Statistica Sinica **8**(1998), 1233-1247

BAYESIAN INFERENCE OF POPULATION SIZE FOR BEHAVIORAL RESPONSE MODELS

Shen-Ming Lee and Cathy W. S. Chen

Feng-Chia University

Abstract: The primary goal of this paper is to estimate population size associated with the capture-recapture model when the capture probability vary with behavior response and time (or trapping occasion). We cast the capture-recapture model in a Bayesian framework and make inference by using the Gibbs sampler, a Markov Chain Monte Carlo method. Using the method of maximum likelihood estimation, certain assumptions on the relationship between the capture and recapture probabilities are required in order to make inference of population size for the behavior response model. The major advantage of this approach is that no assumption is needed in our proposed procedure. The proposed methodology is illustrated with real data and a simulation study. The results show that the Gibbs sampler provides sound inference of population size.

Key words and phrases: Behavior response, capture-recapture model, Gibbs sampler, Markov Chain Monte Carlo method, population size, time variation.

1. Introduction

The capture-recapture model is widely used in the estimation of population sizes. We will consider the problem of estimating population size under a closed population in the capture-recapture model when the capture probabilities vary with behavior response and time (or trapping occasion). This model is known as Model M_{tb}. There are three special models of Model M_{tb}: Model M_t, Model M_b, and Model M_0. Model M_t and Model M_b consider the capture probabilities varying with time and behavior response, respectively. If the capture probability is a constant, we call it Model M_0. Laplace (1786) used the Petersen method to estimate the total population size of France from a register of births for the whole country in 1783 (see p.16 in Otis, Burnham, White, and Anderson (1978) for more detail). This approach received its main impetus in the context of a estimating the size of a wildlife population. However, no breakthrough has been made in this area until the last three decades. For examples, see the work by Otis, Burnham, White, and Anderson (1978), White, Anderson, Burnham, and Otis (1982), Seber (1982, 1986, 1992), and Pollock (1991). The problem of estimating population size also has been encountered in epidemiology (Wittes (1974), Wittes, Colton, and Sidel (1974)), computer science (Jewell (1985), Lanbgberg and Singpurwalla

(1983), Nayak (1991), Chao, Ma and Yang (1993)), and demography (Wolter (1986)).

Model M_b and Model M_{tb} are practically important and useful in biological and ecological applications because animals frequently exhibit a behavioral response to capture. Little work has been done on Model M_{tb} in the literature (see Lloyd (1994), Rexstad and Burnham (1991), and Lee (1996) for references). Seber and Whale (1970) show that the maximum likelihood estimate (MLE) of population sizes exists only under some conditions for Model M_b. From a Bayesian viewpoint, Castledine (1981), Smith (1991) and George and Robert (1992) make inference of population size for Model M_t. When there is a behavioral response to capture, the inference of population size for Model M_t may have either a positive or negative bias according to whether the animals become trap shy or trap happy. Therefore, we concentrate on the behavioral response models. In this paper, we generalize a Bayesian analysis of Model M_b and Model M_{tb} by using the Gibbs sampler, a Markov Chain Monte Carlo method.

The aim of this paper is to show that the Gibbs sampler, as a viable alternative to both analytical calculation and numerical approximation, can facilitate Bayesian calculations for capture-recapture behavioral response models, thereby enhancing their scope. Moreover, since the Model M_{tb} involves more parameters than the minimal sufficient statistic, not all parameters can be estimated and maximum likelihood estimation of the population size proves to be impossible. Consequently, in order to make the population size N an identifiable parameter under maximum likelihood estimation, one has to make certain assumptions on the relationship between the recapture probability and first capture probability. One assumption often used is that of proportionality. It is possible in a Bayesian approach to estimate more parameters than observations at hand. (For example, see McCulloch and Tsay (1994).) A great advantage of the proposed procedure is that the unidentifiability problem can be resolved under Model M_{tb} in the Bayesian approach using the Gibbs sampler. We shall not repeat the details of the Gibbs sampler which can be found elsewhere (e.g. Geman and Geman (1984), Tanner and Wong (1987), Gelfand and Smith (1990), and Gelfand, Hills, Racine-Poon, and Smith (1990)). It suffices to say that what we need are conditional distributions of subsets of parameters given the others. The Gibbs sampler is iterated many times in order to obtain a sample of draws from the posterior distribution. The empirical distribution of this sample converges weakly to the true joint distribution. (For more details of convergence results, see Tierney (1994).) Interested readers are also referred to Casella and George (1992) and Tanner (1994) for a general comprehensive review of the Gibbs sampler. Section 2 presents the behavioral response models and the setup of Bayesian framework. Section 3 illustrates the methodology using a real data set and simulation study. We give concluding remarks in Section 4.

2. Bayes Estimates for the Behavioral Response Model

Let N be the unknown size of the population of interest and t be the total number of trapping samples. The animals can be indexed by $1,\ldots,N$ and P_{ij} is the capture probability of the ith animal in the jth trapping sample, $i=1,\ldots,N$; $j=1,\ldots,t$. Animals are assumed to act independently. If the animals exhibits behavior response , P_{ij} depends on the capture history of the first $j-1$ samples and P_{ij} can be expressed as

$$P_{ij}=\begin{cases} P^*_{ij} & \text{if the } i\text{th animal is not caught in the first } j-1 \text{ samples;} \\ b^*_{ij} & \text{if the } i\text{th animal has been caught in the first } j-1 \text{ samples.}\end{cases} \quad (1)$$

Let X_{ij} be equal to 1 if the ith animal is caught in the jth sample, and 0 otherwise. The underlying general probability structure of the capture-recapture experiments is as follows:

$$\begin{aligned} L(N,\boldsymbol{P}|\mathcal{D}) &= \prod_{i=1}^{N}\prod_{j=1}^{t} P_{ij}^{X_{ij}}(1-P_{ij})^{1-X_{ij}} \\ &= \prod_{i=1}^{N}\prod_{j=1}^{t} P_{ij}^{*X_{ij}I[(\sum_{k=1}^{j-1}X_{ik})=0]} b_{ij}^{*X_{ij}I[(\sum_{k=1}^{j-1}X_{ik})>0]} \\ &\qquad \times(1-P^*_{ij})^{(1-X_{ij})I[(\sum_{k=1}^{j-1}X_{ik})=0]}(1-b^*_{ij})^{(1-X_{ij})I[(\sum_{k=1}^{j-1}X_{ik})>0]}, \quad (2)\end{aligned}$$

where $I(\cdot)$ is the usual indicator function, $\boldsymbol{P}=(P_{ij}, i=1,\ldots,N; j=1,\ldots,t)$, $\mathcal{D}=\{X_{ij},\ i=1,\ldots,N;\ j=1,\ldots,t\}$, and $L(N,\boldsymbol{P}|\mathcal{D})$ denotes the likelihood function.

There are too many parameters in the general model in (2), so the information about N can not be extracted from data. Therefore, the parameter space of the general model in (2) must be restricted. When the animals do not exhibit behavior response, the most common restrictions used are $P_{ij}=P$ or $P_{ij}=P_j$ (see Darroch (1958), Castledine (1981), George and Robert (1992)). These models are designated as Model M_0 or Model M_t respectively in Otis, Burnham, White, Anderson (1978). If there is behavior response for the captured animals, the restrictions become

$$P_{ij}=P_j I\Big(\sum_{k=1}^{j-1}X_{ik}=0\Big)+b_j I\Big(\sum_{k=1}^{j-1}X_{ik}>0\Big) \quad (3)$$

or

$$P_{ij}=PI\Big(\sum_{k=1}^{j-1}X_{ik}=0\Big)+bI\Big(\sum_{k=1}^{j-1}X_{ik}>0\Big), \quad (4)$$

where b_j is the recapture probability in the jth sample and b is the recapture probability for any sample. We denote (3) and (4) by Model M_{tb} and Model M_b, respectively. In this paper, we will focus on Model M_{tb} and Model M_b.

2.1. Bayesian inference about N for Model M_{tb}

In this subsection, we consider all animals which are not caught in the first $j-1$ samples having the same capture probability P_j in the jth sample. The recapture probability for all animals in the jth sample is b_j. The structure of P_{ij} has the same explicit formula as (3). The likelihood function for this model is a special case of (2) and is obtained as follows:

$$L(N,\boldsymbol{P},\boldsymbol{b}|\mathcal{D}) \propto \frac{N!}{(N-M_{t+1})!}\prod_{j=1}^{t} P_j^{u_j}(1-P_j)^{N-M_{j+1}}\prod_{j=2}^{t} b_j^{m_j}(1-b_j)^{M_j-m_j}, \quad (5)$$

where $\boldsymbol{P}=(P_1,\ldots,P_t)$, $\boldsymbol{b}=(b_2,\ldots,b_t)$, $M_{j+1}=u_1+\cdots+u_j$ is the number of distinct animals captured before the first j samples, and u_j and m_j are the number of unmarked and marked animals captured in the jth sample, respectively.

The likelihood function can be exhibited by the product of some binomial distributions. The conditional distributions of m_j given M_j and u_j given $N-M_j$ are, respectively,

$$m_j \mid M_j \sim B(b_j, M_j), \quad j=2,\ldots,t$$

and

$$u_j \mid N-M_j \sim B(P_j, N-M_j), \quad j=1,\ldots,t.$$

Consequently, the explicit formula for the likelihood function is

$$L(N,\boldsymbol{P},\boldsymbol{b}|\mathcal{D})=\Big\{\prod_{j=1}^{t}\binom{N-M_j}{u_j}P_j^{u_j}(1-P_j)^{N-M_j-u_j}\Big\}\Big\{\prod_{j=2}^{t}\binom{M_j}{m_j}b_j^{m_j}(1-b_j)^{M_j-m_j}\Big\}. \quad (6)$$

We consider priors of the form $\pi(N,\boldsymbol{P},\boldsymbol{b}) = (\prod_{j=1}^{t}\pi(P_j))(\prod_{j=2}^{t}\pi(b_j))\pi(N)$. Such priors lead to posterior conditionals of the forms:

$$\pi(N \mid \boldsymbol{P},\boldsymbol{b},\mathcal{D}) \propto \frac{N!}{(N-M_{t+1})!}\Big(\prod_{j=1}^{t}(1-P_j)^N\Big)\pi(N), \quad (7)$$

$$\pi(\boldsymbol{P} \mid N,\boldsymbol{b},\mathcal{D}) \propto \prod_{j=1}^{t} P_j^{u_j}(1-P_j)^{(N-M_{j+1})}\pi(P_j), \quad (8)$$

$$\pi(\boldsymbol{b} \mid N,\boldsymbol{P},\mathcal{D}) \propto \prod_{j=2}^{t} b_j^{m_j}(1-b_j)^{(M_j-m_j)}\pi(b_j). \quad (9)$$

The conditional posterior of N in Equation (7) is the same as proposed by George and Robert (1992) under Model M_t.

We take the priors of $\boldsymbol{P}$ and $\boldsymbol{b}$ to be $\pi(\boldsymbol{P}) = \prod \pi(P_j)$ and $\pi(\boldsymbol{b}) = \prod \pi(b_j)$ respectively , where $\pi(P_j) = Be(\gamma_1, \gamma_2), \pi(b_j) = Be(\gamma_3, \gamma_4)$ with Be(x, y) denoting a beta distribution. It follows that Eqns. (8) - (9) can then be reduced to

$$\pi(\boldsymbol{P} \mid N, \boldsymbol{b}, \mathcal{D}) \propto \prod_{j=1}^{t} Be(u_j + \gamma_1,\ N - M_{j+1} + \gamma_2) \tag{10}$$

and

$$\pi(\boldsymbol{b} \mid N, \boldsymbol{P}, \mathcal{D}) \propto \prod_{j=2}^{t} Be(m_j + \gamma_3,\ M_j - m_j + \gamma_4). \tag{11}$$

If the prior of N is Jeffrey's prior $\pi(N) = 1/N$, the conditional posterior of N is

$$P(N = n \mid \boldsymbol{P}, \boldsymbol{b}, \mathcal{D}) = \binom{n-1}{M_{t+1}-1}\Big(1 - \prod_{j=1}^{t}(1 - P_j)\Big)^{M_{t+1}}\Big(\prod_{j=1}^{t}(1 - P_j)\Big)^{N-M_{t+1}},$$

where $n = M_{t+1}, M_{t+1} + 1, \ldots$ It is easy to recognize the conditional posterior of N is negative binomial with parameter $(M_{t+1}, 1 - \prod(1 - P_j))$. Alternatively, for the constant prior of N, the conditional posterior on N is negative binomial with parameter $(M_{t+1} + 1, 1 - \prod(1 - P_j))$. Starting with an initial value of $N^{(0)}$ for N, we can produce a 'Gibbs sequence' $\{\boldsymbol{P}^{(k)}, N^{(k)}, \boldsymbol{b}^{(k)}\}(k = 0, 1, \ldots)$ with simulated sampling from (10), (7) and (11).

We can also consider a logit model on the P_j and b_j, that is, $\alpha_j = \log(P_j/(1-P_j)) \sim N(\mu_j, \sigma^2)$ and $\beta_j = \log(b_j/(1 - b_j)) \sim N(\nu_j, \sigma^2)$. In this structure, the conditional posterior of $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_t)$ is

$$\pi(\boldsymbol{\alpha} \mid N, \boldsymbol{\beta}, \mathcal{D}) \propto \prod_{j=1}^{t} \frac{\exp(\alpha_j u_j - \frac{1}{2}(\frac{\alpha_j - \mu_j}{\sigma})^2)}{(1 + e^{\alpha_j})^{N-M_j}}, \tag{12}$$

where $\boldsymbol{\beta} = (\beta_2, \ldots, \beta_t)$. It is easy to check $\pi(\boldsymbol{\alpha} \mid N, \boldsymbol{\beta}, \mathcal{D})$ is log concave in α_j so that α_j can be simulated by adaptive rejection sampling. (For more details of this sampling, see Gilks and Wild (1992).) Note that the inference of N does not depend on β_j when $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, N) = (\prod \pi(\alpha_j))\ (\prod \pi(\beta_j))\pi(N)$. Therefore, we omit the conditional posterior of $\boldsymbol{\beta}$ here. Note further that the inference of N depends only on $u_1, \ldots, u_t$. Consequently, our method can also be extended to the removal model.

2.2. Bayesian inference about N for Model M_b

Suppose all animals have the same capture probability P in the first capture, and the same recapture probability b after the first capture. This structure of capture probability P_{ij} is reduced to (4) and the likelihood function becomes

$$L(N,P,b|\mathcal{D}) \propto \frac{N!}{(N-M_{t+1})!} P^{M_{t+1}}(1-P)^{tN-M_{\bullet}-M_{t+1}} b^{m_{\bullet}}(1-b)^{M_{\bullet}-m_{\bullet}}, \quad (13)$$

where $M_{\bullet} = M_2 + \cdots + M_t$ and $m_{\bullet} = m_2 + \cdots + m_t$. Taking N, P, and b to be a priori independent, the conditional posterior distributions are

$$\pi(N \mid P,b,\mathcal{D}) \propto \frac{N!}{(N-M_{t+1})!}(1-P)^{tN}\pi(N), \quad (14)$$

$$\pi(P \mid N,b,\mathcal{D}) \propto P^{M_{t+1}}(1-P)^{tN-M_{\bullet}-M_{t+1}}\pi(P), \quad (15)$$

and

$$\pi(b \mid N,P,\mathcal{D}) \propto b^{m_{\bullet}}(1-b)^{M_{\bullet}-m_{\bullet}}\pi(b). \quad (16)$$

We choose $\pi(P) = Be(\gamma_1,\gamma_2)$ and $\pi(b) = Be(\gamma_3,\gamma_4)$. Subsequently, (15) and (16) can be reduced to

$$\pi(P|N,b,\mathcal{D}) = Be(M_{t+1}+\gamma_1, tN - M_{\bullet} - M_{t+1} + \gamma_2) \quad (17)$$

and

$$\pi(b|N,P,\mathcal{D}) = Be(m_{\bullet}+\gamma_3, M_{\bullet} - m_{\bullet} + \gamma_4). \quad (18)$$

Taking the prior of N to be constant or Jeffrey's prior in (14), the conditional posterior of N follows a negative binomial distribution with respective parameter $(M_{t+1}+1, 1-(1-P)^t)$ or $(M_{t+1}, 1-(1-P)^t)$.

Based on these conditional posterior distributions, the Gibbs sampler can be readily implemented.

3. Illustrative Examples

In this section, we illustrate the proposed methodology with a real example and a brief simulation study, focusing on inference about population sizes. The chosen prior of N is the Jeffrey's prior.

3.1. Real example

We consider the cotton rat data in White, Anderson, Burnham, and Otis (1982). In a Florida sugar cane field, 76 traps were placed along 6 parallel transects and baited with apples. Traps were placed 15.4 m apart on a transect, transects were an average 80 m apart, and trapping was done for eight consecutive days. As shown in Table 1, it consists of $t = 8$ capture occasions from a population of cotton rats (*Sigmodon hispidus*). Notice that the total number of animals captured (not counting recaptures) is $M_{t+1} = M_9 = 82$ for this cotton rat data. Consequently we know the population size of cotton rats is above 82.

The model selected by White, Anderson, Burnham, and Otis (1982) as the appropriate model for population estimate is the behavioral response Model M_b. We are interested in making inference of the number of cotton rats for Model M_b as well as Model M_{tb}. The hyper-parameters used are $(\gamma_3, \gamma_4) = (3.0, 3.0)$ and 9 specifications for (γ_1, γ_2) are given in Table 2. The choice $(\gamma_1, \gamma_2) = (5.3, 17.6)$ maximizes the likelihood as empirical Bayes. Note that $(106, 352)$ is proportional to $(5.3, 17.6)$, and other choices are informative. We do not vary the hyper-parameters (γ_3, γ_4). The reason is that the value of (γ_3, γ_4) only affects $\boldsymbol{b}$ in (11) which does not alter the estimate of population sizes. For each prior, Gibbs sampler is run for 3500 iterations. We record every 5th value in the sequence of the last 1500 in order to have more nearly independent contributions. For each sequence, Table 2 lists median, mean, standard error of N, and a 95% credible interval for N is obtained from the 2.5% and 97.5% quantiles.

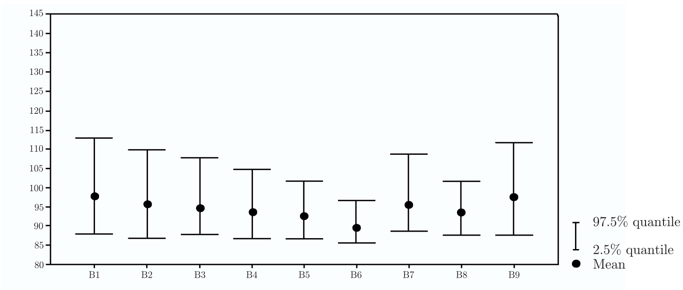
Table 1. Capture-recapture counts of cotton rat data

occasion		1	2	3	4	5	6	7	8
animals caught		19	26	33	27	33	37	27	28
total caught	(M_i)	0	19	36	52	60	66	74	81
newly caught	(u_i)	19	17	16	8	6	8	7	1
marked animals caught	(m_i)	0	9	17	19	27	29	20	27

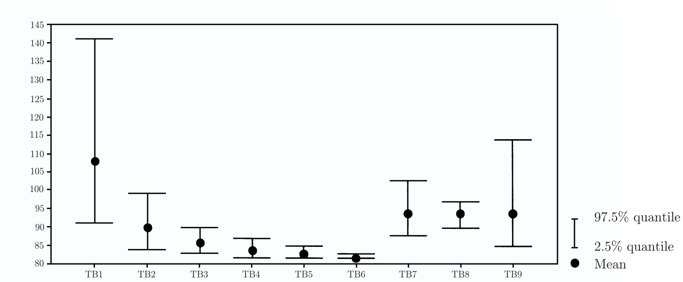
Table 2. Results of cotton rat data under Model M_b.

		M_b			M_{tb}	
$Beta(\gamma_1, \gamma_2)$	Median	Mean	95% CI	Median	Mean	95% CI
(1.0,5.0)	97	98 (7.03)	(88,113)	105	108 (13.95)	(91,141)
(2.0,5.0)	95	96 (5.78)	(87,110)	89	90 (4.09)	(84,99)
(3.0,5.0)	94	95 (5.21)	(88,108)	85	86 (1.70)	(83, 90)
(4.0,5.0)	94	94 (4.57)	(87,105)	84	84 (1.12)	(82, 87)
(5.0,5.0)	92	93 (3.92)	(87,102)	83	83 (0.79)	(82, 85)
(10.0,5.0)	90	90 (3.09)	(86, 97)	82	82 (0.17)	(82, 83)
(5.3,17.6)	95	96 (4.97)	(89,109)	93	94 (3.72)	(88,103)
(106,352)	94	94 (3.0)	(88,102)	94	94 (1.7)	(90, 97)
(1.0,3.3)	97	98 (6.15)	(88,112)	92	94 (7.73)	(85,114)

Although the shape of beta distributions is quite different for each hyperparameter (γ_1, γ_2), the estimator associated with Model M_b reveals only minor difference in mean as compared with that of Model M_{tb}. Using the method of maximum likelihood estimation, White, Anderson, Burnham, and Otis (1982) obtained $\hat{N} = 93$ with a standard error of 6.69 and (79, 107) as the 95% confidence interval , which are similar to our estimates. However, the 95% lower confidence bound for the population sizes in White, Anderson, Burnham, and Otis (1982) is 79 which underestimates the actual population size. Table 2 shows that the 95% lower credible bounds are all above 82 and the range of 95% credible intervals for Model M_b are shorter than the range of White, Anderson, Burnham, and Otis. To compare the 95% credible intervals for each prior for Model M_b and Model M_{tb}, Figures 1(a) and 1(b) are given. Figure 1(a) displays similar variability across priors while Figure 1(b) shows different variability across priors. The wide variability in posterior characteristics shows sensitive dependence on the choice of (γ_1, γ_2). The largest realization for $(\gamma_1, \gamma_2) = (1, 5)$ tends to be much larger than those for any other prior for Model M_{tb}. The value for $(\gamma_1, \gamma_2) = (1, 5)$ is the most dispersed in Figure 1(b). In short, the proposed method works reasonably well for Model M_b but less well for Model M_{tb} with regard to the cotton rat data.



(a)



(b)

Figure 1. (a) 95% credible intervals for Model M_b of cotton rat data. The x-axis represents priors in the order given in Table 2. (b) 95% credible intervals for Model M_{tb} of cotton rat data. The x-axis represents priors in the order given in Table 2.

3.2. Simulation study

In this subsection, we carried out a limited simulation study to investigate the performance of the proposed inference procedure. We fixed the population size $N = 100$ and $t = 6$. Let the first capture probability of animals in the jth sample be P_j and δ be the behavior response factor. Capture and recapture data were generated from a population where the recapture probability is $b_j = \delta P_j$. Since the recapture probability b_j does not alter the estimate of population size under the proposed method for Models M_b and M_{tb}, we fix δ at 0.36 and consider $(P_1, P_2, P_3, P_4, P_5, P_6)$ as follows.

Case 1. (0.14, 0.14, 0.14, 0.14, 0.14, 0.14);
Case 2. (0.18, 0.18, 0.18, 0.18, 0.18, 0.18);
Case 3. (0.22, 0.22, 0.22, 0.22, 0.22, 0.22);
Case 4. (0.18, 0.10, 0.10, 0.22, 0.12, 0.12);
Case 5. (0.22, 0.14, 0.26, 0.14, 0.14, 0.18);
Case 6. (0.20, 0.16, 0.16, 0.36, 0.20, 0.24),

where the underlying models are Model M_b for cases 1-3 and Model M_{tb} for cases 4-6.

We apply the Bayesian approach to all simulation. The hyper-parameters used are (γ_3, γ_4)=(3, 3) and 5 specifications for (γ_1, γ_2). The choices $(\gamma_1, \gamma_2) = (1, 5)$, $(2, 5)$, $(3, 5)$ and $(5, 5)$ can be motivated as informative prior. For each data set, we make inference of N via the proposed method for Model M_b and Model M_{tb}, respectively. It is possible to obtain a beta distribution with either of the parameters γ_1 or γ_2 being zero in Equation (10) when we adopt noninformative priors for (γ_1, γ_2) for Model M_{tb}. Therefore we do not consider noninformative priors for (γ_1, γ_2) in our simulation study. For each data set, the Gibbs sampler was run for 3500 iterations but collected every 5th of the last 1500 iterations for making inference. Two hundred data sets were generated and analyzed for each of the two Models M_{tb} and M_b. Moreover, Seber and Whale (1970) proposed the failure criterion for Model M_b

$$\sum_{j=1}^{t}(t+1-2j)u_j \le 0. \tag{19}$$

When satisfied, it is impossible to obtain valid estimation of N by using the method of maximum likelihood estimation. In the simulation study, we estimate N for Model M_b when (19) is not satisfied, otherwise we do not estimate N. We denote this estimator as "$M_{b(mle)}$" in our simulation. For each data set, we calculated the mean, median, standard deviation, and the 2.5% and 97.5% quantiles of the marginal posterior distribution. Tables 3-8 list the following:

the average of mean, median, and standard deviation, the medians of 2.5% and 97.5% quantiles , and the coverage of the 95% credible intervals for 200 data sets. Notice that we can not obtain valid maximum likelihood estimates for some data sets; therefore, the inference of $M_{b(mle)}$ is based only on the remaining data sets with valid maximum likelihood estimates.

Table 3. Simulation results for 200 runs, $N = 100, t = 6$, $(P_1, P_2, P_3, P_4, P_5, P_6) = (0.14, 0.14, 0.14, 0.14, 0.14, 0.14)$, $\delta = 0.36$.

Prior	Method	Median	Mean	Std	95% CI	Coverage
$Beta(1,5)$	M_b	149	274	366.3	(78 - 318)	0.83
	M_{tb}	90	91	9.6	(76 - 113)	0.92
	$M_{b(mle)}$	–	113	75.3	(42 - 142)	0.85
$Beta(2,5)$	M_b	103	115	43.0	(75 - 181)	0.93
	M_{tb}	71	72	3.5	(66 - 80)	0.00
	$M_{b(mle)}$	–	108	63.0	(44 - 141)	0.83
$Beta(3,5)$	M_b	92	98	21.5	(71 - 144)	0.88
	M_{tb}	66	66	2.0	(63 - 70)	0.00
	$M_{b(mle)}$	–	106	50.0	(39 - 149)	0.83
$Beta(5,5)$	M_b	84	86	12.3	(70 - 111)	0.70
	M_{tb}	63	63	1.0	(62 - 66)	0.00
	$M_{b(mle)}$	–	111	69.7	(43 - 145)	0.82

Table 4. Simulation results for 200 runs, $N = 100$, $t = 6$, $(P_1, P_2, P_3, P_4, P_5, P_6) = (0.18, 0.18, 0.18, 0.18, 0.18, 0.18)$, $\delta = 0.36$.

Prior	Method	Median	Mean	Std	95% CI	Coverage
$Beta(1,5)$	M_b	125	203	163.5	(84 - 185)	0.88
	M_{tb}	104	106	10.9	(88 - 131)	0.97
	$M_{b(mle)}$	–	103	29.2	(62 - 124)	0.83
$Beta(2,5)$	M_b	105	112	26.6	(82 - 155)	0.91
	M_{tb}	82	83	3.8	(77 - 91)	0.10
	$M_{b(mle)}$	–	103	26.1	(61 - 132)	0.84
$Beta(3,5)$	M_b	97	101	15.6	(80 - 130)	0.90
	M_{tb}	77	77	2.1	(73 - 82)	0.00
	$M_{b(mle)}$	–	101	25.8	(62 - 121)	0.87
$Beta(5,5)$	M_b	92	94	10.7	(80 - 116)	0.79
	M_{tb}	73	73	1.1	(71 - 76)	0.00
	$M_{b(mle)}$	–	107	32.7	(61 - 137)	0.88

First, we consider cases 1-3 in Tables 3-5. Most of the coverage probabilities for Model M_b are above 80%. The 95% lower confidence bounds for the method of maximum likelihood estimation are underestimated when compared with that of M_b. In Tables 6-8, the estimator associated with Model M_{tb} is sensitive to the priors selection. It has been known in the literature that the estimator associated with Model M_{tb} can cause technical problems. For example, Pollock et al. (1990) point out that “Model M_{tb} has no estimators in CAPTURE, but Burnham (Colo. Coop. Fish and Wildl. Res. Unit, pers. commun.) has derived an estimator that often does not perform very well $\cdots$”.

Table 5. Simulation results for 200 runs, $N = 100$, $t = 6$, $(P_1, P_2, P_3, P_4, P_5, P_6) = (0.22, 0.22, 0.22, 0.22, 0.22, 0.22)$, $\delta = 0.36$.

Prior	Method	Median	Mean	Std	95% CI	Coverage
$Beta(1,5)$	M_b	114	163	111.9	(90 - 154)	0.82
	M_{tb}	115	117	11.9	(98 - 144)	0.69
	$M_{b(mle)}$	–	103	18.3	(74 - 122)	0.88
$Beta(2,5)$	M_b	105	108	16.7	(87 - 135)	0.84
	M_{tb}	91	91	4.0	(84 - 100)	0.49
	$M_{b(mle)}$	–	101	15.8	(74 - 126)	0.87
$Beta(3,5)$	M_b	102	105	13.0	(88 - 127)	0.86
	M_{tb}	85	85	2.2	(82 - 90)	0.02
	$M_{b(mle)}$	–	103	16.8	(74 - 122)	0.88
$Beta(5,5)$	M_b	95	97	8.0	(85 - 113)	0.84
	M_{tb}	81	81	1.1	(79 - 83)	0.00
	$M_{b(mle)}$	–	100	14.6	(73 - 119)	0.89

Table 6. Simulation results for 200 runs, $N = 100$, $t = 6$, $(P_1, P_2, P_3, P_4, P_5, P_6) = (0.18, 0.10, 0.10, 0.22, 0.12, 0.12)$, $\delta = 0.36$.

Prior	Method	Median	Mean	Std	95% CI	Coverage
$Beta(1,5)$	M_b	134	294	370.1	(77 - 296)	0.91
	M_{tb}	90	91	9.7	(76 - 113)	0.92
	$M_{b(mle)}$	–	105	68.5	(45 - 132)	0.75
$Beta(2,5)$	M_b	97	106	33.2	(72 - 157)	0.86
	M_{tb}	71	71	3.4	(65 - 78)	0.00
	$M_{b(mle)}$	–	103	59.1	(47 - 130)	0.74
$Beta(3,5)$	M_b	88	93	19.2	(70 - 128)	0.79
	M_{tb}	66	66	1.9	(64 - 71)	0.00
	$M_{b(mle)}$	–	95	34.0	(48 - 127)	0.75
$Beta(5,5)$	M_b	114	209	241.9	(71 - 197)	0.80
	M_{tb}	63	63	1.9	(60 - 68)	0.00
	$M_{b(mle)}$	–	102	56.6	(47 - 124)	0.70

Table 7. Simulation results for 200 runs, $N = 100$, $t = 6$, $(P_1, P_2, P_3, P_4, P_5, P_6) = (0.22, 0.14, 0.26, 0.14, 0.14, 0.18)$, $\delta = 0.36$.

Prior	Method	Median	Mean	Std	95% CI	Coverage
$Beta(1, 5)$	M_b	100	113	38.4	(80 - 130)	0.79
	M_{tb}	103	105	10.7	(88 - 129)	0.99
	$M_{b(mle)}$	–	89	13.1	(65 - 105)	0.59
$Beta(2, 5)$	M_b	93	96	15.2	(78 - 118)	0.77
	M_{tb}	81	82	3.6	(76 - 90)	0.05
	$M_{b(mle)}$	–	89	13.9	(65 - 107)	0.58
$Beta(3, 5)$	M_b	89	92	10.2	(79 - 112)	0.72
	M_{tb}	77	77	2.0	(74 - 82)	0.00
	$M_{b(mle)}$	–	89	12.5	(66 - 108)	0.63
$Beta(5, 5)$	M_b	85	87	7.3	(76 - 102)	0.52
	M_{tb}	73	73	1.0	(72 - 76)	0.00
	$M_{b(mle)}$	–	90	13.4	(65 - 104)	0.59

Table 8. Simulation results for 200 runs, $N = 100$, $t = 6$, $(P_1, P_2, P_3, P_4, P_5, P_6) = (0.20, 0.16, 0.16, 0.36, 0.20, 0.24)$, $\delta = 0.36$.

Prior	Method	Median	Mean	Std	95% CI	Coverage
$Beta(1, 5)$	M_b	178	284	296.6	(104 - 344)	0.39
	M_{tb}	118	120	12.6	(100 - 149)	0.54
	$M_{b(mle)}$	–	139	63.1	(62 - 181)	0.99
$Beta(2, 5)$	M_b	134	146	45.1	(99 - 215)	0.55
	M_{tb}	93	93	4.3	(87 - 103)	0.70
	$M_{b(mle)}$	–	139	57.9	(63 - 183)	0.99
$Beta(3, 5)$	M_b	123	128	25.2	(96 - 179)	0.67
	M_{tb}	87	87	2.4	(83 - 93)	0.05
	$M_{b(mle)}$	–	142	76.2	(65 - 181)	0.99
$Beta(5, 5)$	M_b	109	112	14.5	(92 - 144)	0.86
	M_{tb}	82	82	1.2	(80 - 85)	0.00
	$M_{b(mle)}$	–	136	63.0	(65 - 173)	0.99

When the prior mean of P_j, $\gamma_1/(\gamma_1 + \gamma_2)$, is approximately equal to $\bar{P} = (1/t)\sum P_j$, the performance of inferences for Model M_{tb} can be improved and have better coverage probabilities. In the simulation study, the average of P_j ($\bar{P} = (1/t)\sum P_j$) in cases 4-6, are 0.14, 0.18, and 0.22 respectively. Consider $\bar{P} = 0.22$ in case 6. The prior means are between 0.167 and 0.286. Therefore, the inference performance with priors $Be(1, 5)$ and $Be(2, 5)$ for Model M_{tb} are better than the others in Table 8. The remaining coverage probabilities for 95% credible intervals are small since the prior mean is far away from $\bar{P}$. Therefore, the estimator associated with Model M_{tb} behaves nicely when the prior mean is close to the average, $\bar{P}$. The result is summarized as follows:

	$\bar{P}$	$Be(\gamma_1,\gamma_2)$	$\frac{\gamma_1}{\gamma_1+\gamma_2}$	Coverage for Model M_{tb}
case 4	0.14	(1,5)	0.167	0.92
case 5	0.18	(1,5)	0.167	0.99
case 6	0.22	(2,5)	0.286	0.70

We repeated the simulation several times with different capture probabilities and different hyper-parameters for the prior distributions. Due to limited space, the detailed results are omitted. The results indicate that the posterior median for Model M_b gives reasonable inference. In summary, in most of the cases the performance of inference for Model M_b is much better than the performance inference for Model M_{tb}. Moreover, the estimates for Model M_b are less sensitive than estimates for Model M_{tb} when the priors are selected. The estimator associated with Model M_{tb} behaves nicely when the prior mean is close to the average of capture probabilities.

4. Concluding Remarks

The conventional capture-recapture model is often used to estimate population size; however, estimation becomes troublesome when problems arise such as when the MLE does not exist or when there is an unidentifiability problem if the animals exhibit behavior response after they have been captured. The unidentifiability problem can be overcome in the proposed Bayesian approach. We apply the Gibbs sampling technique to make inference of population size for two kinds of behavior response models, Model M_{tb} and M_b.

When it is not possible to obtain valid estimation of the population size for Model M_b by using the method of maximum likelihood estimation, we propose Bayesian estimation procedures. The results show that the performance for Model M_b gives sound inference of the population size and the performance for Model M_b is less sensitive than that for Model M_{tb} with the priors selected. The conventional estimation of the population size for Model M_{tb} requires certain assumptions on the relationship between the recapture probability and first capture probability. The major result of this work is that we have shown no such assumption is needed in our proposed approach. That is, we do not need to know any information about recaptures. Finally, since we make inference of the population size without using the recapture information, we can extend the application of this procedure to the removal model.

Acknowledgement

The authors thank two referees and the associate editor for their constructive comments which led to significant improvement in this paper. S.-M. Lee and C. W. S. Chen acknowledge research support from the National Science Council of Taiwan grants NSC86-2115-M-035-012 and NSC86-2115-M-035-017.

References

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *Amer. Statist.* **46**, 167-174.

Castledine, B. (1981). A Bayesian analysis of multiple-recapture sampling for a closed population. *Biometrika* **68**, 197-210.

Chao, A., Ma, M.-C. and Yang, M. C. K. (1993). Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* **80**, 193-201.

Darroch, J. (1958). The multiple-recapture census, I: Estimation of a closed population. *Biometrika* **45**, 343-359.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approach to calculating marginal densities. *J. Amer. Statist. Assoc.* **85**, 398-409.

Gelfand, A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Amer. Statist. Assoc.* **85**, 972-985.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pat. Anal. and Mach. Intel.* **6**, 721-741.

George, E. I. and Robert, C. P. (1992). Capture-recapture estimation via Gibbs sampling. *Biometrika* **79**, 677-683.

Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Appl. Statist.* **41**, 337-348.

Jewell, W. (1985). Bayesian estimation of undetected errors. In *Bayesian Statistics* 2 (Edited by J. Bernardo, M. DeGroot, D. V. Lindley and A. F. M. Smith), 663-672. Amsterdam: North-Holland.

Langberg, N. and Singpurwalla, N. D. (1983). A unification of same software reliability model. *SIAM J. Sci. Statist. Comput.* **6**, 781-790.

Laplace, P. S. (1786). Sur les naissances, les mariages et les morts. Histoire de ℓ'Académie Royale des Sciences, Année 1783, Paris.

Lee, S.-M. (1996). Estimating population size for capture-recapture data when capture probabilities vary by time, behavior and individual animal. *Comm. Statist. Simulation Comput.* **25**, 431-457.

Lloyd, C. J. (1994). Efficiency of martingale methods in recapture studies. *Biometrika* **81**, 305-315.

McCulloch, R. E. and Tsay, R. S. (1994). Bayesian analysis of autoregressive time series via the Gibbs sampler. *J. Time Ser. Anal.* **15**, 235-250.

Nayak, T. K. (1991). Estimating the number of component processes of a superimposed process. *Biometrika* **78**, 75-81.

Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. R. (1978). Statistical inference for capture data on closed animal populations. *Wildlife Monographs* **62**, Louisville: University of Kentucky.

Pollock, K. H. (1991). Modeling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present, and future. *J. Amer. Statist. Assoc.* **86**, 225-238.

Pollock, K. H., Nichols, J. D., Brownie, C. and Hines, J. E. (1990). Statistical inference for capture-recapture experiments. *Wildlife Monographs* **107**, 1-97.

Rexstad, E. and Burnham, K. P. (1991). *User's Guide for Interactive Program* CAPTURE. Colorado cooperative Fish and Wildlife Research Unit, Fort Collins.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance*, 2nd Edition. Griffin, London.

Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics* **42**, 267-292.

Seber, G. A. F. (1992). A review of estimating animal abundance II. *Internat. Statist. Rev.* **60**, 129-166

Seber, G. A. F. and Whale, J. F. (1970). The removal method for two and three samples. *Biometrics* **26**, 393-400.

Smith, P. J. (1991). Bayesian analysis for a multiple capture-recapture model. *Biometrika* **78**, 399-407.

Tanner, M. (1994). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions.* Springer-Verlag, New York.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation (with discussion). *J. Amer. Statist. Assoc.* **82**, 528-550.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22**, 1701-1762.

White, G. C., Anderson, D. R., Burnham, K. P. and Otis, D. L. (1982). Capture-recapture and removal methods for sampling closed populations. Los Alamos National Lab. Publication LA-8787-NERP.

Wittes, J. T. (1974). Applications of a multinomial capture-recapture to epidemiological data. *J. Amer. Statist. Assoc.* **69**, 93-97.

Wittes, J. T., Colton, T. and Sidel, V. W. (1974). Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *J. Chronic Disease* **27**, 25-36.

Wolter, K. (1986). Some coverage error models for census data. *J. Amer. Statist. Assoc.* **81**, 338-346.

Department of Statistics, Feng-Chia University, Taichung, Taiwan.

E-mail: smlee@stat.fcu.edu.tw

E-mail: wschen@stat.fcu.edu.tw

(Received February 1996; accepted October 1997)