POPULATION ESTIMATION FROM MARK-RECAPTURE EXPERIMENTS USING A SEQUENTIAL BAYES ALGORITHM[1]

W. J. GAZEY

LGL Environmental Research Associates, Limited, Sidney, British Columbia V8L 3Y8, Canada

AND

M. J. STALEY

ESSA Limited, Vancouver, British Columbia V5Z 1G6, Canada

Abstract. Traditional analyses (e.g., Schnabel 1938 or Chapman 1954) of sequential mark-recapture experiments (Petersen and Schnabel type) yield population estimates with substantial negative bias and overly large confidence intervals if the combination of the number of animals marked and examined falls too low. To address these problems, sequential mark-recapture experiments are cast into a Bayesian framework using a "noninformative" discrete uniform improper prior (a priori theoretical) distribution. Some properties of the posterior distribution (probability of each population size given the data) are briefly and informally discussed (inference, convergence, mean, mode, median, and treatment of nuisance parameters). A sequential Bayes computational algorithm, suitable for microcomputers, is given. Several examples are presented as a practical guide to computing estimates.

For relatively small sample sizes, the Bayesian approach yields larger mean abundance estimates than traditional methods. There is little difference in these estimates for larger sampling efforts. Advantages of the approach include the following: the probability of observing the data at all feasible population sizes is calculated exactly; the method works for all cases regardless of sample size or sampling procedure; a plot of successive posterior distributions can be used as a visual diagnostic of conformity with basic assumptions; and finally, inferences can be made directly, since the end product completely describes the uncertainty of the population size given the data.

Key words: Bayesian inference; capture-recapture; improper prior; likelihood; mark-recapture; nonparametric; population estimation; sequential Bayes.

INTRODUCTION

The idea of obtaining information about a population by marking or labeling some of its members can be traced back several centuries. The simplest mark-recapture procedure is the well-known two-stage Petersen experiment, which can be extended to multiple recapture census in which consecutive samples are taken from the population. The theory to develop population estimates from sequential samples has been discussed by Schnabel (1938), Schumacher and Eschmeyer (1943), Chapman (1948, 1951, 1954), Bailey (1951), DeLury (1951, 1958), Cormack (1968), Seber (1973, 1982), and others.

While the published methodologies perform well if large samples are obtained relative to the population level, they have difficulty when the sample is small (see Cormack 1968 or Seber 1973). For example, the lower bounds of calculated confidence intervals of a population estimate are sometimes less than the unique number of animals sampled, but for reporting purposes are rounded upwards to the number of animals actually marked. Chapman (1951), Robson and Regier (1964), Ricker (1975), and others have pointed out that a pronounced negative bias will occur if the combination of

the number of animals marked and the total number later examined falls too low.

To overcome the above difficulties, Bayesian arguments were used to obtain an estimate of the number of Illinois mud turtles (*Kinosternon flavescens spooneri*) utilizing a relict prairie habitat, Big Sand Mound, located in Iowa (J. Bickham and B. J. Gallaway, *personal communication*). The turtle (since found by Houseal et al. [1982] and Berry and Berry [1984] to be synonymous with the yellow mud turtle, *Kinosternon flavescens flavescens*) had been proposed to be listed as rare and endangered based in part on the assumption that there were no more than 650 living representatives spread among three habitats in Iowa and Illinois, most living at Big Sand Mound, Iowa. The issue of the actual number of turtles occupying the Big Sand Mound habitat became one of the focal points of controversy surrounding the proposed listing.

A sequential mark-recapture experiment was conducted during the spring and summer of 1979 (M. D. Springer and B. J. Gallaway, *personal communication*) to address the question; however, only limited numbers of turtles could be captured. (From 88 marks applied only five recaptures were made, with never more than one recapture in a sampling period.) Strong criticisms of subsequent population estimates were leveled, mainly because of the extremely broad confidence in-

tervals (88 to 9411 turtles) associated with the traditional estimates. In any case, orthodox estimation procedures could not cope with small sample sizes, even though the basic assumptions for a mark-recapture experiment had been well met (J. Bickham and B. J. Gallaway, *personal communication*).

While seemingly novel, the Bayesian approach to obtain population estimates was first used by LaPlace in the eighteenth century (Stigler 1975). More recently, Freeman (1973*a, b*), Gaskell and George (1972), and Carle and Strub (1978) have applied Bayesian methods to a variety of population estimation problems. Johnson (1977) provides a useful review of elementary applications of Bayesian statistics to problems faced by wildlife biologists.

Subsequent study and extension of the method has revealed that it is broadly applicable to the mark-recapture experiment. While the approach is intensive in computation, the advent of cheap, universally available computing power has made the algorithm both tractable and convenient. The purpose of this paper is to cast mark-recapture experiments into a Bayesian framework from which estimation properties may be determined from the well-developed theory of Bayesian inference in statistical analysis. In addition, a simple, efficient algorithm is presented for the calculation of the posterior probability distribution and is illustrated with some examples.

AVAILABLE INFORMATION AND ASSUMPTIONS

If indistinguishable marks or labels are applied to individuals from a population with consecutive samples, and these samples are examined for recaptured marks, then the following information is available:

M_t = total marked animals at large at the start of the tth sampling interval;

C_t = total number of individuals sampled during interval t; and

R_t = number of recaptures in the sample C_t.

To make any form of mathematical expression tractable, a number of assumptions must be satisfied. These assumptions have been presented in various forms and discussed by a number of authors: Cormack (1968), Overton (1971), Seber (1973), and Ricker (1975). However, from a practical viewpoint they amount to the following:

1) The population is closed, so the population size does not change over the period of the experiment.
2) The probability of capturing a marked individual at any given time is equal to the proportion of marked members in the population at that time.
3) Animals do not lose their marks over the period of the study.
4) All marks are reported on recovery.

For brevity, assumption (1) was stated strongly. Historically, a number of alternatives have been proposed

such that the population size may change over the period of the experiment:

1a) There is neither recruitment nor immigration but death and emigration affect marked and unmarked individuals equally.
1b) There is recruitment and/or immigration but neither death nor emigration.
1c) Knowledge is available from other sources such that adjustments can be made for migration, recruitment, and death prior to analysis of the data.

Under these assumptions, the probability that C_t contains R_t individuals given a particular population size (N_i) can be described by either of three distributions (hypergeometric, negative binomial, and binomial):

$$P(R_t \mid N_i) = \binom{M_t}{R_t}\binom{N_i - M_t}{C_t - R_t} \Big/ \binom{N_i}{C_t} \tag{1a}$$

$$P(C_t \mid N_i) = \binom{C_t - 1}{R_t - 1}\left(\frac{M_t}{N_i}\right)^{R_t}\left(1 - \frac{M_t}{N_i}\right)^{C_t - R_t} \tag{1b}$$

$$P(R_t \mid N_i) = \binom{C_t}{R_t}\left(\frac{M_t}{N_i}\right)^{R_t}\left(1 - \frac{M_t}{N_i}\right)^{C_t - R_t} \tag{1c}$$

The hypergeometric distribution (Eq. 1a) is appropriate if the sample C_t is taken without replacement (i.e., C_t is taken and then searched for marks) while the binomial (Eq. 1c) holds if the sample C_t is taken with replacement (i.e., the animals are taken one at a time, examined for a mark, and then released). The negative binomial distribution (Eq. 1b) is a special case appropriate for inverse sampling (i.e., the number of recaptures to be obtained is fixed in advance, and sampling is stopped as soon as that number is obtained). Note that a slight change in notation must be made since the sample size (C_t) is the variable of interest.

In general, we will assume that sampling is conducted with replacement (Eqs. 1b or 1c); however, since C_t is usually small in relation to the total population, (Eq. 1c) is usually a very good approximation to (Eq. 1a) regardless of the sampling procedure.

BAYESIAN ESTIMATION OF POSTERIOR PROBABILITY

According to the Bayesian view, all quantities are of two kinds: those known to the person making the inference and those unknown to the person. The uncertainty of the latter is described by a probability distribution. The fundamental reason for adopting the Bayesian position is that, once the basic step of describing uncertainty through probability is admitted, we have a formal procedure for solving all inference problems.

The only known information is the set of M_t, C_t, and R_t obtained over T sampling intervals. The probability of observing all the R_t's given some population level N_i (assuming Eq. 1a or 1c) is then:

$$P(R_1, R_2, R_2, \ldots, R_T | N_i) = \prod_{t=1}^{T} P(R_t | N_i) \quad (2)$$

since by assumption (2) each sample is independent of the other. Eq. 2 is termed the sampling distribution or likelihood of the data given N_i.

On the other hand, the distribution of the population size is unknown before the experiment is conducted. Mathematically, this is represented by a "noninformative" discrete uniform distribution, i.e.,

$$P(N_i) = \frac{1}{K} \quad (3)$$

where K is the number of discrete population levels (N_i, $i = 1, 2, \ldots K$) considered possible a priori. The only condition that needs to be imposed on Eq. 3 is:

$$N_i \geq \text{Max } M_t,$$

i.e., the smallest feasible population size must be equal to or greater than the number of animals marked. The upper bound (N_K) of the feasible population size can be finitely large. Eq. 3 is termed the prior distribution.

The sampling and prior distributions are combined to form the posterior distribution (probability of each N_i given the data) by using Bayes' theorem. This theorem states that if $B_1, B_2, \ldots,$ and B_K constitute a set of mutually exclusive events of which one must occur and none has a zero probability, then for any event A for which $P(A) \neq 0$

$$P(B_r | A) = \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^{K} P(B_i)P(A | B_i)}$$

for $r = 1, 2, \ldots, K$. For the event "population size is N_i," Eqs. 1c, 2, and 3 can be substituted into Bayes' theorem to give:

$$P(N_i | R_1, R_2, \ldots, R_T)$$

$$= \frac{P(N_i)P(R_1, R_2, \ldots, R_T | N_i)}{\sum_{i=1}^{K} P(N_i)P(R_1, R_2, \ldots, R_T | N_i)}$$

$$= \frac{\prod_{t=1}^{T} P(R_t | N_i)}{\sum_{i=1}^{K} \prod_{t=1}^{T} P(R_t | N_i)}$$

$$= \frac{\prod_{t=1}^{T} \left(\frac{1}{N_i}\right)^{R_t} \left(1 - \frac{M_t}{N_i}\right)^{C_t - R_t}}{\sum_{i=1}^{K} \prod_{t=1}^{T} \left(\frac{1}{N_i}\right)^{R_t} \left(1 - \frac{M_t}{N_i}\right)^{C_t - R_t}}. \quad (4)$$

If the negative binomial sampling distribution (Eq. 1b) is assumed, the same result if obtained; however, the notation in the first two lines in Eq. 4 must be changed by replacing the R's with C's. In other words, the above scheme is valid for either direct or inverse sampling. In fact, it does not matter how M_t and C_t are obtained (e.g., fixed vs. random variables) within the conditions under which Eq. 1 was derived.

Eq. 4 describes the uncertainty of the population size given the data. The distribution is defined completely by the numerator (the likelihood function), with the denominator serving merely as a normalizing constant to ensure that the distribution sums to one. In addition, the discrete uniform prior distribution only affects the magnitude of this normalizing constant. In other words, the posterior distribution will be proportional to the "true" distribution over the set ($N_1, N_2, \ldots, N_K$) evaluated.

The prior used in Eq. 4 is termed an improper prior because it is not defined over the entire parameter space feasible for N_i (i.e., $M_t \leq N_i < \infty$). The error between the actual posterior and the computed posterior is illustrated in Fig. 1. This error can be made arbitrarily small by increasing the range from N_1 to N_K. It can be shown that the probability of the difference is at most $3A/(1 - A)$, where A is the shaded area in Fig. 1 (DeGroot 1970:201).

PROPERTIES OF THE POSTERIOR DISTRIBUTION

The posterior distribution can be calculated for all N given the data from a mark-recapture study. The properties of this distribution have been studied intensively in the Bayesian literature. Some of the properties pertinent to the mark-recapture experiment are given below without proof (see DeGroot 1970, Chapters 10 and 11, for an axiomatic treatment).

Inference

Every decision problem concerning the population size and the data from a mark-recapture study requires only the posterior distribution as the contribution from the data. Thus, inference is available for any decision about the population size. For example, the probability that N is smaller or larger than a specified value can be obtained directly by summation over the domain of interest. Notice that this is a direct probability statement about the population size (unlike the easily misunderstood statements provided by confidence intervals). Comparisons between populations can also be made directly (cf. Chapman 1951, Chapman and Overton 1966, Skalski et al. 1983) by obtaining the compound distribution of the difference in population size, i.e.,

$$U = N_1 - N_2$$

where U is the difference between populations N_1 and N_2. The smallest value that U can obtain is

$$U_1 = N_{1,1} - N_{2,K_2}$$

and the largest

$$U_{K_1 + K_2 - 1} = N_{1,K_1} - N_{2,1},$$

where K_1 and K_2 are the number of evaluations for populations 1 and 2, respectively. If the evaluation points of the two populations have equal increments, i.e.,

$$(K_2 - 1)(N_{1,K_1} - N_{1,1}) = (K_1 - 1)(N_{2,K_2} - N_{2,1})$$

then the compound posterior distribution, $P(U_j)$, can be computed by:

$$P(U_j) = \sum_{i=L}^{H} P_1(N_{1,i})P_2(N_{2,i-j+K_2}) \quad (5)$$

for $j = 1, 2, 3, \ldots, K_1 + K_2 - 1$

where $P_1(\bullet)$ and $P_2(\bullet)$ are the posterior probabilities (Eq. 4) for the two populations and

$$L = \text{Max} \, (j - K_2 + 1, 1)$$

and

$$H = \text{Min} \, (j, K_1).$$

Finally, minimum population estimates and their associated precision are often of critical interest for rare species. Once the posterior distribution of population size has been determined, the probability that some population level V_j is less than the true population size can be simply constructed as the complement of the cumulative density, i.e.,

$$P(V_j < N) = 1 - \sum_{i=1}^{j} P(N_i). \quad (6)$$

Convergence

The posterior distribution of the population size tends to become more and more concentrated about the true population size with a normal distribution as T (number of sequential samples) or C_t (sample size) increases. Therefore, all statistics formed from the posterior distribution are consistent.

The mean

The mean is calculated as:

$$\bar{N} = \sum_{i=1}^{K} N_i P_T(N_i) = E[N|R]$$

where $P_T(N_i) = P(N_i|R_1, R_2, \ldots R_T)$ is the posterior distribution.

The mean itself is a random variable which is the expectation of the posterior and minimizes the mean square error. Since

$$E[N] = E[E[N|R]],$$

the expectation of the mean equals the expectation of the prior distribution to the extent that the computed and actual posterior are equivalent.

The mode

The mode of the posterior distribution is the maximum likelihood estimate of the sampling distribution. It is equivalent to the iterative methods recommended

by Schnabel (1938) and DeLury (1951) to obtain a maximum likelihood population estimate.

The median

The value $\tilde{N}$ is a median of the posterior distribution if $P(N \geq \tilde{N}) \geq 0.5$ and $P(N \leq \tilde{N}) \leq 0.5$. The median minimizes the absolute value of the error.

Nuisance parameters

While the focus of this paper is upon satisfaction of the four basic assumptions, additional parameters (termed nuisance parameters) can be incorporated into the sampling distribution to represent a failure of one or more of the basic assumptions. In theory, these parameters could be integrated from the posterior distribution (e.g., Carle and Strub 1978) or estimated from the available data. We suspect additional research will yield improved estimates of mortality, recruitment, and population size for experiments where individuals have heterogeneous capture probabilities.

On the other hand, the suspected presence of nuisance parameters may be ignored. If the nuisance parameters do not markedly affect population size, then they can be treated as irrelevant. For example, a common violation of assumption (1) is that the population size changes over the period of the study. It turns out that the probability $P(N_i|R_1, R_2, \ldots, R_T)$ is insensitive to the change in the parameter of interest (i.e., population size) provided that the variance of the posterior distribution is large in comparison to the variance of the population size. Therefore, alternative assumptions (1a) and (1b) are only acceptable if the change in population size is small in relation to the precision obtained from the study. Nevertheless, there may be some utility in proceeding with the analysis even with a substantial change in population size, as point estimates of population level with some measure of precision would be applicable at some (usually unknown) period during the study (assumption 1a implies a monotonically decreasing population, while 1b implies a monotonically increasing population). For example, if all marks were applied before any sampling for recaptures commenced (i.e., $t = 0$) then the computed posterior would be applicable at time $t = 0$.

COMPARISON OF POINT ESTIMATES

Much effort has traditionally been expended to find estimates that are approximately unbiased. If an estimator, $\hat{N} = f(x)$, is unbiased, then

$$E[f(x)] = \sum_{x} f(x)P(x) = N^* \quad (7)$$

where x is the data used in the estimator and N^* is the true population size. The Bayesian points out that the summation occurs over data not observed in the sample. Expressed in another fashion, the concept of an unbiased estimator demands

$$N = \sum_x f(x)P(x|N) \qquad (8)$$

for all N. A series of values of N are taken one at a time, whereas the Bayesian implies that many values should be evaluated simultaneously, since the observed data could arise from a variety of "true" population levels, i.e.,

$$P(x) = \sum_N P(N, x) = \sum_N P(x|N)P(N).$$

Setting $f(x) = \sum NP(N|x)$ we see, using the above, that

$$\sum_x f(x)P(x) = \sum_N NP(N)$$

and the posterior mean, $f(x)$, satisfies Eq. 7 if $E[N] = N^*$.

The point of the above discussion is that the comparison of point estimates between the two approaches cannot be done in the absence of a common base. For example, Chapman (1951) demonstrates that Eq. 8 holds for the Petersen experiment ($T = 1$) when

$$\hat{N} = \frac{(M_1 + 1)(C_1 + 1)}{R_1 + 1} - 1 \qquad (9)$$

for $N < M_1 + C_1$. Since this is less than the maximum likelihood estimator and Eq. 4 will usually be skewed right ($M_t/N < 0.5$), the traditional estimate Eq. 9 and the point estimates derived from the Bayesian posterior recommended in this paper will usually be in the order:

$$\text{Traditional} \le \text{Mode} \le \text{Median} \le \text{Mean}$$

Since $N \gg M_1 + C_1$ for small samples, the bias of the traditional estimate is of order $-N \exp[-E(R_1)]$ (Seber 1982), which can become large if $E(R_1)$ is small.

SEQUENTIAL BAYES ALGORITHM

Eq. 4 can be written as a sequential Bayes relationship, i.e.,

$$P_{t+1}(N_i) = \frac{P(R_{t+1}|N_i)P_t(N_i)}{\displaystyle\sum_{i=1}^{K} P(R_{t+1}|N_i)P_t(N_i)} \qquad (10)$$

for $t = 0, 1, 2, \ldots, T - 1$ and $i = 1, 2, 3, \ldots, K$ where $P_0(N_i) = 1/K$ and $P_t(N_i) = P(N_i|R_1, R_2, R_3, \ldots, R_t)$.

This recursive relationship leads to the following simple algorithm for computation of the posterior distribution:

1) Set the lower and upper bounds and the number of points to be evaluated (N_1, N_K, K).
2) Let $N = \{N_1 + (i - 1)(N_K - N_1)/(K - 1)\}$.
3) Let $P = 1$.
4) Let $t = 0$.
5) Let $t = t + 1$.
6) Let $P = \{(M_t/N_i)^{R_t}(1 - M_t/N_i)^{C_t - R_t}P_i\}$.

7) Let $a = \text{SUM}(P)$.
8) Let $P = P/a$.
9) Plot P vs. N.
10) If $t < T$ go to step 5.
11) Calculate statistics of interest from P and N.

where,

$N = \{N_i, i = 1, 2, 3, \ldots, K\}$ = population size;
$P = \{P_i, i = 1, 2, 3, \ldots, K\}$ = posterior distribution;
$M = \{M_t, t = 1, 2, 3, \ldots, T\}$ = number of marks;
$C = \{C_t, t = 1, 2, 3, \ldots, T\}$ = sample size; and
$R = \{R_t, t = 1, 2, 3, \ldots, T\}$ = number of recaptured marks.

It is assumed that the data from the mark-recapture experiment have been read into vectors M, C, and R (all dimensioned T). Two additional vectors are required (dimensioned K) for the population sizes to be evaluated (N) and the posterior distribution (P). Steps (1) to (4) in the algorithm set the population vector to K equally spaced values for evaluation, initialize the posterior distribution, and set the sequential loop counter to 0. The recursive loop begins at step (5) by incrementing the counter. Next, the numerator of Eq. 10 is calculated. Note that the binomial factor can be written more concisely as $(N_i - M_t)^{C_t - R_t}/N_i^{C_t}$, but this formulation may lead to machine overflow. Steps (7) and (8) normalize the posterior distribution for plotting in step (9). After the final sequence any statistics or inferences of interest can be calculated.

This algorithm is appropriate for the binomial and negative binomial sampling distributions. The hypergeometric sampling distribution (1a) can be included by simply replacing step (6) with:

$$\text{Let } P = \{f_iP_i\} \qquad (6a)$$

where

$$f_1 = \left[\prod_{i=1}^{C_t - R_t} \frac{N_1 - M_t - i + 1}{N_1 - i + 1}\right]\left[\prod_{i=1}^{R_t} \frac{C_t - i + 1}{N_1 - C_t + i}\right]$$

$$f_{i+1} = f_i \prod_{j=1}^{I} \left(\frac{N_i - M_t + j}{N_i + j}\right)\left(\frac{N_i - C_t + j}{N_i - C_t - M_t + R_t + j}\right)$$

$$I = \frac{N_K - N_1}{K - 1}.$$

The plot of the posterior distribution at each iteration can be used as a visual diagnostic. The distribution should tend to stabilize or mass about a single value. Thus, a continuous trend towards a larger or smaller population size is strong evidence that assumption 1 has been violated, i.e., an increase or decrease of the population over the period of the study. The change in the distribution from one sequence to the next is a measure of the information added by the sample. In addition, a relative measure of the effect of the prior on the posterior distribution is the distance from the axis at N_1 and N_K.

TABLE 1. Simulated example of a study of a population of 10 000 animals in perfect compliance with the four basic assumptions. In $t = 0$, 50 animals were marked and returned to the population.

A. The generated data.

Sampling interval (t)	No. animals sampled in interval (C_t)	Marked animals at large at start of interval (M_t)	Recaptures in sample C_t (R_t)
1	34	50	0
2	42	84	1
3	43	125	0
4	40	168	1
5	32	207	0
6	56	239	1
7	42	294	1
8	44	335	4
9	56	375	3
10	44	428	1

B. Population estimates using traditional methods.

Estimation method	Estimated N	95% Range
Original Schnabel (1938)	8688	5256–25 035
Modified Schnabel (Chapman 1954)	8019	4964–16 814
Schumacher and Eschmeyer (1943)	8498	5596–17 652
Delury (1951) weighting formulae	8681	

C. Bayesian estimates.

Parameters	Value
Mean	10 355
Median	9700–9750
Mode	8650–8750
2.5 and 97.5% quantiles	5650–18 600
95% Highest Probability Density (HPD)	5000–17 100

The number of evaluations (K) should reflect the number of significant digits desired for reporting purposes. For example, quantiles (e.g., median) can be reported to the nearest interval between points, while the mode can only be reported to be in the region of two intervals (an exception is when every integer between the bounds is evaluated). The resolution of the mean will be approximately $I/\sqrt{K}$, where I is the interval between points. Since these rules are conservative and approximation through linear interpolation improves as K increases, the use of very large K is unwarranted. Therefore, we recommend $100 < K < 1000$.

We also recommend that an interactive computer environment be used for analysis, since a suitable range of the prior can be found quickly by trial and error. As long as the posterior probability at N_1 and N_K remains small, there will be negligible difference between posteriors computed with alternative priors (see Fig. 1). A convenient strategy is to locate an appropriate range of the prior using few evaluations (e.g., $K = 100$) and then increase K over the established range to obtain the desired number of significant digits.

A standardized probability region (analogous to a confidence interval), i.e., $P(a \leq N \leq b) = 1 - \alpha$, where a and b are bounds and $1 - \alpha$ a fixed probability, can be defined in two possible ways. First, the appropriate quantiles can be reported, i.e., the a and b such that $P(N < a) = \frac{\alpha}{2}$ and $P(N > b) = \frac{\alpha}{2}$, or second, the highest probability density (HPD), i.e., the a and b such that $b - a$ is minimal and $P(a \leq N \leq b) = 1 - \alpha$. The methods produce identical results when the distribution is symmetrical.

EXAMPLES

A simulated example

A sequential mark-recapture experiment was simulated to comply exactly with the four basic assumptions (Table 1A). In 10 subsequent random samples the recaptures were computed as a binomial deviate of the proportion of marked animals (i.e., Eq. 1c). Approximately 5% of the population was sampled. Table 1 shows the population estimates derived from traditional and Bayesian methodologies.

The posterior distribution was computed with 501 evaluations between the bounds of 3000 and 28 000, i.e., an interval size of 50. The successive distributions are plotted in Fig. 2 over the range of the prior. Note that the last few iterations can be recognized by their increased concentration about the true population level of 10 000, and the final posterior (heavy line) cannot be distinguished from the axis at the tails since the probabilities at the tails were on the order of 10^{-20}.

This example, where sample size with respect to the population level is relatively small, was generated to illustrate three points. First, there can be substantial differences between the traditional point estimates and those derived from the Bayesian approach ($\approx 20\%$ for this example), although as the sample size increases the difference will become minimal. Second, the posterior distribution resembles a Poisson distribution if
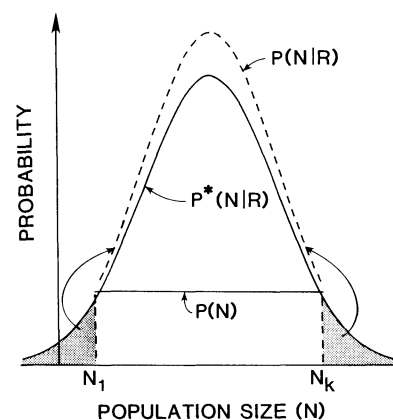

FIG. 1. Effect of the improper prior on the posterior distribution [$P^*(N|R)$ = actual posterior, $P(N|R)$ = computed posterior, $P(N)$ = prior].
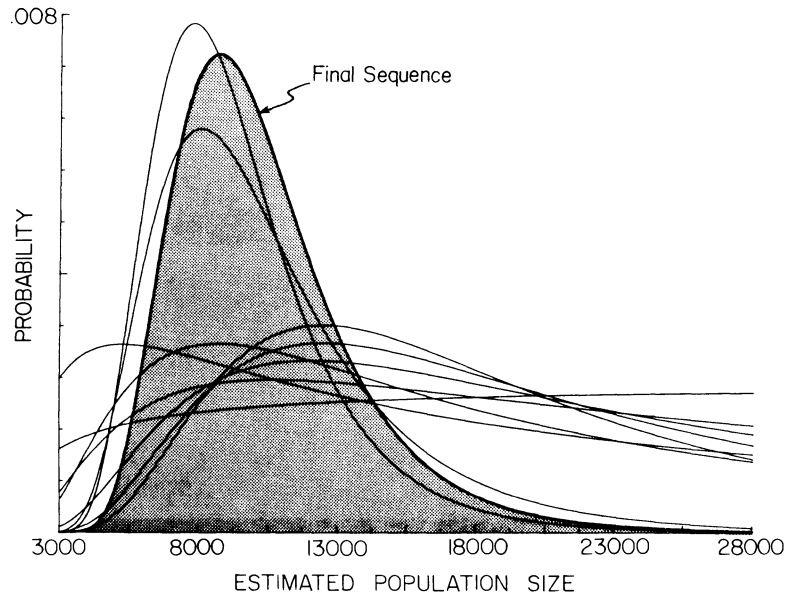
FIG. 2. Sequential posteriors for a simulated example in which the true population size is 10 000 (N_1 [smallest population size considered] = 3000, N_K = 28 000; K [number of discrete population levels] = 501).

sample size is small (compare the Bayes HPD region with the modified Schnabel confidence region, which assumes a Poisson distribution for total recoveries). Finally, the binomial simulation of recoveries does induce shifts in the sequential posterior; therefore, some caution must be used when it is applied as a diagnostic for a closed population.

The best evidence for a closed population is usually biological rather than statistical. However, for large sample sizes the ability to detect real change in the population is enhanced. For example, Fig. 3A displays the sequential posteriors from an initial population of 1000 with 200 marks applied. In 10 subsequent fixed samples ($C_t = 50$), the recaptures were computed as a binomial deviate of the proportion of marked animals in a population expanding (i.e., immigration or recruitment) at the rate of 5% per period. It was again assumed that the entire sample was marked and returned to the population. Similarly, Fig. 3B displays a scenario for a decrease (i.e., mortality or emigration) of 5% for both the marked and unmarked population using the same initial conditions. Note that sensitivity of the posterior is greater for an increasing population than a decreasing population since only the unmarked population is affected in the former.

Illinois mud turtle

Table 2 displays the mark-recapture data and Bayesian population estimates that motivated this paper for the Illinois Mud Turtle at the Big Sand Mound. Compliance with the basic assumptions for a mark-recapture experiment was believed to be excellent. Nevertheless, because of extremely small sample sizes and

the sparsity of recaptures, the standard methods are not applicable (Otis et al. 1978). Blind application of those methods leads to strong negative bias and calculated standard errors in excess of the associated statistics. For example, the lower bound of the traditional confidence interval (88–9411 turtles) was rounded upward to the number of animals actually marked. On the other hand, the application of the sequential Bayes algorithm is straightforward and gives reasonable results (Fig. 4 and Table 2B display the 39 iterations.)

Fig. 5 presents minimum population estimates and their associated precision (i.e., Eq. 6), which, for rare species, is of critical interest. For example, from the curve we can determine that there is a 0.95 probability that the abundance of turtles was at least 640.

Nuttall's cottontail

Skalski et al. (1983) presented a methodology for computing confidence interval estimators of proportional abundance of two populations. To illustrate their computations they used a comparative census of a population of Nuttall's cottontail (*Sylvilagus nuttallii*) conducted during 1974 and 1975. In August of each year cottontails were live-trapped, marked with picric acid (which stains the fur yellow), and released. In the following month a drive census was conducted in which the possibility for multiple sightings of animals existed. While the authors acknowledge that the possibility of multiple sightings invalidates their procedure, this census procedure makes the use of the binomial sampling distribution formally correct. In 1974, 87 animals were marked (M) and 14 animals were sighted during the subsequent drive (C), of which 7 were marked (R). In
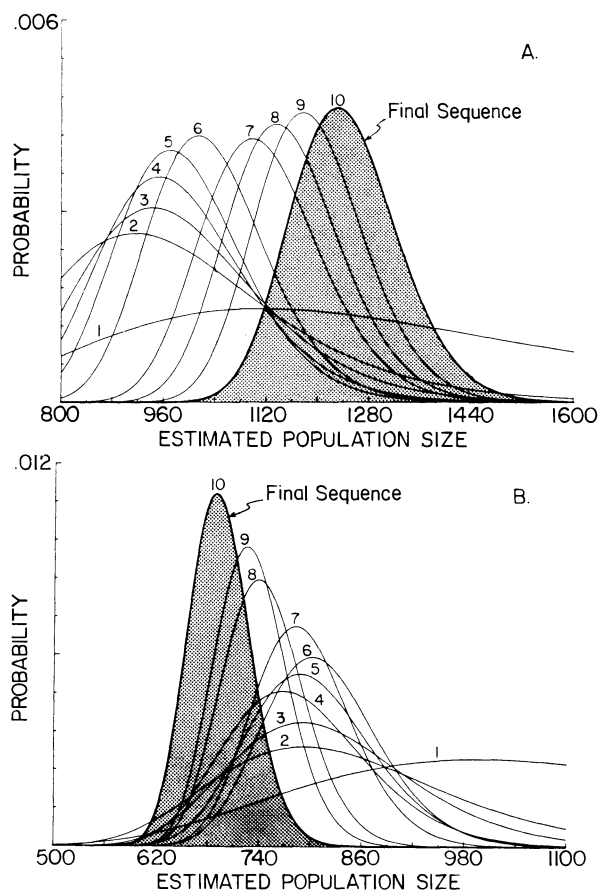
FIG. 3. Simulated sequential posterior distributions for: (A) an increasing population ($N_1 = 800$, $N_K = 1600$, $K = 801$) and (B) a decreasing population ($N_1 = 500$, $N_K = 1100$, $K = 601$). Symbolism as in Fig. 2. The order of sampling sequences is indicated.

1975, the comparable census produced the following results: $M = 101$, $C = 27$, and $R = 17$. Fig. 6 displays the computed posterior distributions for each year.

The distribution for the ratio of 1974 abundance to 1975 abundance (Skalski's K_1) is displayed in Fig. 7A. The distribution was computed by summing the joint probability of the two years for all combinations of population levels evaluated into discrete ratio intervals 0.005 apart.

Fig. 7A demonstrates that the distribution of the ratio is not smooth. Moreover, the probability values are not exact. For example, the probability that the population increased in 1975 is 0.29 (shaded area in Fig. 8A); this is only accurate to the extent that a ratio of 0.9975 to 1.0025 is acceptably close to 1.

A better method to examine the change in population level is simply to compute the compound distribution of the difference between 1974 and 1975 (i.e., Eq. 5). Fig. 7B displays this distribution. It is important to note that since the same basis (i.e., identical prior distributions) was used to form the posteriors, the com-

pound distribution is exactly the probability of differences between the two years given the data. The probability that the population increased in 1975 (given the data) is again 0.29.

If it is judged that there has not been a substantial change in the population level in the two years, an appropriate procedure for describing the uncertainty of the population level after the 1975 census is to use the 1974 census as prior knowledge and update it with the 1975 census (Fig. 6).

TABLE 2. Illinois mud turtle example (J. Bickham and B. J. Gallaway, *personal communication*).

A. Field data.

Sampling occasion (t)	No. animals sampled in interval (C_t)	Marked animals at large at start of interval (M_t)	Recaptures in sample C_t (R_t)
1	1	2	0
2	3	3	0
3	2	6	0
4	9	8	0
5	3	17	0
6	4	20	0
7	1	24	0
8	2	25	0
9	1	27	1
10	3	27	0
11	2	30	0
12	5	32	0
13	3	37	1
14	5	39	1
15	2	43	0
16	3	45	0
17	3	48	0
18	5	51	0
19	1	56	0
20	3	57	0
21	1	60	0
22	1	61	0
23	2	62	1
24	6	63	0
25	2	69	0
26	1	71	0
27	4	72	1
28	1	75	0
29	1	76	0
30	6	76	0
31	3	76	0
32	3	77	0
33	1	79	0
34	1	80	0
35	1	81	0
36	4	82	0
37	1	86	0
38	1	87	0
39	1	88	0

B. Bayesian estimates

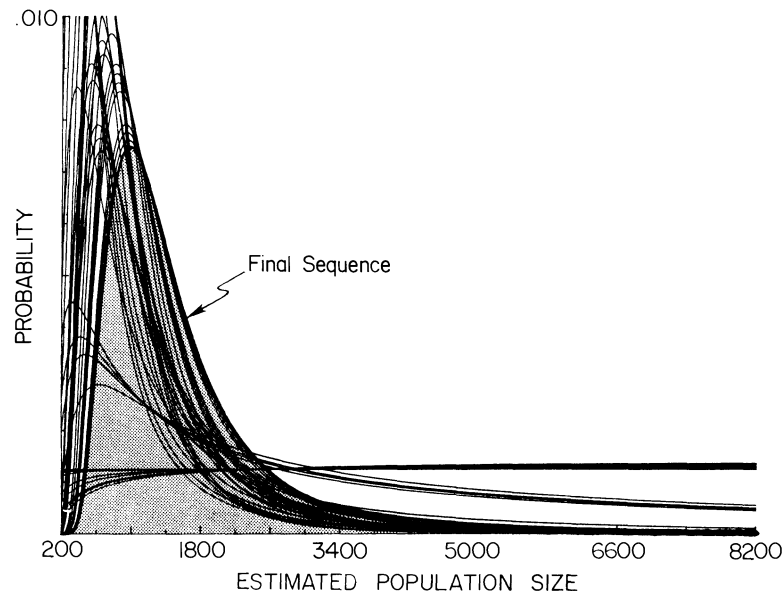| Parameters	Value
Mean	1574
Median	1300–1310
Mode	960–980
2.5 and 97.5% quantile	560–4210
95% Highest Probability Density (HPD)	410–3470

FIG. 4. Sequential posterior distributions for the Illinois mud turtle at Big Sand Mound ($N_1 = 200$, $N_K = 8200$, $K = 801$). Symbolism as in Fig. 2. There were 39 iterations of the Bayesian algorithm.

DISCUSSION

In this paper we have noted that, for mark-recapture experiments with relatively small sample sizes, there will be differences in point estimates obtained from Bayesian vs. traditional methods. These differences reflect the degree to which there is a lack of unity at the very foundation of statistics. A large body of literature comparing sampling theory and Bayesian theories of inference and decision is available. Comparative discussions of the two approaches (as well as other schools of inference) can be found, for example, in Kempthorne and Folks (1971), Barnett (1973), and Cox and Hinkley (1974).

From a practical viewpoint, the approach advocated in this paper has many attractive features. First, the

population distribution is obtained directly by calculating the probability of observing the data at all feasible population sizes. Further, the sequential Bayes algorithm has an exact correspondence to the physical sequence of marking and recapturing animals. While this "blunt nose" approach is intensive in computation, widely available computing power (e.g., microcomputers) makes the computational algorithm both tractable and convenient. Second, the method works for all cases regardless of the sample size, number of sequences, sampling procedures, or whether some quantities are regarded as fixed or random (e.g., number of marks applied). In other words, the user may focus only upon the validity of the basic assumptions
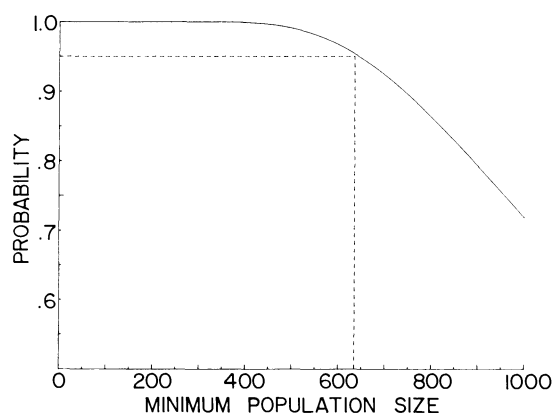

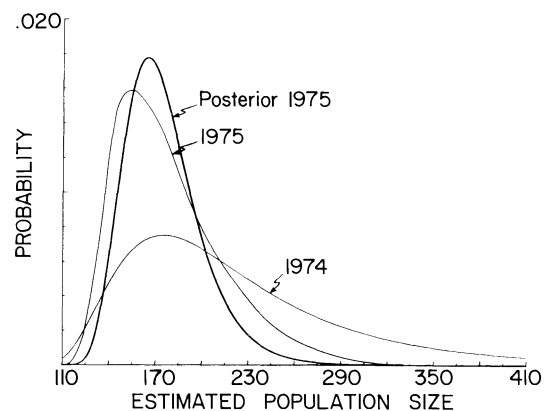FIG. 5. Minimum population size of the Illinois mud turtle at Big Sand Mound.

FIG. 6. Posterior distributions for Nuttall's cottontail during 1974, 1975, and after 1975 (combined data from 1974 and 1975).
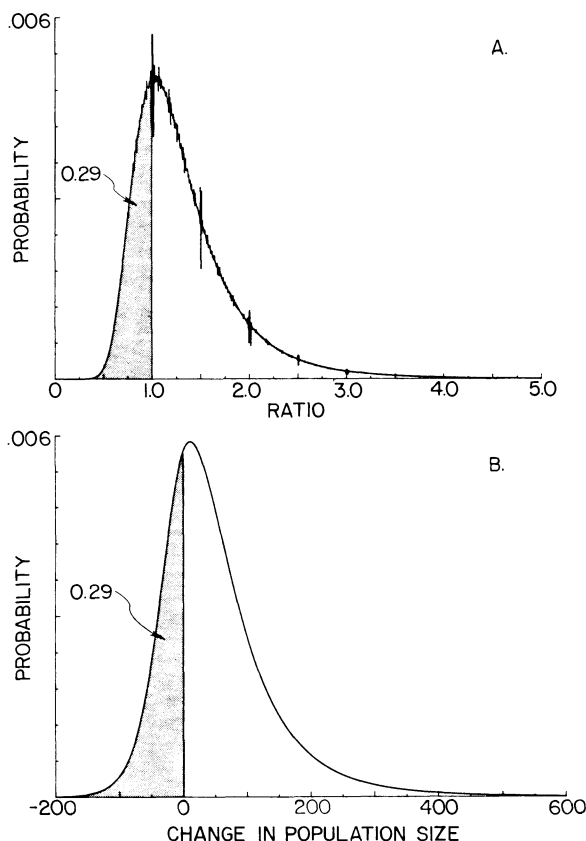
FIG. 7. Compound probability distribution of the change in population of cottontails between 1974 and 1975. (A) As a ratio (1974/1975). (B) as a simple difference.

used to generate the sampling distribution (Eq. 1) without concern for other approximations. Third, the plot of the posterior distribution at each iteration can be used as a visual diagnostic of closure, the information added by each sampling sequence, and the effect of the prior distribution. Finally, the end product completely describes the uncertainty of the population size given the data. This becomes important when inferences are to be made about the population size. For example, the statement that the difference in the number of animals found in two populations is statistically significant is biologically vacuous, since the ecologist knows a priori that any populations separated by space or time are different. Statistical significance is only a comment on the level of precision that the experiment was able to generate. Ultimately, the ecologist requires the compound distribution of the difference between the two populations (i.e., how different are they and what is the uncertainty associated with those differences). Once the terminal sequential posterior distributions have been calculated, the compound distribution can easily be obtained.

In conclusion, the Bayesian framework offers a coherent method for obtaining information about a population from any mark-recapture experiment. Indeed, the method is even applicable to experiments in which very sparse data are obtained, as was the case in the mud turtle study. Such data cannot be sensibly analyzed using more traditional approaches. Given that the sequential Bayes algorithm can deal with small sample sizes, it should prove quite useful in the initial estimation and monitoring of the population trends of rare species.

LITERATURE CITED

Bailey, N. J. J. 1951. On estimating the size of mobile populations from recapture data. Biometrika 38:293–306.
Barnett, V. 1973. Comparative statistical inference. Second edition. Wiley, New York, New York, USA.
Berry, J. F., and C. M. Berry. 1984. A reanalysis of geographic variation and systematics in the yellow mud turtle, Kinosternon flavescens (Agassiz). Annals of the Carnegie Museum 53:185–206.
Carle, F. L., and M. R. Strub. 1978. A new method for estimating population size from removal data. Biometrics 34:621–630.
Chapman, D. G. 1948. A mathematical study of confidence limits of salmon populations calculated by sample tag ratios. International Pacific Salmon Fisheries Commission, Bulletin 2:67–85.
———. 1951. Some properties of the hypergeometric distribution with applications to zoological sample censuses. University of California Publications in Statistics 1:131–160.
———. 1954. The estimation of biological populations. Annals of Mathematical Statistics 25:1–15.
Chapman, P. G., and W. S. Overton. 1966. Estimating and testing differences between population levels by the Schnabel estimation method. Journal of Wildlife Management 30:173–180.
Cormack, R. M. 1968. The statistics of capture-recapture methods. Oceanography and Marine Biology, an Annual Review 6:455–506.
Cox, D. R., and D. V. Hinkley. 1974. Theoretical statistics. Chapman and Hall, London, England.
DeGroot, M. H. 1970. Optimal statistical decisions. McGraw-Hill, New York, New York, USA.
DeLury, D. B. 1951. On the planning of experiments for the estimation of fish populations. Journal of the Fisheries Research Board of Canada 8:281–307.
———. 1958. The estimation of population size by a marking and recapture procedure. Journal of the Fisheries Research Board of Canada 15:19–25.
Freeman, P. R. 1973a. Sequential recapture. Biometrika 60:141–153.
———. 1973b. A numerical comparison between sequential tagging and sequential recapture. Biometrika 60:499–508.
Gaskell, T. J., and B. J. George. 1972. A Bayesian modification of the Lincoln index. Journal of Applied Ecology 41:377–384.
Houseal, T. W., J. W. Bickham, and M. D. Springer. 1982. Geographic variation in the yellow mud turtle, Kinosternon flavescens. Copeia 1982:567–580.
Johnson, D. H. 1977. Some Bayesian techniques useful in estimating frequency and density. United States Fish and Wildlife Service Special Scientific Report, Wildlife Number 203.

Kempthorne, O., and J. L. Folks. 1971. Probability, statistics, and data analysis. Iowa State University Press, Ames, Iowa, USA.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference for capture data from closed populations. Wildlife Monographs 62.

Overton, W. S. 1971. Estimating the number of animals in wildlife populations. Pages 403–455 in R. H. Giles, editor. Wildlife management techniques. Edwards Brothers, Ann Arbor, Michigan, USA.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada 191.

Robson, D. S., and H. A. Regier. 1964. Sample size in Petersen mark-recapture experiments. Transactions of the American Fisheries Society 93:215–226.

Schnabel, Z. E. 1938. The estimation of the total fish population of a lake. American Mathematical Monthly 45:348–352.

Schumacher, F. X., and R. W. Eschmeyer. 1943. The estimate of fish population in lakes or ponds. Journal of the Tennessee Academy of Science 18:228–249.

Seber, G. A. F. 1973. The estimation of animal abundance. Hafner, New York, New York, USA.

———. 1982. The estimation of animal abundance and related parameters. Second edition. Macmillan, New York, New York, USA.

Skalski, J. R., D. S. Robson, and M. A. Simmons. 1983. Comparative census procedures using single mark-recapture methods. Ecology 64:752–760.

Stigler, S. 1975. Napoleonic statistics: the work of Laplace. Biometrika 62:503–517.