independently applied the same method to an insect population. Karl Pearson (1928) examined the theory of this type of sampling scheme. He critized Laplace's formulation in terms of an urn model, and proposed a new solution based on inverse probability on the a prior assumption that every value of $N \geq n$ is equally likely.

If we assume that: the marks are permanent, sampled animals are classified correctly as marked and unmarked, population is closed and that every animal has the same chance of capture, then m_2 is distributed as hypergeometric, given N, n_1, n_2. The properties of $\hat{N}$ are discussed fully by Chapman (1948, 1951) and an equation given for confidence limits for N. Although $\hat{N}$(or strictly its integer part) is the maximum likelihood estimate of N, so that its large sample properties are assured, its behaviour in small samples may be less satisfactory. Notably, because of the non-zero probability that $m_2=0$, it has an infinite bias. Chapman (1951) proposes $\hat{N}_c=[(n_1+1)/(m_2+1)] - 1$ as a more satisfactory estimate, and gives a table of the sample size required, for various N, to ensure that the bias be certainly less than 1. A table of exact and approximate percentage bias of $\hat{N}_c$ for $N=100$, $n_1 = n_2$, is given by Robson and Regier (1964).

Chapman shows also that, for values encountered in Practice, $\hat{N}_c$ has a smaller expected mean square error than $\hat{N}$.

However, since N is usually fairly large, the hypergeometric distribution may be approximated by a binomial, Poisson, or normal distributions. Chapman (1948) suggest the following criteria for approximating the hypergeometric distribution by binomial, Poisson, and normal distributions:

for $n_2 \leq 500$: $m_2/n_2 \leq 0.1$: Poisson
$m_2/n_2 > 0.1$: Binomial
for $500 < n_2 \leq 1000$: $m_2/n_2 \leq 0.075$: Posson
$m_2/n_2 > 0.075$: Normal
for $n_2 > 1000$: $m_2 \leq 100$ or $m_2/n_2 \leq 0.05$: Poisson
Otherwise : Normal.

Robson and Regier (1964) use the normal approximation when N>100. Admas (1951) suggests the use of the Poisson approximation when $N \geq 25$, and provides charts for reading off confidence limits, based on theory developed by Ricker (1937). Since Poisson and normal distributions are themselves approximations to the binomial, it is usual for theoretical discussion to be based on a binomial distribution for the number of recaptures of marked animals. Confidence intervals for estimates from the binomial distribution may be obtained by reference to the charts by Clopper and Pearson (1934). The use of normal approximations was also suggested by DeLury (1951), who derived confidence intervals for N on the assumption that m_2 is normally distributed with mean n_1n_2/N, variance $(n_1n_2/N)(1-n_1/N)$: thus $(n_1n_2/\{\sqrt{m_2(1-m_2/n_2)}\}$, $n_1n_2/\{m_2-1.96\sqrt{m_2(1-m_2/n_2)}\})$ is the 95 percent confidence interval for N. Gaskell and Parr (1966) introduce Bayesian methods to the binomial model. Having shown that what they regard as the ideal prior distribution of N, $f(N) \propto N^B e^{-\alpha v}$, leads to intractable algebra, they consider f(N) constant and $f(N) \propto 1/N^2$ as 'extreme' prior distributions 'between' which their optimal prior must lie. For these distributions $\hat{N} = n_1(n_2-1)/(m_2-2)$ and $n_1(n_2+1)/m_2$ respectively, so that they recommend the use of the 'intermidiate' $\hat{N}=n_1n_2/(m_2-1)$. Questions of whether the prior distributions are really 'extreme', and what inference is to be drawn if one recapture is made, render this estimate unacceptable (Cormark-1968); whatever the merits of Bayesian inference.

Bailey (1952) shows that $\hat{N}=n_1n_2/m_2$ has a positive bias of order $1/m_2$. Thus in long run the size of the population will be overestimated. He proposes (1951,1952), the modified estimate $\hat{N}_B=n_1(n_2+1)/(m_2+1)$ with bias of order e^{-m_2}, its variance is given by $n_1{}^2(n_2+1)(n_2-m_2)/(m_2+1)^2(m_2+2)$. The difference between $\hat{N}_B$ and N_C is negligible.

Bailey suggests an estimate $\hat{N}_I = n_2(n_1+1)/m_2-1$, with variance $(n_1-m_2+1)(N+1)(N;n_2)/(m_2+2)m_2$ if sampling without replacement is assumed. Assuming binomial model, n_1n_2/m_2 is an unbiased estimate of N, with variance

$N(N-m_1)/m_2$ which may be estimated unbiasedly by $n_1{}^2 n_2(n_2-m_2)/m_2{}^2(m_2+1)$ (Chapman (1952)). Chapman uses the normal approximation to set up confidence intervals and tests for N, but suggests that the Poisson approximation will be more useful if both m_2 and n_1/N are small. For inverse sampling from a Poisson distribution it is well known that $2n_1n_2/N$ is distributed as χ^2 with $2m_2$ degrees of freedom. This well-tabulated distribution permits confidence limits for N to be easily constructed.

Chapman (1952) shows that the inverse sampling method gives a more efficient estimate of N with less average effort than can be obtained by direct sampling. However, if the experimenter knows absolutely nothing about N, he may, by an improper choice of n_1, n_2, give himself a sampling scheme which in practice cannot be carried out: the variation is extremly large. This difficulty may be partly overcome by devising the inverse sampling to stop when a predetermined number of unmarked individuals have been caught. Chapman (1952) shows that no strictly unbiased estimate exists: $\hat{N}=n_2\{(n_1+1)/(m_2+1)\}-1$ has a bias less than unity for samples for which $n_1(n_2-m_2) > N \log N$. The variation in N is much reduced by this scheme. Despite these theoretical advantages, inverse sampling has been little used in practice (Ricker, 1958). Czen Pin (1962) shows that, for a loss function of the form $(\hat{N}-N)^2/N^2$, a minimax estimator of N exists, given by $n_1 n_2/(m_2+1) + b$, where (if N_o is the smallest allowable value of N)

$$N_o(1-\sqrt{1-m_2 n_1/N_o})/(m_2+1) \leq b \leq N_o(1+\sqrt{1-m_2 n_1/N_o})/(m_2+1).$$

Zubrzycki (1963) shows that such estimators with $b < N_o(m_2+1)$ are inadmissible.

The decision as to when to stop sampling may be made according to a rule other than a fixed n_2 and m_2 discussed above. Chapman (1954) considers a series of samples of predetermined sizes n_i (which are not returned to the population), sampling being stopped as soon as a total number, m, of marks have been recovered. If N is large, and a Poisson distribution is assumed, $\hat{N} = n_1 \Sigma n_i/m_2$ is asymptotically a minimum variance unbiased estimator with variance N^2/m_2. Knight (1965) discusses the feasibility of estimating 1/N if sampling stops when either m_2 or n_2 attain a pre-assigned value, whichever happens first: he gives rules for choosing these values in such away that the variance of $1/\hat{N}$ is less than any assigned value.

One theoretical difficulty in estimating N is that the distribution of $n_1 n_2/m_2$, or the modifications proposed by Bailey or Chapman, is far from symmetrical. Thus the confidence limits obtained from Clopper-Pearson curves will be biased. One way out of this difficulty is to estimate the reciprocal 1/N. As Leslie (1952) point out, under the binomial assumption m_2/n_1 is an unbiased maximum likelihood estimate of 1/N, and the confidence limits may be obtained for $1/\hat{N}$ by Clopper-Pearson charts or a normal approximation. Since the distribution of $1/\hat{N}$ is more symmetric than that of $\hat{N}$, this procedure should lead to confidence limits of N less baised than the methods described earlier in this section. One disadvantage of using 1/N is seen if sub-populations are estimated separately, and it is desired to add the estimates together.

If we assume that (i) there is neither recruitment nor immigration both of which affect unmarked animals only and death and emmigration affect marked animals equally and (ii) if there exists strata within the population such that, by size, behaviour or any other variation, different strata have different chances of recapture, then the marked animals belong to these strata in exactly the same proportions as the occurence of the strata in the whole population, then the estimates considered above retain their properties of consistency and unbiasedness. However their variances are now dependent on further unknowns, death rates and strata sizes, about which the investigation provides no information. Chapman (1952) shows that the modification is slight unless mortality is excessive. Chapman and Junge (1956) assert that, under the binomial assumption, a death rate, identical for marked and unmarked animals, does not affect the variance of the Petersen estimates. This is, however, true only if both marked and unmarked populations are large enough for the death rate to be truly deterministic. With death rate $(1-\phi)$, the population to be sampled must be assumed to contain $N\phi$ members of which $n_1\phi$ are marked.

Chapman and Junge (1956) have investigated a possible modification of the last assumption. The population is assumed to consist of a number of distinct strata which do not mingle uniformly. These may be 'tribes' differentiated by geographical locality It is known to which stratum any sampled individual belongs at the moment of sampling, but its history is unknown unless it is already marked. Estimates are now required for population migration between strata as well as for total population size. Using suffixes i,j to represent the strata at the times of marking and sampling respectively, Chapman and Junge show that $\sum_j \hat{N}_{.j}$, where $\sum_j m_{ij} \hat{N}_{.j}/n_{.j}=n_i$ for all i, is a consistent estimator of $N_{..}$ if it is assumed that sampling is random within each stratum, individuals in each stratum are properly mixed after moving, individuals move independently one stratum to another, and the probability of such a move is independent of marking. (A suffix replaced by a. (period) has been summed over: thus, for example, while m_{ij} is the number of individuals marked in stratum i, recaptured in stratum j, $m_{.j} = \sum_i m_{ij}$ is the total number of individuals recaptured in stratum j). Under the same assumptions neither the standard Petersen estimate nor an estimate $\sum_i \sum_j n._i\, n_j.m_{ij}/m_i.m_j$, proposed earlier for this situation by Schaefer (1951b), is consistent unless the assumption that there exist strata within the population, such that by behaviour, size or any other variation, different strata have different chances of recaptures holds strictly. Estimates are also given for migration between strata: $\hat{N}_{ij} = m_{ij}\, \hat{N}_i.\hat{N}._j/n_i.n._j$. This situation was studied further by Darroch (1961). If there are at least as many strata in the population at the time of marking, as at the time of recapture, maximum likelihood estimates are obtained without any assumption as to the movement of unmarked animals. If not, then it is necessary to assume that unmarked animals move between strata with the same probabilities as marked ones. If the movement of individuals is not independent, the estimate remain consistent.

More important than these theoretical problems of trying to extract from the data, under the assumptions, the last scrap of information, is the problem of how departures from the assumptions affect the estimate of N, and how, if at all, the estimate may be adjusted to allow for such departures. Indeed, as Schnabel (1938) says: "since the assumptions of random sampling and constant populations are only rough estimates to the actual situation in taking fish census, small differences between the results of various methods are not important."

Report on actual experiments with fish populations illustrating the breakdowns of these assumptions which occur in practice will be mentioned later.

Rupp (1966) has pointed out that the Petersen procedure can be regarded as a particular instance of a survey removal method of estimating population size. In survey-removal study, originally suggested by Kelker (1940, 1944), the change in ratio of the observed frequencies of occurence of two distinguishable classes of individuals, before and after a period during which known numbers of the two classes are removed from the population, provides information about the size of the population if markedly different numbers are removed from the two classes. Theory of this method, allowing for mortality is developed by Chapman (1954), Lander (1962), Hanson (1963) and Chapman and Murphy (1965). In Petersen-type study the initial ratio of marked: unmarked is zero. The final sample ratio is $m_2/(n_2+m_2)$,-n_1 marked animals having been removed before the final sample. Paulik and Robson (1969), in a unified treatment of the methods, study the effect of N of an unobserved removal of C_m, C_u animals from the two classes during the period before the final sample. These C_m, C_u cover mortality, immigration, and emigration (not necessarily the same for each class), and this formulation permits any knowledge of these unobserved removals obtained from other sources to be used in the estimation of population size.

The effects of recruitment may be eliminated in some cases by restricting the counts to suitable age groups - if these are recognizable and non-overlapping (Ricker (1958)). Even if age groups overlap, subsidiary information on growth rate can be used to eliminate the effect of recruitment. The possibility that the process of marking in itself introduces an extra cause of mortality to marked fish may be investigated by using different types of mark. If one of these involves more mutilation than another, and yet both types are recovered equally in the subsequent sample, this provides evidence that marking does not contribute directly to the mortality. This does not cover additional mortality due to purely handling the animals, which often have to be removed from their natural environment.

The assumption that there is no loss of marks, may be investigated by fixing two marks to some individuals. If it is assumed that losses of single marks are independent then the number of individuals in the sample which have lost none or one of the two marks fixed on them provides information on the rate of loss. Thus, if all animals released bear tow marks and m_D are recovered still bearing both marks, m_s with a loss of one mark, the loss rate may be estimated by $m_s/(m_s+m_D)$, and the population size by $4\, m_D m_s n_2/(2m_D+m_s)^2$. A more general model than this is examined by Gulland (1963). Data on plaice recorded by Beverton and Holt (1956) suggest that the rate of detachment of external tags increases with time initially (with increasing wear), but then decreases (as the tags become imbedded).

If the recapture sampling is continued over a period of time, any dilution of the population - by recruitment or immigration - should become apparent through a progressive decrease in the proportion of marked animals in the recaptured samples. Jackson (1937) in his 'positive method' adopted this procedure, and Bailey (1951) provides a mathematical formulation by which dilution, if assumed to have a specific mathematical form, could be estimated. Parker (1955) suggested plotting, as a function of time, m/n, or $\log(m/n)$ or $\arcsin(\sqrt{m/n})$, whichever provides the best straight line, and was more satisfactory.

Knowledge of the effort expended in sampling the population allows the estimate of the population size to be obtained in a different way, since the numbers obtained per unit effort will diminish in successive samples (Leslie and Davis, (1939)). If the effort is the same at each sample, the expected catch in the sample, $E[C_i]$, is related to the probability p of an individual being caught, by the relation $E[C_i] = Np(1-p)^{i-1}$. A weighted regression of log C_i against $(\log(Np)+(i-1)\log p)$ gives estimates of p and N. Moran (1951) pointed out that C_i's are not independent and derived maximum likelihood estimates for N and p. Skellam (1962) proposed direct numerical extrapolation to the curve of catch against time, suggesting that a transformation of the time variable t to a form $b/(b+t)$ (where b is arbitrary) enables the extrapolation to be carried out more accurately. Chapman (1954), following DeLury (1951), proposes an unweighted regression of C_i - in this case the catch per unit effort - on K_i, the total catch removed before the ith sample. This give $\hat{N} = \hat{K} - \bar{C}\ \ (K_j-\bar{K})^2/\Sigma C_i(K_i-\bar{K})$. A comprehensive study of these removal methods is given by Zippin (1956). Chapman shows further that this idea can be suitably combined with capture-recapture experiment with a single release of n_o individuals, and successive periods of recapture yielding n_i individuals of which m_i are marked. If n_i and m_i given n_i, are both assumed to have a Poison distribution, maximum likelihood equations are given for N and q the probability that a unit of effort captures one member of the population, in terms of the efforts f_i expended (Chapman (1954)).

The Petersen estimate or Lincoln Index, perhaps modified for bias, has been much used as providing a simple and intuitively reasonable estimate of population size. However, the universal lack of faith in the assumptions together with the lack of internal evidence as to their applicability led to consideration of models for more complicated sampling procedures.

Discussion of the study of a closed population by the use of marked individuals would be incomplete without mention of some proposed methods based on the number of times individuals are recaptured. Craig (1953) suggested that if the total sampling period is subdivided into a large number of short intervals and each individual is equally likely to be caught in any short interval, the number of recaptures should have a Poisson distribution, truncated at zero since it is not known how many individuals are never captured. From this, N and the Poisson parameter can be estimated by maximum likelihood or by moments. Darroch (1958) shows that Craig's use of truncated Poisson distribution can not serve as a probability distribution of any capture-recapture experiment since it implies that both the total effort expended and the number of different individuals seen be fixed in advance, was impossible. Taylor (1966) reports that for bird population the number of times an individual is recaptured is not well fitted by a trancated Poisson distribution. He suggests a negative binomial distribution.

Mc Donald and Palanacki (1989) considers the problem of estimating the size of a small population based on the results of a certain type of capture-recapture experiment. They give seven methods of constructing confidence intervals for the population size. Among these methods is the 'adhoc' method which for N = 500 tends to give actual confidence levels which are close to the desired level and tends to give shorter intervals than the other methods when the probability of capturing individuals is small.

We now consider a situation where marked animals are released into the population on more than one occasion. As with Petersen method described earlier these marked animals are usually themselves samples from the population under study, but this need not be the case provided due assumptions about marked and unmarked individuals, are satisfied. The earlier studies of this situation took no account of the possibility that a particular individual may be recaptured on more than one occasion. At sampling time i, $i = 0,1,2,\ldots,k$, the data recorded are n_i, the size of the sample, and m_i, the number of previously marked animals in the sample. The (n_i-m_i) unmarked individuals are then marked and all the n_i returned to the population. The first sample serves only to provide a pool or $n_o(=M_1)$ marks in the population. There are k subsequent recapture samples.

The first study (Schnabel, 1938) assumed that M_i, the total number of marked animals in the population immediately before the ith sample in taken, are known parameters of the population. The situation is then identical with a series of Petersen estimates which have to be combined to yield a single estimate of the population size N. The situation remains to decide with what weights the estimates should be combined. Under the assumption of binomial sampling on each occasion, Schnabel (1938) proposed the estimate $\hat{N} = (\Sigma n_i m_i)/\Sigma m_i$ as an approximation to the solution of the maximum likelihood equation:

$$\Sigma m_i = \Sigma \frac{(n_i - m_i) M_i}{(\hat{N} - M_i)}$$

but gave no consideration to the precision of her estimate. If M_i/N is small, and m_i is assumed to be a Poisson variable, $\hat{N}$ is the exact maximum likelihood estimate. Schumacher and Eschmeyer (1943) proposed $\Sigma n_i M_i^2/\Sigma m_i M_i$ and suggested that its variance be obtained from the mean square error about the regression line m_i/n_i against M_i. The estimate of variance has the advantage of referring to $1/\hat{N}$, which is more symmetrically distributed than $\hat{N}$ itself. Hyne (1949) proposed the same method, apparently independently of Schumacher and Eschmeyer, commenting that it has an advantage over the removal method, in that it is not so severely upset by a day-to-day fluctuation in the probability of capture. The different weightings for each point on a graph of n_1/n_2 against n_1 through the origin which are implied by the various estimates were studied by DeLury (1958). The weights for the maximum likelihood solution are $n_2\hat{N}/n_1(1-n_1/\hat{N})$ those for the Schumacher and Eschmeyer solution, preferred by DeLury, are simply n. Ricker (1945b) asserts that Schumacher and Eschmeyer's estimate attains maximum efficiency when half the population is marked; Schnabel's maximum occurs when a negligible proportion is marked. They have equal efficiency when the proportion of marks is 1/4. In an earlier paper DeLury (1951) had given an iterative solution for the maximum likelihood equation. Using Schnabel's estimate as a first approximation, a new weighted estimate $\Sigma W_i n_i M_i / \Sigma W_i m_i$, is constructed with weights $W_i = 1/(1 - M_i/N)$. Gilbert (1956) suggests that the difference between the hypergeometric and binomial distributions can be allowed for by dividing each term in the binomially based likelihood equation by a finite population factor $(1-n_i/N)$. Thus instead of solving the equation $\underset{i}{\Sigma}(m_i n_i - Nm_i)/(N_i - M_i) = 0$, one solves the equation $\underset{i}{\Sigma}(M_i n_i - Nm_i)/(N-n_i)(N-M_i) = 0$. A small example gives results very similar to the binomial model.

Chapman (1952) points out that, although the M_i are known at the interest of taking the ith sample and are thus conditional parameters, they are not known except for M_1 when the 'a priori' probability model is constructed. The above models are thus incorrect. Using the hypergeometric model, Chapman shows that Σm_i is a sufficient statistic for N, and that the maximum likelihood estimate of N is the solution of the equation

$$\prod_{i=0}^{k}(1 - n_i/N) = 1 - M_{k+1}/N$$

where M_{k+1} is the number of different individuals seen during the experiment. The mean square error of this estimate is given by Dorroch (1958) as,

$$\left[1/(n-\rho) + K/_N - \Sigma\, 1/(N-n_i)\right]^{-1},$$

where $\rho = E(M_{k+1})$, this error being of order N.

An alternative model regards the n_i (including n_o) as random dependent on parameters p_i, the probability that any individual is caught in the ith sample. From this model, Darroch (1958) derives the same maximum likelihood equation for $\hat{N}$, and gives an approximate formula for the variance of $\hat{N}$ as

$$N/\left|1/Q + k - \sum_{i=0}^{k} 1/(1-p_i)\right| \quad \text{where} \quad Q = \pi_{i=0}^{k} (1-p_o)$$

is the probability that an animal is never captured. This variance can not be obtained from asymptotic maximum likelihoods theory since the sample size is not constant and the likelihood not the product of likelihoods of the individuals. It is obtained by a standard δ-technique. Dorroch's new model requires that the effort to be expended in each sample be pre-assigned. In practice the applicable model will be determined by whether it is the difficulty of catching animals or the labour of marking them which limits the numbers caught. Since the same numerical estimate of N is obtained, and the variances are both of order N, it cannot make much practical difference which model is used. Confidence limits may be obtained by assuming that M_{k+1} is normally distributed. Seierstad and Mysterud (1965) suggest a version of Darroch's model with $p_i = p$. Their proposed estimator of p is insufficient in that it does not count the number of samples in which these individuals are not seen. No reason is given why the fully efficient estimator for p, based on the total number of sightings is not used.

The maximum likelihood estimate, to which

$\sum_{i=1}^{k}\sum_{j=i+1}^{k} n_i n_j / \sum_{i=1}^{k} m_i$ is the first approximation (Chapman (1952)), is also moment estimator. Czen Pin and Dzan Dzai (1961) show that Σm_i has a limiting Poisson distribution, and suggest the use of (Σm_i+1) as the denominator of the moment estimator $\sum_{i=1} \sum_{i=i+1} n_i n_j / \sum_{i=1}^{k} m_i$, instead of Σm_i. The basis for this suggestion, proposed also by Chapman (1954) in a different context, is that the reciprocal of a binomial or Poisson variate + 1 is an almost unbiased estimator of the reciprocal of the prameter. Czen Pin and Dzan Dzoi give confidence limits for N based on normal approximation to this Poisson variate - An unbiased estimator based on the sufficient statistic will have a smaller variance than the almost unbiased estimator obtained by merely averaging the successive modified Petersen estimates $[(n_i+1)(m_i+1)/(m_i+1)]-1$. However, in situations where the binomial assumption is appropriate, the Petersen estimate, despite the loss of information from the lack of weighting, has the conciderable merit of simplicity. It remains almost unbiased whether the M_i are regarded as parameters or as random variables. If $n_i M_i/N$ are too small for the binomial assumption to be valid, Chapman (1952) suggests $\hat{N}_p = \Sigma n_i m_i/(\Sigma m_i+1)$ as a suitable modification reducing the bias, of Schnabel's original estimate. The variance of $\hat{N}_p$ is approximately

$$(N^2/\Sigma n_i m_i)(1+2N/\Sigma n_i m_i)$$

Chapman also points out that, under the strict assumptions of closed population and equal catchability of marked and unmarked animals, knowledge of the history of capture of an animal contribute no additional information towards the estimation of N. Such knowledge does, however, allow these assumptions to be tested. If expected values are sufficiently large, a standard χ^2 - goodness of fit test can be carried out, although Chapman points out that there is a lack of independence. He conjectures that this will have negligible effect. He suggests as an alternative, the use of non-parametric test proposed by Moore and Wallis (1943) which examines the signs in a sequence of successive series of differences of observations. Apparently, no example of the use of this test is recorded in the literature.

Recently Chapman and Overton (1966) have extended consideration of Chapman's (1952) nearly unbiased estimate $\hat{N}_p$. Each m_i has approximately a Poisson distribution with parameter $n_i m_i/N$, at least if the number of recaptures is fairly small. To test the difference between two different populations we may use the fact that, if x_o and x_1 are Poisson variates, x_o is a binomial variable for fixed (x_o+x_1). An example based on data of Nelson (1960) illustrates the manner in which such a significance tests may be carried out. It is possible to do appropriate calculations on the power of the test before any samples are taken so that the size of the experiment required to detect, with appropriate significance level, a pre-assigned difference in size of population can be calculated beforehand. In view of all the assumptions required, I doubt whether it is wise to attempt to discuss the difference between two populations in terms of a significant test.

As in a single stage census, sampling may be continued at each stage until a predetermined number of marked animals have been captured. The m_i are fixed and the n_i are random variables. An unbiased estimate is now easily found, being the unweighted mean of the corresponding estimates from the inverse sampling. Thus, $\hat{N}_1=1/K\Sigma(n_iM_i+1/m_i-1)$, with approximate variance $\{N^2\Sigma 1/m_i\}/k^2$. Chapman (1952) derives these results and goes on to discuss how to choose the parameters at one's disposal, $n_o,k,m_1 \ldots\ldots\ldots,m_k$. The obvious aim at achieving a fixed precision of estimation while minimizing the effort expended, minimizing $E(\sum_{i=O}^{k} n_i)$ subject to constant $\hat{\sigma}_N/N$, leads to a very complex algebraic problem. Chapman provides guidance in this problem in the form of a table of properties of a number of simple designs. Unfortunately interesting cases with small m_i break the mathematical assumptions and so cannot be considered. An increasing sequence of m_i spreads the effort most evenly; constant m_i seems to provide the maximum precision for the same expected sample size.

Since direct censuses contain the awkward possibility that not enough marked animals are captured to permit reasonable estimates of N, and inverse censuses have a similar physically imposed restriction, namely, that it may not be possible to continue sampling until the pre-assigned number of marked animals have been caught, it is clear that optimal sampling procedures must be sequential. Urn models for different sequential schemes - one at a time, several at a time, single and multiple markings - were introduced by Cox (1949). Chapman (1952) considers in the direct case, the number of recapture samples k as a random variable to be determined by the course of the sampling. He takes n_i as fixed. Under the assumption that the m_i are Poisson variates, a standard type of sequential probability ratio test can be constructed. For any valid study of optimality and expected number of samples required to be taken the m_i have to be independent, a consideration not satisfied in this case. Sequential approach was extended by Goodman (1953). A series of samples of predetermined sizes n_i are to be drawn, sampling to stop as soon as a total of m marked individuals has been captured. A further extension by Chapman (1954) did not restrict the replacement of marked individuals in the population to those captured in previous samples. Chapman showed that $2\ \Sigma n_i M_i/N$ is asymptotically distributed as χ^2 with 2m degrees of freedom, so that $\Sigma n_i M_i/m$ with variance N^2/m, is the asymptotic minimum variance unbiased estimate of N. Darroch (1958) considers a special case of Goodman's sequential census in which each sample consists of a single individual. For this case a unique unbiased estimator, with minimum variance, exists for N. If n samples have to be taken to achieve the recapture of m marked individuals, the estimate is given by the ratio $\sigma_n^{(n-m)}/\sigma_{n-1}^{(n-m)}$ where $\sigma_s^{(r)} = \Delta^r(o^s)/r!$, a Starling number of the second kind. Other stopping rules for this one-at-a-time census have been considered by Samuel (1943). He suggests as a working approximation to Darroch's estimate the value n/w, where w is the solution of the equation $(1-e^{-w})/w = (1-m/n)$. Tables of the function $(1-e^{-w})/w$ are given, for example, by Pearson (1934). Boguslavsky (1956) discusses estimation of N for small populations in which a number of successive observations have yielded only marked animals.

Overton (1965) discussed the modifications to be made to the Schnabel estimator when known numbers of animals, both marked and unmarked, are removed from the population during the course of the experiment. This removal may be deliberately caused by the experimenter or the result of accidential damage to the sampled individuals. The modification takes the form of a term, which has to be computed iteratively, and is then added to the usual Schnabel estimator.

The assumptions under which the above theory of Schnabel type estimators is valid are the same as for the Petersen case. Since the sampling is usually continued over a longer period than the Petersen-type studies, Ricker (1958) considers the recruitment and natural mortality (and fishing mortality if the population is subject to this pressure) as errors of special importance. Undetected natural mortality seriously affects the multiple sample census (Chapman (1952)). One method of testing for natural mortality, particularly adapted for entomological studies, is to compare for any one recapture sample the Petersen estimate obtained by considering only individuals marked on different occasions (Southwood and Saudden (1956)). By considering the multiple census as a time sequence of single cencuses, each of which gives an estimate of the population size at that time, any change in the population over time may be investigated. This is the basic principle of the models to be considered in the later sections of this dissertation, which includes mortality and recruitment as parameters. Mortality causes the population to decline, and thus the Schnabel estimate of the population size will be less than the Petersen estimate from the first sampling. This was used to measure mortality by DeLury (1951). With greater sophistication, natural mortality may be estimated as that which eliminates any time trend from succesive daily estimate of 1/N (Ricker (1958)). This type of estimation is closely analogous to the analysis of catch curves to give estimates of mortality.

Fienberg (1972) considers the problem from a different angle. The resulting data can be put in a form of an incomplete 2^k contingency table, with one missing cell, that displays the full multiple recapture history of all individuals in the population. Log linear models are fitted to this incomplete contigency table and the simplest possible model that fits the observed cells is projected to cover the missing cell, thus yielding an estimate of the population size.

If sampling contines over an appreciable period of time, the population cannot be assumed closed. Other population parameters for recruitment and mortality (and possibly immigration and emigration) must be included in the model. What is meant by an appreciable period of time depends on the population under study. Insect studies with daily samples for a week have to allow for mortality. Surveys of large mammals over period of a month may not. The early models assumed a deterministic death rate, constant over different periods. With death rates $(1-\phi)$, N individuals become exactly ϕN. The simplest estimate of mortality over a period during which dilution may be ignored is the ratio of two Petersen or Schnabel estimates of population size at the beginning and end of the period.

In a series of papers, Jackson (1937, 1939, 1940, 1948) suggested two sampling schemes which he termed 'positive' and 'negative' methods. The positive method is release on a single occasion, a large number of marked animals, recapture (and re-release) being effected frequently on several occasions. The negative method called for the release of marked individuals on several occasions. The number of recaptures being noted only at one final intensive sampling. This second method was deemed most suitable when unskilled workers were used to carry out the marking. Jackson stated that either the first capture or the recapture should be carried out non-selectively since 'dispersal ....might not be complete in the period between marking and recapture, or individual flies might return to places to which they were specifically attached'. Jackson tried to standardize the sampling effort by analysing not the basic number m_{01} of recaptures, but $y_{01} = m_{01}/n_0 n_1$, where m_{01} is the number of individuals among the n_1 caught in week 1, which were marked in week 0 when n_0 were caught.

The ideas behind the negative method is that the samples released early have been exposed to natural mortality for longer periods than samples released at a later date, and therefore will be represented by fewer individuals in the recapture sample. The death rate can be estimated and used to give an estimate of the number of marked animals alive in the population at the time of recapture sample. An estimate of the population size at this stage follows as usual from $\hat{N}=n_1 n_2/m_2$. Since it is the population at this final time which is being estimated, immigration during the period of sampling is an integral part of the population. There is no problem of allowing for it. A very simple example is given by Ricker (1944). If s_1, s_2 fish are marked and released immediately before the fishing season in two successive years, and during the second year's fishing, m_{12}, m_{22} respectively of the s_1, s_2 are caught, then the mortality between the years, inclusive of that due to fishing in the first year, is estimated by $s_2 m_{12}/s_1 m_{22}$. In the positive method when all animals are released in week, 0, the y_{0t} will decrease with time on account of dilution of the population by unmarked animals. This curve may be extrapolated to time 0 to provide an estimate of population size, and the rate of fall of the curve gives an estimate of dilution rate β. Jackson's estimate of population size, which he attributes to Fisher is

$$\frac{\{y_{01}+ y_{02} + \ldots\ldots + y_{0(k-1)}\}^2}{y_{02} + \ldots\ldots + y_{0k}} - \{y_{01}+\ldots+y_{0(k-2)}\},$$

and a variance formula, due to W.L. Stevens, is also quoted. The dilution rate is estimated by $(y_{01}+\ldots\ldots+y_{0k-1})/(y_{02}+\ldots+y_{0k})$. If β does not appear to be constant, Jackson later (1940), suggested using the estimate provided by y_{02}/y_{01} to perform the extrapolation to provide an estimate of the population size. Identical consideration apply to the negative method.

Bailey (1951) comments on the lack of proper weighting factors in Jackson's (1937) estimates, and develops a maximum likelihood solution. For the 'negative' method with recapture only on the final day, day k, assuming a constant death rate $(1-\phi)$, $s_j e^{-(1-\phi)(k-j)}$ of the s_j animals released on day j will be still alive. The likelihood of the parameters ϕ, N_k is thus proportional to

$$\left[\frac{N_k - \sum_j s_j e^{-(1-\phi)j}}{N_k}\right]^{(N_k - \sum_j m_{jk})} \prod_{j=1}^{k-1} \left[\frac{s_j e^{-(1-\phi)j}}{N_k}\right]^{m_{jk}}$$

From this maximum likelihood equation, estimates $\hat{\phi}$, $\hat{N}_k$ are derived by an iterative solution. Bailey also gives the variance of the estimates. He shows also how maximum likelihood estimates, although more complicated, can be obtained from Jackson's positive method. For an experiment with only 3 sampling occasions (2 releases and 1 recapture, or 1 release and 2 recaptures) Bailey (1952) gives explicit formulae for the estimates and modifications of those which make them almost unbiased.

Jackson suggests two estimates of survival rate, ϕ, namely:

$(y_{0k}+\dots+y_{(k-1)k})(y_{1k}+\dots+y_{kk})$ and

$$\{(y_{0k}+\dots+y_{(k-2)k})/(y_{(2k}+\dots+y_{kk})\}^{\frac{1}{2}}$$

Chapman and Robson (1960) point out that the information used by Jackson's negative method is virtually the age distribution of a single sample, which, under the name of 'catch curve' has long been used by marine biologists. Both of Jackson's estimators have the undesirable property that they can be greater than unity. A modification of Jackson's estimators, $(m_{ok}+\dots+m_{(k-1)k})/(m_{ok}+\dots+m_{kk})$, which was suggested by Heincke (1913), avoids this difficulty, and is infact an unbiased estimator of ϕ in the study of catch curve. To allow for variations in effort the m_{ij} should presumably be replaced by Jackson's y_{ij}. Chapman and Robson give a detailed discussion of the regression techniques behind graphical estimates of ϕ from the relation between the logarithm of the number captured and the time since marking. These remarks apply strictly to the 'catch curve' situation but have considerable relevance to such techniques as used in marking experiments,(Beverton and Holt (1956)).

In most recapture experiments the 'age' distribution of the recaptures - in the sense of time since marking-is truncated at the upper and sometimes fairly severely, since these experiments are not usually continued over long periods. For catch curve for which k is the maximum age on which estimation is to be based, Chapman and Robson (1960) show that the maximum likelihood estimate of ϕ is the solution of:

$$\bar{x} = \phi/(1-\phi) - (k+1)\phi^{k+1}/(1-\phi^{k+1}),$$

where $\bar{x}$ is the mean 'age' of the recaptures. Solution of this is facilitated by a table of the function of k and ϕ, given in Chapman and Robson (1960) and extended in Robson and Chapman (1961). Examples of this method both for single release and multiple release surveys, are given by Paulik (1962). For earlier work on this subject see Haldane (1955). An example is given by Murton (1966).

A model based on the Poisson distribution is given by Parker (1963). In this, removal of captured marks is specifically taken into account. If there is a single release of M_o marked fish into a population, subject to constant absolute recruitment B- compare Chapman's comments below - and instantenious mortality rate x, from which samples of n_i, of which m_i are marked, are subsequently taken, then at time t, the population will consist of

$$N_o e^{-xt} - \sum_{i=1}^{t-1} n_i \, e^{x(i-t)} + B(1-e^{-xt})/x$$

among which the expected number of marked individuals is

$$M_o e^{-xt} - \sum_{i=1}^{t-1} m_i \, e^{x(i-t)} .$$

If m_i may be assumed to be a Poisson variable conditional on $m_1, m_2, \ldots, m_{i-1}$ and n_i are taken as parameters, $\hat{N}_o$, $\hat{x}$ and $\hat{B}$ can be determined by iterative solution of maximum likelihood equations.

Further entomological studies on similar lines to Jackson's were carried out by Dowdeswell, Fisher and Ford (1940, 1949) and by Fisher and Ford (1947). These introduce a new method of grouping and displaying the observation as in the form of a trellis diagram.

$$
\begin{array}{lcccccc}
 & & \multicolumn{4}{c}{\text{Releases on day}} & \text{Total recaptures} \\
 & & \multicolumn{4}{c}{\overbrace{0 \qquad 1 \qquad 2 \qquad 3}} & \\
\text{Recaptures} & & & & & & \\
\text{on day} & 1 & m_{01} & & & & n_1 \\
 & 2 & m_{02} & m_{12} & & & n_2 \\
 & 3 & m_{03} & m_{13} & m_{23} & & n_3 \\
 & 4 & m_{04} & m_{14} & m_{24} & m_{34} & n_4 \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Total} & & & & & & \\
\text{releases} & & s_o & s_1 & s_2 & s_3 \ldots &
\end{array}
$$

These m_{ij} include all individuals seen on both day i and j. The average survival rate ϕ is estimated from the average time interval separating release from observed recapture. From this, the population on day j is estimated by $n_j m_j / \hat{M}_j$, where $\hat{M}_j$ is an estimate of the number marked animals alive at time j. The estimate given by Fisher and Ford is $\hat{M}_j = \sum_{i=1}^{j=1} s_i \phi^{j-1}$, which counts each individual, at any recapture, as often as it has previous marks. Using only the last previous recapture of each mark would be more valid. No estimate of the precision is available. The logic of this estimate of N is the same, given ϕ, as Jackson's estimate as derived by Bailey (1951). A modification to Fisher and Ford's calculation of available marked animals is given by MacLoad (1958) for the case in which recapture sampling is continued until no hope remains of survivors being caught.

Jackson (1948) also considers this type of sampling with varying survival rate ϕ_i. Leslie (1952) gives the estimate of ϕ_i from Jackson's modified recaptures y_{ij} as

$$(\sum_{j=i+2}^{k} y_{ij})/(\sum_{j=i+1}^{k} y_{i+1,j}), \text{ with variance}$$

$$\phi_i(1-\phi_i)/\sum_{j=i+2}^{k} m_{i+1,j}$$

If a model with varying ϕ_i is adopted and an average estimate of survival rate is required, the geometric mean of the $\hat{\phi}_i$ should be given.

Maximum likelihood techniques for estimating death rates, initial population size and dilution rates, were developed in three papers by Leslie and Chitty (1951), Leslie (1952) and Leslie, Chitty and Chitty (1953). The general assumption is that at each instant of sampling several distinct classes of animals are at risk, these classes being distinguished by their previous marking history. If the size of each sample is small it can be assumed that the observations at each instant come from a multinomial distribution, the probabilities of the different classes being expressible in terms of the basic parameters of the model. The overal likelihood is the product of a set of such probabilities.

To understand the reasoning behind this method, consider first, following Leslie and Chitty (1951), the estimation only of the survival rate ϕ from the recaptures of M_o marked individuals released at time 0, recaptured at two further times. Using the notations defined earlier, the expected number of marked individuals falling into various classes are:

	Times	1	2	
	02	$\phi M_0 (m_{01})$	$\phi^2 M_o - \phi m_{01}$	(m_{02})
Marks	12		$\phi(n_1 - m_{01})$	(m_{12})
	012		ϕm_{01}	(m_{012})

The likelihood is thus proportional to:

$$\left[\frac{\phi M_0}{\phi M_0}\right]^{m_{01}} \left[\frac{\phi^2 M_0 - \phi m_{01}}{\phi^2 M_0 + \phi(n_1 - m_{01})}\right]^{m_{02}} \left[\frac{\phi(n_1 - m_{01})}{\phi^2 M_0 + \phi(n_1 - m_{01})}\right]^{m_{12}} \left[\frac{\phi m_{01}}{\phi^2 M_0 + \phi(n_1 - m_{01})}\right]^{m_{012}}$$

for which, by differentiating with respect to ϕ and equating to zero in the usual way, an estimate of ϕ may be obtained. The estimate of ϕ so obtained refers to all time intervals except the last. An estimate of expected life span may also be obtained as $-1/\log\phi$, assuming that the death rate is not age-dependent.

For along chain of samples, the solution of this equation will be extremely difficult   Leslie and Chitty suggest grouping the data to reduce the difficulty, and also so that there shall be a reasonable number of animals in each class. They reject the solution proposed by Jackson (1939, 1948) and by Fisher and Ford (1947) which effectively forms the distribution of marks observed, each individual being counted as many times as it has the marks. In this method, the expected number of marked individuals falling into various classes now become

	Times	1	2
Marks	0	$\phi M_0(m_{01})$	$\phi^2 M_O(m_{O2}+m_{O12})$
	1		$\phi n_1(m_{12}+m_{012})$

This (method A) form of grouping they show to be inefficient in terms of the asymptotic variance of the estimate $\hat{\phi}$. They adopt an alternative (model B) grouping of the recaptures according to the interval since they were last seen. The triple catch survey is treated as

	Times	1	2	
Intervals	1	$\phi M_o(m_{01})$	$(\phi^2 M_0-\phi m_{01})$	(m_{02})
	2		ϕn_1	$(m_{12}+m_{012})$

for which grouping they assert there is no loss of information. For the triple catch

$$\hat{\phi} = \{m_{01}(m_{02}+m_{12}+m_{012})+(n_1-m_{01})m_{02}\}/(m_{12}+m_{012})M_0$$

with variance estimated by

$$m_{02}(m_{02}+m_{12}+m_{012})n_1^{\,2}/(m_{12}+m_{012})^3 M_0^{\,2}.$$

A method of estimating the sample size required to reduce var $(\hat{\phi})$ to a pre-assigned value is given by Sonleitner and Bateman (1963) for a three point sample.

Bailey (1951) analyses a triple-catch sample grouped according to the time of first marking. He considers deterministic birth rate, β, and death rate, γ, per unit time. With three available observations m_{01}, $m_{02}+m_{012} = m_{02}$, and three parameters N, β, Y, the maximum likelihood solution equates observation to expectations. Thus, if the samples are taken at times $0, t_1, t_1+t_2$, the estimates are

$$\hat{N}_1 = (s_1-m_1)n_1(m_{02}+m_{012})/m_{01}m_{12}$$

$$e^{\hat{\beta}t_2} = m_{01}n_2/n_1(m_{02}+m_{012})$$

$$e^{-\hat{\gamma}t_1} = (s_1+m_1)(m_{02}+m_{012})/s_0 m_{12}\ .$$

Variances of these estimates are given, as also are modified estimates which are almost unbiased estimates of variances of these estimates. Gaskell and Parr (1966) from their Bayesian standpoint, recommend always replacing $1/m_{ij}$ by $1/(m_{ij}-1)$.

The problems of estimation of population size were subsequently studied by Leslie (1952). The complete likelihood function, including all the animals not caught at each sample, has to be maximized. Classes were grouped according to method B. For a population with constant death rate and no dilution Leslie used the estimate of survival $\hat{\phi}$ from the earlier paper and obtained an approximate estimate of N by substituting $\hat{\phi}$ for ϕ in the likelihood equations. Leslie also derived maximum likelihood equations for the populations with:

(a) constant death rate, but dilution occuring

(b) changing death rate and no dilution;

and he mentions a possible method of attack on the problem when dilution is occuring and the death rate is varying both in time and between different groups of animals. The estimate of the population size suffer the same disadvantage of bias as the standard Petersen index and Leslie extends Bailey's modifications to the general case. It should be noted that dilution in these models - as in Jackson and Bailey - is a function of the size of the existing population.

Leslie, Chitty and Chitty (1953) develops the theory further with particular reference to the problems discovered during an extensive study on two small mammals. An improved approximate method is used to derive estimates from a method B grouping. Various tests were applied to the data. The assumptions necessary for valid estimation of the population parameters were found to hold for one population but not for the other. He asserts that reasonable, and useful inferences, may be deducible about the sub-population of marked animals, even if not about the population. The point is made again by Pope (1963); "when only mortality rates are to be estimated from marking experiments it is not necessary to consider the whole population, but only the sub-population of marked individuals". This is however, no advantage if marking affects mortality.

Corbet (1952), using Fisher and Ford's method estimated different survival rates for different periods of experiment. He reiterates the warning to analyse sexes separately because of their differential activity. Coulson (1962) modified Bailey's tripple catch method to allow the first recapture to be taken on the day of release, this is necessary because of the short lifespan of the animals studied. Wohlschlag (1954) applied Bailey's method to a continouos experiment by dividing the study into three equal periods and considering all sampling and marking as if done at mid-period. Orians (1958), using Leslie and Chitty's method B, and by studying separately the birds first caught in a particular year, found that the estimated survival rate was low for the first year after marking, then higher for a few years, finally decreasing again because of loss of rings. Various tests led Krebs(1966) to the conclusion that, for the population of mice under study, non of the conditions necessary for valid capture-recapture study were satisfied. Although other workers have come to less drastic pessimistic conclusions, it is clear that all relevant biological and statistical knowledge must be applied together if any satisfactory information is to be gained.

Studies of the particular problems of insect populations have led Iwao and Kuno to various modifications of the general model suggested by Leslie (1952) and Leslie and Chitty (1953). These models include migration between two types of fodder plant on which the insect have different survival rates, and from which different fractions are sampled(Iwao, 1963). Similar models for the dispersion of insects among neighbouring areas are given by Jackson (1939). Any more sophisticated treatment of such problems should take note of the stochastic models for animal movement given by Pyke (1955).

The most general deterministic model is given by where m_{ji} is the number of animals released at j which are next caught at i, and

$$W_{i-1,t+1} = \begin{cases} \overset{i-1}{\underset{k=t+1}{\pi}} W_k & \text{for } (i-1) > t. \\ 1 & \text{for } (i-1) = t. \end{cases}$$

and W_k are the set of positive numbers.

That is a reasonable estimate of M_t can be seen by considering one of the $(M_t - m_t)$ marked individuals not caught at t. Its probability of being captured at some subsequent time i is m_{ti}. The number of such animals is $\sum_{i=0}^{t-1} m_{ji}$ so that $(M_t - m_t)$ is estimated by $(s_t \sum_{j=0}^{t-1} m_{ji})/m_{ti}$. The above general formula is a weighted combination of such estimates for different values of i. By setting up a formal model to this estimate, and considering the asymptotic variances of the $\hat{M}_t$, Jolly develops optimal weights:

$$W_r = \frac{1 + \sum_{i=r+2}^{k} (n_i/N_i) \sum_{j=r+1}^{i-1} \phi_j/(1-n_j/N_j)}{(1-n_r/N_r)\{1 + \sum_{i=r+1}^{k} (n_i/N_i) \sum_{j=r}^{i-1} \phi_j/(1-n_j/N_j)\}}$$

In practice, as we illustrated by an example given by Jolly (1963), these W_r have to be obtained iteratively from the $\hat{\phi}_t$ and $\hat{N}_t$ obtained from unweighted estimates of M_t. One iteration is adequate in most examples.

Jolly (1963). At time t, from the population N_t, of which M_t are marked, a sample of fixed size n_t is taken of which m_t are found marked, with various (known) histories of capture. Animals are given an individually identifying mark and allowing for some accidental deaths in handling - s_t are released. These animals are subject to death rate $(1-\phi_t)$ and dilution rate λ_t defined as $N_{t+1}/\phi_t(N_t-n_t+s_t)$ before next sampling period at time t+1. The unknown parameters of interest, N, ϕ, λ, are all expressible in terms of M and $\alpha = M/N$, the proportion of marked animals in the population. Assuming, at each sample, a multinomial distribution conditional on the results of all previous samples, and that (n_t-s_t) is so small that the essential randomness of the s_t can be neglected. Jolly shows that the maximum likelihood estimate of α_t is m_t/n_t, and gives an equation - a kin to Leslies' - for the maximum likelihood estimate of M_t. All the $\hat{\alpha}_t$ and $\hat{M}_t$ are shown to be asymptotically independent, and explicit formulae are given for the variances of the estimates.

From further study of various forms of the maximum likelihood equation, Jolly (1963) was led to postulate a general form of estimate for M_t, in the form of weighted ratio:

$$\hat{M}_t = \{(s_t \sum_{i=t+1}^{k} W_{i-1,t+1} \sum_{j=0}^{t-1} m_{ji}) / \sum_{i=t+1}^{k} W_{i-1,t+1} m_{ti}\} + m_t.$$

We now discuss the stochastic models. If the population is subject to naturally occuring changes, in the form of birth, death, migration, and growth, any comprehensive theory must take into account the natural variability between individuals in respect of these changes. If not only the population but also the observed samples are large in number, and interest lies in the behaviour of an 'average' individual in the population, this variability contributes negligibly to the variability of the estimates of the behaviour of this hypothetical animal. Every individual can then be assumed to behave identically; a deterministic model, such as we have considered in the previous sections, is adequate. However in most practical studies it is stretching the definition to assume large samples. Since the stochastic models describing capture-recapture studies have recently been shown by Jolly (1965) and Seber (1965) not only to be less complicated to analyse than the corresponding deterministic models, but also to provide more valid results, these methods should totally supersede their deterministic counterparts.

Although some earlier workers had pointed out that the variance of an estimator from a deterministic model would underestimate the true variance of the estimator, Moran (1952) was the first to discuss the implications of deterministic and stochastic models in this context. A truly deterministic model is virtually unworkable in practice since, if the proportion ϕ_o of the population survive between sampling at time t_o and t_1, not only $\phi_o N_o$ but also $\phi_o M_o$ - and the size of any other survival class - is restrained to be an integer. Thus an approximation has to be introduced by assuming that all values such as $\phi_o N_o$ can be regarded as integers. Under the assumption of multinomial, rather than multi-hypergeometric, sampling, the probability distribution, assuming integers, of the set of observed recaptures remains a probability distribution if the $\phi_o N_o$ and so on are integers. This difference defines the 'deterministic model' which can only be an approximation to what really happens. Moran points out that Leslie and Chitty (1951) and Leslie (1952) use a 'semi-probabilistic' model. In this, a fixed proportion ϕ_o of the population survives, but the $N_o \phi_o$ individuals surviving are chosen at random without reference to their markings. A fully probabilistic model assumes that each individual independently has a probability ϕ_o of survival, the number which survive being a random variable binomially distributed with index N_o and parameter ϕ_o. Moran does not attempt to analyse this model, being deterred by the difficulty that the sample numbers n_i can not be prescribed in advance since there is a non-zero probability that fewer than n_i animals are alive at that time.

The way round this difficulty was noted first by Hammersley (1953). The sampling scheme hitherto considered have all assumed that the sample on each occasion is complete when a predetermined number of animals have been captured. In practice, however, sampling usually ceases through a decision of the experimenter determined by the effort expended, and not the number of animals caught. Each individual may be assumed to have the same probability of capture, p_i, during the ith sample period, the size of that sample being thus binomially distributed with index N_i. The p_i are parameters, sometimes to be estimated from the data, sometimes considered as known except for a single constant of proportionality; the n_i are random variables. This simple change of outlook, bringing theory more into line with practice, circumvents the difficulties of a stochastic approach. Unfortunately, as shown by Dorrah (1959), Hammersley's subsequent analysis is based on an incorrect model for the likelihood function. Hammersley does consider the population to be of two different classes, young and adult, with different survival rates, the one growing into other during the course of the experiment.

The distinction between the two models was first brought out by Darroch (1958) in his analysis of a closed population. The earlier model with fixed sample size (Darroch's model B) "is likely to be appropriate when the main limiting factor on the sample size is the trouble involved in marking animals". The new model (Darooch's model A) "when it is the difficulty in catching them". The simplicity of the multinomial distribution implied by the new model, rather than the multihypergeometric model, enables Darroch (1959) to derive the maximum likelihood estimators for the unknown parameters when either death or immigration, but not both, are affecting the population. Seber (1962) derived maximum likelihood estimates for a model appropriate to commercial fisheries. A series of batches of marked individuals are released, after each of which -allowing time for redistribution of the marked animals- a commercial catch is taken and this is immediately followed by the release of more marked fish. Seber shows that if only one release followed by a number of catches is made and no immigration or emigration is assumed to occur, his model becomes a discrete equivalent of Gulland's(1955) model for a continouos fishery.

These methods of Darroch and Seber are special cases of a general theory proposed independently by Jolly and Seber (1965). Since Jolly's model is slightly more general than Seber's in that it allows for the removal, either deliberately or by accident, of some of the animals caught in any sample, it will be discussed here. A part from this modification, the estimates obtained by Jolly and Seber are identical.

At time t_i, a population of N_i individuals, M_i of which are marked by a mark which identifies the last occasion on which they were caught, is sampled in such a way that each individual in the population has probability p_i of being captured. This results in a sample of n_i individuals, m_i of which are marked. After this sample each of the n_i has probability η_i of being marked and released again, s_i individuals being released into the population. Between t_i and t_{i+1} the population of $(N_i-n_i+s_i)$ individuals is altered by the effect of death (or emigration) each individual independently having a probability $(1-\phi_i)$ of leaving the population; and also by immigration of a number of new individuals, B_i of which are alive in the population at time t_{i+1}. Of these parameters $N_o, B_i, p_i, \phi_i, \eta_i$ are unconditional. The others N_i and M_i are conditional on events prior to t_i. This means that the general probability distribution of N_{ji}, M_{ji}, m_{ji}, U_i, u_i, s_i (where M_{ji} and m_{ji} are the numbers in the population and sample at time t_i which were last caught at t_j and U_i, u_i those that have never been caught), conditional on the fixed parameters N_o, B_i, p_i, ϕ_i, η_i, is not identical to the likelihood function of the parameters. Three approaches are possible. One is the orthodox one of summing over all values of the conditional parameters; Darroch (1959) discovered the difficulties of this, which were surmounted by a change of approach in Jolly's work, by treating B_i not as a parameter but as a determinate function of U_i, U_{i+1}, ϕ_i , u_i. This introduction of a deterministic element affects only unmarked animals which do not contribute to the estimate of survival rates. The third method treats the general probability distribution; regarded as a function of the parameters, as if it were a true likelihood. This is proved justifiable by the identity of the estimates provided by it and the previous approach. Seber avoids the difficulty by considering the net increase in the number of marked individuals as a fixed parameter again, identical estimates are obtained.

These estimates, obtained after much complicated algebra, turn out to be remarkably simple in form, each with obvious intuitive meaning:

(i) $\hat{M}_i = \{Z_i s_i / r_i\} + m_i \quad (i=1,2,\ldots,k-1)$

where r_i is the number of the s_i which are recaptured subsequently, Z_i is the number of individuals marked before t_i, not caught at t_i, but which are recaptured subsequently. Intuitively this estimate is obtained by equating $Z_i/(M_i-m_i)$, the proportion of marked animals alive but not seen subsequently, to r_i/s_i, the proportion of marked animals released at time t_i which are seen subsequently.

(ii) $\hat{N}_i = \frac{n_i}{m_i} \hat{M}_i$ $(i=1,2,\dots,k-1)$

Intuitively this is equating the proportion of marked animals caught at time t_i, m_i/M_i to the proportion caught n_i/N_i.

(iii) $\hat{\phi} = \hat{M}_{i+1}/(\hat{M}_i-m_i+s_i)$ $i=1,2,..,k-2$

of the $(M_i-m_i+s_i)$ marked animals known to be alive in the population immediately after t_i, M_{i+1} are alive at t_{i+1}.

(iv) $\hat{B}_i = \hat{N}_{i+1}-\hat{\phi}_i(N_i-n_i+s_i)$, $(i=1,2,\dots,k-2)$

of the $(N_i-n_i+s_i)$ animals alive in the population immediately after t_i, $\phi_i(N_i-n_i+s_i)$ will be alive at time t_{i+1}. The difference between this figure and N_{i+1} must be accounted for by immigration.

(v) $\hat{p}_i = n_i/\hat{N}_i$ $(= m_i/\hat{M}_i)$

The numbers of recaptures contribute to these estimates in two distinct ways. First recaptures after t_i contribute to $\hat{N}_i$ through $\hat{M}_i$ in the form of the ratio z_i/r_i. Also recaptures at t_i contribute directly to $\hat{M}_i$ and also in the estimate m_i/n_i of the proportion of marked animals in the population which is used in (ii) to convert $\hat{M}_i$ to $\hat{N}_i$. The s_i, on the otherhand, contribute to the estimates only in the form of creating a pool of marked animals from which later samples may be taken. There is no direct connection between s_i and n_i,m_i in the estimation procedure. Seber's (1965) estimates for a model in which no losses on capture are allowed, are identical with those given if s_i is put equal to n_i. Moreover Jolly shows that Seber's (1962) estimates can also be expressed by the above relations, because, as Jolly points out, the only relevant information is the occassion on which an individual was last seen. Its history of capture prior to the last previous sighting contains no information about its chance of death or recapture after that sighting. The information provided by a single individual caught at times t_i, t_j, t_k, is precisely the same as if one individual released at t_i were recaptured (and removed) at t_j, a second individual released at t_j and recaptured at t_k. Thus a commercial fishery with single recapture yields the same estimates as a research project with multiple recapture. The name multiple recapture is thus irrelevant, at least as far as the estimation of the population parameters is concerned. If interests lie in testing the assumptions of the model, particularly with regard to the equal catchability of individuals, then information on multiple recaptures does become valuable.

It is interesting to compare these estimates with the corresponding estimates obtained on the basis of

Jolly's (1963) deterministic model. Relation (ii) is identical with the earlier estimate of the proportion which has been marked. In the estimation of M_i the deterministic model used complicated weighting sums of various recapture frequencies. The stochastic M_i has the same form, but is in terms of the unweighted sums r_i and Z_i. As Jolly says "the stochastic solution therefore, far from being complicated, is actually simpler than the other". The variances of these simple estimates, as derived by Jolly (1965) on the assumption that B_i are the basic parameters (Seber's model), the M_{ji}, n_{ji}, and so on random variables arising from partitioning the B_i into multinomial classes, are complex expressions involving the various observations used in the estimates, or strictly their expected values, and also such terms as $N_i(j)$ and $M_i(j)$, respectively the number of individuals and of marked individuals at t_i which were part of the B_j animals which first entered the population between t_j and t_{j+1}. Evaluation of these is carried out by use of relationships:

$$N_{j+1}(j) = B_j$$

$$N_{k+1}(j) = \{(N_{k+1}-B_k)/N_k\}N_k(j) \text{ for all } k>j$$

These terms thus appearing include $N_o(=B_{-1})$ which cannot be estimated from the experiment. Jolly suggests that a value for N_o, which will be needed in calculating estimates of variances of the parameters, be taken as equal to $\hat{N}_1$ if the population is initially stable, or otherwise guessed from the trend of $\hat{N}_i$.

The theoretical and estimated variances of $\hat{N}_i$, $\hat{\phi}_i$, $\hat{B}_i$ may be obtained by substituting respectively expected values and estimated or observed values for the symbols in the expressions below.

$$\mathrm{var}[\hat{N}] = N_i(N_i-m_i)[M_i-m_i+s_i)(1/r_i-1/s_i)/M_i+(N_i-M_i)/m_iN_i]$$

$$+ N_i - \sum_{j=1}^{i-1} N_i^{\,2}(j)/Bj$$

$$\mathrm{var}[\hat{\phi}_i] = \phi_i^{\,2} \quad \frac{(M_{i+1}-m_{i+1})(M_{i+1}-m_{i+1}-s_{i+1})}{M_{i+1}^{\,2}} \; (1/r_{i+1}-1/s_{i+1})$$

$$+ \frac{(M_i-m_i)}{(M_i-m_i+s_i)} \quad (\frac{1}{r_i} - \frac{1}{s_i}) + \frac{1-\phi_i}{M_{i+1}}$$

$$\mathrm{var}[\hat{B}_i] = B_i^{\,2} \frac{(M_{i+1}-m_{i+1})(M_{i+1}+s_{i+1})}{M_{i+1}^{\,2}} (\frac{1}{r_{i+1}} - \frac{1}{s_{i+1}})$$

$$+ \frac{(M_i-m_i)}{(M_i-m_i+s_i)} \quad \frac{\phi_i^{\,2} s_i^{\,2} (N_i-m_i)^2}{M_i^{\,2}} \quad (\frac{1}{r_i} - \frac{1}{s_i}).$$

$$+ \frac{(N_i-n_i)(N_{i+1}-B_i))N_i-M_i)(1-\phi_i)}{M_i(M_i-m_i+s_i)}$$

$$+ N_{i+1}(N_{i+1}-n_{i+1})\frac{(N_{i+1}-M_{i+1})}{M_{i+1}m_{i+1}} + \phi_i^{\,2} N_i(N_i-n_i)$$

$$+ \phi_i^{\,2} N_i(N_i-n_i)(N_i-M_i)/M_im_i$$

Similar expressions for the covariances of the estimates may be found in Jolly. The terms $N_i \sum_{j=1}^{i-1} N_i^2(j)/B_j$ in var $[\hat{N}_i]$, and $(1-\phi_i)\ \phi_i^2/M_{i+1}$ in var$[\hat{\phi}_i]$ are contributions to the variance from the conditional nature of N_i and ϕ_i.

Jolly's initial formulation of the problem is even more general than has been described here. The joint probability distribution from which the estimates are derived, is stated from a heterogeneous population that is, for a population composed of different classes (for example, sexes) between which there is no migration. If all parameters are different for each class, each class may be analysed as a homogenous population. But if some parameters, perhaps survival rate or probability of capture, are common to different classes, more complex estimation equations would have to be constructed, and solved by iteration. Jolly (1965) also analyses the restricted cases of "death, no immigration" and "immigration, no death", showing that, when $s_i=n_i$, his estimates and their variances are identical with those given earlier by Darroch (1959). These restricted models should not be used, now that a simple analysis of the general model is available (Cormack(1968)).

In common with earlier, simpler models, these likelihood estimates can be considerably biased if the numbers of recaptures are small. Seber (1962, 1965) suggests modifications, similar in form to those of both Bailey and Chapman which lead to almost unbiased estimates with a negligible loss of efficiency. For the case of no losses on capture, $n_i = s_i$, the modified estimate of M_i is

$$M_i^* = Z_i(S_i+1)/(r_i+1) + (m_i+1);$$

The estimates N_i^*, ϕ_i^*, B_i^*, p_i^* being the same functions of M_i^* as $\hat{N}_i$, $\hat{\phi}_i$, $\hat{B}_i$, $\hat{p}_i$ are of $\hat{M}_i$. A precisely similar modification, adding unity to each observation except Z is shown by Seber (1962) to yield almost unbiased estimates for the single recapture census.

Alternative description of the properties of dilution and survival are introduced by Jolly (1963) and Seber (1965), respectively. Jolly defines a dilution rate λ_i as the ratio of the population size at t_{i+1} to the survivors at that time of the population at t_i. He gives an expression of λ_i as:

$$\lambda_i = [\phi_i(N_i-n_i+s_i)+B_i]/\phi_i(N_i-n_i+s_i)$$

The estimate of λ_i is merely the corresponding function of other estimates. This representation has the advantage that λ_{k-1} is estimable although B_{k-1} is not. Seber (1965) considers an instantaneous death rate μ_i during the period (t_i, t_{i+1}), defined by the relation $\phi_1 = e^{-\mu_i(t_{i+1} - t_i)}$. Again its estimate is the corresponding function of the estimate of ϕ_i.

One further extension of this model was suggested by Darroch (1958, 1959) but not by the later authors, except Seber (1965) makes use of the idea in constructing at test of the assumptions. In many fields of study a measure of the effort, f_i, put into the sampling at t_i, is available. This usually takes the form of the number of man-hours expended, or in commercial catches, the total number of gear used. If f_i are known, and the catchability (q) of all individuals remains constant throughout the experiment, the p_i are known functions of a single unknown parameter q. Most earlier authors have assumed that $p_i = qf_i$. Darroch (1958) makes the assumption that the probability of any individual being caught when subjected to df units of effort is qdf, which leads to the model $p_i = 1-e^{-qf_i}$. As manageable approximation to this, Darroch uses $p_i = qf_i(1-qf_i/2)$.For the models assuming death or immigration but not both, Darroch derives an approximate maximum likelihood estimator for q, and modifies it in the usual way to provide an almost unbiased estimator. The estimator $\hat{p}_i$ may be replaced by $\hat{q}f_i(1-\hat{q}f_i/2)$, and this used in the subsequent estimation of the other population parameters. Darroch gives a table showing the increase of information about N_i provided by this knowledge of effort.

The only other model to have received a fully stochastic treatment in these terms was proposed by Cormack (1964). This refers to the situation in which recapture information is available on marked individuals only, random sampling of the whole population being impossible. Estimates were obtained for the mortality of the marked population. For these estimates to be applicable to the whole population some assumptions of the representative character of the marked animal is required. In general no estimate of the population size is possible from such data, although Jolly (1965) points out that, "whether or not the survival rates are the same for the two populations, provided they have the same probability of being captured in a sample", an estimate of N_i is possible. If, however, as in the data from Dunnet, Anderson and Cormack (1963) discussed by Cormack (1964), sampling of the marked and unmarked individuals are totally separate, the proviso in Jolly's statement will be impossible to fulfil. Similar data may also be obtained in commercial studies, n_i not being recorded although m_i is. Again N_i cannot be estimated and the survival estimates ϕ_i apply strictly only to the marked population. This situation fits easily into Jolly's frame work within which s_i and n_i need not be composed of the same individuals, and Jolly has shown that both the estimates and their variances as given by Cormack are a special case of the general formulae.

Parr (1965) discusses the various methods which have been proposed for entomological population studies and compares their results with those provided by Jolly's method. Stochastic models of a somewhat different kind are discussed by Chapman (1954). These assume that m_{ij} have a Poisson distribution and that the probability of an individual's survival between samples is a constant ϕ. Model I refers to a population not subject to dilution. In this case m_{ij} conditional on M_{ij} and N_j, has expected value $n_j \; M_{ij}/N_i$. Since $E[m_{i.j}] = s_i \phi^{j-i}$, and $E[N_j] = N_o \phi^j$, for larger population, m_{ij} can be taken to have a Poisson distribution with mean $(s_j n_j \phi^{-i}) N_o$. This leads to the maximum likelihood equations:

$$\hat{N}_o = \sum_{ij}\sum s_i n_j \; \hat{\phi}^{-i} / \sum_{ij}\sum m_{ij} \quad \text{and,}$$

$$\Big(\sum_{ij}\sum m_{ij}\Big)\Big(\sum_{ij>i}\sum \; i s_i n_j \hat{\phi}^{-i}\Big) = \Big(\sum_{ij}\sum i m_{ij}\Big)\Big(\sum_{ij>i}\sum \; s_i n_j \hat{\phi}^{-1}\Big)$$

The second of these equations, being a polynomial in $\hat{\phi}$, can be solved by standard methods. Chapman admits this ignores multiple recapture and thus loses some information.

2.2 STATEMENT OF THE PROBLEM.

The estimation of total population size of animal populations is of great importance in a variety of biological problems which may relate to population growth, ecological adaptation, genetic constitution, natural selection and evolution, and so on. Obvious practical consequences are the maintenance of human supplies of food and control of insect pests. For human communities procedures employing fixed sampling units are available, but for mobile populations other methods must be used. Of the methods available for obtaining information about animal populations, capture-recapture method is the most widely used.

Any animal population is subject to continual change by birth, death, immigration and emigration. In a study carried out over a short period of time, these effects may be negligible and the experimenter may reasonably consider the population as closed. More usually, temporal changes in the population are of primary interest; is the population growing, stationary or heading for extinction? A study over a longer time-scale is required. If sampling continues over an appreciable period of time, the population can no longer be considered closed. Other parameters of recruitment and mortality (and possibly immigration and emigration) must be included in the model.

A detailed study of such parameters as mortality rates, survival rates and total population size, provide the information required to generate enough scientific guidelines and advice for rational exploitation, management and utilization of renewable resources. Hence the need to study them.

Estimation of population size is important in the following areas: the exploitation of fresh water fish for food, the removal of insect pests from crops, the control of Tse-Tse flies, the destruction of verminous small mammals by poisoning, the control of an out break of an desease, and so on.

The aim of this dissertation is to study capture-recapture models for estimation of population size. We shall also give an application of these models to the real-life data. More specific objectives are listed below.

1) To review some capture-recapture models used in the estimation of population size.
2) To apply the selected model(s) to the real life data.
3) To study the departures from the underlying assumptions

CHAPTER THREE

3.1 ESTIMATION

Hypergeometric Model

A simple method, which we shall call the **Petersen method** for estimating N, the number of animals in a closed population, is described as follows. A sample of n_1 animals is taken from the population, the animals are marked or tagged for the future identification and then returned to the population. After allowing time for marked and unmarked to mix, a second sample of n_2 animals is taken and suppose it is found that m_2 are marked. Assuming that the proportion of marked animals in the second sample is a reasonable estimate of unknown population proportion, we can equate the two and obtain an estimate $\hat{N}$ of N. Thus

$$m_2/n_2 = n_1/N \quad \text{or} \quad \hat{N} = n_1 n_2/m_2$$

which is the so called **Petersen estimate** or **Lincoln index**. As this estimate is widely used in ecological investigations, we shall now discuss the above method in some detail.

For $\hat{N}$ to be a suitable estimate of N, the following assumptions must be made .

I. the population is closed, so that N is constant

II. all animals have the same probability of being caught in the first sample.

III. marking does not affect the catchability of an animal

IV. the second sample is a simple random sample that is . each of the $\binom{N}{n_2}$ possible samples has an equal chance of being chosen.

V. animals do not loose their marks in the time between the two samples.

VI. all marks are reported on recovery in the second sample.

Note that the assumptions listed above are not mutually exclusive, for example IV will depend on the validity of II and III, as any variation in the catchability of animals, whether natural or induced by the handling and marking will lead to a non-random second sample.

When the assumptions I, IV, V, and VI are satisfied, then the conditional distribution of m_2 given n_1 and n_2 is the Hypergeometric distribution.

$$f(m_2/n_1,n_2) = \binom{n_1}{n_2}\binom{N-n_1}{n_2-m_2}\Big/\binom{N}{n_2} \qquad (3.1)$$

where $m_2 = 0,1,2,\ldots,$ minimum (n_1, n_2).

Now, regarding the hypergeometric distribution (3.1) as the likelihood function $L(N)$ of N, then

$$\log L(N)=\text{Const.}+\log(N-n_1)!+\log N!+\log(N-n_2)!+\log(N-n_1-n_2+m_2)! \qquad (3.2)$$

An equation for maximum likelihood estimate $\hat{N}$ of N is found by equating $\Delta\log L(N)$ to zero, where Δ is the difference operator. This involves an error of less than unity in the solution and is equivalent to the **ratio method** of maximizing L, which equates $L(N)$ to $L(N-1)$ (Darroch, 1958). Now,

$$\Delta\log N! = \log N \qquad (3.3)$$

Thus N must satisfy the equation

$$\log(N-n_1) - \log(N-n_1-n_2+m_2) - \log N + \log(N-n_2)=0$$

That is

$$\log\{(N-n_1)(N-n_2)\} = \log\{(N(N-n_1-n_2-m_2)\}$$

or

$$(N-n_1)(N-n_2) = N(N-n_1-n_2+m_2)$$

which gives

$$n_1 n_2 = N m_2$$

That is

$$\hat{N} = n_1 n_2 / m_2$$

Which is the **Petersen estimate**. It can be shown that the bias of $\hat{N}$ is

$$(n_1-m_2)(n_2-m_2)/m_2$$

Darroch (1958) derived an approximate bias of $\hat{N}$ resulting from multiple recapture experiment of s samples to be

$$\hat{\beta} \cong \frac{\{(-1)/(\ \hat{N}-n_i\ \}^2 + \{(s-1)/\hat{N}^2 - \sum 1/(\hat{N}-n_i)^2\}}{2\{1/(\hat{N}-\rho) + (s-1)/\hat{N}\ \ - \sum 1/(N-n\ \ )\}^2}$$

where $\rho = E[\ m_i\ ]$

In our case s = 2 , thus the approximate bias is given by

$$\hat{\beta} \cong (\hat{N}-n_1)(\hat{N}-n_2)/n_1 n_2 = (n_1-m_2)(n_2-m_2)/m_2^2 \qquad (3.6)$$

That is, the bias of $\hat{N}$ is:

$$(n_1-m_2)(n_2-m_2)/m_2^2$$

Now consider the estimate

$$N^* = \frac{(n_1+1)(n_2+1)}{(m_2+1)} - 1$$

of N. Robson and Regier (1964) shows that when $n_1+n_2 \geq N$, then N^* is an exactly unbiased estimate of N , while, when $n_1+n_2 < N$ we have to a reasonable degree of approximation,

$$E[N^* | n_1, n_2] = N - Nb \qquad (3.7)$$

where

$$b = \exp\{-(n_1+1)(n_2+1)/N\ \}$$

The bias Nb of N^* in this later case is less than that of $\hat{N}$. Thus N^* is almost an unbiased estimate of N and hence it is more preferable to $\hat{N}$ as an estimate of N. It is easy to check that N^* is the solution of

$$(N - n_1)(N - n_2) = (N + 1)(N - r) \qquad (3.8)$$

where, $r = n_1 + n_2 - m_2$.

Properties of $\hat{N}$.

The properties of $\hat{N}$ with respect to the hypergeometric distribution have been fully discussed by Chapman (1951). He shows that although $\hat{N}$ is a best asymptotically normal estimate of N as $N \longrightarrow \infty$, it is biased and the bias can be small for small samples. However, when $n_1 + n_2 > N$, his modified estimate, N^* is exactly unbiased, while if $n_1 + n_2 < N$, we have, to a reasonable degree of approximation (Robson and Regier(1964))

$$E[N^* | n_1, n_2] = N - Nb$$

where,

$$b = \exp\{-(n_1+1)(n_2+1)/N\}$$

Defining

$$\mu = E\{m_2/n_1, n_2\} = n_1 n_2/N$$

Robson and Regier (1964) recorded that in designing a Petersen type experiment it is essential that $\mu > 4$, so that b is small. They also say if $m_2 > 7$ in a given experiment, then we can be 95 per cent confident that $\mu > 4$. This implies that for 7 or more recaptures we can be 95 per cent confident that the bias of N^* is negligible.

Chapman (1951) shows that N^* not only has a smaller expected mean square error than $\hat{N}$ for values encountered in practice, but also appears close to being a minimum variance unbiased estimate over the range of parameter values for which it is almost unbiased. Using what is essentially a Poisson approximation to the hypergeometric model (3.1), he shows that the variance of N^* is approximately given by

$$V[N^*|n_1,n_2] \cong N(\mu^{-1}+2\mu^{-2}+6\mu^{-3}) \qquad (3.9)$$

An estimate V^* say, of the variance of N^* is obtained by simply replacing N by N^* in (3.9). However, an approximate unbiased estimate has been given by Seber (1970) and Wittes (1972), namely

$$V[N^*|n_1,n_2] = (n_1+1)(n_2+1)(n_1-m_2)(n_2-m_2)/(m_2+1)(m_2+2).$$

which has a positive proportional bias of order $\mu^2 \exp-\mu$ It can be shown that V^* is exactly unbiased when $n_1+n_2>N$. The co-efficient of variation of N^* is approximately given by

$$C[N^*] = 1/\sqrt{\mu}$$

and if a rough estimate of N is available before the experiment, n_1 and n_2 can be chosen before hand to give a desirable value of C. We note that an estimate of C is obtained by replacing μ by m_2, giving $C = 1/\sqrt{m_2}$. This means that the "accuracy" of N^* is almost solely dependent on the number of recaptures m_2.

Bailey's Binomial Model.

Assume that n_2 is sufficiently small compared with N for us to be able to ignore the complications of sampling without replacement. The large sample theory of maximum likelihood approach will be appropriate if we envisage n_1, n_2 and N all tending to infinity while keeping constant ratios to each other. Then using the binomial approximation to hypergeometric distribution (3.1), the likelihood function becomes

$$f(m_2/n_1,n_2) \cong \binom{n_2}{m_2} \left(\frac{n_1}{N}\right)^{m_2} \left(1-\frac{n_1}{N}\right)^{(n_2-m_2)} \qquad (3.10)$$

therefore

$$L = \log\{f(m_2/n_1,n_2)\} \cong \text{constant} + (n_2-m_2)\log(N-n_1) - n_2\log N \qquad (3.11)$$

Differentiating (3.11) with respect to N gives

$$\frac{\delta L}{\delta N} = (n_2-m_2)/(N-n_1) - n_2/N \qquad (3.12)$$

Hence, the maximum likelihood estimate of N satisfies

$$(n_2 - m_2)/(N - n_1) = n_2/N$$

that is

$$N(n_2 - m_2) = n_2(N-n_1)$$

or ,

$$Nm_2 = n_1 n_2$$

Therefore

$$\hat{N} = n_1 n_2/m_2 \qquad (3.13)$$

which is **Petersen's estimate**. Expanding $\hat{N}$ as a taylor series in powers of $t(=m_2-m)$, where $m=E[m_2]$, and taking the expectation of both sides we have

$$\begin{aligned} E[\hat{N}] &= N\{1 + (N-n_1)/n_1 n_2 + o(m^{-2})\} \\ &= N\{1 + 1/m + o(m^{-2})\} \qquad (3.14) \end{aligned}$$

Thus, clearly, the bias of $\hat{N}$ is approximately $Nm-1$.

Now consider the estimate

$$\hat{N}' = n_1(n_2 + 1)/(m_2 + 1)$$

Then,

$$E[\hat{N}'] = \sum_{m_2=0}^{n_2} \frac{n_1(n_2+1)}{(m_2+1)} \begin{pmatrix} n_2 \\ m_2 \end{pmatrix} \left(\frac{n_1}{N}\right)^{m_2} \left(\frac{N-n_1}{N}\right)^{(n_2-m_2)}$$

$$= N \sum_{m_2=0}^{n_2} \begin{pmatrix} n_2+1 \\ m_2+1 \end{pmatrix} \left(\frac{n_1}{N}\right)^{(m_2-1)} \left(\frac{N-n_1}{N}\right)^{(n_2-m_2)}$$

$$= N \left\{1 - \left(\frac{N-n_1}{N}\right)^{(n_2+1)}\right\}$$

Thus the bias of $\hat{N}'$ is approximately $N\exp(-m)$ which is quite small even for moderate m. Thus $\hat{N}'$ may be preferred to $\hat{N}$ as an estimate of N.

To get the variance of $\hat{N}$, we differentiate equation (3.12) with respect to N and taking expectations provides the required information in the case of a single sample. Thus:

$$I_N = -E\,[\partial^2 L/\partial N^2) = n_1 n_2/\{N^2(N-n_1) \qquad (3.16)$$

Using $\hat{N} = n_1 n_2/m_2$

$$\mathrm{Var}[\,\hat{N}\,] = I_N^{-1} = n_1^2\, n_2(n_1-m_2)/m_2^3$$

We note that $\mathrm{var}[\,\hat{N}\,]$ is infinite if we admit the value $m_2 = 0$. On the other hand, this value will occur so rarely in large samples for which $m = E[\,m_2]$ is not small, that we can choose to omit it.

A satisfactory expression for the variance of $\hat{N}'$ is somewhat more difficult to find. However using the δ-method described in chapter one, we can derive the asymptotic series

$$E\,[\,\hat{N}'^{2}\,] \cong \frac{n_1(n_2+1)^2}{m_2}\left\{1 + \frac{1-2p}{m} + \ldots\right\} \qquad (3.18)$$

where $p=n_1/N$. We now obtain the variance from the relation.

$$\mathrm{Var}\,[\,\hat{N}'] = E[\,\hat{N}'^{2}\,] - \{E[\,\hat{N}'\,]\}^2 \qquad (3.19)$$

We could obtain a single series expression for the variance by substituting (3.18) in (3.19) and writing

$$\{E\,[\,\check{N}\,]\}^2 \cong N^2 = \frac{n_1^2\,n_2^2}{m_2} = \frac{n_1^2(n_2+1)^2}{m_2}\,\{1 - 2/n_2 + \ldots\}$$

but this would involve the consideration of inverse powers of n, there seems little gained by this device. Fortunately, there exists a nearly unbiased estimate of the variance of $\hat{N}'$, which is quite convenient for use with samples that are not too small. It is easy to show that

$$E[n_1^2(n_2^2+1)(n_2+2)/(m_2+1)(m_2+2)]$$

$$=N^2\left\{1 - \{(N-n_1)/N\}^{(n_2+2)} - (n_2+2)\left(\frac{n_1}{N}\right)\left(\frac{N-n_1}{N}\right)^{(n_2+1)}\right\}$$

$$\cong N^2(1-me^{-m}) \text{ ,for large } n_2. \qquad (3.20)$$

Therefore, if we write

$$T=n_1^2(n_2+1)/(m_2+1) \;-\; n_1^2(n_2+1)(n_2+2)/(m_2+1)(m_2+2)$$

$$=n_1^2(n_1+1)(n_2-m_2)/(m_2+1)(m_2+2).$$

Then

$$E[T]=E[\hat{N}'^{2}] - N^2(1-me^{-m}) \cong \sigma^2_{\hat{N}'} + N^2me^{-m}$$

Now, we know from equation(3.17) that $\sigma^2_{\hat{N}'}$ is of order N^2/m;

so that

$$E[T] \cong \sigma^2_{\hat{N}'}(1+ m^2e^{-m})$$

Thus the variance of $\hat{N}'$ has a positive proportional bias of order m^2e^{-m}.

If n_1/N is sufficiently small for one to be able to ignore the complications of sampling without replacement, then $\hat{N}'$ may be used as an estimate of N instead of N^*, though in practice there will often be little difference in the two estimates. We saw above that $\hat{N}$ is an intuitively reasonable estimate when the sample proportion of marked animals in the second sample faithfully reflects the population proportion of marked. This means that N can still be used even when assumption IV is false and the second sample is systematic rather than a random sample, provided that

(i) there is uniform mixing of marked and unmarked so the proportion n_1/N of marked throughout the population is constant, and

(ii) given that a certain location in the population area is sampled, all animals at that location, whether marked or unmarked, have the same probability of being caught.

When (i) and (ii) are satisfied, the probability that an animal is found to be marked, given that it is caught in the second sample, is n_1/N , and the binomial model above applies.

Random Sample Size

In practice it is not always possible to fix n_2 in advance as the sample size may depend on the effort or time available for sampling. However, if n_2 is regarded as a random variable rather than a fixed parameter, N^* is still approximately unbiased, since

$$E[N^*/n_1] = E_{n_2} E[N^*/n_1,n_1] \cong E_{n_2}[N] = N.$$

Using a similar argument, it is readily shown that V^* is an approximately unbiased estimate of $\mathrm{Var}[N^*|n_1]$. We know that

$$\mathrm{Var}[x] = E_y\{V[x/y]\} + V_y\{E[x/y]\}$$

thus

$$\mathrm{Var}[N^*|n_1] = E_{n_2}\{V[N^*|n_1,n_2\} + V_{n_2}\{E[N^*|n_1,n_2]\}$$

$$\cong E_{n_2}\{V[N^*|n_1,n_2]\} \cong \mathrm{Var}[N^*|n_1,n_2]$$

where $\mathrm{Var}[N^*|n_1,n_2]$ is evaluated at $n_2=E[n_2/n_1]$. This means that in the former, n_2 is replaced by $E[n_2/n_1]$ in the variance formula. There is therefore little difference between treating n_2 as a fixed parameter or as a random variable as far as estimation is concerned. But it can be argued that once n_2 is known, we are only interested in the distribution of m_2 given n_1 and n_2, and $f(m_2|n_1,n_2)$ is then the appropriate distribution irrespective of whether n_2 is fixed or random.

Confidence Intervals

As $N \longrightarrow \infty$, N^* is asymptotically normally distributed,so that the approximate 95% confidence interval for N is given by

$$N^* \pm 1.96V^*$$

However, according to Ricker (1958), $1/N^*$ is more symmetrically distributed and more nearly normal than N^*, so that in general it is better to base confidence intervals on the probability distribution of m_2. Now, we know that the distribution of m_2 given n_1 and n_2 is hypergeometric. The hypergeometric distribution (3.1) has been tabulated and exact confidence limits for $p = n_1/N$ when N is known and n_1 unknown are available (Chung and Delury (1950)). Unfortunately no such tables are available for the case when N is unknown and n_1 known, so that the approximate methods have to be used. For various values of n_1, n_2 and N the hypergeometric distribution can be satisfactorily approximated by the Poisson, binomial and Normal distributions. But the choice of which approximation to use when N is unknown still needs further investigations, so that the following recommendations should be regarded as a general guide only.

(i) Let $\hat{p} = m_2/n_2$; then when $\hat{p} < 0.1$ and $m_2/n_1 < 0.1$, the Poisson approximation is recommended using m_2 as the *entering* variable in appropriate tables. For example, a confidence interval for $\mu = n_1 n_2/N$ can be read off from tables such as Pearson and Hartly (1966), Crow and Gardner (1959) or from a graph (Adams (1951)).

For $m_2 < 50$, it is simpler to use a table specially prepared by Chapman (1948), giving the shortest confidence intervals for N/λ where $\lambda = n_1 n_2$. This table is reproduced in Appendix 1, and we demonstrate its use with the following example.

Example 1

Suppose $n_1=1000$, $n_2=500$, $m_2=20$ then $\hat{p}=0.04$, $m_2/n_1=0.02$ and the poison approximation is appropriate using m_2 as the entering variable a 95% confidence internal for N/λ is (0.03004, 0.0773) that is

$$Pr[0.03004 \le N/\lambda \le 0.0773] = 0.95$$

and so

$$Pr[15020 \le N \le 38650] = 0.95$$

Thus the 95% confidence interval for N is (15020,38650).

(ii) When $\hat{p} < 0.1$ and $m_2>50$, we can use the normal approximation to hypergeometric distribution given by Cochran (1977) to obtain a 95% confidence interval for p, namely

$$\hat{p} \pm \{1.96\{[1-f)\hat{p}(1-\hat{p})/(m_2-1)^2 + 1/2n_2\} \qquad (3.22)$$

which can be inverted to give a confidence for N. Here $f=n_2/N$, the unknown sampling fraction, can be neglected if its estimate $\hat{f} = m_2/n_1$ is less than 0.1; also $1/2n_2$, the correction for continuity, will often be neglected. We note that, neglecting f when $f > 0.01$ will lead to conservative confidence interval that is intervals which are over wide.

(iii) When $N > 150$, $n_1 > 50$, $n_2 > 50$, m_2 is approximately normally distributed and the most accurate method than the above is to solve the cubic equation in N

$$\frac{N^2(m_2 - n_1 n_2/N)^2(N-1)}{n_1 n_2 (N-n_1)(N-n_2)} = 1.96^2$$

The two largest roots then give an approximate 95% confidence interval for N. The graphical solution is discussed by Schaefer (1961).

(iv) If $\hat{p} > 0.1$, we can use either the binomial approximation or normal approximations named above. A rough guide as to the smallest value of n_2 for which the normal approximation is applicable is given by the following table reproduced from Cochran (1977).

Table 3.1: Smallest value of n_2 for which the normal approximation is applicable.

$\hat{p}$ or $(1-\hat{p})$	0.5	0.4	0.3	0.2	0.1	0.05	~0*
n_2	30	50	80	200	600	1400	∞

When the normal approximation is not applicable, a binomial confidence interval for p can be obtained from the Cloper-Pearson Charts in Pearson and Hartley (1966), or for extensive binomial tables such as those of Harvard Computational Laboratory (1955).

3.2 VALIDITY OF ASSUMPTIONS

Assumption of Closed Population

If the assumption of constant population size is to hold, the experiment should be carried out over a short period of time, in fact, ideally at a single point in time. For this reason the Petersen method is often called a *point census*.

Departures from this assumption can occur in a number of ways and we now discuss these in detail.

Accidental deaths

If there are d accidental deaths through the process of catching and marking the first sample, the general theory in section 3.1, remains unchanged provided that n_1 now refers to the number returned alive to the population, $\hat{N}$ and N^* are estimates of N-d.

Natural Mortality

Suppose that mortality is taking place in the time between the two samples, and let N be the size of the population when the first sample is released. When the assumptions IV,V and VI of section 3.1, are true and the only departure from I is due to mortality, the hypergeometric model (3.1) still holds, provided the mortality process is such that the deaths constitute a simple random sample of known size. This follows from the fact that the survivors will also constitute a simple random sample from which the second sample is a random sub-sample. Since a random sub-sample of a random sample is itself a random sample, the second sample will still represent a simple random sample from the original population.

When deaths do not constitute a simple random sample, the Petersen estimate can still be used provided the marked and unmarked have the same average probability of surviving up till the time of the second sample. This can be seen intuitively from the equation.

$$E\left(\frac{m_2}{n_2}\middle| n_1\right) \cong \frac{\phi n_1}{\phi N} = n_1/N \qquad (3.23)$$

We note that mortality is often selective with regard to the size or age of the animal. However, if the first sample is a simple random sample, the more "vulnerable" individuals will be proportionately represented in both marked and unmarked populations, thus ensuring that marked and unmarked have the same survival probabilities.

To examine the effect of variable mortality in the marked proportion, suppose that there are various subcategories in the population with numbers x, y,...., w where N = x + y +....+w. For the marked members of these categories, let $\phi_x, \phi_y, \ldots, \phi_w$ be the respective survival probabilities and let p_{2x}, $p_{2y}, \ldots, p_{2w}$ be the probabilities of recapture in the second sample. Then using suffix x to denote membership of category X, we have

$$E[m_{2x}|n_{1x}] = \phi_x p_{2x} \qquad (3.24)$$

and we can test the hypothesis

$$H_o: \phi_x p_{2x} = \ldots = \phi_w p_{2w}$$

using a standard chi-squared test based on the following contingency table:

Table 3.2: Contingency table classifying different categories of animals as captured and not captured

	X	Y	...	W	TOTAL
Captured	m_{2x}	m_{2y}		m_{2w}	m_2
Not capt.	$n_{1x}-m_{2x}$	$n_{1y}-m_{2y}$	...	$n_{1w}-m_{2w}$	n_1-m_2
TOTAL	n_{1x}	n_{1w}	...	n_{1w}	n_1

When $\phi_x = \phi_y = \dots = \phi_w$, a test of H_o is a test that the second sample is random with respect to the marked individuals in the various categories. Conversely when a second sample is random so that , $p_{2x} = p_{2y} = \dots = p_{2w}$,this is a test of constant survival probability for marked members.

Catchable Population

It should be noted that N may sometimes refer to the catchable proportion of the population only, and not to the whole population. If an approximately unbiased estimate $\hat{p}_C$ of the catchable proportion p_C of the population is available,then we can estimate the total N_T by

$$N_T = N^*/\hat{p}_C \tag{3.25}$$

If N^* and $\hat{p}_C$ are based on separate sampling experiments, as will usually be the case, they are statistically independent. Therefore.

$$E[\hat{N}] = E[N^*]\,E[1/\hat{p}_C] \tag{3.26}$$

Now, we know that, if x is a random variable with mean, θthen

$$E[g(x)] = g(\theta) + b$$

Where $g(x)$ is some function of x and

$$b = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{2} \ \mathrm{Cov}[x_i, x_j] \frac{\partial^2 g}{\partial x_i \partial x_j}$$

$$= \sum_{i=1}^{n} \frac{1}{2} \ \mathrm{Var}[x_i \ ] \frac{\partial^2 g}{\partial x_i^2} + \sum_{i<j}\sum \mathrm{Cov}[x_i, x_j] \frac{\partial^2 g}{\partial x_i \partial x_j}$$

Therefore

$$E[g(x)] \cong g(\theta) + \sum \frac{1}{2} \ \mathrm{Var}[x_i] \ \frac{\partial^2 g}{\partial x_i^2}$$

It follows that equation (3.6) now becomes

$$E[\hat{N}] = E[N^*]\{ \ 1/p_C + V[\hat{p}_C]/p_C^3 \ \}$$

$$= N\{ \ 1/p_C + \mathrm{Var}[\hat{p}_C]/p_C^3 \ \}$$

$$= N_T\{ \ 1 + \mathrm{Var}[\hat{p}_C]/p_C^2 \ \} \qquad (3.27)$$

and

$$\mathrm{Var}[\hat{N}_T] \cong \mathrm{Var}[N^*]/p_C^2 + \mathrm{Var}[\hat{p}_C]N_T^2/P_C^2 \qquad (3.28)$$

If binomial sampling is used to obtain $\hat{p}_C$ then,

$$v[\hat{p}_C] = p_C q_C/N \ ,$$

where N is the number of animals investigated for catchability. Assuming that $\hat{N}_T$ is approximately normally distributed, an approximate confidence interval for N_T can be calculated in the usual manner. However, unless $\hat{p}_C$ and N^* are accurate estimates, this interval may be too wide to be of much practical use.

Recruitment

Sometimes the time lapse between two samples is sufficient to allow the recruitment of younger animals into the catchable part of the population. These recruits will tend to decrease the proportion of the marked in the second sample, and Petersen estimate $\hat{N}$, would overestimate the initial population size. In the situation where there is recruitment but no mortality, $\hat{N}$ will be a valid estimate of the population number at the time when the second sample is taken. However when both the recruitment and mortality occur, $\hat{N}$ will overestimate both the initial and final population size. This is seen mathematically by noting that if there are r recruits in the population at the time of the second sample, we have the approximate relation

$$E\left[\frac{m_2}{n_2} \mid n_1\right] \cong \phi n_1/(\phi N+r) \qquad (3.28)$$

or

$$E[\hat{N} \mid n_1] \cong N + r\phi^{-1}$$

where $N + r\phi^{-1}$ is greater than N and $\phi N + r$, $(0 < \phi < 1)$. By enlarging the definition of r to include permanent immigration and redefining ϕ as the average probability that an animal in the population at the release of the first sample is alive and still in the population at the time of the second sample, then, provided ϕ is the same for marked and unmarked, the above comments apply to a population in which there is immigration and emigration also.

If an animal becomes immediately catchable as soon as it reaches a certain age, then an age analysis of the second sample would provide an estimate of the ratio of recruits to non-recruits. Using this ratio and the Petersen estimate of non-recruits we could then obtain an estimate of total recruits. But the process of age determination is usually time consuming and the *threshold* for recruitment is not usually well defined so that partial recruitment may occur over a range of younger ages. Usually the more readily available information such as length or weight is used to classify the individual, and such data can be used for carrying out the following tests of recruitment.

Chi-Squared test.

If the individual measurement are not actually recorded and the animals are simply allotted to the particular size-classes, we can test for recruitment as follows. Let x, y,..., w denote both the classes and the numbers in the classes at the beginning of the experiment and suppose that x increases to x+r , and so on, through recruitment. Then if the second sample is random within the classes (but not necessarily between classes), we have

$$E\left(\frac{m_{2x}}{n_{2x}} \middle| n_{1x}\right) = \underset{n_{1x}}{E}\, E \left(\frac{m_{2x}}{n_{2x}} \middle| n_{1x}\right) = \underset{n_{1x}}{E}\left(\frac{n_{1x}}{x+r_x}\right) = p_{1x} \text{ ,say} \quad (3.40)$$

and the hypothesis H_o: $p_{1x} = p_{1y} = \ldots = p_{1w}$ can be tested using a standard Chi-Squared statistic based on the contingency table (3.3) below. When there is no recruitment, so that $r_x = 0$ for each class, then this will be a test that the first sample is random with respect to size-class. Conversely, if the first sample is a simple random sample then $E[n_{1x}/x]$ will be the same for each class, and since there is no recruitment in the classes with larger animals, a test of H_o will then amount to a test of $r_x = 0$ for all the classes. In this later situation, the test will be unaffected by mortality provided that the survival probabilities are the same for marked and unmarked. This follows from the simple relationship.

$$E\left(\frac{m_{2x}}{n_{2x}}\right) = E\left(\frac{\phi n_{1x}}{\phi_x X + r_x}\right) = E\left(\frac{n_{1x}}{X + r_x \phi_x^{-1}}\right) \qquad (3.41)$$

Where ϕ_x is the average survival probability for class X. Finally, it is noted that the above test can still be used, even when animals grow from one class into another with unknown overlap during the course of the experiment.This is because although size-classes are usually determined from the second sample, X may be regarded as the *conceptual* population, existing at the time of the first sample, which grows into the required class; n_{1X}will then be unknown.

Table 3.3

	Size class				
	X	Y	. . .	W	TOTAL
marked	m_{2x}	m_{2y}	. . .	m_{2w}	m_2
unmarked	$n_{2x} - m_{2x}$	$n_{2y} - m_{2y}$	. . .	$n_{2w} - m_{2w}$	$n_2 - m_2$
TOTAL	n_{2x}	n_{2y}	. . .	n_{2w}	n_2

Non-Parametric Test

This test is due to Robson and Flick (1965). In the following discussion, the word "length" will denote some readily available measurement of size. Let us define a variable $u = n_2 - m_2$. We shall assume that n_1 is sufficiently large and the first sample sufficiently random for the length distribution of animals to be the same for both marked and unmarked. Suppose that the lengths of m recaptures are $L_1 < L_2 < \ldots < L_m$ and let u_i

$(i=1,2...,m+1)$ be the number of unmarked animals caught in the second sample with length L in the interval $L_{i-1} < L < L_i$ $(L_0 = 0,\ L_{m+1} = \infty)$. If the second sample is random with respect to mark status and length, the probability that the length of an unmarked animal falls into any one of the above $m + 1$ length classes is $1/(m+1)$ and the expected value of u_i will be $u/(m+1)$. However if recruitment has occurred in the shorter size range, the observed u_i for the intervals in this range will be greater than expected. Thus if recruitment in the length class $[\ 0,L_1\ ]$ has occurred, n_1 will be significantly larger than $u/(m + 1)$. To determine the significance of u_1 we calculate the tail probability

$$\Pr\ [\ U_1 \geq u_1\ ] = \binom{u+m-n_1}{m} \Big/ \binom{u+m}{m} \qquad (3.42)$$

where U_1 is the random variable taking the values u_1 and compare this with the significance level α.

The same procedure can now be applied to the second length class $[\ L_1,\ L_2\ ]$ by eliminating L_1 and the class of u_1 animals from the data. Thus we compare U_2 with the expected value $(u-u_1)/m$ and if this difference is significant, we compare u_3 with $(u-u_1-u_2)/(m-1)$, and so on. Proceeding in this step wise fashion through the larger classes, the recruits, if any, will dwindle in number until the rth step, say, is reached when the recruits no longer make any significant contribution. Thus u_r will not be significantly greater than $(u-a_r-1)/(m-r+2)$, where $a_r = u_1+u_2+...+u_r$; for this step the tail probability is

$$\Pr[U_r \geq u_r \mid u_1, u_2, \ldots, u_{r-1}] = \binom{n-a_r-r+1}{m-r+1} \Big/ \binom{n-a_{r-1}-r+1}{m-r+1} \quad (3.43)$$

This would suggest that the remaining sample of $U_{r+1} + U_{r+2} + \ldots + U_{m+1}$ unmarked animals is free of recruits, and the average

$$\bar{U}_{r+1} = \{U_{r+1} + \ldots + U_{m+1}\}/\{m+1-r\}$$

is therefore an estimate of the number of unmarked non-recruits that should occur between adjacent pair of marked recaptures. Hence the estimated number of non-recruits in the second sample is $(m+1)\bar{U}_{r+1}$ and the modified Petersen estimate of N becomes

$$N^* = \frac{(n+1)\{U_{r+1}(m+1) + m+1\}}{m+1} = (n_1+1)[\bar{U}_{r+1}+1] - 1$$

Mean and Variance of N^*.

In evaluating the mean and variance of N^* we run into the difficulty of r being a random variable. This problem arises, for example, in fitting a polynomial regression where the degree of the final polynomial is strictly a random variable. However, as with the regression problem, treating r as though it were a constant would not seem unreasonable and would perhaps lead to a slight underestimate of $\mathrm{Var}[N^*]$.

Now, under the assumption of non-recruitment after the rth class we have

$$E[U_{r+1} \mid m, r, a_r] = (u-a_r)/(m+1-r)$$

Thus

$$E[N^{*}|m,r,a_{r},n_{1}] = \{(n_{1}+1)(n-r-a_{r}+1)/(m-r+1)\} - 1$$
$$=\{ (n_{1}+1)(n'+1)/(m'+1) \} - 1 \qquad (3.44)$$

Where m' and n' are simply the values of m and n obtained by truncating the second sample of length L_{i}. If this truncation successfully eliminates recruits we would then expect

$$E[N^{*}/n_{1}] \cong N \ .$$

Also

$$V[N^{*}|n_{1}] = (n_{1} + 1)^{2}\mathrm{Var}[U_{r+1}] \qquad (3.45)$$

where the variance of $\bar{u}_{r+1}$ can be estimated robustly from replicated u's, namely

$$\mathrm{Var}[u_{r+1}] = \frac{1}{(m+1-r)(m-r)}\sum_{i=1}^{m+1-r} (u_{r+i} - \bar{u}_{r+2}) \qquad (3.46)$$

Combining classes to Improve the Test Procedure:

As the recruitment will generally tend to decrease with increasing body-length there will be a decrease in the probability of detecting these recruits. Also this decrease in detectability is further accentuated by a decrease in the length interval between marked animals as the test progresses from the lower tail toward the centre of the length distribution, and also by the decrease in sample size resulting from the successive removal of the intervals tested. To overcome this difficulty, we therefore require, a method of pooling intervals as the number of recruits falls off.

A further need for combining intervals arises when length measurements are sufficiently crude to permit ties to occur. In particular, if several marked animals are recorded as having the same body-length, then the resulting degenerate intervals must be combined to include all marked animals having the same recorded length. It transpires that the optimal pooling procedure is simply to combine adjoining intervals giving a new total interval and a new total number of unmarked. If the first k intervals are combined, then a test for recruitment in this total interval has a tail probability of

$$T[S_k] = \Pr[S_k \geq s_k\ ] = \sum_{r=o}^{k-1} \binom{s_k+r-1}{r} \binom{u+m-s_k-r}{m-r} \Big/ \binom{u+m}{m} \qquad (3.47)$$

where S_k $(=u_1+u_2 + \ldots + u_k)$ is the number of unmarked animals in the total interval.

As k gets large, the above tail probability becomes computationally awkward, and for small values of s_k the

recursive relation

$$\Pr[S_k=s_k] = \frac{(s_k+k-1)(u-s_k+1)}{s_k(k+m-s_k-k+1)} \Pr[S_k=s_k-1] \qquad (3.48)$$

where

$$\Pr[S_k=0] = \binom{u+m-k}{u} \Big/ \binom{u+m}{m} .$$

is useful When u is much greater than m, so that

$$(m+1-k)/kn \cong o \quad \text{and} \quad u(n+1)/n^2 \cong 1$$

the napplication of Stirling's formula to (3.48) gives

$$Pr[S_k \geq s_k] \cong \frac{\Gamma(m+1)}{\Gamma(k)\,\Gamma(m+1-k)} \int_p^1 t^{k-1}(1-t)^{m-k}dt.$$

$$= \sum_{i=1}^{k-1} \binom{m}{i} p^i (1-p)^{m-i}$$

where $p=(s_k+k-1)/n$. This incomplete beta approximation to (3.48) holds for all $K \geq 1$.

Suppose now that, the first k_1 intervals, the second k_2 intervals, and so on are combined to give new length class $[0,L_{k_1}),[L_{k_1}, L_{k_2}),\ldots$ with unmarked numbers S_{k}, S_{k2} and so on. Then if

$$Pr[S_{k_1} \geq S_k] < \alpha$$

We reject the hypothesis of no recruitment in the length class $[0, L_{k_1})$ at α level of significance and proceed to consider S_{k_2}. Dropping the date in the first interval from the sample, we now evaluate the tail probability

$$T[S_{k_2},S_{k_1}] = Pr[S_{k_2} \geq s_{k_2} | S_{k_1}=s_{k_1}]$$

$$= \sum \binom{S_{k_2}+r-1}{r} \binom{u'+m'-s_{k_2}-r}{m'-r} \Big/ \binom{u'+m'}{m'}$$

where $m'= m-k_1$ and u' $u-S_{k_1}$ are new values of m and u.We again reject the hypothesis of no recruitment in $[L_{k_1}, L_{k_2}]$ if for S_{k_2}, S_{k_1}, and so on until non-significance is achieved. The number of non-recruits in the second sample can then be estimated as before. One of the problems in combining adjacent intervals is to determine the best sequence $k_1, k_2,\ldots$ since there is a practical possibility that all the recruits are smaller (shorter) than the smallest non-recruit, it would seem reasonable to use $k_1 = 1$. Also, because of the steady reduction in numbers of recruits between successively larger marked animals, the k-sequence should be increasing, so that $1 = k_1 < k_2 \leq \ldots$ unfortunately the optimum sequence can only be determined if length frequency distributions are known for both recruits and non-recruits, although as the statics S_{k_1}, S_{k_2},.... etc are virtually independent, the k-sequence could perhaps abe determined sequentially by regression methods. For example, extrapolating the regression of u on $i=2,3,\ldots,k_2$ to indicate the number k_3 of intervals which must be next combined in order to achieve the relation

$$u_{k_1+k_2+k_3} \cong \frac{u-s_{k_1}-s_{k_2}-s_{k_3}}{m-k_1-k_2-k_3+1} .$$

Further research needs to be done on such methods of finding a suitable k-sequence.

If one wishes to combine intervals still further (e.g the first $k_1 + k_2$ intervals), then, as mentioned above, the optimal procedure is simply to use the sum

$s_{k_1}+s_{k_2}(=S_{k_1+k_2})$ and evaluate the tail probability:

$$T[S_{k_1+k_2}] = Pr\,[S_{k_1+k} \geq S_{k_1+k_2}]$$

However, to avoid this additional computation Robson and Flick (1965) suggest a number of approximate procedures such as using $T[S_{k_1}] + T[S_{k_2};\, S_{k_1}]$ with significance level of $(2\alpha)^{1/2}$,or

$$T[S_{k_1}] + T[S_{k_2};\, S_{k_1}] + T[S_{k_3};\, S_{k_2};S_{k_1}]$$

with significance level of $(6\alpha)^{1/2}$ if three groups are pooled.

Assumption II, III and IV

1. *Practical considerations*

variable catchability.

One of the crucial assumptions underlying the theory of section 3.1, is that the second sample is a simple random sample. Strictly speaking, such a random sample can only be obtained by numbering the animals 1,2,...,N and using a table of random numbers to select n_2 animals. However, in practice, if all the animals have the same catchability, we can approximate to a random sample by arranging that every point of the population area has the same probability of being sampled and that all points selected are sampled with the same effort. If a more even coverage is required one can use stratified random sampling, whereby the population area is divided into equal sub-areas and one or more points are allotted at random within each sub-area. Unfortunately the requirement of constant probability of capture may not hold, either because of an inherent variation in catchability or because catching and handling in the first sample affect future catchability. Very often the probability of capture will vary between various sub-groups defined by age, sex, species, etc, for example certain sub-groups may be more mobile and have different habitat preference, while others may have certain bait and trap preferences.

In fisheries, catchability usually varies with the size of the fish, and considerable research has been carried out on such problems as **gear selectivity** and length-selection curves. When recruitment and mortality are negligible and the marked members are individually identifiable, then the randomness of the second sample with respect to size can be tested by partitioning the first sample into that portion which is ultimately recaptured and into the portion which is not (Robson ,1969).

A Mann-Whitney rank-sum test comparing these two sub-samples with respect to body size will be a test for a monotonic ration between body size and probability of capture in the second sample. Alternatively, if the marked members are allotted to size-classes one can carry out the goodness-of-fit test described earlier. When there is variation in the inherent catchabilities of individuals and the first sample is not random, assumption II will be false and the more catchable individuals will be caught in the first sample. This means that for the second sample, the marked will in general be more catchable than the unmarked, and assumption IV will be false. Unfortunately, apart from a careful choice of catching method and preliminary studies -for example, on activity, feeding habits, length-selection curves, etc -little can be done to overcome this problem of variable catchability in the first sample. However we shall see below that the bias in the Petersen estimate $\hat{N}$ due to variation in catchability can be reduced by using different trapping methods for the two samples.

If the catchability is constant within certain well-defined sub-groups and there are sufficient recaptures from each sub-group, the numbers in each sub-group should be estimated separately.

Systematic Sampling:

We have seen that the Petersen estimate can still be used, even when assumption IV is false and a systematic rather than a random sample is taken, provided there is uniform mixing of marked and unmarked and all animals are catchable in the second sample. But in many populations uniform mixing is unlikely because of territorial behaviour on animals and the presence of well-defined home ranges. Another situation where it is difficult to obtain a uniform mixing is when the population is not randomly distributed throughout the population area and animals are relatively immobile. For example Hancock (1963) suggested that the excessive variability in the monthly returns of marked Whelks may have been due to the random distribution of marked individuals among essentially non-randomly distributed unmarked individuals.

It would seem that where possible the experimenter should aim for a random sample rather than rely on the assumption of uniform mixing. However in many population studies it is helpful to arrange the release of the first sample so that mixing can take place as much as possible. For example, one can divide up the total area into sub-areas,sample each sub-area with same effort, and then release the marked animals back into the same area from which they were taken. If catchability is independent of sub-area, one would hope that this method produces roughly the same proportion of marked in each sub-area. To check this, a standard goodness-of-fit test can be carried out to see whether the proportion of marks recorded in the second sample from the various sub-areas are significantly different.

Catching and Handling:

Departures from assumption III that trapping and marking do not affect catchability can be minimized if the following points are observed by the experimenter.

i. Type of trap.

It is essential that a trapping method be used which will not harm the animal in any way. For example, in small mammal populations such things like placing the trap under cover of vegetation, drugging the bait, and visiting traps frequently can reduce trap mortality. If several types of traps are available one would endeavour to choose the type which is most efficient, as the accuracy of Petersen estimate increases with n_1 and n_2. Also to increase trap efficiency, the bait or lure should be selective for the species under investigations and some consideration should be given to the spacing and distribution of the traps.

ii. Method of handling.

Care is needed in handling the captured animals so that they quickly recover on their return to the population. Another problem arises with small mammals and birds where *trap addiction* or *trap shyness* can alter an animals′ pattern of behaviour after it has been caught for the first time. The effect of trap shyness can be minimized by prebaiting the traps for a suitable period of time before the census, thus allowing the animals to get used to the presence of traps. This however does not always work and of course trap addiction is not helped by prebaiting. Sometimes trap addiction can be reduced by altering the trap positions throughout the trapping period, thus, for example, preventing individuals from building their runways up to the mouth of the trap. Another way of minimizing the effect of trap response is to use a different trapping method for taking the second sample. For example animals can be live-trapped for marking and then shot for recapture; if the mark is conspicuous, sight of the mark itself could be the means of *recapture*. In the later case, observing animals and noting the proportion of marked amounts to sampling with replacement, so that Bailey's binomial model is appropriate.

In some circumstances the tag itself may affect the longevity and behaviour of the animals. For example, jaw tags on fish can interfere with feeding and thus affect growth rate, while Petersen disk tags can make fish more vulnerable to gill nets through the net catching under the disk. Another aspect of marking particularly relevant to insects is that presence of conspicuous marks may well destroy an animal's natural camouflage and make it more or less liable to predation. Also when animals are marked by a method which relies on the sight of the collector then the marked animals may tend to be collected more than the unmarked. On the other hand, if the tags are not conspicuous enough they may be overlooked, particularly if one relies on huntsmen, fishermen, farmers, and so on to return the tags.

iii. Method of release:

Animals often show a high level of activity immediately on release, and efforts should be made to minimize this. For example, birds and insects could perhaps be restrained from flying immediately by covering them with small cages until the effects of handling wear off. While tagged fish could be held in tanks so as to reduce and measure initial tagging mortality. If animals have activity during the day they could perhaps be released during an active period.

Theoretical analysis of catchability

For the ith member of the population (i=1,2,..., N), let x_i be the probability that it is caught in n_1, let y_i be the conditional probability that it is caught in n_2 given that it is caught in n_1 and let z_i be the conditional probability that it is caught in n_2 given that it is not caught in n_1. Assuming that the population represents a random sample of N triples (x_i,y_i,z_i) with regard to the species as a whole and the particular method used, then (x_i, y_i,z_i) may be regarded as a random observation from a trivariate probability density function $(f_1(x,y,z)$, say. If we set $a_i=n_i-m_2$ (the number of animals caught in the ith sample, only, i=1,2), then the probability of a given outcome of the whole experiment.

$$P_w=\left[\prod_{i=1}^{a_1} x(1_i-y_i)\right]\left[\prod_{i=a_1+1}^{a_1+a_2}(1-x_i)z_i\right]\left[\prod_{k=a_1+a_2+1}^{a_1+a_2+m_2} x_k y_k\right]$$

$$\times\left[\prod_{l=a_1+a_2+m_2+1}^{N}(1-x_l)(1-z_l)\right]$$

Therefore the conditional probability function of a_1, a_2, m_1 is

$$f_2(a_1,a_2,m_1|x_i,y_i,z_i) = \sum_w P_w$$

where $\sum_w$ denotes summation over all possible groups of the N animals such as a_1, a_2, m_2 are the numbers in each of the three categories, that is w represents a permutation of N objects such that a_1 fall in the first category, and so on. Since the triple (x_i, y_i, z_i) are independent they can be integrated out and the unconditional probability function is given by

$$f_3(a_1,a_2,m_1)=Ef_2(a_1,a_2,m_1|x_i,y_i,z_i)$$

$$= \sum_w E\ [P_w]= \sum (p_1-p_{12})^{a_1}(p_3-p_{13})^{a_2}(1-p_1-p_3+p_{13})^{N-a_1-n_2}$$

where

$$\sum_w 1 = N!/\{a_1!a_2!m_1!(N-a_1-n_2)!\}$$

and $p_1 = E[x]$, $p_3 = E[z]$, $p_{13} = E[xz]$, and $p_{12} =E[xy]$, all expectations being with respect to f_1. Now

$$f_5(a_2,m_2|n_1) = f_3(a_1,a_2,n_1) \;/\; f_4(n_1)$$

$$= \begin{pmatrix} n_1 \\ m_2 \end{pmatrix} \left(\frac{p_{12}}{p_1}\right)^{m_2} \left(1-\frac{p_{12}}{p_1}\right)^{a_1} \begin{pmatrix} N-n_1 \\ a_2 \end{pmatrix} \left(\frac{p_3-p_{13}}{1-p_1}\right)^{a_1}$$

$$\times \left(\frac{1-p_1-p_3-p_{13}}{1-p_1}\right)^{N-n_1-a_2}$$

and approximating these two binomial distributions by two independent Poisson distribution we have (Feller,1957)

$$f_\sigma(m_2|n_1,n_2) = \binom{n_2}{m_2} P^{m_2}(1-P)^{n_2-m_2}$$

where

$$P= n_1 / \{(N-n_1)k + n_1\} \ , \ k=(p_3-p_{13})p_1 / \{p_{12}(1-p_1)\}$$

Therefore, it can be shown that

$$E[N^*|n_1,n_2] = \{ (n_1+1)/P \}\{ (1-(1-P)^{n_2+1} \} -1$$

$$\cong n_1/P =(N-n_1)k + n_1$$

so that N^* is approximately unbiased estimate of N if and only if $k = 1$. Suppose that assumption III is true. Let

$$k = (B-p_1)/(1-p_1)$$

where

$$B=p_1(p_3-p_{13}+p_{12})/p_{12}$$

obviously $k = 1$ if and only if $B = 1$ and $k < B$ when $B <1$. Since assumption III is true, that is marking does not affect catchability, we have

$$y_i = z_i \ (i=1,2,...N)$$

so that $p_{13}= p_{12}$, $p_3 = p_2$. and

$$1-B = 1-p_1p_2/p_{12}$$

$$= (p_{12}-p_1p_2)/p_{12}$$

$$= Cov[x,y]/E[xy]$$

suppose, now that $B = 1$, then, $Cov[x,y]/E[xy] = 0$ or $Cov[x,y] = 0$ and $Cov[x,y] = 0$ if and only if x and y are uncorrelated. Thus $B= 1$ if and only if x and y are uncorrelated.

A positive correlation will lead to an underestimate of N and a negative correlation will lead to an overestimate of N. We note that if a correlation exists we would generally expect it to be positive. This accounts for the persistent underestimation observed by Buck and Thoits (1965) who checked on several estimates of fish population numbers by draining the ponds containing the populations.

We conclude from the above that variation in catchability due to, say trapping selectivity could exist for both samples without introducing biase of the source of selectivity in two samples were independent. This supports the statements by a number of authors that bias due to difference in catchability can be reduced by using a different sampling method for each sample.

Special case:

Let $y_i = bx_i$ ($b > 0$). Here the correlation is unity and we find that

$$B = 1 - \mathrm{Var}[x]/\{E[x]\}^2 \le 1,$$

with equality if and only if $\mathrm{Var}[x] = 0$ or x is constant.

If $x \in [c,d]$ ($0 \le c \le d \le 1$) then, provided $W=c/d$ is not too small, it can be shwon that B will be nearer unity and insensitive to the shape of $f(x)$, the probability density of x.

On the other hand if $c=0$, B may be significantly less than unity; if in addition $f(x)$ is concentrated near $x=0$, that is the probabilities of capture are near zero for a substantial proportion of the population, then B could still be much smaller. In this case x and y are uncorrelated if either x or y is constant, that is $B=1$ if at least one of the two samples is a random sample. In particular if the first sample is random, the second sample need not be random and in fact could be highly selective, provided the selectivity was independent of mark status.

When x is constant we find that $p_{13}=p_1p_3$, $p_{12}=p_1p_2$ and $k=p_3/p_2$. This means that $k=1$ if and only if the average probability of capture of the marked in the second sample is the same as the average for the unmarked.

We see from the above discussion that if Assumption III is true and $y_i=bx_i$, then $k=1$ if and only if Assumption II is true. A test for this case based on taking two samples from a known (for example, marked) population is given below. Conversely, if Assumption II is true then k=1 if and only if Assumption III is true.

Test for constant catchability

We shall now consider the problem of testing Assumption I given that the assumptions I, II, III, V and VI are true, by taking 3 samples and using the first sample as an identifiable population.

Suppose that m_2 tagged animals in the second sample are given another tag and the second sample then returned into the population. If a third sample of size n_3 is now taken then on the basis of the tagging information obtained and assuming that catching and tagging do not effect future catchability, Cormack (1966) gives two procedures, one of which we discuss below, for testing the hypothesis of constant catchability.

In the following discussion we note that n_1, n_2 and n_3 play the same role as N, n_1, n_2 in theoretical analysis of the previous section above.

The probability p_j $(j=1,2,...,n_1)$ that the jth members of the first sample is captured in a later sample will be proportional to its inherent catchability and to the intensity of sampling or sampling effort. Assuming the sampling effort f to be the same for each individual, we therefore define $c_j=P_j/f$ as the catchability of the jth individual for the particular catching method used. We shall assume that c_j's to be a random sample of size n_1 from a probability density function $g(c)$ with moments μ'_r about the origin and moments μ_r about the mean. If we standardise c_j so that the domain of g is $[o,1]$, then f will be uniquely determined and $0 \leq f \leq 1$. The catchability of an animal may be regarded as the probability with which it places itself in a position where the experimenter is able to catch it; and the sampling intensity is then the probability that an animal in this position will be caught. Alternatively, if c_j is not standardised, we can regard it as the probability that one unit of sampling effort catches the jth individual. Then considering f as the number of units of effort to be additive, $p_j = fc_j$ as before. Let f_2 and f_3 be the sampling efforts for samples two and three respectively. Then if x_j and y_j are the probabilities that the jth member of sample one is caught in samples two and three respectively, and assuming assumption III is true, we have:-

$$x_j=f_2c_j,\ y_j=f_3c_j,\ \text{and hence}\ y_j=bx_j$$

Let m_{10} be those individual caught in the first sample, m_{12} those caught in both samples one and two only, m_{13}, those caught in both samples one and three only, and m_{123}

those caught in all the three samples. Then, the joint probability function of m_{12}, m_{13}, and m_{123} is given by

$$\frac{n_1!}{m_{12}!\ m_{13}!\ m_{123}!\,m_{10}!}\quad\left[\alpha_2(1-\alpha_3\lambda)\right]^{m_{12}}\quad\left[\alpha_3(1-\alpha_2\lambda)\right]^{m_{13}}$$

$$\times\ \left[\alpha_2\alpha_3\lambda\right]^{m_{123}}\ \left[1-\alpha_2-\alpha_3+\alpha_2\alpha_3\lambda\right]^{m_{10}}\qquad(3.49)$$

where

$$\alpha_2 = f_2\mu_1'\ ,\ \alpha_3 = f_3\mu_1' \ \text{ and }\ \lambda = \mu_2'/(\mu_1')^2$$

To test the hypothesis of constant catchability it is sufficient to test whether the variance of c_j is zero. Note that $\mathrm{var}[c_j] = 0$ if and only if c_j is a constant. This is equivalent to testing the hypothesis.

$$H_0: d = 0$$

where d, the square of the co-efficient of variance, is given by

$$d = \mu_2/(\mu_1')^2$$

$$=\lambda - 1 \qquad (3.50)$$

Now, for the multinomial distribution (3.49) the maximum likelihood estimates of α_2, α_3, and λ are simply the moment estimates.

$$\hat{\alpha}_1 = (m_{12}+m_{123})/n_1$$

$$\hat{\alpha}_2 = (m_{13}+m_{123})/n_2$$

$$\hat{\lambda} = n_1 m_{123}/\{(m_{12}+m_{123})(m_{13}+m_{123})\}$$

Therefore, writing $\hat{d} = \hat{\lambda}$, we have from the delta method described earlier, the asymptotic expressions are

$$E[\ \hat{d}\ ] = d+d(1+d)/n_1 \qquad (3.51)$$

and

$$\mathrm{Var}[\ \hat{d}\ ] = (d+1)\{1-(\alpha_2+\alpha_3)(d+1) + \alpha_2\alpha_3(d+1)(2d+1)\}/n_1\alpha_2\alpha_3$$

$$(3.52)$$

The asymptotic bias and variance of $\hat{d}$ can be estimated as usual by replacing each unknown parameter by its estimate; thus

$$\mathrm{Var}[\hat{d}]=\frac{m_{123}^2\, n_1^2}{(m_{12}+m_{123})^2\,(m_{13}+m_{123})^2}\left[\frac{1}{m_{123}}-\frac{1}{n_1}-\frac{m_{12}+m_{13}}{(m_{12}+m_{123})(m_{13}+m_{123})})\right] \tag{3.53}$$

Under the null hypothesis H_0 we have, assuming approximate normality, that $\hat{d}$ is $N(0,(1-\alpha_2)(1-\alpha_3)/n_1\alpha_2\alpha_3)$. Therefore a one sided test statistic for H_0 is given by

$$z = \hat{d}/\sqrt{n_1\hat{\alpha}_2\hat{\alpha}_3/\{(1-\alpha_2)(1-\alpha_3)} \tag{3.54}$$

which is approximately distributed as the standardized normal when H_0 is true. If Z is negative we accept $d=0$ as the most reasonable hypothesis. It is readily seen that the power of this test will be maximized when f_2, f_3 and n_1 are as large as possible.

Suppose we relax the assumption of a closed population to the extent of allowing mortality. Let ϕ_1, ϕ_2, be the probabilities of survival for a tagged animal between the first two and the second two samples respectively. To allow for the estimation of the ϕ_i , we require additional information provided by releasing a further r_2 tagged animals in addition to the m_2 into the population after the second sample and noting the recaptures, m_{23} say, from this group is in the third sample. Under the assumptions of the above test, and assuming the members of r_2 to represent a random sample of catchabilities from $g(c)$, the joint probability of m_{12}, m_{13}, m_{23}, m_{123} is given by

$$\frac{n_1!}{m_{12}!m_{13}!m_{123}!m_{10}!}\left[\phi_1\alpha_2(1-\phi_2\alpha_3\lambda)\right]^{m_{12}}\left[\phi_1(1-\alpha_2\lambda)\alpha_3\phi_2\right]^{m_{13}}$$

$$\times\left[\phi_1\alpha_2\phi_2\alpha_3\lambda\right]^{m_{123}}\theta^{m_{10}}\begin{bmatrix}r_2\\m_{23}\end{bmatrix}(\phi_2\alpha_3)^{m_{23}}(1-\phi_2\alpha_3)^{r_2-m_{23}}$$

where $\theta = 1-\phi_1\alpha_2-\phi_1\phi_2\alpha_3(1-\alpha_2\lambda)$.

In the above probability function we have four independent observations but five parameters, and we find that ϕ_2 and α_3 cannot be estimated separately; only the product $\phi_2\alpha_3$ can be estimated. The maximum likelihood estimates are

$$\hat{\phi}_1 = (m_{12}+m_{123})r_2/m_{23}n_1$$
$$\hat{\alpha}_2 = m_{23}(m_{12}+m_{123})/\{r_2(m_{13}+m_{123})\}$$
$$\hat{\phi}_2\hat{\alpha}_3 = m_{23}/r_2$$
$$\hat{\lambda} = m_{123}r_2/\{m_{23}(m_{12}+m_{123})$$

setting $\hat{d} = \hat{\lambda}-1$ we find as before the asymptotic mean and variance to be

$$E[\hat{d}] = d + \frac{\lambda}{r_2}\left(\frac{1}{\phi_1\alpha_3}-1\right)$$

and

$$\mathrm{Var}[\hat{d}] = \frac{\lambda}{n_1\phi_1\alpha_2\phi_2\alpha_3}\left\{1+\lambda(\phi_1\alpha_2n_1r_2^{-1}-\phi_2\alpha_3-\phi_1\alpha_2\phi_2\alpha_3n_1r_2^{-1})\right\}$$

To test the hypothesis $d = 0$ we again use a one tailed test based on the statistic

$$z = (\hat{d}+\hat{b})/\sigma$$

Where $\hat{b} = \hat{\lambda}(1/m_{23}-1/r_2)$, and

$$\hat{\sigma} = \frac{r_2^2\,m_{123}^2}{m_{23}^2(m_{12}+m_{123})^2}\left\{1/m_{123}+1/m_{23}-1/(m_{12}+m_{123})-1/r_2\right\}$$

In testing $H_0{:}d=0$, there is unfortunately a very considerable loss in power having to estimate ϕ_1. Cormack (1966) states that even when the death-rate is actually zero ($\phi_1 = \phi_2 = 1$), five to ten times the number of tagged animals n_1+r_2 are required to give discriminatory power as the test for a closed population. Because of this lack of sensitivity and the need for such large numbers of tagged animals Cormack (1966) suggests that the experiment should be arranged so that the probability of death can be neglected.

If information on ϕ_1 is required we can use $\hat{\phi}_1$ and the approximate variance formula

$$\mathrm{Var}[\hat{\phi}_1] = \frac{\phi_1}{n_1 r_2 \phi_2 \alpha_3}\{r_2(1-\phi_1\phi_2\alpha_3) + n_1\phi_1(1-\phi_2\alpha_3)\}$$

to obtain an approximate confidence interval for ϕ_1. We note that if $f_3/f_2 = \alpha_3/\alpha_2$ is known, can obtain the estimates

$$\hat{\alpha}_3 = \hat{\alpha}_2 f_3/f_2$$
$$\hat{\phi}_2 = m_{23}/r_2\hat{\alpha}_3$$

When $\phi_1=1$ and f_3/f_2 is known, the second release of r_2 animals is unnecessary as the parameters α_2 , ϕ_2 , and λ can be estimated from the joint multinomial distribution of m_{12}, m_{13},m_{123}, m_{10}. In this case the maximum-likelihood estimates are again the moment-estimates, and their asymptotic means and variances can be derived using the delta method.

Assumption of No loss of tags

If animals lose their tags or marks, the observed recaptures will be smaller than the expected and N^* will be overestimated N. Therefore considerable thought should be given to the choice of tag, and some experiments should be carried out either before or during the census experiment period to check for the tag losses or tag deterioration. The type of tag chosen will depend on such factors as the species studied, the information required by the experimenter, time and personnel available for the tagging, and the method of tag return- whether by hunter, fisherman, or research worker. obviously tags should be durable so that they are not lost through the effects of weather or physical changes in the animal, such as molting.

Estimation of tag loss.

One simple method Of detecting tag loss is to give all the n_1 animals in the first sample two types of tags and then to note those recaptures with just one tag and those with both tags intact. Denoting the two types of tags by A and B, we define

π_x =Probability that a tag of type x is lost by the time of the second sample (x=A,B).

Π_{AB}=The probability that both tags are lost.

m_x = number of tagged animals in the second sample with tag x only (x = A,B).

m_{AB}= number of animals in the second sample with both tags.

and

m_2 = members of n_1 caught in n_2 .

Assuming that the tags are independent of each other that is $\pi_{AB} = \pi_A\pi_B$,the joint probability function of m_A ,m_B, $m_{AB}$and m_2 is given by

$$f(m_A, m_B, m_{AB}, m_2 | n_1, n_2) = f(m_A, m_B, m_{AB} | m_2) f(m_2 | n_1, n_2)$$

Where

$$f(m_A, m_B, m_{AB}/m_2) = \frac{m_2!}{m_A!, m_B!, m_{AB}!, m_O!} \{(1-\Pi_A)\Pi_B\}^{m_A} \{\Pi_A(1-\Pi_B)\}^{m_B}$$

$$\times \{(1-\Pi_A)(1-\Pi_B)\}^{m_{AB}} \{\Pi_A \Pi_B\}^{m_O}$$

$$m_O = m_2 - m_A - m_B + m_{AB}$$

and

$$f(m_2/n_1, n_2) = \binom{n_1}{m_2} \binom{N-n_1}{n_2-m_2} \Big/ \binom{N}{n_2}$$

The maximum - likelihood estimates of N, m_2, π_A, π_B which are also the moment estimates are given by

$$\hat{N} = n_1, n_2 / \hat{m}_2$$

$$m_A = \hat{m}_2 (1-\hat{\pi}_A)\hat{\pi}_B$$

$$m_B = \hat{m}_2 (1-\hat{\pi}_B)\hat{\pi}_A$$

$$m_{AB} = \hat{m}_2 (1-\hat{\pi}_A)(1-\hat{\pi}_B)$$

Which have solutions

$$\hat{\pi}_A = m_B/(m_B+m_{AB})$$

$$\hat{\pi}_B = m_A/(m_A+m_{AB})$$

$$\hat{m}_2 = (m_A+m_{AB})(m_B+m_{AB})/m_{AB}$$

$$= c\ (m_A+m_B+m_{AB}), \text{say.}$$

This means that the observed recaptures $m_A+m_B+m_{AB}$ must be corrected by a factor

$$c = \left[1 - \frac{m_A\ m_B}{(m_A+m_{AB})(m_B+m_{AB})}\right]^{-1} = \frac{1}{(1-\hat{\pi}_A)(1-\hat{\pi}_B)}$$

to give an estimate of the actual number of recaptures m_2. For large samples, $\hat{N}_{AB}$ is an approximately unbiased estimate of N ,and defining $\hat{N} = n_1 n_2/m_2$, we have from the fact that if x and y are two random variables ,then

$$V[x] = E_y[V[x|y]] + V_y[E[x|y]]$$

that is

$$V[\hat{N}_{AB}|n_1,n_2] = E_{m_2}\{V[\hat{N}_{AB}|n_1,n_2,m_2]\} + V_{m_2}\{E[\hat{N}_{AB}|n_1,n_2,m_2]\}$$

$$\cong E_{m_2}\{V[\hat{N}_{AB}|n_1,n_2,m_2]\} + V[\hat{N}|n_1,n_2]$$

which by delta method and the relation

$$V[N^* \ |n_1,n_2] = N^2(\mu^{-1} + 2\mu^{-2} + 6\mu^{-3})$$

is approximately equal to

$$\frac{N^3\pi_A\pi_B}{n_1 n_2}\left\{\frac{1}{(1-\pi_A)(1-\pi_B)}\right\} + \frac{N^3}{n_1 n_2}\left\{1 + \frac{2N}{n_1 n_2} + \frac{6N^2}{n_1 n_2}\right\} \qquad (3.55)$$

In some situations the only information recorded is the number of tags from each tagged individual,so that just the number m_{AB} and $m_C=m_A+m_B$ are available. For this case we can still estimate m_2 if we can assume that $\pi_A=\pi_B=\pi$,say. We then have

$$f(m\ ,m\ \ |m\ \ )=\frac{m_2!}{m_C!\,m_{AB}!\,m_O!}\left\{2\pi(1-\pi)\right\}^{m_C}\left\{(1-\pi)\right\}^{m_{AB}}\ \pi^{2m_O}$$

and the maximum-likelihood estimates (and moment estimates)of m_2 and π are

$$\tilde{m}_2 = (m_C + 2m_{AB})^2/4m_{AB}$$

and (3.56)

$$\tilde{\pi} = m_C/(m_C + 2m_{AB})$$

Now setting

$$\tilde{N}_{AB} = n_1 n_2/\tilde{m}_2 \qquad (3.57)$$

The estimate(3.57) is asymptotically unbiased and its asymptotic variance is given by

$$\frac{N^3\pi^2}{n_1 n_2(1+\pi)^2} + \frac{N^3}{n_1 n_2}\left\{1 + \frac{2N}{n_1 n_2} + \frac{2N^2}{n_1^2\ n_2^2}\right\}$$

When information on m_A and m_B is available for the whole or perhaps a part of the second sample,we can test the hypothesis $\pi_A=\pi_B$ as follows:

Let $\pi_B=k\pi_A$; then,the conditional probability function of m_A given m_A+m_B is given by

$$f(m_A|m_A+m_B\ ) = \binom{m_A+m_B}{m_A} p^{m_A} q^{m_B} \qquad (3.58)$$

Where

$$p = \frac{k(1-\pi_A)\pi_A}{k(1-\pi_A)\pi_A + \pi_A(1-\pi_A)} \quad \text{and } p=1-q$$

Testing k=1 is therefore equivalent to testing $p=1/2$ for the binomial distribution(3.8).

It is noted that when tag losses are small, which will be the case for many populations, the above variance terms involving π_A,π_B and π will be negligible, so that the effect of tag loss can be neglected. For example, when

π_A and π_B are both less than 0.1 ,the contribution of the first expression to equation (3.55) is still less than 3.5per cent. In such cases, the general theory of section 3.1 can still be used with m_2 replaced by $\hat{m}_2$ or $\tilde{m}_2$.

Note that the above theory can be extended to the case of more than two marks by defining π_A as the probability that a particular mark is lost and π_B as the probability of loosing at least one of the other marks.

Assumption that all tags are reported:

When there are incomplete tag returns the observed value of m_2 will be too small and N^* will overestimate N . This problem arises when tags are returned by hunters,commercial fishermen, local inhabitants and so on, who may or may not be interested in the experiment. It is found that the percentage returns is usually related to such factors as the training of observers, size of reward, publicity given to the experiment, and ease of visibility of the tag. However, if the second sample can be classified into two categories, one which has a known reported ratio of unity or nearly so, and the other with an unknown reported ratio, then Paulik's(1961) method described below can be used to test whether the unknown ratio is significantly less than unity.

Let $n_2=n_{2a}+n_{2b}$, where the suffixes a and b denote the categories respectively.

Let m_{2a} , m_{2b} , be the number of of recaptures in the two groups and let r_{2a}, r_{2b} be the number of recaptures actually reported that is $r_{2a}=m_{2a}$. Then if n_2

is large, the ratio n_1/N small, and the sample ratios the same for both groups , that is $n_{1a}/N = n_{1b}/N$, we can use the Poisson approximation to the hypergeometric distribution and assume that the recaptures m_{2i} in each group have independent Poisson distributions with parameters $n_1 n_{2i}/N$ $(i=a,b)$. If ρ is the constant probability that a member of m_{2a} is reported, then the conditional probability function of r_{2b} given m_{2b} is

$$f(r_{2a}|m_{2b})=\binom{m_{2b}}{r_{2b}} \rho^{r_{2b}}(1-\rho)^{(m_{2b}-r_{2b})} \qquad (3.59)$$

We know that if x_1 and x_2 are independent Poisson random variables with means θ_1 and θ_2 respectively, then it is readily shown that the distribution of x_1 conditional on $y = x_1 + x_2$ is binomial, namely

$$f(x_1|y) = \binom{y}{x_1} p^{x_1} q^{x_2}$$

Where $p=\theta_1/(\theta_1+\theta_2)$. Conversely, if x_1 and y are a pair of variables such that the conditional distribution of x_1 given y is binomial with parameters y and p , and y is Poisson with mean λ , then the conditional distribution of x_1 is Poisson with mean λP (Feller, 1957). therefore , it follows that , the unconditional distribution function of r_{2b} is Poisson with parameter $n_1 n_2 \rho/N$. As r_{2a} and r_{2b} are independent Poisson variables, it follows that

$$f(r_{2a}|r)=\binom{r}{r_{2a}} p^{r_{2a}} q^{r_{2b}} \qquad (3.60)$$

Where, $r = r_{2a}+r_{2b}$ and $p = n_{2a}/(n_{2a}+n_{2b})$. Therefore the estimate of ρ is given by

$$r_{2a}/r = n_{2a}/(n_{2a}+\hat{\rho}n_{2b}) \quad ,$$

or

$$\rho = \frac{n_{2a}r_{2b}}{n_{2b}r_{2a}}$$

and a test of $H_0:\rho=1$ against the one sided alternative $H_1:\rho<1$ is equivalent to testing $H_0:p=n_{2a}/n_2$ $=(p_0,\text{say})$ against $H_1:p > p_0$ for the binomial model (3.60).

Note that an estimate of m_2 ,the actual number of tagged individuals recaptured is given by

$$\hat{m}_2 = m_{2a}+(r_{2b}/\rho) = m_{2a}n_2/n_{2a} \, .$$

and

$$\hat{N} = n_1 n_2/\hat{m}_{2a} = n_1 n_{2a}/m_{2a} \, .$$

This means that we base the Petersen estimate on the recapture data for which we have 100 per cent reporting rate.

If the tag ratio n_1/N is not small, the m_{2i} may be more approximately represented by a binomial law, namely

$$f_1(m_{2i}) = \binom{n_{2i}}{m_{2i}} \left(\frac{n_1}{N}\right)^{m_{2i}} \left(1-\frac{n_1}{N}\right)^{(n_{2i}-m_{2a})}$$

Assuming model (3.59), the above equation leads,to

$$f_2(r_{2i}) = \binom{n_{2i}}{m_{2i}} \left(\frac{\rho_i n_1}{N}\right)^{r_{2i}} \left(1-\frac{\rho_i \; n_1}{N}\right)^{(n_{2i}-r_{2i})}$$

(i=a,b). (3.6)

Where $\rho_a = 1$, $\rho_b = \rho$. Therefore testing H_0 is now equivalent testing for homogeneity in the 2×2 contingency table below.

r_{2a}	r_{2b}	r
$n_{2a} - r_{2a}$	$n_{2b} - r_{2b}$	$n_2 - r$
n_{2a}	n_{2b}	n_2

As the alternative is one sided, the usual test is modified slightly by only rejecting H_o when $r_{2a}/n_{2a} > r_{2b}/n_{2b}$, and the Chi-Squared statistic is significant at 2α level of significance, where α is the size of the test. When the r_{2i} and n_{2i} are small, Fisher's exact test should be used.

It is important not only to detect incomplete reporting after the experiment, but also to decide before the experiment how many tags should be released and how much of the sample should be inspected to be reasonably sure of detecting non-reporting of a given magnitude. Such an information can be obtained by examining the power of the test H_o for particular alternative $\rho < 1$. When the smaller of rp or rq is greater than 5 , and the correction for continuity is used, the normal approximation to binomial distribution can be used. In this case the most powerful one sided test of H_o is to reject H_o when $r_{2a} > d$, where

$$d = z_\alpha \sqrt{(rp_o + q_o)} + \frac{1}{2} + rp_o$$

Here z_α satisfies $\Pr[Z \geq z_\alpha] = \alpha$,where z has a unit normal distribution. If β is the type II error for a given alternative ρ ,that is

$$\Pr[r_{2a} \geq d \;|\rho] = 1-\beta$$

then the number of recoveries needed to test H_o at pre-set values of α and β for different values of ρ

can be expressed as

$$r = (z_{\beta}\sqrt{pq} + z_{\alpha}\sqrt{p_o q_o})^2/(p-p_o)^2$$

Where $p = n_{2a}/(n_{2a}+\rho n_{2b}) = p_o/(p_o+\rho(1-p_o))$.

Paulik has tabulated r as a function of β, ρ, and p_o for α = 0.10, 0.05, 0.01, and the tables are reproduced in table A.2 of the appendix. He points out that if the binomial model (3.60)is more approximate than the Poisson model which led to (3.61), then the value of r obtained from the tables is conservative in that the true true power of the test exceeds $1-\beta$.

The method of Paulik is a very flexible one and can be applied to a number of different experimental situations. In particular it can be used for the case when p_o refers to the proportion of the total effort inspected rather than the proportion of the total catch. We shall not give the theoretical justification of the above extension here.

Note:

The incomplete reporting can arise either through (i)tags being accidentally overlooked, or (ii)tags being deliberately with held. For larger animals a tag can usually be designed so that the chance of being overlooked is negligible. In fisheries, however, where large numbers are rapidly handled and tags are small this source of error can usually be minimised by extensive advertising and the offering of an *adequate* reward for a returned tag.

In conclusion we see that the above methods of Paulik can still apply even if natural mortality is operating, provided that the tagged and untagged have the same mortality rate, so that the tag ratio remains constant throughout.

3.3 ESTIMATION FROM SEVERAL SAMPLES:

Let n_1 animals be captured, marked and released throughout the population and let $p=n_1/N$, where N is the total population size. Suppose that the total population area can be divided up into R sub areas of which k are selected at random for further sampling. Let

N_i=number of animals in the ith sub area $(i=1,2,...R)$

$\overline{N}=N/R$

n_{1i}=number of marked animals in the ith sub area

$\{n_{1j}\}$=set of $n_{1j}(j=1,2,..,R)$

$\{n_{1ik}\}$ =sample set of n_{1i} $(i=1,2,...R)$

n_{2i}=number of animals caught in the ith sub are $(i=1,..,R)$

n_{2i}=number of marked in n_{2i}

$p_i = n_{1i}/N_i$

$\hat{p} = m_{2i}/n_{2i}$

$f_1=k/R$

and

$f_{2i}=n_{2i}/N_i$

Note that

$$p = n_1/N = \sum_{i=1}^{R} N_i p_i \Big/ \sum_{i=1}^{R} N_i$$

Ratio Estimate.

With the above experimental set up we effectively have two-stage sampling, in which we choose k unequal size units or clusters at the first stage and then sub sample from each sub unit, noting the proportion p_i of the marked in each sub sample. Therefore, if the cluster sizes N_i were known, p could be estimated by the ratio type estimator

$$\sum N_i \hat{p}_i / \sum N_i \ .$$

However, as the N_i are known, one possibility is to use weights proportional to the sample sizes n_{2i}, so that p is estimated by

$$p = \sum_{i=1}^{k} n_{2i}\hat{p}_i / \sum_{i=1}^{k} n_{2i} = \sum m_{2i} / \sum n_{2i}$$

$$= m_2/n_2 \quad \text{,say.}$$

and N is estimated by $n_1/\hat{p}$, the Petersen estimate once again.It is noted that

$$E\left[\hat{p} \,|\, \{n_{2i}\}, \{n_{1i}\}_k\right] = \sum n_{2i}p_i \;/\; n_2 \qquad (3.62)$$

We now consider two cases where $\hat{p}$ is unbiased and approximately unbiased respectively.

Constant Mark Ratio:

When $p_i = p$ $(i=1,2,...,k)$ the right-hand side of (3.62) is reduced to p and $\hat{p}$ is unbiased. In this case $\mathrm{Var}[\hat{p}|\{n_{1i}\}]$, the variance of $\hat{p}$ can be estimated by

$$V_1[\hat{p}] = \hat{p}(1-\hat{p})/(n_2-1)$$

The hypothesis $H_o{:}p_i=p$ can be tested using the standard goodness-o-fit statistic

$$Q = \sum_{i=1}^{k} (m_{2i}-n_{2i}\hat{p}\ )^2/n_{2i}\hat{p}(1-\hat{p}).$$

Which when H_o is true is approximately distributed as Chi-Squared with k-1 degrees of freedom. When H_o is true and n_2 is large (say, greater than 30). $\hat{p}$ is approximately normal with mean p and variance estimate

$V_1[\hat{p}]$. An approximate confidence interval for p can be calculated in the usual manner.

Constant Sampling effort:

In general, if H_o is not true $\hat{p}$ will be biased. If however, the same sampling effort is used within each sub area and the expected fraction caught (θ,say) is proportional to the sampling effort, then

$$E[n_{2i}|n_{1i}] = E[n_{2i}] = N_i\theta$$

and

$$E[m_{2i}|n_{1i}] = n_{1i}\theta$$

Hence, for large n_{2i},

$$E[\hat{p}|\{n_{1i}\}] \cong \sum_{i=1}^{k} E[m_{2i}|n_{1i}] \Big/ \sum_{i=1}^{k} E[n_{2i}/n_{1i}]$$

$$\cong \sum_{i=1}^{k} n_{1i} \Big/ \sum_{i=1}^{k} N_i \qquad (3.63)$$

Since k sub areas are chosen at random, equation (3.63) represents a ratio estimate of p with respect to the first stage in the sampling. Therefore, taking expectations with respect to this first stage, we have, for large k

$$E[\hat{p}|\{n_{1i}\}] \cong \underset{k}{E}\left[E[\hat{p}|\{n_{1i}\}]\right] \cong \sum_{i=1}^{k} n_{1i} \Big/ \sum_{i=1}^{k} N$$

$$= n_1/N = p.$$

Thus $\hat{p}$ is approximately unbiased for large n_{2i} and large k. Noting that

$$\theta = E[n_{2i}]/N_i = E[n_2]/\sum N_i \cong E[n_2]/k\bar{N} = E[\bar{n}_2]/\bar{N} ,$$

Where $\bar{n}_2 = n_2/k$, $N_i/\bar{N}$ can be estimated by $n_{2i}/\bar{n}_2$. Therefore,from Cochran (1977) an approximate estimate of $Var[\hat{p}|\{n_{1i}\}]$ is given by

$\mathrm{Var}[\hat{p} \mid \{n_{1i}\}]$ is given by

$$V_2[\hat{p}] = \frac{1-f_1}{k(k-1)} \sum \left(\frac{n_{2i}}{n_2}\right)^2 (\hat{p}_i - \hat{p}) + \frac{f_1(1-f_2)}{kn_2} \sum \left(\frac{n_{2i}}{n_2}\right) \frac{n_{2i}}{n_{2i}-1} \hat{p}_i (1-\hat{p}_i) \tag{3.64}$$

Where f_2 $(\cong \bar{n}_2/\bar{N})$ can either be ignored, or estimated using $n_1/\hat{p}$ as an estimate of N.

When $k = R$, that is $f_1 = 1$, we have the special case of stratified random sampling. The expression (3.63) then reduces to p and the first term of (3.64) is zero.

When f_1 is small the second term in (3.64) can be neglected, and $\mathrm{Var}[\hat{p} \mid \{n_{1i}\}]$ is now estimated by

$$V_3 = k \sum (m_{2i} - n_{2i} p)/(k-1) n_2^2$$

Using the delta method(explained earlier), the variance of the Petersen estimate $\hat{N} = n_1/\hat{p}$ is approximately given by

$$n_1^2 p^{-4} \mathrm{Var}[\hat{p} \mid \{n_{1i}\}] \tag{3.65}$$

which can then be approximated by

$$\mathrm{Var}[\hat{N}] = n_1^2 p^{-4} V_3[\hat{p} = \frac{\hat{N}^2 k}{m_2^2 (k-1)} \sum_{i=1}^{k} (m_{2i} - n2i\hat{p})^2 \tag{3.66}$$

Mean Estimates:

An alternative estimate of p is the average

$$\bar{\hat{p}} = \sum_{i=1}^{k} \hat{p}_i / k$$

and from the general theory of Cochran (1977), setting

$$y_i = \hat{p}_i / k \quad , \quad y_i' = p_i / k \quad , \quad \pi_i = k/R$$

an unbiased estimate of the variance is given by

$$\mathrm{Var}[\bar{\hat{p}}] = \frac{(1-f_1)}{k(k-1)} \sum (\hat{p}_i - \bar{\hat{p}})^2 + \frac{f_1}{k^2} \sum (1-f_{2i}) \frac{\hat{p}_i \hat{q}_i}{(n_{2i}-1)} ,$$

which reduces to the usual sample estimate of the variance when f_1 is ignored. As

$$E\,[\hat{\bar{p}}\,|\{n_{1i}\}] = \underset{k}{E}\,E[\hat{\bar{p}}\,|\{n_{2i}\}] = \underset{k}{E}\left(\sum p_i\;/k\right) = \sum_{i=1}^{R} p_i/R$$

$$= \bar{p}\text{ , say}$$

Then $\hat{\bar{p}}$ is unbiased when either $H_0\!:p=p_i$ for all i, is true or when $\bar{p}=p$. If H_0 is true $\hat{\bar{p}}$ seems to be preferable to $\hat{p}$ because of the general robustness of a mean with regard to normality and because $\mathrm{Var}[\hat{\bar{p}}]$ is more robust than $V_1[\hat{p}]$ with regard to the departures from H_0. On the other hand, if H_0 is rejected by the goodness-of-fit test and the sampling effort is uniform, then $\hat{p}$ can be used with $V_2[\hat{p}]$ of equation (3.64).

In some experiments the numbers n_{1i} of marked animals in the individual sub ares are known and N can be estimated by

$$N' = R\sum_{i=1}^{k} N_i/k = R\sum\left\{\left[(n_{1i}+1)(n_{2i}+1)/(m_{2i}+1)\right] - 1\right\}/k$$

Then

$$E[N'] = \underset{k}{E}\,E[N^*|k] \cong R\,\underset{k}{E}\sum N_i/k = R\bar{N} = N$$

and using the general theory of Cochran (1977), setting

$$y_i^* = N_i'/k\ ,\ y_i^* = N_i/k\ ,\ \pi_i = k/R$$

it can be shown that an approximate unbiased estimate of the variance of N' is

$$\mathrm{Var}[N'] = \frac{R(R-k)}{k(k-1)}\sum_{i=1}^{k}(N_i^*-N')^2 + \frac{R}{k}\sum_{i=1}^{k}\mathrm{Var}[N_i^*]\ ,$$

Where

$$\mathrm{Var}[N_i^*] = \frac{(n_{1i}+1)(n_{2i}+1)(n_{1i}-m_{2i})(n_{2i}-m_{2i})}{(m_{2i}+1)^2(m_{2i}+2)}$$

When $n_{1i}+n_{2i} \geq N_i$ for each i, then N' and $\mathrm{Var}[N']$

are exactly unbiased.

3.4 INVERSE SAMPLING METHODS.

We now consider an inverse sampling method for the second sample which in contrast to *direct* Petersen method considered so far, provides an unbiased estimate of N with an exact (rather than large sample) expression for the variance, and a co-efficient of variation which is almost independent of N.

The method is to tag or mark n_1 animals as before and then continue taking the second sample until a prescribed number m_2 of marked animals have been recovered. This means that n_1 and m_2 are now considered fixed parameters and n_2 is a random variable. As in the direct Petersen method, the second sample can be taken with or without replacement and we shall consider these two cases separately.

<u>Sampling Without Replacement.</u>

When the assumptions I,IV,V, and VI of section 3.1 are satisfied, we shall show that the probability of n_2 conditional on n_2 , m_2 , is negative hypergeometric.

The probability of obtaining n_1 is the probability of drawing first a sample of size (n_1-1) containing (m_2-1) marked animals, followed by the drawing of just one further marked animal. Having regard to the existence of sampling without replacement the likelihood function is clearly

$$f(n_2|n_1,m_2)=\frac{n_1-m_2+1}{N-n_2+1}\begin{pmatrix}n_1\\m_2-1\end{pmatrix}\begin{pmatrix}N-n_1\\n_2-m_2\end{pmatrix}\Big/\begin{pmatrix}N\\n_2-1\end{pmatrix} \qquad (3.67)$$

Which is the negative hypergeometric distribution. Where $n_2 = m_2, m_2+1, \ldots, N+m_2-n_1$. Now the distribution (3.67) can be written as

$$f(n_2|n_1,m_2) = \frac{n_1}{N}\binom{n_1-1}{m_2-1}\binom{N-n_1}{n_2-m_2}\Big/\binom{N-1}{n_2-1} \qquad (3.68)$$

Where $0 \leq m_2 \leq n_1$, $m_2 \leq n_2+m_2-n_1$. It follows from (3.68) that

$$\sum_{n_2=m_2}^{N+m_2-n_1}\binom{N-n_1}{n_2-m_2}\Big/\binom{N-1}{n_2-1} = \frac{N}{n_1}\sum_{n_2=m_2}^{N+m_2-n_1} f(n_2|n_1,m_2)\Big/\binom{N-1}{m_2-1}$$

$$= N/n_1\binom{n_1-1}{m_2-1} \qquad (3.69)$$

The expectations of n_2 and $n_2(n_2+1)$, which we shall have an occasion to use below, are readily derived as follows:

$$E[n_2] = \sum_{n_2=m_2}^{N+m_2-n_1}\frac{n_2 n_1}{N}\binom{n_1-1}{m_2-1}\binom{N-n_1}{n_2-m_2}\Big/\binom{N-1}{n_2-1}$$

$$= n_1\binom{n_1-1}{m_2-1}\sum_{n_2=m_2}^{N+m_2-n_1}\binom{N-n_1}{n_2-m_2}\Big/\binom{N}{n_2} \qquad (3.70)$$

The summation on the right-hand side of (3.70) is obtained immediately from expression (3.69) on writing $N+1$, n_1+1, n_2+1, and m_2+1 for N, n_1, n_2, and m_2 respectively. therefore

$$E[n_2] = n_1(N+1)\binom{n_2-1}{m_2-1}\Big/(n_1+1)\binom{n_1}{m_2} = m_2(N+1)/(n_1+1) \qquad (3.71)$$

similarly,

$$E[n_2(n_2+1)] = \sum_{n_2=m_2}^{N+m_2-n_1}\frac{n_2(n_2+1)}{N}\binom{n_1-1}{m_2-1}\binom{N-n_1}{n_2-m_2}\Big/\binom{N-1}{n_2-1}$$

$$= (N+1)n_1\binom{n_1-1}{m_2-1}\sum\binom{N-n_1}{n_2-m_2}\Big/\binom{N+1}{n_2+1} \qquad (3.72)$$

To evaluate the summation on the right-hand side of(3.72) we write $N+2$, n_1+2 ,n_2+2 and m_2+2 for N ,n_1 , n_2 , and m_2 in equation (3.69). Therefore

$$E[n_2(n_2+1)] = n_1(N+1)(N+2)\binom{n_1-1}{m_2-1}\Big/(n_1+2)\binom{n_1+1}{m_2+1}$$

$$= \frac{m_2(m_2+1)(N+1)(N+2)}{(n_1+1)(n_2+2)} \qquad (3.73)$$

Now, strictly speaking, N can take only integral values. Thus the greatest value attained by the likelihood is derived by considering the ratio

$$\frac{f(n_2|N+1)}{f(n_2|N)} = \frac{(N-n_1+1)(N-n_2+1)}{(N+1)(N-n_1-n_2+m_2+1)} \qquad (3.74)$$

Using (3.68), it follows from (3.74) that

$$f(n_2/N+1)\begin{cases} > f(n_2|N) & \text{if } N < \dfrac{n_1n_2}{m_2} - 1 \\ < f(n_2|N) & \text{if } N > \dfrac{n_1n_2}{m_2} - 1 \end{cases} \qquad (3.75)$$

Unless n_1n_2/m_2 is an integer, the likelihood attains its greatest value when N is equal to the integral part of n_1n_2/m_2 . If n_1n_2/m_2 is integral, the greatest value is attained at both $(n_1n_2/m_2 - 1)$ and n_1n_2/m_2 . In practice N is usually sufficiently large for us to be able to ignore its discreteness. In any case we are led to consider the estimates of the general type n_1n_2/m_2 . It follows from (3.71) that

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

is an unbiased estimate of N for any values of N , n_1 , n_2 and m_2 . In large samples $\hat{N}_2$ approximates to the maximum likelihood solution. The exact sample variance of

$\hat{N}_2$ is also easily evaluated using (3.71) and (3.73) . We find

$$Var[\hat{N}_2] = (n_1-m_2+1)(N+1)(N-n_1)/m_2(n_1+2) \cong N^2/m_2 \ .$$

Assuming $N+1$ and $N-n_1$ to be approximately equal to N , the co-efficient of variation is close to

$$C(\hat{N}_2) = \left[(n_1-m_2+1)/m_2(n_1+2)\right]^{1/2}$$

and since n_1 is known, m_2 can be chosen before hand so that this coefficient has a prescribed value.

Properties of $\hat{N}_2$:

Chapman (1952) mentions several useful properties of $\hat{N}_2$ relative to the direct sampling. The inverse sampling method provides an estimate of N with an expected sample size

$$E[n_2] = (N+1)m_2/(n_1+1)$$

so that

$$\hat{N}_2 = \frac{n_2(n_1+1)}{m_2} -1 \qquad (3.76)$$

is unbiased estimate of N . In contrast to the direct sample estimate, this unbiasedness does not depend on the parameters n_1 and n_2 . Further more

$$Var[\hat{N}_2]=(N+1)(N-n_1)(n_1-m_2+1)/m_2(n_1+2) \cong N^2/m_2 \qquad (3.77)$$

This approximate formula (which exaggerates the actual variance) may be useful in the choice of m_2 By an appropriate choice of m_2 , $Var[\hat{N}_2]/N^2$ can be fixed at any desired level.

For testing purposes we note that $\hat{N}_2$ is approximately normal with mean N and variance N^2/m_2 for large m_2. A model similar to that of David (1938) can be used to prove that as m_2 tends to infinity n_2 is asymptotically normally distributed.

On the average, the inverse sampling procedure is better than direct sampling procedure. For if n_2 , m_2 denote the fixed sample and the random number of tags recovered in the direct procedure and m_2 is chosen equal to $n_2'n_1/N$, then

$$E[n_2] = \frac{1 + 1/N}{1 + 1/n_1} n_2' < n_2'$$

While

$$\mathrm{Var}[\hat{N}_2] < N^2/m_2 = N^2N/n_2'n_1 < \sigma_{\hat{N}}^2$$

Where

$$N = (n_2'+1)(n_1+1)/(m_2'+1)$$

is the almost unbiased estimate in the direct sampling case . Hence a more efficient estimate is obtained with less average effort.

On the other hand if the experimenter knows absolutely nothing about the possible population size, then by an improper choice of n_1 and m_2 , $E[n_2]$ may be extremely large. Moreover

$$\mathrm{Var}[\hat{N}_2] \cong m_2N^2/n_1^2$$

is very large. This may be undesirable feature of this procedure.

These difficulties may be partly overcome by a modification of the inverse sampling plan as follows: the number of untagged individuals to be recaptured is predetermined, rather than the number of tagged individuals. In other words, n_2-m_2 is chosen in advance of sampling:m_2 and n_2 are now both random variables, though, completely dependent. For convenience write u_2=n_2- m_2, the number of unmarked individuals caught in n_2. Then , the conditional distribution of n_2 given n_1 and m_2

$$f(n_2 / n_1,m_2)= \frac{(n_2-1)!n_1!(N-n_1)!(N-n_2)!}{(n_2- u_2)!(n_1-n_2+u_2)!(n_2-1)!(N-n_1-n_2)!N!} \qquad (3.78)$$

is derived in the usual manner. However the obvious estimate again is no longer strictly unbiased, for

$$E\left[\frac{n_2}{n_2-u_2+1}\right]= E\left[\frac{n_2}{m_2+1}\right]$$

$$= \sum_{n_2=u_2}^{n_2+u_2}\left[\frac{n_2!(n_1+1)!(N+1-n_1-1)!(N+1-n_2-1)!}{(n_2-u_2+1)!(n_1+1-n_2-1+n_2)!(N-1-n_1+1-n_2)!(N+1)!(u_2-1)}\right]$$

$$\times \frac{N+1}{n_1+1}$$

$$= \frac{N+1}{n_1+1}\left[1- \frac{(N+1-u_2)!(N-n_1)!(N-n_1)!}{(N+1)!(N-n_1-u_2)!}\right] \qquad (3.79)$$

It follows that that the second term in the parentheses is negligible provided $n_1u_2/ N > \log N$. For such values of u_2, n_1 , N the estimate

$$\hat{N}_3=\frac{n_2(n_1+1)}{m_2+1} - 1$$

has bias less than 1 in absolute value .

To determine the variances of $\hat{N}_3$, we note that

$$\hat{N}_3-1)^2=(n_1+1)^2\left[n_2(N_2+1)-n_2\right]\left(\frac{1}{(m_2+1)(m_2+2)}\right.$$

$$\left.+\frac{1}{(m_2+1)(m_2+2)(m_2+3)}+\frac{2}{(m_2+1)(m_2+2)(m_2+3)(m_2+4)}+ R\right]$$

To evaluate the expectation of this later series let

$$n_{(i)}= n_2(n_2+1)(n_2-2)\ldots(n_2-i+1)$$

and observe that for any $i \le j$ $(j-i=\varepsilon)$

$$E\left[\frac{(n_2+i-1)_{(i)}}{(m_2+j)_{(j)}}\right] = \frac{(N+1)_{(i)}(N-n_1)_{(\varepsilon)}}{(n_1+j)_{(j)}(u_2-1)_{(\varepsilon)}}(1-\eta_j) \qquad (3.80)$$

Where

$$\eta_j = (N+1-u_j)!(N-n_1)!/(N+1)!(N-n_1-u_j)!$$

If the η_j is neglected and the remaining terms on the right hand side of (3.80) are denoted by q_{ij}, that is

$$q_{ij} = \frac{(N+i)_{(i)}(N-n_1)_{(\varepsilon)}}{(n_1+j)_{(j)}(u_2-1)_{(\varepsilon)}}$$

Then

$$\mathrm{Var}[\hat{N}_3]=E[\hat{N}_3]-N^2=(n_1+1)^2\Big[(q_{22}+q_{23}+2q_{24}+\ldots)-(q_{12}+q_{13}+2q_{14}\ldots)\Big]$$

$$-2N-2-N^2.$$

The average sample size required by this procedure can be determined from (3.80). In fact by direct summation

$$E[(n_2+i-1)_{(i)}] = \frac{(u_2+i-1)_{(i)}(N+i)_{(i)}}{(N-n_1+i)_{(i)}}$$

so that

$$E[n_2] = \frac{u_2(N+1)}{(N-n_1+1)}$$

and

$$\mathrm{Var}[n_2]= \frac{u_2 n_1(N+1)(N-u_2-n_1+1)}{(N-n_1+1)^2(N-n_1+2)^2} \cong \frac{n_1}{N-n_1} \qquad (3.81)$$

Either of these formulae can be obtained from (3.76) and (3.78)by an interchange of $N-n_1$ and n_2 and by replacing m_2 by u_2. It is now seen immediately that the tremendous variation in the earlier inverse sampling model is now eliminated.

Since u_2 will usually be reasonably large the most appropriate approximation to use for the testing purposes is the normal distribution. In particular, it is desirable to work with n_2 using formula

$$E[n_2] = \frac{u_2(N+1)}{(N-n_1+1)}$$

and (3.81). Writing $n_1/(N-n_1+1)=p'$ and using a trivial modification of the approximation of (3.81)

$$z = \frac{n_2 - u_2/(1-p')}{\sqrt{u_2 p'(1-p')}}$$

is approximately N(0,1) and confidence limits for p' (and hence for N) are obtained from the quadratic equation

$$\left[n_2(1-p')-u_2\right]^2 = k_\alpha^2 u_2^2 p'(1-p') \quad ,k_\alpha=z_{\alpha/2}$$

The approximation may be improved slightly if the exact formula for $Var[n_2]$ is used but this involves solving a higher degree equation.

Sampling With Replacement

We now consider the less common situation in which members of the second sample are caught one at a time and returned immediately to the population. For example this model apply when the animals are merely observed and not actually captured. In the inverse method, sampling is continued until a prescribed number m_2 of marked animals have been caught and released, so that the probability function of n_2 is now negative binomial;

$$f(n_2|n_1,m_2) = \binom{n_2-1}{m_2-1} p^{m_2}(1-p)^{(n_2-m_2)}$$

Where $n_2=m_2,m_2+1,\dots$, $p = n_1/N$. As

$$E[n_2|n_1,m_2] = Nm_2/n_1$$

the obvious estimate for N is the Petersen estimate

$$\hat{N} = n_1n_2/m_2$$

This maximum likelihood estimate is unbiased with variance

$$\mathrm{Var}[\hat{N}|n_1,m_2] = (N^2-Nn_1)/m_2$$

and an unbiased estimate of the variance is

$$\mathrm{Var}[\hat{N}] = \frac{n_2n_1^2(n_2-m_2)}{m_2^2(m_2+1)} \; .$$

Confidence Intervals.

Consider the statistic

$$z = \left\{n_1n_2/m_2 - N\right\} / \sqrt{(N^2/m_2 - Nn_1/m_2)}$$

The moment generating function of z is

$$M_z(\theta)=p^{m_2}\, e^{\left(pm_2/\sqrt{m_2(1-p)} - \sqrt{m_2\theta/(1-p)} \right)}$$

$$\times \left\{1-(1-p)e^{p/\sqrt{m_2(1-p)}}\;)\right\}^{-m_2}$$

By routine algebra and the usual manipulations it may be seen that $\log M_z(\theta)$ tends to $\theta^2/2$ as m_2 tends to infinity. Thus z is approximately distributed as $N(0,1)$ for large m_2 . This fact may be used to set up confidence intervals and tests for N .

For m_2 and n_1/N both small as will be more frequently the case,the second limiting distribution given above will be more useful. For this range of parameters $2n_1n_2/N$ has approximately χ^2-distribution with $2m_2$ degrees of freedom. This is equivalent to the Poisson approximation to the binomial and is related to the results of Sandelius (1950) concerning inverse sampling with random variables distributed according to a Poisson distribution.

If $\chi^2_{2m_2}(\epsilon)$ denotes the ϵ-th quartile of the χ^2 distribution with $2m_2$ degrees of freedom that is:

$$\Pr\left[\chi^2_{2m_2} \le \chi^2_{2m_2}(\epsilon)\right] = \epsilon$$

Then $(1-\epsilon)$ confidence intervals for N are given by

$$\underline{N} = 2n_1n_2/\chi^2_{2m_2}(1-\epsilon/2) \quad ; \quad \overline{N} = 2n_1n_2/\chi^2_{2m_2}(\epsilon/2)$$

3.5 COMPARING TWO POPULATIONS.

Suppose we have two populations of unknown sizes N_a and N_b, and on the basis of a single Petersen experiment in each population we wish to test the null hypothesis

$$H_o: N_a = N_b$$

such a situation could arise, for example, where a control area and an experimental area are under observation and one wishes to test any difference in population size due to experimental management practice. Alternatively, N_a and N_b could refer to the same population area but at different times. A third possibility is when N_a and N_b refer to the same population at the same time but two different sampling methods are used, for example seine and gill-net fishing. A test of H_o would then be indirectly provide a test for the hypothesis that marked animals are equally vulnerable to both methods of sampling. In notation of section 3.1 , we can test H_o by assuming that

$$z = \frac{N^*_a - N^*_b}{\sqrt{V^*_a + V^*_b}}$$

is approximately unit normal when H_o is not true. A more sensitive test is given by Chapman (1951) as follows: using suffixes a and b to denote the two populations respectively, let

$$\tilde{N} = \frac{\lambda_a^2 m_{2b} u_{1b} u_{2b} - \lambda_b^3 m_{2a} u_{1a} u_{2a}}{m_{2a} m_{2b}(\lambda_a^2 u_{1b} u_{2b} + \lambda_b^2 u_{1a} u_{2a})}$$

where $\lambda_a = n_{1a} n_{2a}$, $u_{1a} = n_{1a} - m_{2a}$, $u_{2a} = n_{2a} - m_{2a}$,and so on.

Define

$$T_1 = \sum \frac{\tilde{N}}{\lambda_o} \left[m_{2c} - \lambda_c / \tilde{N} \right] / \left[1 - \frac{n_{1c}}{\tilde{N}} \right] \left[1 - \frac{n_{2c}}{\tilde{N}} \right]$$

$$T_2 = \sum \lambda_c \left[m_{2c} - \lambda_c / \tilde{N} \right]^2 / m_{2c} n_{1c} n_{2c}$$

$$T_3 = \sum \frac{2(m_{2c} - \lambda_c / \tilde{N})^2}{\frac{\lambda_c}{\tilde{N}} \left[1 - n_{1c} / \tilde{N} \right] \left[1 - n_{2c} / \tilde{N} \right] + m_{2c} n_{1c} n_{2c} / \lambda_c}$$

Where $\sum$ denotes summation over two values $c = a, b$. Then when H_o is true, T_1, T_2, T_3 are each approximately distributed as Chi-Squared with one degree of freedom when N is large, and the three statistics are candidates for testing H_o . Chapman (1952) suggests using T_2 when $\lambda_a = \lambda_b$, T_3 when these quantities differ moderately, and T_1 when λ_a is widely different from λ_b .

To test H_o against the two sided alternative $N_a \neq N_b$, the criterion is to reject H_o at 100α per cent level of significance when T_i is greater than α critical value of χ_1^2 .

Using Poisson Approximation.

When the experimental circumstances are such that the hypergeometric distributions of m_{2a} and m_{2b} can be approximated by Poisson distributions, the following technique of Chapman and Overton (1966) can be used for testing H_o . Let

$$m_2 = m_{2a}+m_{2b} \quad \text{and} \quad N_b = kN_a \ .$$

then the conditional probability function of m_{2a} given m_2 is

$$f(m_{2a}/m_2) = \begin{bmatrix} m_2 \\ m_{2a} \end{bmatrix} p^{m_{2a}} q^{m_{2b}} \ ,$$

Where

$$p = 1-q = k\lambda_a/(k\lambda_a+\lambda_b)$$

Setting $p_o = \lambda_a/(\lambda_a+\lambda_b)$,we note that p is greater than or less than p_o if and only if k is greater or less than unity, so that testing H_o against the two sided alternative $N_a \neq N_b$ is equivalent to testing $p=p_o$ against $p \neq p_o$. Therefore, given m_2 , m_{2a} and p_o , we can test H_o by evaluating the exact tail probabilities, using such tables as those of Harvard computational laboratory (1955), or we can obtain a confidence interval for p and reject H_o if p_o lies outside this interval. Confidence interval for p can be determined from Clopper and Pearson charts of Pearson and Heartly (1966) , using $\hat{p} = m_{2a}/m_2$ as the entering variable, or from tables such as Owen (1962).

Special cases.

(i) When $\hat{p} < 0.1$ and for any m_2 (Raff,1956), the Poisson approximation to the binomial can be used to find a confidence interval for m_2p , using m_{2a} as the entering variable in Poisson tables (for example Pearson and Heartly(1966)).

(ii) When $0.1 \leq \hat{p} \leq 0.9$, $m_2\hat{p}$ and $m_2\hat{q}$ are both greater than 5 , and a correction for continuity is used, the normal approximation is applicable (Raff,1956) and we can use the following statistic for testing H_o

$$z = \frac{|m_{2a}-m_2p_o|-1/2}{\sqrt{m_2p_oq_o}}$$

(iii) When $\lambda_a=\lambda_b$ (> 2.5) ,another procedure is given by Sichel (1973).

We note that the above methods can be used to obtain a confidence interval for k . In particular when N_a and N_b refer to the same population at different times, and the population is closed except for mortality, k can be interpreted as the proportion of N_a surviving, that is as a survival probability. If we are interested in just one sided alternative, say $N_a > N_b$, then we could reject H_o if m_{2a} lay in the upper tail of the binomial distribution with $p=p_o$; the approximate probability could be evaluated using binomial tables or normal approximation.

3.6 ESTIMATION BY LEAST SQUARES.

In commercially exploited populations the second sample in the Petersen method may consist of a sequence of samples each being permanently removed from the population. For this situation N can be estimated by the following least-squares method due to Paloheimo (1963). Let

N_o= initial size of the total population

M_o= initial size of the marked population.

$U_o=N_o-M_o$

n_i=size of the ith sample removed from the population (i=1,2,..,s)

m_i=number of marked individuals in the ith sample,

$u_i=n_i-m_i$

$y_i=m_i/n_i$

$M_i = M_o - \sum_{j=1}^{i-1} m_j$ and,

$N_i = N_o - \sum_{j=1}^{i-1} n_j$.

Then if the assumptions underling the Petersen method of section 3.1 hold for each sample,

$$E\left[y_i | M_i, N_i\right] = M_i/N_i \qquad i=1,2,\ldots,s$$

Paloheimo suggests estimating N_o by minimizing $\sum W_i(y_i - M_i/N_i)^2$ with respect to N_o , where the W_i's are appropriate weights, customarily taken to be proportional to the inverse of the variance of y_i. When the sampling is random, or the marked and unmarked are randomly mixed, these variances may be calculated by assuming Poisson or binomial sampling. For example assuming Poisson sampling, the variance of y_i equals its expected value M_i/N_i and the weights would have to be estimated iteratively as they contain unknown N_o . Not only are such weights awkward to compute, but very often , in practice, the y_i vary more than expected on the assumption of random fluctuations. Under these circumstances DeLury (1958) argues that one should preferably choose weights equal to the sample sizes. Also if the marked and unmarked are removed at the same rate, we have approximately

$$M_i/N_i = M_o/N_o$$

so that

$$E[y_i/M_i, N_i] = M_o/N_o \ (=\beta_o, \text{say}).$$

Therefore assuming the y_i's to be approximately independently and normally distributed with variances σ^2/n_i , and setting $W_i = n_i$, we can use the general theory of chapter 1 (Regression models) to obtain an estimate and a 100(1-α) per cent confidence interval for β_o, namely

$$\tilde{\beta}_o = \sum W_i y_i / \sum W_i = \sum m_i \;/\; \sum n_i$$

and

$$\tilde{\beta}_o - t_{s-1}[\alpha/2]\left(\tilde{\sigma}_o^2 / \sum n_i\right)^{1/2}$$

Where

$$(s-1)\tilde{\sigma}_o^2 = \sum m_i^2/n_i - \left(\sum m_i\right)^2 / \sum n_i$$

The least squares estimate of N_o is then

$$\tilde{N}_o = M_o/\tilde{\beta}_o = M_o\, \Sigma\, n_i/\Sigma\, m_i$$

This is simply the usual Petersen estimate based on pooling the data from all catches. The above interval for β_o gives the following 100(1-α) per cent confidence interval for N_o, namely

$$\left[\frac{M_o \sum n_i}{\sum m_i \pm t_{s-1}[\alpha/2](\tilde{\sigma}_o^2 \sum n_i)^{1/2} \sum n_i}\right]$$

In the same way we can obtain an estimate and confidence interval for $\sum n_i/N_o$,the rate of exploitation.

As the first step in examining the underlying assumptions of the above least squares method we can plot y_i against i as a visual check on the constancy of M_i/N_i. If necessary, a test of $\beta=0$ for model $E[y_i]=\beta_o+\beta i$ could be carried out using the theory of chapter 1 (Regression models). Also by drawing the line $y_i=\tilde{\beta}_o$, an examination of the deviation of each y_i from this line would provide a rough check on the reliability of the weights $W_i=n_i$. When there is no mortality taking place the above method can still be used, provided that the mortality rates for marked and unmarked are the same, so that M_i/N_i remains approximately constant. However, if recruitment and immigration into the population are appreciable then more complex methods which we shall not discuss here are required.

CHAPTER 4

4.1 SCHNABEL CENSUS.

Notation.

A simple extension of the Petersen method to a series of s samples of sizes $n_1, n_2, \ldots\ldots, n_s$ is the so called Schnabel census (Schnabel,1938). In this method each sample captured (except the first) is examined for marked members and then every member of the sample is given another mark before the sample is returned to the population. If different marks or tags are used for different samples, then capture-recapture history of any animal caught is known. For the closed population, a variety of theoretical models have been suggested, but before we discuss these, we shall need the following notation:

N =total population size,

s =number of samples

n_i =size of the ith sample $(i=1,2,\ldots\ldots,s)$

m_i = number of marked individuals in n

$u_i = n_i - m_i$

$M_i = \sum m_j \quad , i\ 1,2,\ldots\ldots s+1$

= number of marked individuals in the population just before the ith sample is taken

Since there are no marked animals in the first sample, we have $m_1\ 0$, $M_1=0$, $M_2=u_1=u_1$ and define M_{s+1}(=r,say) as, the total number of marked animals in the population at the end of the experiment that is, the total number of different animals caught throughout the experiment.

FIXED SAMPLE SIZES

The Generalised Hypergeometric Model.

Let a_w be the number of animals with a particular capture history w, where w is a nonempty subset of integers $\{1,2,\ldots s\}$: thus a_{124} represents those animals caught in the first, second and fourth samples only, also $r = \sum_w a_w$. If P_w, the probability that an animal chosen at random from the population has history w, is the same for each animal, and animals act independently, the animals may be regarded as N independent **trials** from a multinomial experiment. Therefore the joint probability function of the random variables $\{a_w\}$ is

$$f(\{a_w\}) = \frac{N!}{\prod_w a_w!(N-r)!} Q^{N-r} \prod_w P_w^{a_w} , \qquad (4.1)$$

where $Q = 1 - \sum_w P_w$. We shall assume that:

(i) all individuals have the same probability $p_i = 1 - q_i$ of being caught in the ith sample, and

(ii) for any individual the events "caught in the ith sample $(i = 1,2,3, \ldots,s)$" are independent. Then

$Q = \prod_{i=1}^{s} q_i$, $P_{124} = p_1 p_2 q_3 p_4 q_5 \ldots\ldots q_s = p_1 p_2 p_4 Q/q_1 q_2 q_4$ and so on

We notice that

$$Q^{N-r} \prod_w P_w^{aw} = Q^N \prod_w \left(\frac{P_w}{Q}\right)^{a_w} = Q^N \prod_i \left(\frac{p_i}{q_i}\right)^{\sum_{w \geqq i} a_w}$$

$$= \left(\prod_i q_i\right)^N \prod_i \left[\frac{p_i}{q}\right]^{n_i} = \prod_i p_i^{n_i} q_i^{N-n_i}$$

Thus equation (4.1) reduces to

$$f\left[\{a_v\}\right] = \frac{N!}{\prod_v a_v!(N-r)!} \prod_{i=1}^{s} p_i^{n_i} q_i^{N-n_i} \qquad (4.2)$$

It is obvious from our assumptions and it is easily deduced from (4.1) that n_i´s are independent binomial variables $B[n,p_i]$, so that

$$f(\{n_i\}) = \prod_{i=1}^{s} \binom{N}{n_i} p_i^{n_i} q_i^{N-n_i}$$

and the joint function of $\{a_v\}$ conditional on fixed sample sizes $\{n_i\}$ (that is the sample sizes are chosen in advance) is

$$f\left[\{a_v\}|\{n_i\}\right] = \frac{N!}{\prod_v a_v!(N-r)!} \prod_{i=1}^{s} \binom{N}{n_i}^{-1} \qquad (4.3)$$

Equation (4.3) represents a generalizes hypergeometric density.

__Estimation using the generalized Hypergeometric Model.__

Regarding equation (4.3) as representing likelihood L(N) of N and omitting the constant terms,

$$\log L(N) = \sum_i \log(N-n_i)! - (s-1)\log N! - \log(N-r)!$$

an equation for the maximum likelihood estimate $\hat{N}$ of N can be found by equating $\Delta\log(N)$ to Zero. This involves an error of less than unity in the solution and is equivalent to the **ratio method** of maximizing L , which equates L(N) to L(N-1) . Since $\Delta\log N! = \log N$, N must be one of the roots of the equation

$$\left(1 - \frac{r}{N}\right) = \prod_{i=1}^{s} \left(1 - \frac{n_i}{N}\right) \qquad (4.4)$$

This equation has a single finite root greater than r which maximizes the likelihood, except when r takes one of its extreme values:

(i) If $r = \sum_i n_i$, no individual is observed more than once and $\hat{N}$ is infinite

(ii) If $r = \max\{n_i\} = n_m$ say, no individual is observed which does not appear in mth sample and $\hat{N} = r = n_m$. It is of course in the nature of capture-recapture experiment that (i) and (ii) are extremely unlikely to occur.

When s = 2, equation (3.4) is of first degree and we find that $\hat{N} = n_1 n_2/m_2$,the Petersen estimate. For s = 3 , we have the quadratic equation.

$$N^2(m_2+m_3)-N(n_1n_2+n_1n_3+n_2n_3)+n_1n_2n_3=0$$

which can readily be solved for the larger root $\hat{N}$. When $s > 3$ we require some iterative method of solution. The most widely used is the Raphson iterative method. However, since we are only interested in finding $\hat{N}$ to the nearest integer, Robson and Regier's (1964) technique will be used as it is the easiest method for a desk calculator. Let

$$g(N) = \prod_{i=1}^{s} (1 - n_i/N)$$

and let h(N)= N-r-Ng(N), then the ith step of the iteration is given by

$$N_{(i+1)}= N_{(i)}-h(N_{(i)})/\nabla h(N_{(i)})$$

where

$$\nabla h(N_{(i)}) = h(N_{(i)}) - h(N_{(i)}-1)$$

$$= 1 - N_{(i)}g(N_{(i)}) - (N_{(i)}-1)g(N_{(i)}-1)$$

To begin the iterations we require first of all a trial solution $N_{(1)}$, where $N_{(1)} > r$. If N is large, then expanding $g(N)$ in powers of $1/N$, neglecting powers greater than the second and using

$$\sum_{i=1}^{s} n_i - r = \sum_{i=2}^{s} m_i \quad \text{since} \quad m_1 = 0$$

we find that equation (4.4) yields the approximate solution

$$N_B = \sum_{i=1}^{s} \sum_{j=i+1}^{s} n_i n_j \Big/ \sum_{i=2}^{s} m_i = R_2/m, \text{ say}$$

where R_2 can be expressed in the form

$$R_2 = \frac{1}{2}\left\{ \left(\sum_{i=1}^{s} n_i\right)^2 - \sum_{i=1}^{s} n_i^2 \right\}$$

However, if the cubic terms are retained, equation (4.4) reduces to the quadratic

$$N^2 m - NR_2 + R_3 = 0 \tag{4.5}$$

where

$$R_3 = \sum_{i=1}^{s} \sum_{j=i+1}^{s} \sum_{k=j+1}^{s} n_i n_j n_k$$

$$= \frac{1}{3}\left\{ \left(\sum_{i=1}^{s} n_i\right)\left(R_2 - \sum_{i=1}^{s} n_i^2\right) + \sum_{i=1}^{s} n_i^3 \right\}$$

and the desired solution of the quadratic is the larger root, N_A say. Chapman (1952) shows that under certain conditions, which are often satisfied when N is much larger than Σn_i, $N_A < \hat{N} < N_B$. Chapman also gives another pair of numbers

$$N_c = \text{maximum}\left\{ r, \min_{2 \le i \le s} (n_i M_i / m_i) \right\}$$

$$N_D = \underset{2\le i\le s}{\text{maximum}} \left\{ n_i M_i / m_i \right\}$$

and in general this pair will be satisfactory provided that no m_i (except m_1) is zero.

In solving equation (4.4) we see that the only recapture information required is r, the number of different animals caught during the experiment. This follows from the fact that r is a sufficient statistic for N, and means that as far as the estimation of N is concerned, distinguishing marks are not needed for each sample. In fact, at each stage, we need only mark the unmarked members of the sample. However, if the tags have sufficient information (for example numbered) then we can record all recapture histories. This information is useful in testing some of the underlying assumptions as shall be seen later.

Properties of $\hat{N}$

Using the generalised hypergeometric distribution derived above, Darroch (1958) proves that asymptotically (that is $N \longrightarrow \infty$, $n_i \longrightarrow \infty$ such that n_i/N remains constant)

$$E[\hat{N}] = N + b$$

where b, the bias, is estimated by

$$\frac{\left(\frac{s-1}{\hat{N}} - \sum 1/(\hat{N}-n_i)\right)^2 + \left(\frac{s-1}{\hat{N}^2} - \sum 1/(\hat{N}-n_i)^2\right)}{2\left[1/(\hat{N}-r) + (s-1)/\hat{N} - \sum 1/(\hat{N}-n_i)\right]^2}$$

and the asymptotic variance of $\hat{N}$ is estimated by

$$\mathrm{Var}[\hat{N}]=(\hat{N}-r)/h'(\hat{N}) = \left[1/(\hat{N}-r) + (s+1)/\hat{N} - \Sigma\, 1/(\hat{N}-n_i) \right]^{-1}$$

All summations are for $i=1,2,\ldots..,s$. Obviously the last step in the Newton-Raphson method which requires $h'(\hat{N})$, can be used for evaluating $\mathrm{Var}(\hat{N})$. Also $\nabla h(\hat{N}) \cong h'(\hat{N})$, so that the last step of Robson and Regiers' method will provide a reasonable approximation for $\mathrm{Var}(\hat{N})$. It can be shown that

$$\hat{b} = -(\hat{N}-r)h''(\hat{N})/2\{h'(\hat{N})\}^2$$

so that $\hat{b}$ can be approximated by

$$-(\hat{N}-r)\nabla^2 h(\hat{N})/2\{\nabla h(\hat{N})\}^2$$

Although this approximation may not be accurate, it does at least indicate the order of magnitude of $\hat{b}$.

Confidence Interval for N

Assuming $\hat{N}$ to be asymptotically normal, we have the approximate 95 per cent confidence interval for N, namely

$$\hat{N}-\hat{b} \pm 1.96/\mathrm{Var}[\hat{N}], \qquad (4.6)$$

where $\hat{b}$ can be neglected if it is less than one tenth of $\sqrt{\mathrm{var}[\hat{N}]}$ (cochran, 1977). However, the statistic r is more nearly normally distributed than $\hat{N}$ (Darroch, 1958). Therefore, we can use r as a basis for a confidence interval as follows:

We know that $\hat{N}$ is a solution of the equation.

$$\prod_{i=1}^{s}(N-n_i) = N^{s-1}(N-r)$$

Therefore

$$(N-r) = \prod_{i=1}^{s} (N-n_i)/ N^{s-1}$$

Now, let $E[r]=\rho$ Then using the identity

$$\prod_{i=1}^{s} \begin{bmatrix} N \\ n_i \end{bmatrix}^{-1} \sum_{r\{a_i\}} \frac{N!}{(N-r)!\prod_w a_w!} = 1$$

We have

$$E[N-r] = N-\rho = \left\{ \prod_{i=1}^{s} (N-n_i)/N^{s-1} \right\} \prod_i \binom{N-1}{n_i}^{-1} \sum_{r\{a_w\}} \frac{(N-1)!}{(N-1-r)!\ \prod_w a_w!}$$

$$= \prod_i (N-n_i)/N^{s-1}$$

which implies that, the expected value of r, regarded as a function of N is

$$\rho(N) = N - \prod_{i=1}^{s} (N-n_i)/N^{s-1}$$

The variance of r, expressed as a function of N, is (Lee,1972)

$$\sigma^2(N\ ) = \Big[N-\rho(N)\Big]\Big[\rho(N)-\rho(N-1)\Big]$$

and we have

$$P_r\Big[r-1.96\sigma(N) < \rho(N) < r+1.96\ \sigma(N)\Big] \cong 0.95 \qquad (4.7)$$

$$\cong P_r\Big[r-1.96\sigma(\hat{N}) < \rho(N) < r + 1.96\sigma(\hat{N})\Big]$$

$$\cong P_r\Big[r_1 < \rho(N) < r_2\Big]$$

$$\cong P_r\Big[\ \rho^{-1}(r_1) < N < \rho^{-1}(r_2)\Big]$$

$$\cong P_r\left[N_1 < N < N_2\right] \text{, say} \tag{4.8}$$

since $\rho(N)$ is a monotonic increasing function of N. The confidence limits N_1 and N_2 can be calculated by setting $h(N)=\rho(N)-r_i$ and solving $h(N)=0$ iteratively as above. Alternatively we could deal with the interval (4.7) directly by solving the equations $r \pm 1.96\ \sigma(N)=\rho(N)$ iteratively on a computer.

Random Sample Sizes

We now mention briefly the more common situation in which the sample sizes n_i are random variables rather than fixed parameters. Darroch (1958) has investigated model (4.2) in some detail and shows that as far as the point and interval estimation of N is concerned, there is no difference (asymptotically) between the two cases of fixed and random sample size. This is because, the maximum-likelihood estimate $\hat{N}$ is almost the same in both cases, and in estimating the variance of $\hat{N}$, one effectively replaces n_i by $E[n_i]$, when n_i is random.

Sample of Size one

Putting $n_i=1$ in equation (4.4), we find that $\hat{N}$ is the solution of

$$(1 - r/N) = (1 - 1/N)^s \tag{4.9}$$

and this equation can be solved in the same way as equation (4.4). However, since s is generally large, Craig (1953) suggests taking logarithms and solving

$$H(N) = (s-1)\log_{10}N + \log_{10}(N-r) - s\log_{10}(N-1) = 0$$

Using a good table of logarithms for example Spencely (1952) and a suitable first approximation N_o such as the following:

Let f_x be the frequency of cases in which the same individual is caught x times ($x = 1,2,\ldots\ldots$) and let $s_2 = \sum x^2 f_x$; then craig suggests

$$N_o = s^2/(s_2-s)$$

Alternatively, following Darroch (1958) and letting $N \longrightarrow \infty$, $s \longrightarrow \infty$ subject s/N (=D say) remaining constant, equation (4.9) becomes

$$1 - r/N = e^{-D} \qquad (4.10)$$

A first approximation to $\hat{N}$ (an upper bound) is then $N_o' = s/D_o'$ where D_o' , the solution of

$$(1-e^{-D})/D = r/s \text{ (= a say)} \qquad (4.11)$$

is obtained by linear interpolation in table A3 of the Appendix. Samwel (1969) suggests a further approximation $N_o'' = sD_o''$, where D_o'' is the solution of (4.11) with

$$a = r/(s + D_o').$$

For the limiting process mentioned above, Darroch (1959) shows that asymptotically

$$E[\hat{N}] = N + b$$

where

$$b = D^2(e^D-1-D)^{-2}/2$$

and

$$\mathrm{Var}[\hat{N}] = N(e^{D}-1-D)^{-1}$$

Both b and Var[$\hat{N}$] can be estimated by replacing D by $\hat{D} = s/\hat{N}$. Extensive tables of e^{D} are given, for example, in Becker and Van Orstrand (1924), and Comrie (1959).

Confidence interval for N based on r can be calculated as in equation (4.7) above using

$$\rho(N) = N(1-e^{-D}).$$

and

$$\sigma^2(N) = Ne^{-2D}(e^{D}-1-D).$$

If the interval (4.8) is used, we have to solve two equations of the form

$$\rho(N) = r_i \quad \text{or} \quad (1-e^{-D})/D = r_i/s$$

Which, as for the case $r_i=r$ above in equation (4.11) can be solved for D by interpolating linearly in table A3 of the Appendix.

Mean Petersen Estimate

At each stage of sampling a modified Petersen estimate of N can be calculated as explained in section 3.1

$$N_i^* = \frac{(M_i + 1)(n_i + 1)}{(m_i + 1)} - 1 \qquad (i = 1,2,\ldots,s)$$

With variance estimate

$$V_i^* = \frac{(M_i + 1)(n_i + 1)(M_i - m_i)(n_i - m_i)}{(m_i + 1)^2(m_i + 2)}$$

Therefore, a natural estimate of N, suggested by Chapman

(1952) is the average:

$$\bar{N} = \sum_{i=2}^{s} N_i^{*}/(s-1)$$

Since the covariances of the N_i^{*} are asymptotically negligible compared with their variances, we have approximately

$$\mathrm{Var}\left[\bar{N} \mid \{n_i, M_i\}\right] = \sum_{i=1}^{s} \mathrm{Var}\left[N_i^{*} \middle| n_i, M_i\right]/(s-1)^2$$

This can be estimated by either

$$V^{*} = \sum_{i=1}^{s} V_i^{*}/(s-1)^2$$

which is unbiased if and only if each V_i^{*} is almost unbiased, or by

$$\mathrm{Var}[\bar{N}] = \sum(N_i^{*} - \bar{N})^2/(s-1)(s-2).$$

Which is almost unbiased when N_i^{*} have the same mean . When these conditions of unbiasedness are not satisfied, both estimates are conservative in that they tend to overestimate the true variance.

Schanabel's Binomial Model

An alternative approach to the Schnabel census can be made assuming that the M_i are fixed parameters and then using the binomial approximation of section 3.1 .This leads to Schnabel´s (1938) model

$$f(m_2, \ldots, m_s \mid \{n_i, M_i\}) = \prod_{i=2}^{s} \binom{n_i}{m_i} \left[\frac{M_i}{N} \right]^{m_i} \left[1 - \frac{M_i}{N} \right]^{n_i - m_i} \tag{4.12}$$

and the maximum-likelihood estimate of N is now the appropriate root of the equation.

$$\sum_{i=2}^{s} \frac{(n_i - m_i)M_i}{N-m_i} = \sum_{i=2}^{s} m_i \qquad (4.13)$$

This model assumes that at each stage, n_i/N is sufficiently small (say less than 0.1) for one to be able to ignore the complications of sampling without replacement. If each M_i/N is also small, a first approximation to the solution of equation (4.13) is

$$N' = \sum_{i=2}^{s} n_i m_i \Big/ \sum_{i=2}^{s} m_i = \lambda/m \text{ ,say}$$

We note that, irrespective of any assumptions concerning the probability function of m_i or the magnitudes of various parameters, N' has a certain intuitive appeal, being simply a weighted average of Petersen estimates $n_i M_i/m_i$. When n_i/N and M_i/N are both less than, say, 0.1 for each i, a modification of N' which is almost unbiased is

$$N'' = \lambda/(m + 1)$$

Noting that when the above conditions hold, m_i is approximately a Poisson with parameter $M_i n_i/N$ and the sum, m of independent Poisson variables is also Poisson with parameter λ/N . Chapman (1952) shows that

$$E\left[N'' \,|\, \{n_i, M_i\}\right] = N(1 - \exp{-\lambda/N})$$

and

$$V\left[N'' \,|\, \{n_i, M_i\}\right] = N^2\left(N/\lambda + 2N^2/\lambda^2 + 6N^3/\lambda^3\right) .$$

A study of Raff (1956) would suggest that Poisson approximation still applies, even if $0.1 < M_i/N < 0.2$, provided that n_i/N is less than 0.1, so that the hypergeometric distribution of m_i is well approximated by binomial. But errors in these approximations have a cumulative effect on the sum m, so that for m to be approximately Poisson we would not have more than one or two samples with M_i/N greater than 0.1 .

Confidence Interval

Assuming N^* to be asymptotically normal we can calculate a confidence interval for N in the usual manner. However, as in Petersen method of section 3.1 , it is recommended to base confidence intervals on the distribution of m . For $m \leq 50$, we can use Chapman's Poisson Table A1 in the Appendix to obtain the shortest interval for N/λ, and hence for N.

When $N > 50$ we can use the normal approximation to Poisson, and the 95 per cent confidence interval for N is given by the roots of the quadratic equation.

$$N(m - \lambda/N)^2/\lambda = 1.96^2$$

That is

$$N/\lambda = \frac{2m + 1.96^2(1-\delta) \pm 1.96 \sqrt{(1-\delta)(4m+1.96^2(1-\delta))}}{2m^2} \qquad (4.15)$$

where

$$\delta = \sum_{i=2}^{s} n_i M_i^2/\lambda N'$$

Regression Methods

The maximum-likelihood method described in section 4.1 will give the most efficient estimate $\hat{N}$ of N, provided the assumptions underlying the model are satisfied $\hat{N}$ will, however, tend to be sensitive to departures from underlying assumptions, particularly those relating to constant N and the random behaviour of marked animals. Therefore, in practice, $\hat{N}$ may sometimes be an inefficient estimate and $\mathrm{Var}[\hat{N}]$ may be unreliable. For this reason less efficient but more robust estimate of N, like $\bar{N}$ are desirable. In particular a useful regression method has been suggested by Schumacher and Eschmeyer (1943) and we now discuss this technique in detail.

Schumacher and Eschmeyer's Method

In Schnabel's Model (4.12) each m_i is assumed to be binomially distributed, so that $y_i = m_i/n_i$ has mean M_i/N and variance

$$\sigma_i^2 = \frac{M_i}{N}\left(1-\frac{M_i}{N}\right)\frac{1}{n_i}$$

We may therefore write

$$y_i = \beta m_i + e_i \qquad i = 1,2,\ldots.,s$$

where $\beta = 1/N$ and the "error" e_i has mean zero and variance σ_i^2.

If we plot y_i against M_i, the plotted points should lie approximately on a straight line of slope β passing through the origin. Since the variance of e_i is not constant, the least squares fitting of a straight line should be done using weights W_i, say, as in section 1.3.

Thus $\tilde{N}$, the least squares estimate of N, is given by

$$1/\tilde{N} = \tilde{\beta} = (\Sigma w_i y_i M_i)/\Sigma w_i M_i^2$$

Where all summations throughout this section are for i=1,2,...,s. If the weights are chosen in the usual manner, namely proportional to the reciprocal of variances, then the above equation becomes the maximum likelihood equation (4.13). However, although, these weights will give the most efficient estimate of N when sampling is truly random, we are computationally no better off than before, as those unknown weights have to be estimated iteratively. DeLury (1958) also points out that "owing to the tendency of fishes to stratify and for other reasons that lead to similar effects, the proportion, of marked individuals available to the sampling at any one time is likely to differ widely from the "true" proportion, and the weights are therefore likely to be seriously wrong. In these circumstances, weighting by sample size alone is preferable to the weighting according to the proportions tagged". In support of this last statement we note that $(M_i/N)(1-M_i/N)$ does not vary much as M_i/N varies from 0.2 to 0.8. Therefore putting $W_i = n_i$, is now given by

$$\tilde{\beta} = (\Sigma m_i M_i)/(\Sigma n_i M_i^2) \qquad (4.16)$$

which is equivalent to the formula given by Schumacher and Eschmeyer (1943).

Properties of $\tilde{N}$

The mean and variance of $\tilde{N}$ could be calculated using the δ-method described in section 1.3 and an approximate confidence interval for N obtained in the usual manner.

However, following Delury (1958) it seems preferable to assume that e_i's are independently normally distributed with variances σ^2/n_i and the $100(1-\alpha)$ per cent confidence interval for N is given by

$$\left\{\frac{\sum n_i m_i^2}{\sum m\ M \ + \ t_{s-2}[\alpha/2](\tilde{\lambda}^2(n\ m\ )^{1/2}}, \frac{\sum n_i m_i^2}{\sum n_i M_i + t_{s-2}[\alpha/2](\tilde{\lambda}^2(n_i M_i^2)^{1/2}}\right\} \qquad (4.17)$$

where

$$(s-2)\tilde{\sigma}^2 = \sum m_i^2/n_i - (\sum n_i M_i)^2/\sum n_i M_i^2$$

From bead sampling experiments, DeLury showed that the above confidence interval compared favourably with the confidence interval based on the more efficient binomial weights. We expect the interval (4.17) to be robust with regard to departures from underlying assumptions and this model should therefore be used in conjunction with the other methods mentioned so far in this chapter. In particular a graph is always a useful indicator of any marked departures from assumptions underlying the model.

Tanaka's Model

Sometimes a plot of y_i versus m_i as indicated above yields a graph which is definitely curved. For this situation Tanaka (1951, 1952) has prepared a non-linear relationship of the form $y = (M/N)^\gamma$ or, taking logarithms, the linear regression model

$$E\left[-\log_{10} y_i\right] \cong \gamma(\log_{10} N - \log_{10} M_i\ ) \qquad (i=1,2,\ldots,s)$$

Least squares estimates and confidence intervals for γ

and $\phi = \log_{10}N$ can be obtained by setting.

$$Y_i = \log_{10}y_i = = \log_{10}(n_i/m_i) \text{ , } x_i = \log_{10}M_i$$

and using the regression methods of Chapter 1 . A visual estimate of ϕ can be obtained by drawing the regression line by free hand and extending this line to meet the x-axis . However, before actually looking up the logarithms it is simpler to plot y_i versus M_i on log-log paper first. The parameter γ can be interpreted as an index of trap response. For example, if $\gamma < 1$, then

$$E[m_i/n_i] > M_i/N$$

or, rearranging,

$$E[m_i/n_i] > E[n_i/N]$$

and the marked individuals have a higher probability of capture than the unmarked. However, care should be taken in interpreting the graph of y_i versus M_i as several interpretations are possible.

If for example, the graph curves downwards, Hayne (1949) argues that the fall-off in the proportion of marked in the sample could be due to the immigration of unmarked into the trapping area. But if the graph is interpreted in the light of Tanaka´s model we have $\gamma < 1$ and the curvature is due to the marked animals having a higher probability of capture than the unmarked. In this case the fall off is simply due to the curve settling down to its "correct" position instead of dropping away from its "correct" position as suggested by Hyne. Obviously both interpretations are possible, and one could perhaps distinguish between the two by an analysis of the recaptures to see whether any individuals were being captured more often than expected. Alternatively Marten´s regression model discussed below may be applicable.

Marten´s Model

One of the difficulties in using Tanaka´s regression model above is the interpretation of parameter γ. Although one may obtain a good straight-line fit to the graph of Y_i versus x_i, the model lacks, a simple "physical" interpretation and other regression curves may give just as good a fit.

For example, suppose that the average catchability of the m_i marked individuals in each sample bears a constant ratio to the average catchability of the u_i unmarked members, then we have approximately

$$\frac{m_i}{N_i} = \frac{u_i}{k(N-M_i)} = \frac{n_i}{k(N-M_i) + M_i}$$

and , since $y = m_i/n_i$, then

$$E[m_i/n_i|M_i) \cong \frac{M_i}{k(N-M_i)+M_i}$$

That is

$$E[y_i|M_i] \cong \frac{m_i}{k(N-M_i)+M_i}$$

This means that the plot of y_i versus M_i will be curved upwards or downwards, depending on whether k is less than or greater than unity. Instead of fitting Tanaka´s model we can rearrange the above equation, apply a bias correction, and obtain linear regression model

$$E[y_i/m_i M_i) \cong k(N-M_i) \qquad (4.18)$$

where

$$y_i = u_i(M_i+1)/(m_i+1)$$

This model was first suggested by Marten (1970)

We now give a derivation of Marten's model. A closed population of size N is sampled at times 0, 1, ..., T. Each sample is observed to contain m_t marked animals and u_t unmarked animals ($m_o=0$). At each sample unmarked animals are marked and all animals returned to the population. The total number of unmarked animals in the population at sample t is known to be

$$M_t = \sum_{i=0}^{t-1} m_i$$

and the remainder of the population

$$U_t = N-M_t \qquad (4.19)$$

is unmarked.

Catchability may vary from sample to sample, but let us assume the average catchability of unmarked animals estimated by U_t/U_t is in constant and unknown ration k to the average catchability of marked animals estimated by m_t/M_t. That is

$$\frac{u_t/U_t}{m_t/M_t} = k \qquad (4.20)$$

for all t. The estimate of the unmarked population of sample t, under the assumption of equal catchability (that is $k=1$), may be obtained from equation (4.20) as

$$\hat{U}_t = u_t m_t/M_t \qquad (4.21)$$

combining equations (4.20) and (4.21) we get

$$\hat{U}_t = k\ u_t \tag{4.22}$$

Thus $\hat{U}_t$, the estimate of the unmarked population under the assumption of equal catchability, consistently overestimates or underestimates the actual unmarked population by a constant proportion k .

In order to illustrate how removal can exploit this error of constant proportion, suppose the unmarked population is estimated on two occasions.

$$\hat{U}_1 = kU_1 \tag{4.22}$$

and

$$\hat{U}_2 = kU_2 \tag{4.23}$$

subtracting (4.23) from (4.22) we get

$$\hat{U}_1-\hat{U}_2 = k(U_1-U_2) \tag{4.24}$$

That is, the difference between estimates of the unmarked population at two different samples, under the assumption of equal catchability, is also an overestimate or underestimate of the actual difference by a constant proportion k . If the actual difference in the unmarked population is know, by removing a known number of animals from it, then k may be estimated by rearranging equation (4.24) to

$$k = (\hat{U}_1-\hat{U}_2)/(U_1-U_2)$$

If the population is closed, the number of animals marked a new at each sample represents a known difference in the unmarked population. The total numbers of marked animals M_t at successive samples therefore provide a succession of known differences in the unmarked population which may be exploited to obtain a composite estimate of k over all samples. This is accomplished by substituting equation (4.19) in equation (4.22), which gives

$$U_t = kN-kM_t \qquad (4.25)$$

Note that equation (4.25) has a form suitable for linear regression analysis with M_t as the independent variable and U_t as the dependant variable. The ratio k is the negative of the slope, and N may be obtained by extrapolating the line to $U_t = 0$ that is N=M when $\hat{U} = 0$.

The U_t are not independently normally distributed with equal variance over all samples, as assumed in regression analysis; but this is of small practical significance. The procedure, then, is to compute the M_t from the fact that $M_t = \sum m_t$ and the $\hat{U}_t$ from $\hat{U}_t = U_t\ M_t/m_t$ The appropriateness of the regression method which depends up on whether the catchability ratio k is in fact constant over all samples, may be evaluated graphically by the extent to which points $(M_t,\hat{U}_t)$ fall along a straight line.

If judged appropriate equation (4.25) is then fitted to the M_t and $\hat{U}_t$ by conventional linear regression analysis. The estimates of k and N following Bennet and Franklin (1954) are

$$\hat{k} = -\sum_{t=1}^{T} U_t(M_t-\bar{M}) \ / \ \sum_{t=1}^{T} (M_t-\bar{M})^2 \qquad (4.27)$$

and

$$\hat{N} = \bar{M} + \bar{\hat{U}}/\hat{k}$$

The standard error of $\hat{K}$ is then

$$S_{\hat{k}} = \sqrt{\frac{\sum_{t=1}^{T}\left[U_t-(\hat{k}\hat{N}-\hat{k}M_t)\right]^2}{(T-2)\sum_{t=1}^{T}(M_t-\bar{M})^2}} \qquad (4.27)$$

A first approximation for the standard error of $\hat{N}$ (following Bennet and Franklin, 1954) is

$$S_{\hat{N}} = \sqrt{\frac{\sum_{t=1}^{T}(U_t-(\hat{K}\hat{N}-M_t))^2}{T(T-2)\hat{k}^2} + \frac{(S\hat{k}\ \hat{U})^2}{\hat{k}^4}}$$

The confidence limits of $\hat{k}$ may be obtained by multiplying $S_{\hat{k}}$ by the appropriate value from t-table with T-2 degrees of freedom. To test for equal catchability of marked and unmarked animals we test the hypothesis.

$$H_o: k = 1$$

This is rejected if the confidence interval does not cover 1 .

Allowing for Known Removals

The Hypergeometric Model

In many population experiments there are accidental deaths due to trapping and handling and some animals may be deliberately removed from further study. If the percentage of such removals is appreciable, some allowance must be made for them in the particular model used. For example, the removal could form a major part of the sample as in commercially exploited populations, with the remainder of the sample being tagged (or retagged) and returned to the population.

Suppose, then, that d_i members of the ith sample are not returned to the population and let M be the number of marked animals alive in the population before the ith sample is taken. Then, assuming the n_i to be fixed parameters, and noting that

$$f(m_1,m_2,\ldots..m_s/\{n_i\}) = \prod_{i=2}^{s} \begin{bmatrix} M_i \\ m_i \end{bmatrix} \begin{bmatrix} N-M_i \\ u_i \end{bmatrix} / \begin{bmatrix} N \\ n_i \end{bmatrix}$$

as given by Chapman (1952), then the above model becomes

$$f(m_1,m_2,\ldots..m_s/\{n_i\}) = \prod_{i=2}^{s} \begin{bmatrix} M_i \\ m_i \end{bmatrix} \begin{bmatrix} N-M_i-D_i \\ n_i-m_i \end{bmatrix} / \begin{bmatrix} N-D_i \\ n_i \end{bmatrix}$$

where

$$D_i = \sum_{j=1}^{i-1} d_j$$

is the total number removed up to but not including the ith sample, and N is now the initial population size.

It is readily shown that $\hat{N}_D$, the maximum likelihood estimate of N, is the unique root greater than r of the polynomial

$$(N-r)/N = \prod_{i=1}^{s} \left\{\frac{N-D_i-n_i}{N-D_i}\right\} \qquad (4.28)$$

where r is the total number of different animals caught during the whole experiment, including the ones not returned.

By setting g(N) equal to the right-hand side of equation (4.28) and defining

$$h(N) = N - r - Ng(N),$$

equation (4.28) can be solved in exactly the same way as equation (4.4).

If N is large, so that each n_i/N is less than about 0.1 , a reasonable first approximation of (4.28) suggested by Robson and Regier (1968) is

$$N_{(1)} = \frac{\left[\Sigma(n_i+D_i)\right]^2 - \Sigma(n_i+D_i)^2 - (r+\sum_i D_i)^2 + r^2 + \Sigma D_i^2}{2\Sigma m_i} \qquad (4.29)$$

Where all summations are for $i = 1,2,\ldots,s$ $(m_1=0,\ D_1=0)$. This approximation is obtained by cross-multiplying in equation (4.28), dividing both sides by N^s , expanding the products and neglecting powers of $1/N$ greater than the second.

Another possible first approximation is the mean estimate

$$\bar{N}_D = \sum_{i=1}^{s} (N_i^* + D_i)/(s-1)$$

where

$$N_i^* = \left\{(n_i + 1)(M_i + 1)/(m_i + 1)\right\} - 1 \qquad (4.30)$$

is the modified Petersen estimate.

The mean and the variance of $\hat{N}_D$ can be evaluated using the method outlined in Darroch (1958). Asymptotically,

$$E[\hat{N}_D] = N + b$$

where b , the bias, is estimated by

$$\hat{b}=\frac{\left[\frac{1}{\hat{N}_D}+\sum_{i=1}^{s}\left(\frac{1}{\hat{N}_D-D_i-n_i}-\frac{1}{N_D-D_i}\right)\right]^2-\left[\frac{1}{\hat{N}_D^2}+\sum_{i=1}^{s}\frac{1}{(\hat{N}_D-D_i-n_i)}-\sum_{i=1}^{s}\frac{1}{(\hat{N}_D-D_i)^2}\right]}{2\left[\frac{1}{\hat{N}_D-r}-\frac{1}{\hat{N}_D}-\sum_{i-1}^{s}\frac{n_i}{(\hat{N}_D-D_i-n_i)\ (N_D-D_i)}\right]^2}$$

and asymptotic variance of $\hat{N}_D$ is estimated by

$$\mathrm{Var}\left[\hat{N}_D\right] = \left[\hat{N}_D-r\right]/h'\left[\hat{N}_D\right]$$

$$=\left[\frac{1}{\hat{N}_D-r}-\frac{1}{\hat{N}_D}-\sum_{i=1}^{s}\frac{n_i}{(\hat{N}_D-D_i-n_i)\ (\hat{N}_D-D_i)}\right]^{-1}$$

Note that

$$\hat{b} = -\frac{1}{2}(\hat{N}_D-r)h''(\hat{N}_D)/\left[h'(\hat{N}_D)\right]^2$$

Therefore the bias and the variance of $\hat{N}_D$ can be estimated from the last steps in Robson and Regier's iterative procedure as for the case of no removals described earlier.

Overton's Method

Overton (1965) has given the following method for modifying Schnabel's estimate N´ to allow for known removal. Since

$$E\left[m_i|M_i,n_i,D_i\right] = M_i n_i/(N-D_i)\ ,\quad (i=1,2,.....,s)$$

then summing this equation for i=1,2,...,s and setting

$m = \sum_{i=2}^{s} m_i$ leads to

$$E\left[m \mid \{ M_i, n_i, D_i \}\right] = \frac{1}{N}\sum_{i=2}^{s} \left\{ \frac{n_i \; M_i (N - D_i + D_i)}{N - D_i} \right\}$$

$$= \sum_{i=}^{s} \frac{n_i m_i}{N} + \sum_{i=2}^{s} \frac{n_i m_i D_i}{N(N - D_i)}$$

Equating m to its expected value leads to an estimate N'_D of N given by

$$N'_D = \sum \frac{n_i M_i}{m} + \sum_{i=1}^{s} \frac{n_i m_i D_i}{(N'_D - D_i)m} = N' + A \text{ say.} \qquad (4.31)$$

Where A is to be added to the usual Schnabel estimate N'. Equation (4.31) must be solved iteratively for N_D and overton suggests the following first approximation which will usually be close unless removal is heavy, namely,

$$N_{(1)} = N' + A_{(1)}$$

$$A_{(1)} = \sum n_i M_i D_i / mN' = (\sum n_i M_i D_i) / \sum n_i M_i$$

so that $N_{(1)} < N'_D$. Another first approximation suggested by Robson and Regier (1968) is obtained from equation (4.31) directly by neglecting D_i in the denominator of the right-hand side and solving for N'_D , namely

$$\frac{1}{2}\left\{N' + \sqrt{N'^2 + 4\sum n_i M_i D_i / m}\right\}$$

whichever first approximation is used, however, subsequent a approximations are

$$N_{(j)} = N' + A_{(j)} \qquad (4.32)$$

where,

$$A_{(j)} = \sum n_i M_i D_i / (N^*_{(j)} - D_i)m$$

and $N^*_{(j)}$ is to be determined by $N_{(j-1)}$. In determining a suitable method for choosing $N^*_{(j)}$, Overton points out that the iterative process is not necessarily convergent if we set $N^*_{(j)} = N_{(j-1)}$. But if $N^*_{(j)} < N'_D$, then $N_{(j)} > N'_D$ and vice versa, so that N'_D will be between any pair $(N^*_{(j)}, N_{(j)})$. Overton therefore suggests the reasonable procedure of choosing

$$N^*_{(j+1)} = \frac{1}{2}\left[N^*_{(j)} + N_{(j)}\right]$$

and taking $N^*_{(2)}$ as the integer nearest to $N_{(1)}$.

Confidence limits for N can be obtained as in equation (4.14) using

$$\lambda = \sum n_i M_i + \sum n_i M_i D_i / (N'_D - D_i)$$

which can be obtained from the last step of equation (4.32) and either Chapman's table (Table A1, in the Appendix) or normal approximation can be used with

$$\delta = \frac{N'_D}{\lambda} \sum_{i=2}^{s} \frac{n_i M_i^2}{(N'_D - D_i)^2}$$

Regression Methods

All the regression methods discussed earlier depend heavily on the assumption of N remaining constant, and these methods cannot be used unless removal is negligible. However, the mean estimate $\bar{N}_D$ given above will provide a robust estimate of N, provided the number of recaptures in each sample is not too small. The variance of $\bar{N}_D$ can be estimated by

$$\mathrm{Var}\left[\bar{N}_D \right] = \sum_{i=2}^{s} (N_i^* - D_i - \bar{N}_D)/(s-1)(s-2)$$

or when the assumptions underlying the Petersen method hold for each sample, more efficiently by

$$V_D^* = \sum_{i=2}^{s} \frac{(M_i=1)(n_i+1)(M_i-m_i)(n_i-m_i)}{(m_i+1)(m_i+2)(s-1)^2}$$

Testing the Underlying Assumptions

Validity of the Models

MultinomialModel

Using the notation of section 4.1, we wish to test the 8null hypothesis

H_0 that P_w can be written as a product of the $\{p_i\}$ and $\{q_i\}$ if H_0 holds, then the distribution

$$f(\{a_w\}) = \frac{N!}{\prod_w a_w!(N-r)!} Q^{N-r} \prod_w p_w^{a_w} \quad ,$$

reduces to Darroch's multinomial model

$$f(\{a_w\}) = \frac{N!}{\pi_w a_w!(N-r)!} \prod_{i=1}^{s} p_i q_i^{N-n_i}$$

We now derive a test statistic for the above null hypothesis. Let $x_1, x_2, \ldots, x_n$ be a random sample from the binomial distribution.

$$\binom{N}{x} p^x q^{N-x} \qquad (q=1-p \ , \ x=0,1,\ldots,N)$$

Suppose that x_i takes a value x with frequency $f_x (\Sigma_x f_x = n)$, Then it is easily seen that the joint distribution of f_x is multinomial with $N+1$ categories. The goodness-of-fit statistic for testing the appropriateness of the binomial model is

$$T = \sum_{x=0}^{N} (f_x - E_x)^2 / E_x$$

Where E_x is the expected frequency.

In our case then the estimate of the expected frequency is $\hat{N}\hat{p}_w$, so that the appropriate goodness of fit test statistics is

$$T = \sum_{w} \left[a_w - \hat{N}\hat{p}_w\right]^2 / \hat{N}\hat{p}_w \quad .$$

Where, for example

$$\hat{p}_{124} = \hat{p}_1 \hat{p}_2 \hat{p}_4 Q / \hat{q}_1 \hat{q}_2 \hat{q}_4 \quad \text{and} \quad \hat{p}_i = 1 - \hat{q}_i = n_i / \hat{N}$$

When H_0 is true, T is asymptotically distributed as chi-squared with $d-s-1$ degrees of freedom, where d is the number of different recapture histories w. If any of the groups are too small they can be pooled in the usual manner.

Hypergeometric Model

Chapman (1952) has suggested a non-parametric test for the validity of hypergeometric model.

$$f(\{n_i\}) = \frac{N!}{\prod_w a_w!(N-r))!} \prod_{i=1}^{s} \begin{bmatrix} N \\ n_i \end{bmatrix}^{-1}$$

Let b_{ij} $(i < j)$ be the number of marked animals in the jth sample which were first caught and marked in the ith sample. When sampling is random we have

$$E[b_{ij}/n_i | n_j, u_i] = u_i/N = \theta_i \quad \text{,say}$$

and an array

$$\begin{array}{ccccc}
\frac{b_{12}}{u_1u_2}, & \frac{b_{13}}{u_1u_3}, & \frac{b_{14}}{u_1n_4}, & \ldots, & \frac{b_{1s}}{u_1n_s} \\
 & \frac{b_{23}}{u_2u_3}, & \frac{b_{24}}{u_2n_4}, & \ldots, & \frac{b_{2s}}{u_2n_s} \\
 & & & \cdots & \cdots \\
 & & & & \frac{b_{(s-1)s}}{u_{(s-1)}n_s}
\end{array}$$

may be formed, in which each element is a random variable with expectation $1/N$. These random variables are independent within each row, but are dependent between rows as $b_{i_1 j}$ and $b_{i_2 j}$ belong to the same sample and are therefore correlated. For large N. however, the correlation is small and Chapman suggests testing for the validity of the underlying model by testing whether the $t = \{ s(s-1)\}/2$ elements formed by putting the rows one after another is a sequence of random observations from a common distribution. The test suggested is the sign test of Moore and Wallis (1943) based on D, the number of negative signs in the sequence of successive differences of observations (i.e first observation minus second observation, e.t.c).

When the hypothesis of a common distribution is true,

$$E[D] = (t - 1)/2 \quad , \quad \sigma^2[D] = (t + 1)/2$$

and $(D - E[D])/\sigma[D]$ is approximately unit normal for $t \geq 12$; Moore and Wallis (1943) have tabled the exact distribution of D for small values of t.

In many cases the alternative hypothesis to randomness are essentially one sided. For example, possible alternatives are:

(a) Some marked individuals die off more rapidly or disappear, so that they are not available for sampling;

(b) the marked individuals disperse from the tagging location slowly and are more likely to be recaptured in the samples taken soon after marking than later; and

(c) the population size N is increasing through recruitment.

If any of these alternatives is true, the numbers in each row of the array will tend to decrease from left to right. In this case a test based on the whole array as a single sequence has the following defect: In each row the probability of a negative difference between successive elements is less than 1/2 , but the probability of negative difference between the last element of any row and the first element of the next row will be much greater than 1/2 . Also another disadvantage of considering the whole array as a single sequence is that the variances of the elements will vary from row to row. However, if the sample sizes n_i are approximately the same, the elements within a given row will have approximately the same distribution when the underlying model is valid . Therefore to overcome the above objections, Chapman recommends treating each row separately so that the array may be considered as (s-1) sequences of observations decreasing in length from s-1 to 1 . A test of randomness may be made using the statistic

$$X = D_1 + D_2 + \ldots + D_{s-2}$$

Where D_i is the number of negative differences in row i.

Note that no difference is obtained from the last row. Then

$$E[X] = (S-1)(S-2)/2 \quad , \quad \sigma^2[X] = (s+3)(s-2)/4 \ ,$$

and X is asymptotically normal.

An alternative test can be carried out using table 4.5 below. Neglecting the complications of sampling without replacement, the columns of Table 4.5 represent independent multinomial distributions.

Let p_{ij} be the probability of being in the class containing b_{ij} individuals, then from the fact that

$$E\left(b_{ij}/n_j \,|\, n_j, u_i\right) = u_i/N = \theta_i \ , \text{ say}$$

it follows that $p_{ij} = \theta_i$ for $j = i+1 , \ldots, s$ and the likelihood function for Table 4.5 is proportional to

$$\prod_{j=2}^{s} \left\{ \left(\prod_{i=1}^{j-1} \theta_i^{b_{ij}} \right) \left(1 - \theta_1 - \theta_2 - \ldots - \theta_{j-1}\right)^{u_j} \right\}$$

$$= \theta^{b_1} \theta^{b_2} \ldots (1-\theta_1)^{u_2} (1-\theta_1-\theta_2)^{u_3} \ldots$$

The maximum-likelihood estimate $\hat{\theta}_i$ of θ_i are the solutions of the equations

$$\frac{b_1}{\theta_1} - \frac{u_2}{(1-\theta_1)} - \frac{u_3}{(1-\theta_1-\theta_2)} - \cdots - \frac{u_s}{(1-\theta_1-\theta_2-\ldots-\theta_{s-1})} = 0$$

$$\frac{b_2}{\theta_2} - \frac{u_3}{(1-\theta_1-\theta_2)} - \cdots - \frac{u_s}{(1-\theta_1-\theta_2-\ldots-\theta_{s-1})} = 0$$

$$\ldots \qquad \ldots$$

$$\frac{b_{s-1}}{\theta_{s-1}} - \frac{u_s}{(1-\theta_1-\theta_2-\ldots-\theta_{s-1})} = 0$$

The expected frequencies corresponding to the observed frequencies b_{ij} are $n_j \hat{\theta}_i$ and the expected frequencies for the u_j are obtained by subtraction. The goodness-of-fit statistic based on comparing the observed frequencies with the expected frequencies in Table 4.1 is Chi-squared with (s-1)(s-2)/2 degrees of freedom.

Table:4.1 Contingency table for carrying out Leslie´s test for dilution

able3.5

					TOTAL
	b_{12}	b_{13}	...	b_{1s}	$b_{1.}$
		b_{23}	...	b_{2s}	$b_{2.}$
			...	...	...
				$b_{s-1,s}$	$b_{s-1,.}$
	u_2	u_3	...	u_s	$u_.$
TOT.	u_2	n_3	...	n_s	$n_.$

Unfortunately the above maximum likelihood equations do not seem to have explicit solutions and alternative estimates of θ_i are desirable. For example if there were no blanks in Table 4.1 the estimate of θ_i would be $b_{i.}/n_{.}$ Another problem that arises in the use of the above test is that the expected frequencies are often small and pooling may be needed. Leslie (1935) suggest pooling the b_{ij} (and their expected frequencies) in each column, thus reducing the table to s-1 pairs of frequencies (m_j, u_j).

The above method can still be used when mortality is taking place, provided that all sub classes of marked and unmarked have the same mortality rates between successive samples. In this case the proportion of the population first marked at the ith sample will remain constant and equal to u_i/N, so that $E[b_{ij}/n_j | n_j, u_i] = u_i/N$ is still satisfied.

Regression Model

The best evidence for the validity of the Schumacher and Eschmeyer's regression method is obviously the linearity of the graph. Any change in N through recruitment, mortality, etcetera or any variation in catchability will affect the basic equation $E[m_i/n_i] = M_i/N$ and this in turn will show up in the graph provided the m_i are not too small. In practice, point and interval estimates of N should be obtained using as many different methods as possible, as any departures from underlying assumptions will usually affect different models in different ways. A substantial agreement among the estimates would then give support for the validity of the models concerned.

Tests for Random Sampling

A part from poor experimentation and inadequate experimental design, there are three basic sources of non-randomness:

(a) There may be sub categories in the population due to size, sex, species, etcetera; for which sampling is random within each sub category, but not between the sub categories. In this case, if there is no mortality, the chi-squared goodness-of-fit test based on the following contingency table

	SUBCATEGORY				TOTAL
	X	Y	...	W	
Recaptured	m_{2x}	m_{2y}	...	m_{2w}	m_2
Not recap.	$n_{1x}-m_{2x}$	$n_{1y}-m_{2y}$	...	$n_{1w}-m_{2w}$	n_1-m_2
No.releas.	n_{1x}	n_{1y}	...	n_{1w}	n_1

can be applied to each sample except the first using the pairs (m_{ix} , $M_{ix}-m_{ix}$) . These s-1 Chi-Squared are independent and can be added together to give a total Chi-Square.

(b) Catching and handling may affect catchability, so that marked and unmarked have different probabilities of capture in a given example. However, in some populations, once an animal has been caught its catchability remains fairly constant irrespective of future recaptures. In this case the ratio of k of the average probability of capture of an unmarked animal may remain approximately constant from sample to sample, so that Marten′s method can be used for testing k=1.

(c) If catching and handling affects the catchability of marked individuals after their first capture, then the sampling will not be random within the marked population. Such randomness can be detected using the following technique due to Leslie (1958) based on the frequency of recapture of individuals.

Suppose a multiple-recapture experiment consisting of t samples is carried out in a closed population containing an identifiable group of animals and let G denote both the group and the number in the group. If g_j of this group are caught in the jth sample (j=1,2,...,t), then, on the assumption of simple random sampling, the probability $P_j = 1 - Q_j$ that an individual member of G bears the recovery mark j is g_j/G . Suppose a particular member of G is caught x times, then from Kendall and Stwart (1969) we obtain

$$E[x|\{g_i\}] = \sum_{j=1}^{t} P_j = \mu \text{ , say.}$$

$$\mathrm{Var}[x|\{g_j\}] = \Sigma P_j Q_j = \Sigma P_j(1-P_j) = \Sigma P_j - \Sigma P_j^2 = \mu - \Sigma g_i^2/G^2$$

If f_x members of G are caught x times, then

$$\mu = \Sigma P_i = \Sigma g_j/G = \Sigma x f_x/\Sigma f_x = \bar{x}$$

and

$$T = \Sigma\, f_x(x-\mu)^2/\sigma^2$$

is approximately distributed as chi-squared with $G-1$ degrees of freedom when sampling is random. Leslie suggests that the approximation is satisfactory when $G > 20$ and $t \geq 3$. We note that any samples for which $g_i=0$ are ignored in the above analysis.

Applications

The above test can be applied to a Schnabel census of s samples. We first of all define $G = n_1(=u_1)$, the animals tagged in the first sample. In this case g_1, g_2, ... are the members of this group caught in the second and third samples, etcetera, and $t=s-1$. We can then apply this procedure to the newly tagged individuals in each sample, so that G successively represents u_2, u_3, ..., u_{s-3} with corresponding t values $s-2$, $s-3$, ..., s respectively. If there are accidental deaths through catching, and handling, then G refers to the members of u_i which are still alive at the end of the experiment. Since the Test Statistics thus obtained are based on different individuals they are independent and can be combined to give a total chi-square. In practice, G will often be greater than the degrees of freedom tabled, so that we must use the normal approximation

$$Z = \sqrt{2T} - \sqrt{(2G-3)}$$

which is approximately distributed as unit normal when sampling is random (Carothers (1971)).

Note that any group of identifiable animals can be used for G . In particular if the population size is known, we can put $G=N$, and H_o is now a test that sampling is random with respect to the whole population and not just the marked population.

One advantage of the above method is that it can be adapted to open populations in which there are natural death and recruitment. In this case the group G consists of the members of u_i known to be alive over a certain sequence of samples, say $i+1$, $i+2$,..., $i+t$, through having been caught after sample $i+t$. This method, however, does not apply of there is migration, as some marked animals may be out if sampling area for several sampling occasions.

Approximate Tests.

If $g_j = g$, say $(j = 1,2,\ldots,t)$, we have

$$P_j = \Sigma P_j/t = \bar{P} \ (=\bar{x}/t)$$

and x is the outcome of t binomial trials. The test for randomness is then a test that G values of x constitute a random sample of size G from a binomial distribution with parameters t and $\bar{P}$. In this case $\sigma^2 = t\bar{P}\bar{Q}$ and T reduces to the standard binomial Index of Dispersion

$$T' = \sum f_x (x-\bar{x})^2 / \bar{x}(1-\bar{x}/t)$$

which may be regarded as an approximation for T when the g_j are not too different. In fact

$$t\bar{P}\bar{Q} - \Sigma P_j Q_j = -\Sigma(P_i-\bar{P})(Q_i-\bar{Q}) = \Sigma(\bar{P}_i-P)^2 \geq 0$$

so that $T' \le T$ with equality only in the unlikely event of the g_j being equal. Therefore, if T_1 is significant then T' will be significant, and T_1 is a *conservative* approximation.

When P is small (< 0.05, say) we can use the Poisson approximation to the binomial with $\sigma^2 = \mu = t\bar{P}$. The statistic T then becomes the Poisson index of dispersion, $\sum f_x(x-\bar{x})^2/\bar{x}$ and since $\bar{P} \ge \overline{PQ}$ this statistic will be smaller than T' . The Poisson approximation is particularly relevant to the situation where the sampling is a continuous process and the animals are caught one at a time (that is for the $\Sigma x f_x$ samples in which a member of G is caught, we have $P_j = 1/G$).

Comparing Observed and Expected Frequency

We note that the statistic T is based on comparing the observed variance of x with the theoretical variance, calculated on the assumption of random sampling. In general these tests will be more sensitive than a goodness-of-fit test based on comparing the observed frequencies f_x with the expected frequencies (Cochran (1977)). However, if T is significant, a comparison of the observed and expected frequencies can be helpful in detecting where departures from random sampling occur. Unfortunately, using Leslie's method, the expected frequencies require length calculations, particularly for large values of x . For example

$$\Pr[x=0] = Q_1Q_2\ldots Q_z$$

$$\Pr[x=1] = Q_1Q_2\ldots Q_t\ \Sigma P_i/Q_i$$

$$\Pr[x=2] = Q_1Q_2\ldots Q_t \Sigma\ \Sigma P_iP_j\ /Q_iQ_j$$

using the binomial approximation, however, the expected frequencies are more readily calculated, namely

$$E_x = G\binom{t}{x}\bar{P}^{x}\,\bar{Q}^{t-x}$$

Models Based On Constant Probability of Capture

General Theory

Suppose that the trapping effort is the same for each sample, so that p_i , the probability of capture in the ith sample (i = 1,2,...,s) is constant (= p,say), then model (4.2) reduces to

$$f(\{a_w\}) = \frac{N!}{\prod_w a_w!(N-r)!}\, p^{\Sigma n_i}\, q^{sN-\Sigma n_i}$$

and $\hat{N}_p$, the maximum-likelihood estimate of N is now the unique root greater than r of the equation

$$\left(1 - r/N\right) = \left(1 - \Sigma n_i/sN\right)^s$$

Darroch (1958) shows that as $N \longrightarrow \infty$ the asymptotic variance of $\hat{N}_p$ is

$$\mathrm{Var}[\hat{N}_p] = N(1/q^s + s-1 - s/q)^{-1}$$

which may be compared with

$$\mathrm{Var}[\hat{N}] = N(1/\prod_{i=1}^{s} q_i + s-1 - \Sigma 1/q_i)^{-1}$$

the corresponding expression when the p_i are unequal.

Since $\mathrm{Var}[\hat{N}_p]$ follows from Var[N] by putting $p_i = p$, Darroch concludes that asymptotically, no information is gained by using the knowledge that p_i is constant. It is therefore recommended that the methods of section 4.1 be used irrespective of whether p_i is constant or not, except possibly for small samples.

Frequency of Capture

Several models based on the frequency of capture have been developed recently, and these have been used mainly for detecting any variation in trap response. For example, if p is constant, the probability that an animal is caught x times ($x = o,1,2,\ldots,5$) is given by the binomial probability

$$f(x) = \begin{pmatrix} s \\ x \end{pmatrix} p^{x} q^{s-x} \qquad (4.33)$$

If N is known, we can regard the animals as representing N independent observation from the above distribution and carry out a standard goodness-of-fit test to test for constant p. When p is small (< 0.05), the Poisson approximation to the binomial (above) can be used.

It should be noted that the above theory is not just a repetition of the approximate methods given previously. In the above theory, p_i is the probability of catching any individual in the ith sample, while P_i in the Leslie´s method is the conditional probability that an individual from an identifiable group is caught in the ith sample, given that at least one member of this group is caught in the sample.

Truncated Models: Constant Probability of Capture

Binomial Model.

When N is unknown we can test whether sampling israndom with respect to just the marked population by using Leslie´s method described above. When p_i is constant an alternative approach is to truncate the model (4.33) by ignoring the group of N-r animals not captured during the experiment. Thus x , the number of times an animal is captured given that it is captured at least once, has probability function

$$f(x) = \binom{s}{x} p^{x} q^{s-x}/(1-q^{s}) \quad ,x=1,2,\ldots,s$$

For this model the maximum-likelihood estimate $\hat{q}$ of q is the unique root of

$$h_s(q) = \left\{(1-q^{s})/(1-q)\right\} - s/\bar{x} = 1+q+\ldots+q^{s-1}-s/\bar{x}$$

where $\bar{x} = \Sigma x_i/r$ is now the mean number of captures per animal for the r animals actually captured. For $s > 3$,this equation can be solved using the Newton-Raphson method. The ith step is given by

$$q_{i+1} = q_i - h_s(q_i)/h_s'(q_i) \ ,$$

and a possible first approximation q_1 is given by the positive root of the quadratic equation $h_3(q) = 0$. Alternatively we can use the following technique given by Hartley (1958) for handling truncated distributions. Beginning with a first approximation N(1) of N , we carry out the chain iterations

$$p(i) = \frac{\bar{x}}{s} \cdot \frac{r}{N(i)}$$

and

$$N(i+1) = r/(1-q^{s}(i))$$

This procedure not only gives us $\hat{p}$ but, as a bonus, we also get $\hat{N}_p$, the solution of

$$\left(1 - \frac{r}{N}\right) = \left(1 - \frac{\Sigma n_i}{sN}\right)^{s}$$

This follows from the fact that $r\bar{x} = \Sigma n_i$. Once $\hat{p}$ is calculated, a standard goodness-of-fit test for the above truncated binomial model can be carried out.

Poisson

When p is small we can use the Poisson approximation to the binomial and consider the truncated distribution (Craig ,1953)

$$f(x) = \frac{e^{-\lambda}}{(1-e^{-\lambda})} \frac{\lambda^{x}}{x!} \tag{4.34}$$

David and Johnson (1952) show that the maximum-likelihood estimate $\hat{\lambda}$ of λ is the solution of

$$(1-e^{-\lambda})/\lambda = 1/\bar{x}$$

which can be solved by interpolating in Table A3 in the Appendix. One can then carry out a Chi-Squared goodness-of-fit test by comparing observed and expected frequencies in the usual manner.

David and Johnson (1952) suggest the alternative procedure of using the usual Poisson Dispersion Test, but with the class of zero recaptures left out, that is

$$T = \sum_{x=1}^{X} f_x (x-\bar{x})^2/\bar{x}$$

where

$$\bar{x} = \sum_{x=1}^{X} x f_x / \sum_{x=1}^{X} f_x$$

and X is the largest observed value of x . They show that treating T as Chi-Squared with r-1 degrees of freedom leads to a conservative test, for if T is significant then the Poisson Dispersion Test derived from complete data is also significant.

It should be noted that the distribution (4.43) should be truncated on the right at x=s as no more than s recaptures are possible. However, if s is sufficiently large for Pr(X ≤ s) to be almost 1, the effect of truncation on the data will be negligible.

An estimate of N, the population size, is given by

$$r/(1-\exp(-\hat{\lambda})) \quad \text{or} \quad r\bar{x}/\hat{\lambda}$$

Truncated Models: Allowing for Trap Response

Geometric Model

Eberhardt, Peterle and Schofield (1963) found that the capture frequencies for a rabbit population are well fitted by the geometric distribution:

$$f(x) = PQ^{x} \quad x = 0,1,2,\ldots, \ (\ 0 < P < 1,\ Q = 1-P)$$

and we now outline one of the two derivations that they give for this model.

Suppose that conditional on λ, the *average capture rate*, x has a Poisson distribution

$$f(x|\lambda) = e^{-\lambda}\lambda^{x}/x! \quad x = 0,1,2,\ldots$$

Then assuming a circular home range of radius R, we would expect the average capture rate to be proportional to the area of the home that is $\lambda = d\Pi R^{2}$, where d is a constant depending on such factors as the density of traps and the probability of recapture, given that there is a *contact* with one or more traps. Following Calhoun and Casby (1958), it is assumed that R has a density function

$$f_{1}(R) = \frac{R}{\sigma^{2}}\, e^{R^{2}/2\sigma^{2}}$$

so that if $C = 2d\Pi\sigma^{2}$,

$$f_{2}(\lambda) = c^{-1}e^{-\lambda/c} \quad , \quad \lambda \geq 0$$

Hence

$$f(x) = \int_{0}^{\infty} f(x|\lambda)f_{2}(\lambda)d\lambda = PQ^{x} \ .$$

where $P = 1/(1 + c)$.

When the size of the zero class is known, this distribution can be truncated at the origin as in the previous models. If trapping is carried out on s occasions then the distribution should also be truncated on the right, so that we have

$$f(x) = PQ^{x-1}/(1-Q^{s}), \; x = 1,2,\ldots,s. \qquad (4.35)$$

For a sample of r observations from this distribution, the maximum-likelihood estimate $\hat{P}$ is the unique solution of

$$\bar{x} = \frac{sQ^{s+1} - (s+1)Q^{s} + 1}{Q^{s+1} - Q^{s} - Q + 1}$$

which can be solved by interpolating linearly in Table A4 in the Appendix.

When s is large, the effect of truncation on the right is negligible and $\hat{P} = 1/\bar{x}$; or allowing for bias (Eberhardt,1969),

$$\hat{P} = (r-1)/(r\bar{x}-1)$$

which, from Chapman and Robson (1960), is the minimum variance unbiased estimate of P . In this case the total population size can be estimated by $\hat{N} = r/\hat{Q}$ (Edwards and Eberhardt ,1967). To determine when the truncation can be neglected, we enter $1/\bar{x}$ at the top of table A4 of the Appendix and in the nearest column we note when the entry becomes independent of s .

If f_x animals are caught x times, then, truncating the distribution of x on the right only,

$$E[f_x] = NPQ^{x}/(1-Q^{s+1}), \; x = 1,2,\ldots, s$$

and taking logarithms we have

$$E\left[\log f_x\right] \cong \log\left[NP/(1-Q^{s+1})\right] + x\log Q \;.$$

$$= \beta_0 + \beta x \text{ , say } \quad (x=1,2,\ldots,s)$$

which neglecting Q^{s+1} is the regression model suggested by Edwards and Eberhardt (1967). We can use least squares method to estimate N and P. One method of obtaining a confidence interval for N is to calculate the confidence limits for $NP/(1-Q^{s+1})$, the expected number of animals not caught, and add r to both limits.

Example: *Squirrel (Sciurus carolinensis, and so on):Nixon, Edwards and Eberhardt (1967)*

The aim of the study was to investigate the accuracy of Schnabel and Schumacher-Eschmeyer estimates of population size for squirrel populations, and consider the population application of the above geometric model. The study area occupied 237 acres of continuous forest habitat in 1250 acre Waterloo Wildlife Experiment Station, Athens County, Ohio. Both fox and grey squirrels occurred in the area, with the grey squirrels comprising about 95% of the squirrel population. The area was gridded on a 3x3 chain interval with a trap placed at the discretion of the trapper within a 1/5-acre plot surrounding each point of intersection. This yielded a trap density of about one (0.96) trap per acre.

Prebating was used for 10 days before the experiment and trapping was carried out for 11 consecutive days just before the hunting season. All squirrels captured were ear-tagged and released at their points of capture. squirrels killed by the hunters on the study area during the hunting season provided an estimate of proportion of tagged, from which N could be estimated using Petersen method.

Recapture data for the year (1962) are given in table (4.3) and a plot of $\bar{m}_i/n_i$ versus M_i is not linear, thus suggesting that the Schnabel estimate and its modifications will not give reliable estimate of N. This was borne out from a more detailed analysis by the authors, who felt that the Schnabel and Schumacher-Eschmeyer methods led to underestimation of population size.

Using the capture frequencies f_x as given in table (4.4) a goodness-of-fit test for the truncated geometric model

$$f(x) = PQ^{x-1}/(1-Q^s) \qquad x = 1,2,....$$

can be carried out. We have $1/\bar{x} = 72/223 = 0.323$ and from table A4 of the Appendix, we find that the truncation of the distribution at $s=11$ must be taken into account. Therefore, entering the table with $s=11$ and $\bar{x} = 3.097$ we find that $\hat{P} = 0.300$. The expected frequencies are then given by

$$E_x = 72\hat{P}\hat{Q}^{x-1}/(1-\hat{Q}^{11})$$

and

$$\sum (f_x-E_x)^2/E_x = 2.3 ,$$

which at 4 degrees of freedom indicates a close fit. The authors fitted a truncated Poisson which gave a very poor fit to the observed frequencies (χ^2_4=26.1)

They concluded that the probability of capture did not seem to be the same for all individuals and that the geometric model gave a reasonable fit to the observed frequencies. However, a good fit to this model provides no conclusive evidence that the assumptions used in deriving the model actually occur in nature: different sets of assumptions can give rise to the same model. For example, suppose that radio-tracking data indicate that the number of visits y of an animal to some small regular area around the trap follows a geometric distribution

$$f(y) = w(1-w)^{y} \qquad , \; y=1,2,\ldots$$

If x the number of captures, is conditionally binomial with parameters (y,p), then the unconditional distribution is once again geometric, namely

$$f(x) = w(1-w)^{x} \; , \quad x = 0,1,2,\ldots\ldots \qquad (4.36)$$

where, $w = \theta/\{\theta+(1-\theta)p\}$

Table 4.3: capture-recapture data from a Schnabel census: From Nixon et al (1967: Tables, 1963)

Trap day	Sample size	Marked		
i	n_i	m_i	M_i	m_i/n_i
1	38	—	—	—
2	29	19	38	0.66
3	31	23	48	0.74
4	16	13	56	0.81
5	20	19	59	0.95
6	18	17	60	0.94
7	17	14	61	0.82
8	19	13	64	0.68
9	16	14	70	0.88
10	14	14	72	1.00
11	5	5	72	1.00

Table 4.4: Observed capture frequencies fitted to zero-truncated geometric and Poisson distributions:data from Nixon et-al (1967)

No. Captrd	Obser. fre		Expected frequencies		
x	f_x	xf_x	Geometric $s=11$	Geometric $s=\infty$	Poisson $s=\infty$
1	23	23	22.0	23.3	11.9
2	14	28	15.4	15.8	17.4
3	9	27	10.8	10.7	17.2
4	6	24	7.6	7.2	12.6
5	8	40	5.3	4.9	7.4
6	7 } 12	42	10.9	10.1	5.5
7	3	21			
8	0	0			
9	2	18			
10	0	0			
11	0	0			
	72	223	72.0	72.0	72.0

Negative Binomial Model

The above derivation model (4.36) applies to the situation where a single trap is randomly located within a given animal's home range. If however the the traps are closer together, so that k traps fall within the home range, and assuming the traps act independently, then it is readily shown that the sum of k random variables independently sampled from the geometric distribution (4.36) has a negative binomial distribution (Eberhardt ,1969).

$$\frac{k(k+1)\ldots(k+x-1)w^k(1-w)^x}{x!} \quad ,x=1,2,\ldots \qquad (4.37)$$

When k=1 then distribution (4.37) reduces to (4.36). Alternatively this distribution can also be derived by assuming that the poisson with parameter λ , but with λ varying according to a Pearson type III distribution (Kendall and Stwart ,1969). In this case x has probability function

$$\frac{k(k+1)\ldots(k+x-1)}{x!}\ a^{x}(1+a)^{k+x} \qquad ,x=0,1,2.. \qquad (4.38)$$

By putting 1-w = a/(a+1), (4.38) reduces to (4.37). Thus whichever method of derivation is used, the distribution of x, truncated at x = 0, is given by

$$\frac{k(k+1)\ldots(k+x-1)}{x!}\ \frac{w^{k}(1-w)^{x}}{(1-w^{k})} \qquad ,x=1,2,\ldots \qquad (3.39)$$

Where k may not be an integer.

From Sampford (1955), the maximum likelihood estimates of w and k for a sample of r observations from this distribution are solutions of

$$\left\{rk/w(1-w^{k})\right\} - \left\{r\bar{x}/(1-w)\right\} = 0$$

and

$$\left\{r\log w\ /(1-w^{k}\right\} + \sum_{x=1}^{X}\left[1/k + 1/(k+1) + \ldots + 1/(k+x-1)\right]f_{x} = 0$$

Where x is the maximum observed value of x. These equations can be solved iteratively. However, a simpler method of obtaining estimates for k and w has been proposed by Brass (1958), and we describe this briefly below. Let

$$\Pi_{1} = \Pr[x=1] = kw^{k}(1-w)/(1-w^{k})\ ;$$

then if μ and σ^{2} are the mean and variance of x for distribution (4.39), Bras shows that

$$w = \mu(1=\pi)/\sigma^{2}$$

$$k = (w\mu-\pi_{1})/(1-w)$$

Therefore, replacing μ , σ^{2} and π_{1} by their sample estimates

$$\bar{x} = \sum_{x=1}^{X} xf_{x}/r$$

$$s^2 = \sum f_x(x-\bar{x})^2/(r-1)$$

and

$$\hat{\pi}_1 \; f_1/r$$

respectively, we have the simple estimates

$$\tilde{W} = \bar{x}(1-\hat{\pi}_1)/r$$

0 $$\tilde{k} = (\tilde{W}\bar{x}-\hat{\pi}_1)/(1-\tilde{W})$$

Brass gives the efficiency of the above estimation procedure as compared to the maximum-likelihood method for different values of k and m = k(i-w)w. For small M or large k, Brass's procedure is remarkably efficient.

Skellam's Model

Suppose that for a given animal the frequency of capture follows the binomial distribution with parameters s and p . In many experimental situations p may not be same for each animal but will vary according to some distribution $f_1(p)$, namely

$$f_1(p) = \frac{1}{B(\alpha,\beta)}\, p^{\alpha-1}(1-p)^{\beta-1} , \quad 0 \le p \le 1$$

Where

$$B(\alpha,\beta) = \Gamma\alpha \, \Gamma\beta \, / \, \Gamma\alpha+\beta$$

Accordingly Skellam (1948) has shown that

$$f(x) = \int_0^1 f(x|p).f_1(p)dp = \binom{s}{x} \frac{B(\alpha+x,\beta+s-x)}{B(\alpha,\beta)}$$

$$x=0,1,\ldots,s \qquad (4.40)$$

If $s \longrightarrow \infty$, $\beta \longrightarrow \infty$, $\beta/s \longrightarrow c$ and $P \longrightarrow 0$ in such away that sp remains finite, then Skellam showed that sp tends to Pearson Type III distribution

$$f(x) = \frac{1}{\Gamma\lambda} \; x^{\lambda-1} e^{-x} \qquad , \lambda > 0 \; , \quad 0 \leq x \leq \infty$$

as given by Kendall and Stwart (1969), and the limit of f(x) is the negative binomial distribution (4.38), with $k = \alpha$, $a = 1/c$.

The truncated version of the distribution (4.40) which is appropriate when zero class is not observed is given by

$$\binom{s}{x} \frac{B(\alpha+x,\beta+s-x)}{B(\alpha,\beta)-B(\alpha,\beta+s)} \quad ,x=1,2,\ldots,s$$

Unfortunately estimation for this distribution is not easy, which rather produces its use in practice. However, if n_i is the size of the ith sample (i = 1,2), m_2 the number of recaptures in the second sample, and $\hat{N}=n_1 n_2/m_2$, then

$$E[n_i] = NE[p] = N \; \alpha/(\alpha+\beta) \qquad (4.41)$$

$$E[m_2] = NE[p^2] = N\alpha(\alpha+1)/(\alpha+1)(\alpha+\beta+1) \qquad (4.42)$$

and asymptotically

$E[\hat{N}] = E[n_1] \; E[n_2]/E[m_2] = N\alpha(\alpha+\beta+1)/(\alpha+1)(\alpha+\beta) = NB$, say.

Where B is tabulated in Table 4.5 below

Table 4.5: *B = α(α+β)/(α+1)(α+β) for selected values of α and β*

β \ α	1	2	3	5	10	∞
1	0.75	0.89	0.94	0.97	0.99	1.00
2	0.67	0.83	0.90	0.95	0.98	1.00
3	0.63	0.80	0.88	0.94	0.98	1.00
5	0.58	0.76	0.84	0.92	0.97	1.00
10	0.55	0.72	0.81	0.89	0.94	1.00
∞	0.50	0.67	0.75	0.83	0.91	1.00

Note that B is small when a large proportion of the population has a low probability of capture that is α is small.

For the special case $\alpha = 1$ then the limiting distribution of distribution (4.40) is geometric rather than negative-binomial and Eberhardt (1969) uses this special case to derive a new estimate of N when $s = 2$. Thus setting $\alpha = 1$ in equations (4.41) and (4.42), we have

$$E[n_1+n_2] = 2N/(\beta+1)$$
$$E[m_2] = 2N/(\beta+1)(\beta+2)$$

and solving we have moment estimates

$$\hat{\beta}+2 = (n_1+n_2)/m_2$$
$$\hat{N}_\beta = (n_1+n_2)(\hat{\beta}+1)/2 = (n_1+n_2)(n_1+n_2+m_2)/2m_2$$

If in fact p is actually constant, then N is asymptotically unbiased (since $B = 1$) and

$$E[\hat{N}_\beta] = E[n_1+n_2].E[n_1+n_2-m_2]/E[2m_2]$$
$$= 2Np(2Np-Np^2)/2Np^2 = N(2-p)$$

which lies between N and $2N$.

To find the asymptotic variance of $\hat{N}_\beta$, let y_i be the number of animals caught i times, that is $y_1=n_1+n_2-2m_2$ and $y_2=m_2$. Then the joint distribution of y_1 and y_2 is multinomial, namely

$$f(y_1,y_2) = \frac{}{y_1!y_2\,!}\,\frac{N!}{(N-y_1-y_2)!}\; p_1^{y_1}\; p_2^{y_2}\; p_3^{N-y_1-y_1}$$

$$= \frac{N!}{y_1!y_2!(N-y_1-y_2)!}\left[\frac{2\beta}{(\beta+1)(\beta+2)}\right]^{y_1}\left[\frac{2}{(\beta+1)(\beta+2)}\right]^{y_2}$$

$$\times\left[\frac{\beta}{\beta+2}\right]^{N-y_1-y_2}$$

Since from equation (4.40) with $s = 2$, $P_i=Pr[x=1]=f_1(i)$. Then the maximum likelihood estimate of N once again is

$\hat{N}_\beta$ which takes the form

$$\hat{N}_\beta \ (y_1+2y_2)(y_1+y_2)/2y_2 \ .$$

Hence using the delta method, we find, after some algebra, that

$$\mathrm{Var}[\hat{N}_\beta] \cong N\beta(\beta^5+7\beta^4+20\beta^3+29\beta^2+21\beta+6)/2(\beta+1)^2(\beta+2)^2$$

Models Based on Waiting Times Between Captures.

Time for First Recapture

The probability that an animal is caught for the first time in the yth sample, given that it is caught at least once in s samples, is given by the function

$$f(y) \ q^{y-1}p/(1-q^s) \quad ,y=1,2,\ldots,s$$

Where p is the probability of capture in a sample. However, a slightly different model has been suggested by Young, Nees and Emlen (1952) which can still be used when there is migration and mortality.

Suppose an animal is captured for the second time in sample number $y+z$ ($z = 1,2,\ldots s-y$). Then, given y and given that an animal is caught at least twice, z has the probability function

$$f(z|y) = q^{z-1}p/(1-q^{s-y}) \quad ,z=1,2,\ldots$$

If we consider only those animals for which $s-y$ is large, then the truncation of $Z = s-y$ can be neglected. We are led to consider

$$f(z) = q^{z-1}p$$

This model has the simple maximum-likelihood estimate

$p = 1/z$. Young *et al* (1952) points out that once an animal has been recaptured, we are not interested in its subsequent fate, so that we do not need to "correct" the data for those dying or disappearing before the end of the experiment. Also if the tendency to die or emigrate is not related to trap vulnerability, then those animals which die or emigrate before being captured at all will be distributed randomly over the groups that would have been recaptured after 1,2,... samples, and the disappearance of such animals will therefore not bias $\hat{p}$ and associated goodness-of-fit test.

Time of Residence

Suppose that the population under study is such that animals move into the population area, stay for a random number of time, and move out and stay out of the area for the remainder of the investigation. If the experiment is carried out at equally spaced intervals of time with constant trapping effort (that is p constant), then it is not unreasonable to assume that θ , the probability that an animal does not leave the trapping area sometime between two successive trappings, is the same for all animals in the area and for all successive pairs of trappings. On the basis of these assumptions, Holgate (1964b) gives the following method for estimating q and p from the observed values of y , the recorded period of residence, that is, the interval between the first and last occasions when it is actually captured.

Let z denote the true period of residence of an individual in the study area, that is the interval between the first and last occasions when it is exposed to capture; then (ignoring truncation on the right)

$$\Pr[Z=z] = (1-\theta)\theta^{z} \qquad , z=0,1,2,... \qquad (4.43)$$

Now, an animal that remains in the area for z complete intervals is exposed to capture on $z+1$ occasions, so that the probability of its not being caught at all (that is y undefined) is

$$\Pr[Y \text{ undefined} | Z=z] = q^{z+1}$$

Also

$$\Pr[Y=0 \; | Z=z] = (z+1)pq^{z}$$

and noting that as far as y is concerned, it does not matter how often an animal is captured between its first and last capture

$$\Pr[Y=y | Z=z] = (z-y+1)p^{2}q^{z-y} \qquad y=1,2,\ldots,z$$

Now

$$\Pr[Y \text{ undefined}] = \sum_{z=0}^{\infty} \Pr[Y \text{ undefined} | Z=z]\Pr[Z=z] \qquad (4.44)$$

$$= (1-\theta)q/(1-q\theta)$$

and in a similar fashion it is readily shown that

$$\Pr[Y=0] = (1-\theta)p/(1-q\theta)^{2} \qquad (4.45)$$

and

$$\Pr[Y=y] = p^{2}(1-\theta)\theta^{y}/(1-q\theta)^{2} \qquad y=1,2,\ldots \qquad (4.46)$$

Finally dividing both sides of equations (4.45) and (4.46) by $\{1- (1-\theta)q/(1-q\theta)\}$ we have the zero modified geometric distribution

$$\Pr[Y=0 | Y \text{ defined}] = (1-\theta)/(1-q\theta) \qquad (4.47)$$

and

$$\Pr[Y=y | Y \text{ defined}] = (1-q)(1-\theta)\theta^{y}/(1-q\theta) \qquad y=1,2,\ldots \qquad (4.48)$$

If a sample of r observations is taken from the above distribution, then the maximum-likelihood estimates of θ and q are

$$\hat{\theta} = 1 - u/\overline{y} \qquad \text{and} \qquad \hat{q} = (\hat{\theta}-u)/\hat{\theta}(1-u)$$

Where u is the proportion of individuals caught more than once (that is with $y > 0$). Holgate shows that as

$r \longrightarrow \infty$ asymptotic variances and covariances of these estimates are given by

$$\mathrm{Var}[\hat{\theta}|r] = (1-\theta)^2(1-q\theta)/r(1-q) \ ,$$

$$\mathrm{Var}[\hat{q}|r] = (1-q)(1-q\theta^2)(1-q\theta)/r\theta^2(1-\theta)$$

and

$$\mathrm{Cov}[\hat{\theta},\hat{q}|r] = (1-\theta)(1-q\theta)/r\theta$$

It is noted that strictly speaking, distribution (4.43) should be truncated on the right since the number of trappings is finite. Otherwise, at the end of trapping series, the animals still in the area will be ascribed a duration of residence which is too short. Unfortunately, since the time of the first capture varies, each animal will have a different truncation point thus leading to a complicated likelihood function for the estimate of θ and q. However, if the study is long compared with the average time of residence, this effect will be negligible and the truncation can be ignored.

Bivariate Distribution

Let y be the recorded period of residence and w be the number of captures during the intervening period, that is between the first and last capture. Holgate (1966) has utilised the joint distribution of Y and W to obtain more efficient estimates of q and θ as follows (in Holgate´s notation $y=x$, $w=y$).

Since $Y=y$ implies $y+1$ possible captures, the range of W is 0 to $y-1$ and

$$\Pr[W=w|Y=y] = \binom{y-1}{w}(1-q)^w q^{y-w-1} \qquad (y > 0)$$

Hence from equations (4.44) and (4.45) the joint function is given by

$$\Pr[Y=0|Y \text{ defined}] = (1-\theta)/(1-q\theta)$$

and

$$\Pr[Y=y,W=w|Y \text{ defined}] = \binom{y-1}{w}(1-q)^{w+1}q^{y-w-1}(1-\theta)\theta^{y}/(1-q\theta) \ .$$

Where w = 0,1,2,..., y-1 and y = 1,2,... (for Y =0, W is not defined). Let n_o and n_{yw} denote the corresponding sample frequencies and let r be the total number in the sample that is

$$r = n_o + \sum_y \sum_w n_{yw}$$

Then the likelihood function for the sample is proportional to

$$\left(\frac{1-\theta}{1-q\theta}\right)^{n_o} \prod_{y,w} \left\{\frac{(1-q)^{w+1}\ q^{y-w-1}(1-\theta)\theta^{y}}{(1-q\theta)}\right\}^{n_{yw}}$$

and it follows that the maximum-likelihood estimates $\tilde{q}$ and $\tilde{\theta}$ are given by

$$\tilde{q} = (\bar{y}-\bar{w}-u)(1+\bar{w}+u)/\{\bar{y}-(\bar{w}+u)(1-\bar{y})\}$$

and

$$\tilde{\theta} = \{\bar{y}-(\bar{w}+u)(1-\bar{y})\}/\bar{y}(1+\bar{w}+u)$$

where

$$\bar{w} = \sum_y \sum_w wn_{yw}/r = \sum_w wn_{.w}/r \ \text{,say}$$

$$\bar{y} = \sum_y \sum_w yn_{yw}\ /r = \Sigma yn_{y.}/r \ \text{,say}$$

and

$$u = 1 - n_o/r$$

Holgate shows that asymptotic variance covariance matrix of $\tilde{q}$ and $\tilde{\theta}$ is given by

$$\frac{1}{r}\begin{bmatrix} \dfrac{q(1-q)(1-q\theta^2)(1-\theta)}{\theta^2(1-q\theta)} & \dfrac{q(1-\theta)^3}{\theta(1-q\theta)} \\ \\ \dfrac{q(1-\theta)^3}{\theta(1-q\theta)} & \dfrac{q+\theta+q^2\theta^2-3q\theta)(1-\theta)^2}{(1-q)(1-q\theta)} \end{bmatrix}$$

Where the (1,1) element is the asymptotic variance of $\tilde{q}$.

Using the determinant of the variance-covariance matrix as a measure of asymptotic efficiency of the above method with respect to the previous method based on the marginal distribution of Y only ,the asymptotic efficiency is

$$e = q(1-\theta)^2/(1-q\theta)^2 \ .$$

4.2 INVERSE MULTIPLE SAMPLE CENSUS

Consider a Schnabel census in which, for each sample n_i , the sampling is continued until a predetermined number of animals are recaptured. This modification is a generalization of the simple inverse sampling census already discussed in section 3.4.

Using the same notations as in section 4.1 , we have fixed parameters N, s, $n_1(=M_2)$, m_2 , m_3 , ..., m_s , random variables M_3 , M_4 ,..., M_s . r_1 , n_2 ,..., n_s and the joint probability function of the random variables is a straight forward generalisation of the negative hypergeometric distribution derived for simple inverse sampling case, namely

$$f(n_2|n_1,n_2) = \frac{\binom{n_1}{m_2-1}\binom{N-n_1}{n_2-m_2}}{\binom{N}{n_2-1}} \cdot \frac{n_1-m_2+1}{N-n_2+1} \quad ,$$

Where $n_2 = m_2, m_2+1, \ldots, N + m_2 - n_1$.

Thus the joint probability function of the random variables is

$$\prod_{i=1}^{s} \left\{ \frac{\binom{M_i}{m_i-1}\binom{N-M_i}{u_i}}{\binom{N}{n_i-1}} \cdot \frac{M_i-m_i+1}{N-n_i+1} \right\}$$

There is now no no-trivial sufficient statistics for N. However an unbiased estimate is easily found, namely

$$\bar{N}_2 = \sum_{i=2}^{s} \left(\frac{n_i(M_i+1)}{m_i} - 1 \right) / (s-1)$$

for

$$E\left(\bar{N}_2\right) = \sum_{i=2}^{s} E\left\{E\left(\frac{n_i(M_i+1)}{m_i} - 1 \mid n_i, m_i\right)\right\} / (s-1)$$

$$= \sum_{i=2}^{s} E[N]/(s-1) = N$$

Using the approximate formula for the variance of $\bar{N}_2$ in single sample case, we find that, the variance of $\bar{N}_2$ is given by

$$\mathrm{Var}\left(\bar{N}_2\right) = \frac{N^2}{(s-1)} \sum_{i=2}^{S} 1/m_i$$

Thus the coefficient of variation is

$$C(\bar{N}_2) = \frac{1}{(S-1)} \sum_{i=2}^{S} 1/m_i$$

This can be used for choosing the fixed parameters to give a predetermined precision. However, the correct choice of m_i is also important for it was earlier pointed out that a wrong choice of m_i , coupled with unfavourable M_i , could give rise to a large n_i. Therefore a reasonable criterion, suggested by Chapman (1952), for choosing these parameters is to minimize $E\left(\sum n_i\right)$, subject to $C(\bar{N}_2)$ being held constant; unfortunately this does not have a simple solution.

To avoid the possibility of large n_i we can modify the above model and continue the sampling until predetermined number u_i of unmarked individuals is taken in the ith sample (i = 2,3,..., s). This means that our fixed parameters are now N, s, n_1, u_2, ..., u_s, M_3, ..., M_s, r and the random variables n_2, n_3,...n_s, m_2, m_3, ..., m_s. An approximately unbiased estimate of N is the mean

$$\bar{N}_3 = \sum_{i=2}^{s} \hat{N}_{3i}/(s-1)$$

Where,

$$N_{3i} = \frac{n_i(M_i+1)}{m_i} - 1$$

and asymptotically,

$$\text{Var}(\bar{N}_3) = \sum \text{Var}\left[\hat{N}_{3i}\right]/(s-1)^2$$

Where $\text{Var}\left[\hat{N}_{3i}\right]$ can be estimated using equation

$$\text{Var}(\bar{x}) = \Sigma (x_i-\bar{x})^2/n(n-1)$$

and we get

$$\text{Var}\left[\bar{N}_3\right] = \sum_{i=2}^{s}\left[\hat{N}_{3i}-\bar{N}_3\right]/(s-1)(s-2) \quad .$$

4.3 THE MULTI-SAMPLE SINGLE RECAPTURE CENSUS.

The main advantage of this method is that individualsare only captured once and are then removed from the population and is thus specially suited to commercially exploited population such as fisheries where they are permanently removed.

The technique used for this single recapture census is as follows: the experimenter using differentiated marking, releases batches of marked individuals of sizes R_1, R_2, ... into the population he is investigating, and after each batch R_i is released, a commercial catch of size n_i is made and individuals are killed, thus giving a sequence R_1 added, n_1 removed, R_2 added, n_2 removed, and so on. The numbers of marked individuals from the different R_i and unmarked individuals are noted for each catch n_i and passed on to the experimenter.

Ideally, the marked individuals which are to be released, should either be caught before the whole experiment or perhaps taken from a similar population not connected with the one under investigation. In actual practice, however, the experimenter could take the samples R_i from the population during the experiment, because in general, the R_i , although large, will be much smaller than the n_i and therefore the recaptures in the sequence R_i , R_2 , ... will be negligible. Also the overall reduction in the number of unmarked individuals due to marking of $\sum R_i$ will be small compared with the total population size. Let

N = initial population size,

M_j = number marked in n_j $(j = 1,2,..., s)$,

m_j = The number of individuals from R_i caught in n_j $(j=i,i+1,...,s)$

$u_j = n_j - m_j$

$$r = \sum_{j=1}^{s} u_j$$

We shall make the following assumptions:

(1) all marked individuals present in the population at any instant have the same probability of being caught.

(2) unmarked individuals have the same probability of being caught as the marked ones.

Let $p_i = 1 - q_i$ be the probability that an individual is caught in the sample n_i, given that it is in the population at the time, the ith sample is taken. If $\{n_i\}$ are considered as stochastic variables and the $\{R_i\}$ are known constants, then the joint probability distribution of $\{m_j, u_j\}$ is the product of $s = 1$ independent multinomial distributions, namely,

$$P\left(\{\underline{m}_{ij}, \underline{u}_{ij}\} \mid \{\underline{R}_i\}, n\right) = \prod_{i=1}^{s} \left\{ \frac{a_i!}{\prod_{j=1}^{s} m_{ij}!\,(R_j - \sum_{j=1}^{i-1} m_{ij})} \; p_i^{m_{ii}} \right.$$

$$\left. \times (q_i p_{i+1})^{m_{i,i+1}} \ldots (q_i \ldots q_s)^{R_i - \sum_{j=1}^{s} m_{ij}} \right\}$$

$$\times \frac{N!}{\prod_{i=1}^{s} u_i!\,(N-r)!} \, p_1^{u_1} (q_1 p_2)^{u_2} \ldots (q_1 \ldots q_s p_s)^{u_s} (q_1 \ldots q_s)^{N-r}$$

If $\hat{N}$ and $\hat{p}$ are the maximum-likelihood estimates of N and p_i respectively, then we find that

$$\hat{p}_i = \frac{n_i}{\hat{N} + \sum_{j=1}^{i} R_j - \sum_{j=1}^{i-1} n_j} \qquad (i=1,2,\ldots s) \qquad (4.49)$$

and $\hat{N}$ is the root of

$$\left(1-\frac{r}{\hat{N}}\right)= \hat{q}_1\ldots\hat{q}_s=\prod_{i=1}^{s}\left\{1-\frac{n_i}{\hat{N}+\sum_{j=1}^{i}R_j-\sum_{j=1}^{i-1}n_j}\right\} \qquad (4.50)$$

We see intuitively that these equations give reasonable estimates and may be readily explained if one considers the proportion dealt with and the way in which they are multiplied together. A natural estimate of p_i is given by the proportion of existing population caught in the ith sample and this is given by equation (4.48). From the assumption that unmarked individuals have the same probability of being caught as the marked ones, equation (4.49) a rise naturally by equating the total proportion of unmarked individuals which are not captured with the overall proportion of uncaptured individuals.

To find $\hat{N}$ we need to solve a polynomial of degree s-1 ; however, we now show that for a successful experiment that is one in which there is at least one recapture (r=0), equation (4.49) will have a unique finite root greater than r . If

(a) r=0 , that is no marked individuals are caught, then equation (4.49) is not valid. However, we see that$P[\{m_{ij}, u_j\}| \{R_i\}, n]$ is maximized when $\hat{N}=0$, as we would expect intuitively.

(b) If $r = \sum_{j=1}^{s} n_i$ then no marked individuals are caught and $\hat{N} = \infty$.

(c) If $r = \sum_{j=1}^{i} n_j - \sum_{j=1}^{i} R_j$, for i = 1,2,...s, then there are no marked individuals left immediately after sample n_i is taken, and the samples n_{i+1} , ... , n_s consists solely of marked individuals. For the limit process $N \longrightarrow \infty$ and each $R_i \longrightarrow \infty$ in such a way that R_i/N is constant, it can be shown that the probability of (a), (b) or (c) occurring is $o(c'')$ where $0 < c < 1$ and therefore can be neglected. Thus we may assume that

$$0 < r < \sum_{j=1}^{s} n_j \quad \text{and} \quad r \neq \sum_{j=1}^{i} n_j - \sum_{j=1}^{i} R_j \quad \text{for } i=1,\dots,s$$

Let

$$\phi(N) = \frac{N}{N-r} \prod_{i=1}^{s} \left\{N+\sum_{j=1}^{i} R_j - \sum_{j=1}^{i} n_j\right\} \Big/ \left\{N+\sum_{j=1}^{i} R_j - \sum_{j=1}^{i-1} n_j\right\}$$

Then

$$\phi(r+0) = \infty$$

Then as $n \longrightarrow \infty$

$$\phi(N) = 1 - \frac{\sum_{j=1}^{s} n_j - r}{N} + o\left(\frac{1}{N^2}\right) \longrightarrow 1-0$$

Now, $\phi(N)$ is continuous for $N > r$ and therefore the equation $\phi(N) = 1$ has at least one finite root greater than r. Denote any such root by $\hat{N}$. It can be shown by induction that $[d\phi/dN]_{N=\hat{N}} < 0$ for each $\hat{N}$ and since $\phi(N)$ is one valued and $d\phi/dN$ is continuous for $N > r$, we see that $\hat{N}$ is unique. Now

$$L[N] = \log P\left(\{m_{ij}, n_j\} \,|\, \{R_i\}, n_j\right)$$

may be regarded as a function of N, and since $\Delta L(N)=\phi(N)$ is decreasing at $N=\hat{N}$ we see that $\hat{N}$ maximizes the likelihood.

The mean and the variance of $\hat{N}$ can be found by the δ-technique (Darroch ,1958).

Let $r = n_{s+1}$ and $E[n_i]=\beta_i$ $(i=1,2,..,s+1)$, then $\hat{N}=\hat{N}(n_1,n_2,\dots,n_{s+1})$ and $\hat{N}(\beta_1,\beta_2,\dots,\beta_{s+1}) = N$. Thus expanding $\hat{N}$ as a Taylor's series about N, $E[\hat{N}-N] = o(1)$ Hence $\hat{N}$ is approximately an estimate of N and

$$\mathrm{Var}[\hat{N}] = E[\hat{N}-N]^2 + o(1)$$

$$= \sum_{i=1}^{S+1} \mathrm{Var}(n_i)\left(\frac{\partial\hat{N}}{\partial n_i}\right)^2 + 2\sum_{1<r<t<s+1}\sum \left\{\mathrm{Cov}(n_r, n_t)\frac{\partial\hat{N}}{\partial n_r}\frac{\partial\hat{N}}{\partial n_t}\right\} + o(1)$$

Where all the derivatives are evaluated at $E\{b_i\} = \{\beta_i\}$.

Special Case

Now regarding $\{n_j\}$ as fixed parameters and assuming that each sample is a simple random sample, the joint probability function $\{m_{ij}, u_j\}$ is the multi-hypergeometric distribution, namely

$$f(\{m_{ij}, u_j\}|\{R_i, n_j\}) = \prod_{i=1}^{s}\left\{\frac{\begin{pmatrix} R_1 - \sum_{j=1}^{i-1} m_{1j} \\ m_{1i}\end{pmatrix} \cdots \begin{pmatrix} R_i \\ m_{ii}\end{pmatrix}\begin{pmatrix} N - \sum_{j=1}^{i-1} m_{1j} \\ u_i\end{pmatrix}}{\begin{pmatrix} N + \sum_{j=1}^{j} R_j - \sum_{j=1}^{i-1} n_j \\ n_i\end{pmatrix}}\right\}$$

setting $\Delta \log f = 0$, the maximum -likelihood estimate $\hat{N}$ of N is the unique root greater than r, of $h(N) = 0$, Where

$$h(N) = N - r - N\prod_{i=1}^{s}\left\{1 - \frac{n_i}{\hat{N} + \sum_{j=1}^{i} R_j - \sum_{j=1}^{i-1} n_j}\right\}$$

For $0 < r < \sum_{j=1}^{s} n_j$ and $r \neq \sum_{j=1}^{i} n_j - \sum_{j=1}^{i} R_j$ for $i=1,2,\ldots,s$. We have that

$$\hat{N}-N = (r-\rho)\left(\frac{d\hat{N}}{dr}\right)_{\rho} + \frac{1}{2}(r-\rho)^2\left(\frac{d^2\hat{N}}{dr^2}\right)_{r=r'}$$

Where $\rho = E\,[r|\{n_i\}]$ and r' lies between r and ρ.

Therefore

$$E[\hat{N}-N \;|\{n_i\}] = \tfrac{1}{2}\, E\left[(r-\rho)^2 \frac{d^2\hat{N}}{dr^2} \;|\{n_i\}\right]$$

$$= b \;,\text{say.}$$

Thus asymptotically $\hat{N}$ is unbiased estimate for N . The asymptotic variance of $\hat{N}$ is estimated by

$$\mathrm{Var}[\hat{N}] = (\hat{N}-r)/h'(\hat{N})$$

$$= \frac{1}{\hat{N}(\hat{N}-r)} - \sum_{i=1}^{s} \frac{n_i}{(N+\sum_{j=1}^{i} R_j - \sum_{j=1}^{i} n_j)(\hat{N}-\sum_{j=1}^{i} R_j - \sum_{j=1}^{i-1} n_j)}$$

Now , $h(N)=0$ can be solved using Robson and Regier's iterative method, and once again the last iteration provides an approximation for $\mathrm{Var}[\hat{N}]$. By setting

$$D_i = \sum_{j=1}^{i} n_j - \sum_{j=1}^{i} R_j$$

a first approximation to $\hat{N}$ is given by

$$N(1) = \frac{\left[\Sigma(n_i+D_i)\right]^2 - \Sigma(n_i-D_i)^2 - (r+\Sigma D_i)^2 + r^2 + \Sigma D_i^2}{2\Sigma m_i} \;.$$

For the more realistic situation where the mark releases R_i are obtained from the population during the course of the experiment, we have

$$f(\{m_{ij},u_j\}|\{R_i,n_j\}) = \prod_{i=1}^{s} \left\{ \frac{\binom{R_1-\sum_{j=1}^{i-1} m_{1j}}{m_{1i}} \cdots \binom{R_i}{m_{ii}} \binom{N-\Sigma R_j \;\; \sum_{j=1}^{i-1} u_{1j}}{u_i}}{\binom{N \;\; -\sum_{j=1}^{i-1} n_j}{n_i}} \right\}$$

In this case R_i refers to the newly marked individuals released; any recaptures are not tagged but simply returned to the population. The maximum likelihood estimate $\tilde{N}$ is the largest root of

$$\prod_{i=1}^{s}\left[\left\{N-\sum_{j=1}^{i}R_j-\sum_{j=1}^{i}n_j\right\} \Big/ \left\{N-\sum_{j=1}^{i}R_j-\sum_{j=1}^{i-1}n_j\right\}\right]=\prod_{i=1}^{s}\left\{\frac{N-\sum_{j=1}^{i}n_j}{N-\sum_{j=1}^{i-1}n_j}\right\}$$

We note that $\hat{N}$ and $\mathrm{Var}[\hat{N}]$ are the same, irrespective ofwhether the n_i are regarded as fixed parameters or random variables. However the model arising from the later case can be used to provide a goodness-of-fit test to test the hypothesis

H_o: that all marked individuals have the same probability p_i of being caught in the ith sample

Let p_{ij} be the probability that a member of R_i is caught in n_j, then H_o is the hypothesis that

$$p_{ij} = q_i \ldots q_{j-1} p_j$$

$(j > i)$and $p_{ii} = p_i$. This can be tested using the statistic suggested by Mitra (1958):

$$T = \sum_{i=1}^{s}\sum_{j=i}^{s+1}\left(m_{ij}-R_i\hat{p}_{ij}\right)^2/R_i\hat{p}_{ij}$$

Where $\hat{p}_{ij}=\hat{q}_i\ldots\hat{q}_{j-1}\hat{p}_j$, $\hat{p}_{i,s+1}=q_iq_{i+1}\ldots q_s$

$$p_i = m_i/\left(\sum_{j=1}^{i}R_j-\sum_{j=1}^{i-1}m_j\right) , \quad m_{i,s+1}= R_i-r_i$$

When H_o is true and R_i are large, T is approximately distributed as Chi-Squared with $\{s(s+1)\}/2 - s$ degrees of freedom.

4.4 CONCLUSIONS

A wide variety of models have been discussed in this dissertation, and the question now arises as to which method should be used in a given situation. Obviously the choice of the method will depend very much on the nature of the population, its distribution over the population area, and the method of sampling the population. Where possible, the experiment should be designed so that more than one method of estimation can be used.

In the past, little attention has been devoted to the problem of designing an experiment to yield an estimate with a given minimum accuracy or precision. For example, in many of the early applications of the Petersen method, too few individuals were tagged, so that the number of recaptures was too small and resulting confidence intervals too wide. However, for many of the models the variance formulae are complicated, so that it is not easy to plan for a given precision. Clearly more research is needed on the question of design for some methods.

Where possible, the robust but less efficient regression estimates should be calculated along with more efficient maximum likelihood estimates. The regression method is particularly useful when expected values appear to be correct; but the variances predicted by the model underlying the maximum-likelihood theory are open to question because of departures from the underlying assumptions, for example sampling is not strictly random. However, in all cases, the assumptions underlying a particular model should be studied carefully,and where possible appropriate tests carried out. If there is likely to be any question about the validity of the underlying assumptions, the sample data should be collected in such a way that emperical variance estimates are available from replicated samples. A comparison of the sampling variance with the estimated theoretical variance predicted by the model will often throw some light on the validity of the model.

Of the methods considered in this dissertation the Petersen method appears to be useful, provided the assumptions underlying the method are satisfied and there are sufficient recaptures in the second sample. The main assumption underlying the Petersen estimate is that marked and unmarked animals have the same probability of being caught in the second sample. Unfortunately it is not always easy to detect departures from this assumption so that even when all precautions are taken and the assumption appear to be satisfied, the Petersen estimate may appear to be biased. If Petersen estimate is to be used extensively for a given species, then it should be compared with other estimates and where possible tested on a known population.

If the second sample can be taken in stages then regression method can be used for testing the assumption that marked and unmarked animals have the same probability of capture. The problem of variable catchability may be overcome by prebaiting, using different sampling methods for the two samples, changing trap positions etcetera. However, the most promising approach to the problem is to avoid recapturing altogether and to obtain an estimate of the proportion of the marked by simply observing the animals. Tagged animals may also be detected using remote sensing techniques. If sight records are used then the second sample is obtained by sampling with replacement and the binomial model (Bailey ,1952) is applicable.

Sometimes it is not possible to catch enough animals on the first occasion for a satisfactory application of Petersen estimate, so that Schnabel method must be used. In any case the later method should be used if variable is a problem. Moreover, the schnabel estimate is more efficient than the Petersen estimate, though the difference is not great.

It had initially been planned to apply this procedure to real life data especially on fisheries but my efforts were fruitless. This was because, most of the models in literature require well monitored experiments, but the type of data I got from Kenya Marine and Fisheries Research Institute (K.M.F.R.I.) lacked this quality. My future goals are therefore to carry out extensive survey on fisheries based on well monitored experiments and give applications of some of the models discussed here. I also hope to extend my work to open populations.

APPENDIX

A1 Shortest 95% confidence interval for N/λ based on the Poisson distribution

Entering variable m_2 (or m)

m_2	Lower limit	Upper limit	m_2	Lower limit	Upper limit
0	0·088 5				
1	0·072 0	19·489	26	0·024 78	0·056 3
2	0·076 7	2·821	27	0·024 08	0·053 9
3	0·073 6	1·230	28	0·023 42	0·051 6
4	0·069 0	0·738	29	0·022 79	0·049 5
5	0·064 4	0·513	30	0·022 21	0·047 5
6	0·060 0	0·388	31	0·021 65	0·045 7
7	0·056 1	0·309	32	0·021 12	0·044 0
8	0·052 6	0·256	33	0·020 61	0·042 5
9	0·049 5	0·217	34	0·020 14	0·041 0
10	0·046 8	0·188	35	0·019 68	0·039 6
11	0·044 3	0·165	36	0·019 25	0·038 4
12	0·042 0	0·147	37	0·018 83	0·037 2
13	0·040 0	0·133	38	0·018 43	0·036 0
14	0·038 2	0·121	39	0·018 05	0·035 0
15	0·036 5	0·111	40	0·017 69	0·033 96
16	0·035 0	0·1020	41	0·017 33	0·033 00
17	0·033 62	0·0945	42	0·017 00	0·032 10
18	0·032 33	0·0880	43	0·016 68	0·031 24
19	0·031 14	0·0823	44	0·016 36	0·030 43
20	0·030 04	0·0773	45	0·016 06	0·029 66
21	0·029 01	0·0729	46	0·015 78	0·028 92
22	0·028 06	0·0689	47	0·015 50	0·028 22
23	0·027 16	0·0653	48	0·015 23	0·027 55
24	0·026 32	0·0620	49	0·014 98	0·026 91
25	0·025 52	0·0591	50	0·014 75	0·026 25

(Reproduced from Chapman [1948].)

Applications of the above table are given on p. 63 and pp. 139–40.

2 Tag recoveries needed for prescribed probabilities of detecting incomplete tag-reporting with various levels of catch inspection

The parameters are defined in **3.2.4.**

$\alpha = 0.10$

$1-\beta$	p_0									
	.05	.10	.15	.20	.25	.30	.40	.50	.70	.90
.50	6	4	4	3	3	4	4	5	9	28
.80	24	14	11	10	10	9	10	11	17	51
.90	39	23	18	15	14	14	14	15	23	66
.95	54	31	24	20	19	18	18	19	28	80
.99	88	50	37	32	29	27	26	28	40	109
.50	39	23	18	15	14	14	14	15	23	66
.80	136	76	56	47	42	39	38	39	55	147
.90	210	116	85	71	63	58	55	56	76	201
.95	284	156	114	94	83	76	71	73	97	252
.99	453	246	178	146	128	117	108	109	143	365
.50	84	47	36	30	27	26	25	27	38	106
.80	276	151	110	91	80	74	69	71	95	247
.90	421	229	166	136	120	110	101	102	134	344
.95	564	305	221	180	158	144	132	133	172	436
.99	888	478	344	280	243	221	201	200	256	639
.50	197	109	80	66	59	55	52	53	72	191
.80	615	332	240	196	171	156	143	143	186	469
.90	925	498	359	291	253	230	209	208	266	662
.95	1230	660	474	384	333	302	273	271	342	847
.99	1917	1026	734	593	513	464	416	410	513	1256
.50	322	176	128	106	93	86	80	81	108	278
.80	981	527	380	308	268	244	221	220	280	696
.90	1468	787	564	456	395	358	323	319	402	990
.95	1943	1040	744	601	519	470	422	415	519	1272
.99	3016	1609	1148	925	797	720	643	630	781	1895
.50	568	307	222	182	159	145	133	134	173	439
.80	1688	904	647	523	453	410	369	364	457	1121
.90	2512	1341	958	772	667	602	539	529	658	1603
.95	3314	1767	1260	1015	875	789	704	689	853	2065
.99	5120	2725	1939	1558	1340	1207	1073	1046	1285	3093
.50	1130	607	436	354	307	279	252	250	317	787
.80	3290	1754	1251	1007	868	783	699	684	847	2052
.90	4866	2590	1844	1482	1275	1148	1021	997	1225	2950
.95	6401	3403	2420	1943	1670	1502	1333	1298	1590	3814
.99	9845	5227	3712	2975	2554	2294	2031	1972	2402	5733
.50	2832	1512	1079	869	750	677	605	593	736	1789
.80	8074	4289	3048	2444	2100	1887	1672	1626	1985	4748
.90	11881	6305	4475	3585	3076	2761	2442	2369	2880	6859
.95	15579	8261	5859	4691	4022	3608	3186	3086	3743	8894
.99	23864	12643	8959	7166	6138	5501	4850	4691	5669	13423
.50	12548	6658	4724	3785	3246	2914	2576	2499	3036	7227
.80	35080	18573	13151	10513	8999	8060	7097	6855	8262	19515
.90	51367	27181	19237	15369	13149	11770	10352	9989	12013	28320
.95	67153	35522	25132	20072	17167	15362	13503	13020	15641	36822
.99	102459	54175	38312	30586	26148	23389	20540	19791	23732	55778

Table A.2 (*continued*)

$\alpha = 0.05$

p	$1-\beta$	p_0									
		.05	.10	.15	.20	.25	.30	.40	.50	.70	.90
.25	.50	9	6	5	5	5	6	7	8	14	46
	.80	30	19	15	13	13	13	14	15	25	75
	.90	47	28	22	19	18	18	18	20	32	93
	.95	63	37	29	25	23	22	23	25	38	109
	.99	101	57	41	37	34	33	32	35	51	143
.50	.50	63	37	29	25	23	22	23	25	38	109
	.80	180	100	75	63	57	53	51	54	76	207
	.90	264	146	108	90	80	75	71	74	102	271
	.95	346	191	140	116	103	95	90	92	125	331
	.99	530	289	211	174	153	141	131	133	177	458
.60	.50	137	77	58	49	45	42	41	44	63	174
	.80	368	202	148	123	109	101	95	97	132	346
	.90	533	291	212	175	154	142	131	134	178	460
	.95	693	376	273	224	197	180	166	168	221	566
	.99	1048	566	409	334	291	266	243	244	315	794
.70	.50	324	178	131	109	97	90	85	87	119	315
	.80	827	448	325	266	233	213	196	197	257	653
	.90	1183	638	461	375	327	298	272	272	350	879
	.95	1521	820	590	479	417	379	344	343	437	1091
	.99	2281	1223	877	710	615	558	503	498	628	1549
.75	.50	530	289	211	174	153	141	131	133	177	458
	.80	1325	714	515	419	364	332	302	301	386	968
	.90	1883	1011	726	589	511	464	420	416	528	1310
	.95	2417	1296	929	751	651	590	532	525	662	1631
	.99	3600	1924	1375	1110	959	867	777	764	953	2328
.80	.50	935	506	366	299	261	239	219	220	285	722
	.80	2289	1227	880	712	617	560	505	499	630	1554
	.90	3234	1730	1237	999	864	781	701	690	863	2114
	.95	4137	2209	1578	1272	1098	992	888	872	1084	2641
	.99	6132	3268	2329	1874	1614	1455	1297	1269	1567	3788
.85	.50	1862	1000	718	582	506	459	415	412	523	1297
	.80	4173	2388	1705	1374	1185	1070	957	939	1166	2836
	.90	6287	3350	2387	1920	1654	1491	1329	1299	1604	3877
	.95	8016	4267	3037	2441	2100	1891	1682	1641	2018	4859
	.99	11827	6286	4468	3585	3080	2769	2456	2390	2922	7000
.90	.50	4666	2490	1777	1432	1235	1115	997	977	1213	2947
	.80	11012	5854	4162	3340	2871	2582	2291	2230	2730	6544
	.90	15398	8176	5807	4655	3997	3590	3179	3088	3764	8989
	.95	19572	10386	7371	5905	5066	4548	4022	3902	4743	11299
	.99	28752	15243	10808	8651	7415	6649	5870	5684	6886	16345
.95	.50	20671	10967	7782	6234	5348	4799	4243	4116	5000	11905
	.80	47983	25442	18000	14393	12324	11041	9727	9402	11345	26830
	.90	66770	35343	25021	19997	17114	15324	13487	13022	15682	27011
	.95	84619	44776	31689	25317	21660	19389	17053	16456	19791	46652
	.99	123797	65478	46319	36989	31632	28303	24871	23978	28788	67742

Table A.2 (*concluded*)

$\alpha = 0.01$

-β	p_0									
	.05	.10	.15	.20	.25	.30	.40	.50	.70	.90
50	17	12	10	10	10	11	13	16	28	92
80	45	28	23	21	20	20	22	25	42	131
90	65	39	31	28	27	26	28	32	51	155
95	84	50	39	35	33	32	33	37	59	175
99	126	73	57	49	46	44	45	49	75	218
50	126	73	57	49	46	44	45	49	75	218
80	278	157	118	100	90	85	83	88	127	350
90	382	213	158	133	119	112	108	113	159	432
95	479	266	197	164	147	137	130	136	188	506
99	692	381	280	232	205	190	179	184	250	661
50	274	154	116	98	89	84	82	87	125	347
80	578	319	235	195	174	162	153	158	217	579
90	781	428	314	259	230	212	199	204	276	723
.95	971	531	387	319	281	259	242	246	329	855
.99	1385	753	546	448	393	360	332	336	441	1131
.50	648	357	262	218	193	179	169	174	238	629
.80	1312	713	518	425	373	342	316	320	422	1083
.90	1751	948	686	561	490	449	412	414	539	1368
.95	2162	1167	843	687	599	547	500	500	645	1629
.99	3048	1640	1180	959	834	758	688	685	874	2181
.50	1060	578	422	347	305	281	261	266	353	914
.80	2110	1140	823	671	586	535	489	490	633	1597
.90	2802	1509	1087	883	769	700	636	634	811	2029
.95	3446	1852	1331	1080	938	852	772	767	975	2424
.99	4835	2591	1857	1503	1302	1180	1063	1050	1323	3261
.50	1869	1011	732	597	522	477	437	439	570	1444
.80	3661	1966	1413	1145	994	903	817	811	1029	2555
.90	4834	2591	1857	1502	1301	1179	1063	1050	1323	3261
.95	5927	3172	2270	1834	1586	1435	1290	1271	1593	3909
.99	8775	4419	3156	2544	2196	1984	1775	1743	2169	5282
.50	3724	2000	1437	1165	1011	918	830	824	1045	2593
.80	7184	3840	2744	2214	1913	1729	1550	1524	1902	4647
.90	9441	5038	3595	2896	2498	2255	2015	1976	2452	5957
.95	11537	6150	4383	3528	3040	2741	2445	2393	2958	7161
.99	16035	8534	6074	4882	4201	3782	3364	3283	4036	9719
.50	9332	4980	3554	2863	2470	2229	1993	1954	2424	5894
.80	17752	9444	6720	5398	4643	4178	3714	3621	4445	10689
.90	23221	12342	8773	7040	6049	5438	4824	4695	5742	13759
.95	28291	15027	10674	8560	7351	6604	5851	5687	6938	16586
.99	39150	20774	14743	11812	10134	9096	8044	7804	9488	22601
.50	41348	21937	15566	12469	10696	9600	8487	8232	10002	23813
.80	77628	41128	29142	23312	19969	17897	15779	15263	18444	43682
.90	101102	53539	37919	30318	25958	23254	20483	19795	23878	56448
.95	122823	65021	46037	36797	31496	28206	24830	23981	28893	68223
.99	169262	89565	63386	50642	43326	38784	34111	32916	39588	93317

roduced from Paulik [1961].)

A3 Solution of $\exp(-x) + ax = 1$

Note: For $a < 0{\cdot}05$ take $x = 1/a$.

a	x	a	x	a	x	a	x	a	x
0·050	20·0000000	0·100	9·9995458	0·150	6·6581095	0·200	4·9651142	0·250	3·9206904
·051	19·6078431	·101	9·9004937	·151	6·6130295	·201	4·9395120	·251	3·9037180
·052	19·2307691	·102	9·8033798	·152	6·5697225	·202	4·9141487	·252	3·8868671
·053	18·8679244	·103	9·7081477	·153	6·5263771	·203	4·8890207	·253	3·8701303
·054	18·5185184	·104	9·6147420	·154	6·4835823	·204	4·8641246	·254	3·8535241
0·055	18·1818160	0·105	9·5231129	0·155	6·4413272	0·205	4·8394568	0·255	3·8370292
·056	17·8571425	·106	9·4332073	·156	6·3996012	·206	4·8150140	·256	3·8206502
·057	17·5438592	·107	9·3449775	·157	6·3583940	·207	4·7907929	·257	3·8043857
·058	17·2413787	·108	9·2583768	·158	6·3176956	·208	4·7667902	·258	3·7882344
·059	16·9491518	·109	9·1733599	·159	6·2774962	·209	4·7430026	·259	3·7721949
0·060	16·6666657	0·110	9·0898836	0·160	6·2377864	0·210	4·7194272	0·260	3·7562660
·061	16·3934414	·111	9·0079060	·161	6·1985568	·211	4·6960608	·261	3·7404463
·062	16·1290307	·112	8·9273666	·162	6·1597983	·212	4·6729003	·262	3·7247347
·063	15·8730138	·113	8·8482865	·163	6·1215021	·213	4·6499430	·263	3·7091299
·064	15·6249974	·114	8·7705681	·164	6·0836596	·214	4·6271857	·264	3·6936307
0·065	15·3846122	0·115	8·6941952	0·165	6·0462625	0·215	4·6046258	0·265	3·6782358
·066	15·1515112	·116	8·6191320	·166	6·0093024	·216	4·5822605	·266	3·6629441
·067	14·9253682	·117	8·5453466	·167	5·9727714	·217	4·5600869	·267	3·6477545
·068	14·7058763	·118	8·4728044	·168	5·9366617	·218	4·5381024	·268	3·6326657
·069	14·4927463	·119	8·4014745	·169	5·9009657	·219	4·5163044	·269	3·6176767
0·070	14·2857054	0·120	8·3313262	0·170	5·8656758	0·220	4·4946903	0·270	3·6027863
·071	14·0844963	·121	8·2623301	·171	5·8307850	·221	4·4732575	·271	3·5879935
·072	13·8888760	·122	8·1944575	·172	5·7962859	·222	4·4520036	·272	3·5732972
·073	13·6986147	·123	8·1276809	·173	5·7621718	·223	4·4309262	·273	3·5586962
·074	13·5134952	·124	8·0619734	·174	5·7284359	·224	4·4100228	·274	3·5441897
0·075	13·3333117	0·125	7·9973091	0·175	5·6950714	0·225	4·3892910	0·275	3·5297765
·076	13·1578693	·126	7·9336629	·176	5·6620721	·226	4·3687286	·276	3·5154556
·077	12·9869832	·127	7·8710106	·177	5·6294315	·227	4·3483333	·277	3·5012261
·078	12·8204781	·128	7·8093287	·178	5·5971436	·228	4·3281028	·278	3·4870869
·079	12·6581876	·129	7·7485942	·179	5·5652022	·229	4·3080351	·279	3·4730370
0·080	12·4999534	0·130	7·6887851	0·180	5·5336015	0·230	4·2881278	0·280	3·4590756
·081	12·3456253	·131	7·6298800	·181	5·5023357	·231	4·2683789	·281	3·4452017
·082	12·1950603	·132	7·5718581	·182	5·4713992	·232	4·2487864	·282	3·4314143
·083	12·0481222	·133	7·5146991	·183	5·4407865	·233	4·2293482	·283	3·4177125
·084	11·9046814	·134	7·4583837	·184	5·4104922	·234	4·2100622	·284	3·4040955
0·085	11·7646144	0·135	7·4028927	0·185	5·3805110	0·235	4·1909266	0·285	3·3905624
·086	11·6278033	·136	7·3482078	·186	5·3508378	·236	4·1719393	·286	3·3771122
·087	11·4941358	·137	7·2943110	·187	5·3214675	·237	4·1530985	·287	3·3637441
·088	11·3635044	·138	7·2411851	·188	5·2923952	·238	4·1344024	·288	3·3504572
·089	11·2358068	·139	7·1888131	·189	5·2636161	·239	4·1158490	·289	3·3372508
0·090	11·1109150	0·140	7·1371786	0·190	5·2351255	0·240	4·0974366	0·290	3·3241240
·091	10·9888254	·141	7·0862658	·191	5·2069187	·241	4·0791634	·291	3·3110759
·092	10·8693583	·142	7·0360593	·192	5·1789912	·242	4·0610277	·292	3·2981058
·093	10·7524581	·143	6·9865438	·193	5·1513386	·243	4·0430277	·293	3·2852129
·094	10·6380427	·144	6·9377049	·194	5·1239566	·244	4·0251617	·294	3·2723964
0·095	10·5260334	0·145	6·8895283	0·195	5·0968408	0·245	4·0074282	0·295	3·2596555
·096	10·4163548	·146	6·8420002	·196	5·0699872	·246	3·9898254	·296	3·2469895
·097	10·3089347	·147	6·7951072	·197	5·0433917	·247	3·9723518	·297	3·2343976
·098	10·2037037	·148	6·7488362	·198	5·0170503	·248	3·9550058	·298	3·2218790
·099	10·1005954	·149	6·7031744	·199	4·9909591	·249	3·9377858	·299	3·2094331
0·100	9·9995458	0·150	6·6581095	0·200	4·9651142	0·250	3·9206904	0·300	3·1970591

Table A.3 (*continued*)

a	*x*	*a*	*x*	*a*	*x*	*a*	*x*	*a*	*x*
0·300	3·1970591	0·350	2·6566127	0·400	2·2316119	0·450	1·8847348	0·500	1·5936243
·301	3·1847564	·351	2·6471411	·401	2·2240016	·451	1·8784240	·501	1·5882633
·302	3·1725241	·352	2·6377143	·402	2·2164218	·452	1·8721351	·502	1·5829186
·303	3·1603617	·353	2·6283321	·403	2·2088724	·453	1·8658679	·503	1·5775904
·304	3·1482684	·354	2·6189939	·404	2·2013532	·454	1·8596224	·504	1·5722783
0·305	3·1362436	0·355	2·6096996	0·405	2·1938638	0·455	1·8533984	0·505	1·5669823
·306	3·1242866	·356	2·6004487	·406	2·1864042	·456	1·8471958	·506	1·5617022
·307	3·1123968	·357	2·5912409	·407	2·1789740	·457	1·8410145	·507	1·5564381
·308	3·1005735	·358	2·5820758	·408	2·1715732	·458	1·8348542	·508	1·5511898
·309	3·0888160	·359	2·5729530	·409	2·1642015	·459	1·8287149	·509	1·5459572
0·310	3·0771238	0·360	2·5638723	0·410	2·1568586	0·460	1·8225965	0·510	1·5407403
·311	3·0654961	·361	2·5548333	·411	2·1495444	·461	1·8164989	·511	1·5355369
·312	3·0539325	·362	2·5458356	·412	2·1422588	·462	1·8104218	·512	1·5303531
·313	3·0424323	·363	2·5368790	·413	2·1350014	·463	1·8043652	·513	1·5251826
·314	3·0309949	·364	2·5279631	·414	2·1277721	·464	1·7983290	·514	1·5200275
0·315	3·0196198	0·365	2·5190875	0·415	2·1205707	0·465	1·7923131	0·515	1·5148876
·316	3·0083062	·366	2·5102520	·416	2·1133971	·466	1·7863172	·516	1·5097628
·317	2·9970537	·367	2·5014563	·417	2·1062510	·467	1·7803413	·517	1·5046632
·318	2·9858617	·368	2·4927000	·418	2·0991322	·468	1·7743854	·518	1·4995685
·319	2·9747297	·369	2·4839828	·419	2·0920406	·469	1·7684491	·519	1·4944788
0·320	2·9636570	0·370	2·4753044	0·420	2·0849759	0·470	1·7625325	0·520	1·4894139
·321	2·9526432	·371	2·4666645	·421	2·0779381	·471	1·7566355	·521	1·4843637
·322	2·9416876	·372	2·4580629	·422	2·0709268	·472	1·7507578	·522	1·4793282
·323	2·9307898	·373	2·4494991	·423	2·0639420	·473	1·7448995	·523	1·4743074
·324	2·9199493	·374	2·4409730	·424	2·0569834	·474	1·7390603	·524	1·4693011
0·325	2·9091655	0·375	2·4324843	0·425	2·0500510	0·475	1·7332402	0·525	1·4643092
·326	2·8984379	·376	2·4240326	·426	2·0431444	·476	1·7274391	·526	1·4593317
·327	2·8877660	·377	2·4156176	·427	2·0362635	·477	1·7216568	·527	1·4543685
·328	2·8771494	·378	2·4072391	·428	2·0294083	·478	1·7158932	·528	1·4494195
·329	2·8665874	·379	2·3988969	·429	2·0225784	·479	1·7101483	·529	1·4444847
0·330	2·8560797	0·380	2·3905906	0·430	2·0157738	0·480	1·7044219	0·530	1·4395640
·331	2·8456257	·381	2·3823199	·431	2·0089942	·481	1·6987139	·531	1·4346573
·332	2·8352249	·382	2·3740846	·432	2·0022395	·482	1·6930242	·532	1·4297645
·333	2·8248770	·383	2·3658845	·433	1·9955095	·483	1·6873527	·533	1·4248856
·334	2·8145814	·384	2·3577192	·434	1·9888042	·484	1·6816993	·534	1·4200205
0·335	2·8043377	0·385	2·3495885	0·435	1·9821232	0·485	1·6760639	0·535	1·4151691
·336	2·7941455	·386	2·3414922	·436	1·9754665	·486	1·6704464	·536	1·4103313
·337	2·7840042	·387	2·3334300	·437	1·9688339	·487	1·6648467	·537	1·4055072
·338	2·7739134	·388	2·3254016	·438	1·9622253	·488	1·6592647	·538	1·4006965
·339	2·7638728	·389	2·3174068	·439	1·9556405	·489	1·6537002	·539	1·3958994
0·340	2·7538818	0·390	2·3094453	0·440	1·9490792	0·490	1·6481533	0·540	1·3911155
·341	2·7439401	·391	2·3015170	·441	1·9425415	·491	1·6426237	·541	1·3863450
·342	2·7340472	·392	2·2936214	·442	1·9360271	·492	1·6371111	·542	1·3815878
·343	2·7242027	·393	2·2857585	·443	1·9295360	·493	1·6316164	·543	1·3768437
·344	2·7144062	·394	2·2779280	·444	1·9230678	·494	1·6261384	·544	1·3721128
0·345	2·7046573	0·395	2·2701297	0·445	1·9166226	0·495	1·6206774	0·545	1·3673948
·346	2·6949555	·396	2·2623632	·446	1·9102001	·496	1·6152333	·546	1·3626899
·347	2·6853006	·397	2·2546284	·447	1·9038002	·497	1·6098061	·547	1·3579979
·348	2·6756921	·398	2·2469251	·448	1·8974228	·498	1·6043955	·548	1·3533187
·349	2·6661296	·399	2·2392530	·449	1·8910677	·499	1·5990016	·549	1·3486523
0·350	2·6566127	0·400	2·2316119	0·450	1·8817348	0·500	1·5936243	0·550	1·3439987

Table A.3 (*continued*)

a	*x*	*a*	*x*	*a*	*x*	*a*	*x*	*a*	*x*
0·550	1·3439987	0·600	1·1262612	0·650	0·9336939	0·700	0·7614337	0·750	0·6058600
·551	1·3393577	·601	1·1221820	·651	·9300635	·701	·7581691	·751	·6028986
·552	1·3347293	·602	1·1181127	·652	·9264412	·702	·7549111	·752	·5999427
·553	1·3301135	·603	1·1140533	·653	·9228268	·703	·7516598	·753	·5969924
·554	1·3255101	·604	1·1100038	·654	·9192204	·704	·7484150	·754	·5940475
0·555	1·3209191	0·605	1·1059641	0·655	0·9156220	0·705	0·7451767	0·755	0·5911081
·556	1·3163405	·606	1·1019342	·656	·9120315	·706	·7419450	·756	·5881742
·557	1·3117742	·607	1·0979141	·657	·9084489	·707	·7387198	·757	·5852457
·558	1·3072201	·608	1·0939037	·658	·9048742	·708	·7355010	·758	·5823227
·559	1·3026782	·609	1·0899029	·659	·9013073	·709	·7322887	·759	·5794050
0·560	1·2981485	0·610	1·0859117	0·660	0·8977481	0·710	0·7290829	0·760	0·5764927
·561	1·2936307	·611	1·0819301	·661	·8941968	·711	·7258834	·761	·5735858
·562	1·2891250	·612	1·0779581	·662	·8906531	·712	·7226903	·762	·5706842
·563	1·2846312	·613	1·0739955	·663	·8871172	·713	·7195036	·763	·5677880
·564	1·2801493	·614	1·0700424	·664	·8835889	·714	·7163232	·764	·5648970
0·565	1·2756792	0·615	1·0660987	0·665	0·8800683	0·715	0·7131491	0·765	0·5620114
·566	1·2712209	·616	1·0621643	·666	·8765553	·716	·7099813	·766	·5591310
·567	1·2667743	·617	1·0582393	·667	·8730498	·717	·7068198	·767	·5562558
·568	1·2623394	·618	1·0543236	·668	·8695519	·718	·7036645	·768	·5533859
·569	1·2579161	·619	1·0504171	·669	·8660616	·719	·7005154	·769	·5505213
0·570	1·2535043	0·620	1·0465198	0·670	0·8625787	0·720	0·6973725	0·770	0·5476618
·571	1·2491040	·621	1·0426317	·671	·8591032	·721	·6942358	·771	·5448075
·572	1·2447151	·622	1·0387527	·672	·8556352	·722	·6911053	·772	·5419584
·573	1·2403377	·623	1·0348828	·673	·8521746	·723	·6879808	·773	·5391144
·574	1·2359715	·624	1·0310219	·674	·8487214	·724	·6848625	·774	·5362756
0·575	1·2316167	0·625	1·0271701	0·675	0·8452755	0·725	0·6817503	0·775	0·5334419
·576	1·2272730	·626	1·0233272	·676	·8418370	·726	·6786441	·776	·5306132
·577	1·2229405	·627	1·0194933	·677	·8384057	·727	·6755440	·777	·5277897
·578	1·2186192	·628	1·0156683	·678	·8349816	·728	·6724499	·778	·5249712
·579	1·2143089	·629	1·0118521	·679	·8315648	·729	·6693618	·779	·5221578
0·580	1·2100097	0·630	1·0080447	0·680	0·8281552	0·730	0·6662796	0·780	0·5193494
·581	1·2057214	·631	1·0042461	·681	·8247528	·731	·6632035	·781	·5165460
·582	1·2014440	·632	1·0004563	·682	·8213575	·732	·6601332	·782	·5137476
·583	1·1971774	·633	0·9966752	·683	·8179694	·733	·6570689	·783	·5109542
·584	1·1929217	·634	0·9929027	·684	·8145883	·734	·6540104	·784	·5081658
0·585	1·1886768	0·635	0·9891389	0·685	0·8112143	0·735	0·6509579	0·785	0·5053823
·586	1·1844426	·636	·9853837	·686	·8078473	·736	·6479112	·786	·5026037
·587	1·1802190	·637	·9816371	·687	·8044874	·737	·6448703	·787	·4998301
·588	1·1760060	·638	·9778989	·688	·8011344	·738	·6418352	·788	·4970613
·589	1·1718036	·639	·9741693	·689	·7977884	·739	·6388059	·789	·4942975
0·590	1·1676118	0·640	0·9704481	0·690	0·7944493	0·740	0·6357824	0·790	0·4915385
·591	1·1634304	·641	·9667354	·691	·7911171	·741	·6327646	·791	·4887843
·592	1·1592594	·642	·9630310	·692	·7877918	·742	·6297526	·792	·4860350
·593	1·1550988	·643	·9593350	·693	·7844733	·743	·6267462	·793	·4832905
·594	1·1509485	·644	·9556473	·694	·7811617	·744	·6237456	·794	·4805508
0·595	1·1468085	0·645	0·9519679	0·695	0·7778568	0·745	0·6207506	0·795	0·4778159
·596	1·1426788	·646	·9482967	·696	·7745587	·746	·6177612	·796	·4750858
·597	1·1385592	·647	·9446338	·697	·7712674	·747	·6147775	·797	·4723604
·598	1·1344498	·648	·9409790	·698	·7679828	·748	·6117994	·798	·4696398
·599	1·1303505	·649	·9373324	·699	·7647049	·749	·6088269	·799	·4669239
·600	1·1262612	0·650	·9336939	0·700	0·7614337	0·750	0·6058600	0·800	0·4642129

Table A.3 (*concluded*)

a	*x*	*a*	*x*	*a*	*x*	*a*	*x*
0·800	0·4642128	0·850	0·3343447	0·900	0·2145557	0·950	0·1034788
·801	·4615063	·851	·3318552	·901	·2122529	·951	·1013381
·802	·4588045	·852	·3293697	·902	·2099535	·952	·0992005
·803	·4561073	·853	·3268882	·903	·2076576	·953	·0970656
·804	·4534148	·854	·3244107	·904	·2053652	·954	·0949342
0·805	0·4507270	0·855	0·3219371	0·905	0·2030762	0·955	0·0928056
·806	·4480438	·856	·3194675	·906	·2007906	·956	·0906799
·807	·4453652	·857	·3170019	·907	·1985084	·957	·0885573
·808	·4426911	·858	·3145402	·908	·1962296	·958	·0864376
·809	·4400217	·859	·3120824	·909	·1939542	·959	·0843209
0·810	0·4373568	0·860	0·3096286	0·910	0·1916824	0·960	0·0822071
·811	·4346965	·861	·3071786	·911	·1894137	·961	·0800963
·812	·4320407	·862	·3047325	·912	·1871484	·962	·0779885
·813	·4293894	·863	·3022903	·913	·1848866	·963	·0758836
·814	·4267426	·864	·2998520	·914	·1826281	·964	·0737816
0·815	0·4241004	0·865	0·2974176	0·915	0·1803729	0·965	0·0716825
·816	·4214626	·866	·2949869	·916	·1781211	·966	·0695864
·817	·4188292	·867	·2925601	·917	·1758725	·967	·0674932
·818	·4162003	·868	·2901372	·918	·1736274	·968	·0654028
·819	·4135759	·869	·2877180	·919	·1713855	·969	·0633154
0·820	0·4109558	0·870	0·2853027	0·920	0·1691469	0·970	0·0612308
·821	·4083402	·871	·2828911	·921	·1669116	·971	·0591492
·822	·4057290	·872	·2804833	·922	·1646795	·972	·0570704
·823	·4031222	·873	·2780793	·923	·1624508	·973	·0549944
·824	·4005197	·874	·2756790	·924	·1602254	·974	·0529213
0·825	0·3979215	0·875	0·2732825	0·925	0·1580031	0·975	0·0508511
·826	·3953278	·876	·2708897	·926	·1557841	·976	·0487837
·827	·3927383	·877	·2685007	·927	·1535683	·977	·0467191
·828	·3901532	·878	·2661153	·928	·1513558	·978	·0446574
·829	·3875723	·879	·2637337	·929	·1491465	·979	·0425985
0·830	0·3849957	0·880	0·2613557	0·930	0·1469404	0·980	0·0405424
·831	·3824235	·881	·2589815	·931	·1447374	·981	·0384891
·832	·3798554	·882	·2566109	·932	·1425377	·982	·0364386
·833	·3772916	·883	·2542439	·933	·1403412	·983	·0343909
·834	·3747321	·884	·2518806	·934	·1381478	·984	·0323460
0·835	0·3721768	0·885	0·2495210	0·935	0·1359576	0·985	0·0303038
·836	·3696257	·886	·2471649	·936	·1337706	·986	·0282644
·837	·3670787	·887	·2448125	·937	·1315867	·987	·0262278
·838	·3645360	·888	·2424637	·938	·1294060	·988	·0241939
·839	·3619974	·889	·2401185	·939	·1272283	·989	·0221628
0·840	0·3594630	0·890	0·2377769	0·940	0·1250538	0·990	0·0201345
·841	·3569327	·891	·2354389	·941	·1228825	·991	·0181088
·842	·3544066	·892	·2331044	·942	·1207142	·992	·0160859
·843	·3518846	·893	·2307735	·943	·1185490	·993	·0140657
·844	·3493667	·894	·2284461	·944	·1163869	·994	·0120482
0·845	0·3468528	0·895	0·2261222	0·945	0·1142279	0·995	0·0100335
·846	·3443431	·896	·2238019	·946	·1120720	·996	·0080214
·847	·3418374	·897	·2214851	·947	·1099191	·997	·0060120
·848	·3393358	·898	·2191718	·948	·1077693	·998	·0040053
·849	·3368383	·899	·2168620	·949	·1056226	·999	·0020013
0·850	0·3343447	0·900	0·2145557	0·950	0·1034788	1·000	0·0000000

(Reproduced from Barton, David and Merrington [1960]. The above table is also given by David, Kendall and Barton, *Symmetric Functions and Allied Tables* (Cambridge University Press, for the Biometrika Trustees), 1966, Table 8.1.)

A4 Table for finding the maximum-likelihood estimate of P, the parameter of a truncated geometric distribution

In the following table the function

$$f(Q) = \frac{sQ^{s+1} - (s+1)Q^s + 1}{Q^{s+1} - Q^s - Q + 1}$$

is evaluated for different values of P $(= 1 - Q)$ and s. Using linear interpolation, this table can be used for solving the equation $f(Q) = \overline{x}$ as described in **4.1.6** (*4*).

	P										
s	0·001	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80	0·90	0·999
2	1·500	1·474	1·444	1·412	1·375	1·333	1·286	1·231	1·167	1·091	1·001
3	1·999	1·930	1·852	1·767	1·673	1·571	1·462	1·345	1·226	1·108	1·001
4	2·499	2·369	2·225	2·069	1·904	1·733	1·562	1·396	1·244	1·111	1·001
5	2·998	2·790	2·563	2·323	2·078	1·839	1·615	1·416	1·248	1·111	1·001
6	3·497	3·195	2·868	2·533	2·206	1·905	1·642	1·424	1·250	1·111	1·001
7	3·996	3·582	3·142	2·705	2·298	1·945	1·655	1·427	1·250	1·111	1·001
8	4·495	3·953	3·387	2·844	2·363	1·969	1·661	1·428	1·250	1·111	1·001
9	4·993	4·308	3·605	2·955	2·408	1·982	1·664	1·428	1·250	1·111	1·001
10	5·492	4·647	3·797	3·043	2·439	1·990	1·666	1·429	1·250	1·111	1·001
11	5·990	4·969	3·966	3·111	2·460	1·995	1·666	1·429	1·250	1·111	1·001
12	6·488	5·277	4·115	3·165	2·474	1·997	1·666	1·429	1·250	1·111	1·001
13	6·986	5·569	4·244	3·206	2·483	1·998	1·667	1·429	1·250	1·111	1·001
14	7·484	5·847	4·356	3·238	2·489	1·999	1·667	1·429	1·250	1·111	1·001
15	7·981	6·111	4·453	3·262	2·493	2·000	1·667	1·429	1·250	1·111	1·001
16	8·479	6·360	4·537	3·280	2·495	2·000	1·667	1·429	1·250	1·111	1·001
17	8·976	6·597	4·608	3·294	2·497	2·000	1·667	1·429	1·250	1·111	1·001
18	9·473	6·821	4·670	3·304	2·498	2·000	1·667	1·429	1·250	1·111	1·001
19	9·970	7·033	4·722	3·312	2·499	2·000	1·667	1·429	1·250	1·111	1·001
20	10·467	7·232	4·767	3·317	2·499	2·000	1·667	1·429	1·250	1·111	1·001
21	10·963	7·429	4·805	3·322	2·500	2·000	1·667	1·429	1·250	1·111	1·001
22	11·460	7·597	4·836	3·325	2·500	2·000	1·667	1·429	1·250	1·111	1·001
23	11·956	7·763	4·863	3·327	2·500	2·000	1·667	1·429	1·250	1·111	1·001
24	12·452	7·920	4·886	3·329	2·500	2·000	1·667	1·429	1·250	1·111	1·001
25	12·948	8·066	4·905	3·330	2·500	2·000	1·667	1·429	1·250	1·111	1·001
26	13·444	8·204	4·921	3·331	2·500	2·000	1·667	1·429	1·250	1·111	1·001

(Reproduced from Thomasson and Kapadia [1968].)

A5 Tabulation of

$$f(x) = \frac{1}{x} - \frac{1}{\exp(x) - 1}$$

Using linear interpolation, the following table can be used for solving the equation $f(x) = a$ (e.g. see equation (6.5)).

x	·0	·1	·2	·3	·4	·5	·6	·7	·8	·9
0		·4916	·4832	·4750	·4668	·4584	·4504	·4422	·4340	·4260
1	·4180	·4102	·4024	·3946	·3870	·3794	·3720	·3648	·3576	·3504
2	·3434	·3366	·3300	·3234	·3168	·3106	·3044	·2984	·2924	·2866
3	·2810	·2754	·2700	·2648	·2596	·2546	·2496	·2450	·2402	·2358
4	·2314	·2270	·2228	·2188	·2148	·2110	·2072	·2036	·2000	·1966
5	·1932	·1900	·1868	·1836	·1806	·1778	·1748	·1720	·1694	·1668
6	·1642	·1616	·1592	·1568	·1546	·1524	·1502	·1480	·1460	·1440
7	·1420	·1400	·1382	·1364	·1346	·1328	·1310	·1294	·1278	·1262
8	·1246	·1232	·1216	·1202	·1188	·1174	·1160	·1148	·1134	·1122

(Reproduced from Deemer and Votaw [1955].)

A6 Tabulation of $A_K(S) \quad \sum_{k=0}^{K} kS^k \Big/ \sum_{k=0}^{K} S^k$

S \ $K =$	2	3	4	5	6	7	8
.01	.0101	.0110	.0110	.0110	.0110	.0110	.0110
.02	.0204	.0204	.0204	.0204	.0204	.0204	.0204
.03	.0308	.0309	.0309	.0309	.0309	.0309	.0309
.04	.0415	.0416	.0416	.0416	.0416	.0416	.0416
.05	.0523	.0526	.0527	.0527	.0527	.0527	.0527
.06	.0632	.0638	.0638	.0638	.0638	.0638	.0638
.07	.0742	.0742	.0752	.0752	.0752	.0752	.0752
.08	.0854	.0868	.0869	.0869	.0869	.0869	.0869
.09	.0967	.0986	.0989	.0989	.0989	.0989	.0989
.10	.1081	.1107	.1111	.1111	.1111	.1111	.1111
.11	.1196	.1230	.1235	.1236	.1236	.1236	.1236
.12	.1312	.1355	.1363	.1364	.1364	.1364	.1364
.13	.1428	.1483	.1493	.1494	.1495	.1495	.1495
.14	.1545	.1612	.1625	.1627	.1628	.1628	.1628
.15	.1663	.1744	.1761	.1764	.1765	.1765	.1765
.16	.1781	.1879	.1900	.1904	.1905	.1905	.1905
.17	.1900	.2015	.2041	.2047	.2047	.2047	.2047
.18	.2019	.2153	.2186	.2193	.2195	.2195	.2195
.19	.2138	.2294	.2333	.2343	.2345	.2346	.2346
.20	.2258	.2436	.2484	.2496	.2499	.2499	.2499
.21	.2378	.2580	.2638	.2653	.2657	.2658	.2658
.22	.2498	.2727	.2795	.2814	.2818	.2819	.2819
.23	.2618	.2875	.2955	.2978	.2985	.2986	.2987
.24	.2737	.3025	.3118	.3147	.3155	.3158	.3158
.25	.2857	.3177	.3285	.3319	.3329	.3332	.3333
.26	.2977	.3330	.3454	.3495	.3508	.3512	.3513
.27	.3096	.3485	.3627	.3675	.3691	.3696	.3698
.28	.3216	.3641	.3803	.3860	.3879	.3886	.3888
.29	.3335	.3800	.3982	.4049	.4072	.4080	.4083
.30	.3453	.3959	.4164	.4242	.4270	.4281	.4284
.31	.3572	.4120	.4319	.4439	.4474	.4486	.4491
.32	.3690	.4282	.4538	.4642	.4682	.4698	.4704
.33	.3807	.4445	.4729	.4848	.4895	.4914	.4921
.34	.3924	.4610	.4923	.5058	.5114	.5137	.5145
.35	.4041	.4775	.5121	.5271	.5340	.5367	.5377
.36	.4157	.4942	.5321	.5494	.5570	.5603	.5616
.37	.4272	.5109	.5524	.5718	.5806	.5845	.5861
.38	.4387	.5277	.5730	.5948	.6049	.6094	.6114
.39	.4502	.5446	.5938	.6181	.6297	.6350	.6374
.40	.4615	.5616	.6149	.6420	.6552	.6614	.6643
.41	.4728	.5786	.6363	.6664	.6813	.6885	.6920
.42	.4841	.5957	.6579	.6910	.7080	.7164	.7205
.43	.4953	.6128	.6798	.7162	.7353	.7450	.7499
.44	.5064	.6299	.7019	.7418	.7633	.7745	.7801
.45	.5174	.6471	.7242	.7679	.7919	.8047	.8114
.46	.5284	.6644	.7467	.7945	.8212	.8358	.8435
.47	.5392	.6816	.7694	.8214	.8512	.8677	.8767
.48	.5500	.6988	.7923	.8488	.8817	.9005	.9109
.49	.5608	.7161	.8154	.8766	.9140	.9341	.9461
.50	.5714	.7333	.8387	.9048	.9449	.9686	.9824
.51	.5820	.7506	.8621	.9333	.9774	1.0041	1.0198
.52	.5925	.7678	.8857	.9623	1.0106	1.0403	1.0583
.53	.6029	.7850	.9094	.9917	1.0445	1.0775	1.0978
.54	.6132	.8022	.9333	1.0214	1.0789	1.1157	1.1387
.55	.6235	.8193	.9573	1.0514	1.1140	1.1547	1.1806
.56	.6336	.8365	.9813	1.0818	1.1497	1.1946	1.2237
.57	.6437	.8535	1.0055	1.1125	1.1860	1.2354	1.2681
.58	.6537	.8705	1.0297	1.1435	1.2229	1.2772	1.3136
.59	.6636	.8875	1.0540	1.1748	1.2604	1.3198	1.3604
.60	.6735	.9044	1.0784	1.2064	1.2984	1.3634	1.4084
.61	.6832	.9213	1.1028	1.2382	1.3370	1.4077	1.4576
.62	.6929	.9380	1.1273	1.2702	1.3761	1.4530	1.5081
.63	.7025	.9548	1.1518	1.3025	1.4157	1.4991	1.5598
.64	.7119	.9714	1.1763	1.3350	1.4557	1.5461	1.6127
.65	.7214	.9880	1.2009	1.3677	1.4963	1.5938	1.6668
.66	.7307	1.0044	1.2254	1.4006	1.5373	1.6423	1.7220
.67	.7399	1.0208	1.2499	1.4336	1.5787	1.6917	1.7786
.68	.7491	1.0371	1.2743	1.4667	1.6205	1.7417	1.8362
.69	.7581	1.0534	1.2988	1.5000	1.6626	1.7925	1.8950
.70	.7671	1.0695	1.3232	1.5333	1.7051	1.8439	1.9549
.71	.7760	1.0855	1.3476	1.5667	1.7479	1.8960	2.0158
.72	.7848	1.1014	1.3719	1.6002	1.7910	1.9487	2.0778
.73	.7936	1.1173	1.3961	1.6338	1.8343	2.0019	2.1407
.74	.8022	1.1330	1.4202	1.6773	1.8779	2.0557	2.2046
.75	.8108	1.1486	1.4443	1.7009	1.9217	2.1100	2.2694
.76	.8193	1.1641	1.4683	1.7345	1.9656	2.1647	2.3350
.77	.8277	1.1795	1.4921	1.7680	2.0097	2.2199	2.4014
.78	.8360	1.1947	1.5159	1.8015	2.0539	2.2754	2.2486
.79	.8443	1.2099	1.5395	1.8350	2.0981	2.3312	2.5364
.80	.8525	1.2249	1.5631	1.8683	2.1424	2.3872	2.6048
.81	.8605	1.2399	1.5865	1.9016	2.1868	2.4436	2.6737
.82	.8686	1.2547	1.6097	1.9348	2.2311	2.5001	2.7432
.83	.8765	1.2694	1.6328	1.9678	2.2754	2.5567	2.8131
.84	.8843	1.2839	1.6558	2.0007	2.3197	2.6135	2.8833
.85	.8921	1.2984	1.6786	2.0335	2.3638	2.6702	2.9538
.86	.8998	1.3127	1.7013	2.0662	2.4079	2.7270	3.0245
.87	.9075	1.3269	1.7238	2.0986	2.4518	2.7838	3.0953
.88	.9150	1.3409	1.7461	2.1309	2.4955	2.8405	3.1662
.89	.9225	1.3549	1.7683	2.1629	2.5391	2.8970	3.2372
.90	.9299	1.3687	1.7903	2.1948	2.5824	2.9534	3.3080
.91	.9372	1.3824	1.8121	2.2264	2.6255	3.0096	3.3788
.92	.9445	1.3960	1.8337	2.2578	2.6684	3.0655	3.4493
.93	.9517	1.4094	1.8552	2.2890	2.7110	3.1212	3.5196
.94	.9588	1.4227	1.8765	2.3200	2.7533	3.1765	3.5896
.95	.9658	1.4359	1.8975	2.3506	2.7953	3.2315	3.6593
.96	.9728	1.4490	1.9184	2.3811	2.8369	3.2861	3.7285
.97	.9797	1.4619	1.9391	2.4112	2.8783	3.3402	3.7972
.98	.9865	1.4747	1.9596	2.4411	2.9192	3.3940	3.8654
.99	.9933	1.4874	1.9799	2.4707	2.9598	3.4472	3.9332
1.00	1.0000	1.5000	2.0000	2.5000	3.0000	3.5000	4.0000

(Reproduced from Robson and Chapman [1961].)

REFERENCES:

Albers, P.H. (1976). Determining population size of territorial red-winged blackbirds *J. Wildl. Manag. 40,761-8.*

Allen, K.R.(1966). Some methods for estimating exploited populations. *J. Fish Res. Bd. Canada.23, 1553-74.*

Ambrose, H.W.III (1973). An experimental study of some factors affecting the spatial and temporal activity of *Microths pennysylvanicus, J.Mammal. 54,79-110.*

Anderson, J. (1962).Roe-deer census and population analysis by means of modified marking and release technique. In E.D. Le Cren and M.W. Holdgate (eds). *The exploitation of Natural Animal Pop populations, 72-82.* Blackwell: Oxford.

Anderson, D.R., Burnham, K.P. and Crain, B.R. (1978). A log-linear model approach to estimation of population size using the line-transact method. *Ecology 59, 190-3.*

Andrzejewski, R.(1967). Estimation of the abundance of small rodent populations for the use of biological productivity investigations. In K. Petrusewice (ed.),*Secondary Productivity of Ecosystems, Vol 1, 275-81.* Institute of Ecology, Polish Academy of sciences.

Anscombe, F.J. (1948) the statistical analysis of insect counts based on the negative binomial distribution. *Biometics 5,1 165-73.*

Arnason, A.N. (1973) the estimation of population size, migration rates and survival in the stratified population. *Res. Popul Ecol. 15,1-8.*

Ayre, G.L. (1962). Problems in using the Lincoln Index for estimating the size of ant colonies (*Hymenoptera: Formicidae). J.N.Y. Ent.Soc. 70, 159-66.*

Babinska, J. and Bock, E. (1969) The effects of pre-baiting on captures of rodents. *Actatheriologica 14, 267-70.*

Bailey N.T.J. (1951). On estimating the size of mobile populations from capture-recapture data. *Biometrika.38, 293-306.*

Bailey N.T.J.(1952) Improvement interpretation of recapture data. *J.Animal Ecol. 21, 120-7.*

Barnett, V.D. (1966). Evaluation of the maximum-likelihood estimator when the likelihood equation has multiple roots. *Biometrika 53, 151-65.*

Bell. G. (1974). Population estimates from recapture studies in which no recapture have been made. *Nature 248, 616*

Bergs. (1975) Some properties and applications of a ratio of stirling numbers of the second kind. *Scand. J.Statist. 2, 91-4.*

Bergs. (1976). A note on the UMVU estimate in a multiple -recapture census. *Scand. J. statist. 3, 86-8.*

Beverton, R.J.H. (1954). Notes on the use of theoretical models in the study of exploited fish populations. *Misc.contr. No.2, U.S. Fishery Lab., Benfort, N.C. 1-159.*

Boguslavsky, G.W. (1956) statistical estimation of the size of a small population. *Science 124,317-18.*

Brass, W. (1958). Simplified methods of fitting the truncated negative binomial distribution. *Biometrika 45, 59-68.*

Buck, D.H. and Thoits, C.F. 3rd (1965). An evaluation of Petersen estimation procedures employing seines in 1-acre ponds. *J. Wildl. Manag. 29, 598-621.*

Burnham, K.P. (1972). Estimation of population size in multiple capture studies when capture probabilities vary among animals. Ph.D. Dissertation, *Ore, St. Univ Carvallis.*

Bukham, K.P. and overton, W.S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals *Biometrika 65, 625-33*

Carothers, A.D. (1971). An examination and extension of Leslie´s test of equal catchability. *Biometrics 27, 615-30.*

Chapman, D.G. (1948) A mathematical study of confidence limits of salmon populations calculated from sample tag ratios. *Internat. Pac. Salmon Fisheries Comm. Bull.2, 69-85*

Chapman, D.G. (1951) Some properties of the hypergeometric distribution with applications to Zoological Censuses. *Univ.calf.Public. Stat.1, 131-60.*

Chapman, D.G. (1952) Inverse multiple and sequential sample Censuses. *Biometrics 8, 1-15.*

Chapman, D.G. and Overton, W.S. (1986) . Estimating and testing differences between population levels by the Schnabel estimation method. *J.Wildl.Manag. 30, 173-80.*

Chapman and Robson, D.S.(1960). The analysis of catch curve. *Biometrics 16, 354-68.*

Chung, J.H. and De Lury, D.B. (1950). *Confidence Limits for the hypergeometric Distribution, University of Toronto Press.*

Cochran, W.G.(1977). *Sampling techniques,* 3rd ed. John Willey and sons: New York.

Cormack, R.M.(1966). A test for equal catchability. *Biometrics 22, 330-42.*

Cormack, R.M. (1968). The statistics of capture-recapture methods. *Oceanogr. Mar. Bio. Ann. Rev. 6, 455-506*

Darroch, J.N. (1958). The multiple-recapture census. I: Estimation of a closed population. *Biometrika 45, 343-59.*

David, F.N. and Johnson, N.L.(1952). The truncated Poisson. *Biometrics 8, 275-85.*

De Lury, D.B. (1947). On the estimation of biological population. *Biometrics 3, 145-67.*

Eberhardt, L.L. (1969). Population estimates from recapture frequencies. *J. Wildl. Manag. 33, 28-39.*

Eberhardt, L.L. Peterle, T.J. and Schofield, R. (1963) problems in a rabbit population study. *Wildl. Manogr. 10, 1-51.*

Fienberg, S.E.(1972). The multiple recapture census for closed population and incomplete 2^k contingency tables. *Biometrika 59, 591-603*

Goodman, L.A. (1953). Sequential sampling tagging for population size problems. *Ann. Math. Statist. 24, 56-69.*

Gulland, J.A. (1955b). On estimation of population parameters from marked members. *Biometrika 42, 269-70.*

Hyne, D.W. (1949b). Two methods for estimating populations from trapping records. *J. Mammal. 30, 399-411.*

Holgate, P. (1966). Contribution to the mathematics of animal trapping. *Biometrics 22, 925-36.*

Jackson, C.H.N. (1933) on true density of tsetse flies. *J.Animal Ecol. 2, 204-9.*

Jolly, B.M.(1965). Explicit estimates from capture -recapture data with both death and immigration-stochastic model. *Biometrika 52, 225-47.*

Junge, C.D.(1963). A quantitative evaluation of the bias on population estimates based on selective samples. In North Atlantic Fish Marking Symposium, I.C.N.A.F. *Special Publication No.4, 26-8*

Laplace, P.S. (1786). *Sur les naissances, less mariages et les morts. Histoire de l´Acade´mie Royale des sciences, Anne´e 1783 Paris, P.683.*

Le Cran, E.D. (1965). A note on the history of mark -recapture population estimates. *J. Animal Ecol. 34, 453-4.*

Leslie, P.H.(1952). The estimation of population parameters from data obtained by means of capture -recapture method.II: the estimation of total numbers. *Biometrika 39,363-88*

Manly, B.F.J.(1969). Some properties of a method of estimating the size of mobile animal populations. *Biometrika 56, 407-10.*

Marten, G.G.(1970a). A regression method for mark -recapture estimates with unequal catchability. *Ecology 51,291-5.*

Mitra, S.K.(1958). On the limiting power function of the frequency Chi-Squared test. *Ann. Math. Statist.29,1221-33*

Otis, D.L., Burnham, K.P., White, G.L. and Anerson. D.R. (1978). Statistical inference for capture data from closed populations. *Wildl. Monogr. No. 62,135p.p.*

Overton, W.S.(1965). A modification of the schnabel estimator to account for removal of animals from the population. *J. Wildl.Manag.29, 392-5.*

Owen, D.B.(1962).*Handbook of statistical tables.* Addison-Wesley,Reading, Massachusetts.

Paloheimo,J.E.(1963). Estimation of catchabilities and population on sizes of lobsters. *J. Fish Res. Bd. Canada 20,59-88.*

Parr, M.J. Gaskell, T.J. and George, B.J. (1968). capture -recapture methods of estimating animal numbers. *J.Biol. educ 2, 95-117.*

Paulik, B.J.(1961). Detection of incomplete reporting of tags. *J. Fish. Res.Bd. Canada 18, 817-29.*

Pielou, E.C. (1969). *An introduction to Mathematical Ecological,* Wiley-Interscience: New York

Pollock, K.H(1974). The assumption of equal catchability of animals in tag-recapture experiments. PHD. Dissertation, *Cornell Univ. Ithaca,N.Y.*

Pollock, K.N.(1975a). Building models of capture-recapture experiments. *The statistician, 25,253-60.*

Pollock, K.H.(1275b). A k-sample tag-recapture model allowing for unequal survival and catchability. *Biometrika 62, 577-83*

Rafail, S.Z (1972a) A further contribution to the study of fish populations by capture-recapture experiments. *Marine Biology 14, 338-40.*

Regier, N.A. and Robson, D.S. (1967) Estimating population number and mortality rates. In S.D. Gerking (ed.), *The Biological Basis of Freshwater Fish Production, 31-66.* Blackwell Scientific Publications: Oxford.

Ricker, W.E.(1956). Uses of marking animals in ecological studies: the marking of fish. *Ecology 37,665-70.*

Ricker, W.E. (1958). Handbook of computations for biological statistics of fish populations. *Bull.Fish.Bd Canada 119, 1-300*

Robson, D.S.(1969). Mark-recapture methods of population estimation Fh N.L. Johnson and H. Smith J. (Editors) *New Development in survey sampling, 120-40.* Wiley Interscience, Wiley and Sons : New York.

Robson, D.S. and Flick, W.A.(1965) A non parametric statistical method for calling recruits from a mark-recapture experiment. *Biometrics 21,936-47.*

Robson, D.S. and Regier, H.A(1964). Sample size in Petersen mark-recapture experiments *Trans. Amer. Fish. Soc. 93, 215-26*

Samwel, E. (1968). Sequential likelihood estimation of the size of a population. *Ann. Math. Statist. 39,1057-68*

Sanathanan, L.(1972). Estimating the size of a multinomial population. *Ann, Math. Statist. 43, 142-52.*

Seber, G.A.F.(1962). The multi-sample recapture census. *Biometrika 49, 330-49.*

Seber G.A.F (1965). A note on the multiple recapture census *Biometrika 52, 249-259.*

Seber G.A.F. (1970a). The effects of trap response on tag-recapture estimates. *Biometrics 26,13-22*

Seber G.A.F.(1982).*The estimation of animal abundance. Griffin and Company London.*

Tanaka, R.(1967). New regression formula to estimate the whole population for recapture addictive small mammals. *Res. Popul. Ecol. 9, 83-94*

Welch, H.E.(1960) Two applications of a method of determining the error of population estimates of mosquito larvae by the mark and recapture technique. *Ecology 41,228-9.*

Young H. Nees. J. and Emlen, J.T.Jr.(1952). Heterogeneity of trap response in a population of house mice. J. *Wildl.Manag.16, 169-80.*

Zippin, C.(1958). The removal method of population estimation *J. Wildl. Manag. 22,82-90.*