Bayesian Estimation of Population Size via Capture-Recapture Model with Time Variation and Behavioral Response

Xiaoyin Wang[1], Zhuoqiong He[2], Dongchu Sun[2]

[1]Department of Mathematics, Towson Univeristy, Towson, MD, USA
[2]Department of Statistics, University of Missouri-Columbia, Columbia, MO, USA
Email: xwang@towson.edu, hezh@missouri.edu, sund@missouri.edu

Abstract

We consider the problem of population estimation using capture-recapture data, where capture probabilities can vary between sampling occasions and behavioural responses. The original model is not identifiable without further restrictions. The novelty of this article is to expand the current research practice by developing a hierarchical Bayesian approach with the assumption that the odds of recapture bears a constant relationship to the odds of initial capture. A real-data example of deer mice population is given to illustrate the proposed method. Three simulation studies are developed to inspect the performance of the proposed Bayesian estimates. Compared with the maximum likelihood estimates discussed in Chao *et al*. (2000), the hierarchical Bayesian estimate provides reasonably better population estimation with less mean square error; moreover, it is sturdy to underline relationship between the initial and re-capture probabilities. The sensitivity study shows that the proposed Bayesian approach is robust to the choice of hyper-parameters. The third simulation study reveals that both relative bias and relative RMSE approach zero as population size increases. A *R*-package is developed and used in both data example and simulation.

Keywords

1. Introduction

The capture-recapture sampling methods were originally designed to estimate wildlife populations. They are

now widely applied to numerous areas, such as biology and ecology for estimating animal and plant populations, sociology and demography for estimating the size of the population at risks, official statistics for census under-count procedures, economics and finance in credit scoring and fraud detection, and so on. Varieties of methods have been developed to analyse these models, such as maximum likelihood [1]-[6], moment methods based on sample coverage [7]-[9], martingales [10]-[12], and Bayesian methods [13]-[23], and so forth. With the rapid growth in both methodology and application, the literature on capture-recapture analysis is tremendously massive. For a comprehensive literature review on capture-recapture models, see [24]-[29].

In this paper, we focus on inferences for the closed population capture-recapture experiment with capture probabilities being assumed to vary with time (or sampling occasion) and behavioral response. This model was originally proposed by Pollock in his doctoral dissertation (1974), and was referred to as the model M_{tb} in the literature, where the subscript *t* denotes the time (or sampling occasions) and *b* denotes the behavioral response; see Otis *et al*. [5]. The behavioral response means that animals change their behavior after their initial captures. Animals might become "trap-happy" if being caught is rewarding (getting food from a baited trap, for example) or "trap-shy" if being caught is traumatic (having a plastic tag stapled through the ear, for example). The model M_{tb} assumes that on a given sampling occasion all unmarked animals have one probability of capture and all marked animals have another probability of capture, and it assumes that these probabilities vary from one sampling occasion to another.

The model M_{tb} is important in practice because animals often exhibit a behavioural response to capture. This model has been selected as the most likely model in some population-estimation studies obw78, pnbh90. The original M_{tb} model is not identifiable without further restrictions on the parameters because the number of unknown parameters is more than that of the minimal sufficient statistics. In order to make the parameters estimable, one may assume that at least two of the capture probabilities are equal. However, there is no biological justification for making such an assumption [5]. Tanaka [30] [31] suggested a regression procedure, which had serious problems of interpretation [32]. Another assumption considered by Otis *et al*. [5] is that the ratio of the recapture probability to the initial capture probability is a constant. Chao *et al*. [2] derived the unconditional MLE (UMLE), the conditional MLE (CMLE) and the maximum quasi-likelihood (MQLE) estimators under this assumption without putting a minimal constraint on the constant ratio to ensure that neither initial nor recapture probabilities exceed 1.0. Under the same constant ratio assumption and with the minimal constraint of the constant ratio embedding in the MCMC, Lee and Chen [20] and Lee *et al*. [21] derived Bayes estimate by applying Beta prior to the initial capture probability and Uniform prior on the constant ratio. The results show that the performance of the estimate often depends on the choice of the hyper-parameters in the prior distributions. [19] proposed a Bayesian analysis of all eight closed population capture-recapture models within a logistic regression frame work. The method was illustrated through two real data samples, and the results are compared with those from MARK, which is a free software for capture-recapture data using classical MLE approach. Unfortunately, neither data set was determined to be suitable to the model M_{tb}, and the merit of the method with the model M_{tb} remains undiscovered. Also, there is no discussion about the performance of the Bayes estimates with different possible of capture probabilities and/or different population sizes. Regarding the prior specification, [19] performs a limited prior sensitivity analysis in the second example and concludes that the different prior specification has relatively little influence on the final estimation. However, in the same example, it also points out that "the prior specified on the individual variance component appeared to be somewhat restrictive in this case, with values supported by the data having relatively small prior mass". Therefore, the choice of the hyper-parameter remains ambiguous. We are thus motivated to explore more reasonable restrictions to deal with the non-identification, and to find more robust population estimation of the model M_{tb} in the Bayesian framework with more objective prior distribution, and the results are compared to the state-of-art MLE as proposed in [2].

R is a free software programming language and software environment for statistical computing and graphics. The *R* language is widely used among statisticians and data miners for developing statistical software and data analysis. In this paper, a *R*-package named sbmtb is created and used within the data analysis and simulation studies.

In this paper, we assume that the odds ratio of recapture probability and initial probability is a constant, as was initially proposed by Wang [22] in her doctoral dissertation; hierarchical prior distributions with vague information are assigned to the unknown parameters after logistic transformation; and Markov chain Monte Carlo (MCMC) methods are used to implement the Bayesian computation. The MCMC methods allow us to construct an ergodic Markov chain with stationary posterior distribution and, furthermore, to obtain the posterior quanti-

ties of the parameters of interest. Full conditional distributions are used in the MCMC methods to obtain approximations of the posterior distributions and Bayes estimates of the posterior quantities of unknown parameters.

Otis *et al.* [5] provides an interesting data set on the deer mice study that was also used by Chao *et al.* [2] Altogether, there are 110 distinct mice caught out of 283 captures in 5 consecutive days. This same data set is used here to illustrate the proposed Bayesian approach of population estimation.

Simulation studies are conducted to confirm the findings in the example, and to examine the performance of the proposed Bayesian approach. Compared with the state-of-the-art maximum likelihood estimates derived by Chao *et al.* [2], the proposed Bayesian estimates provide a relative comparable population estimation with smaller standard error and mean square error, especially when the overall capture probability is low. In addition, the Bayesian estimate relies little on the underlying relationship between initial and recapture probabilities. An intensive prior sensitivity study shows that the proposed Bayesian inference remains robust with different values of prior distributions. The third simulation study shows some of the asymptotic properties of the proposed Bayes estimate such that both relative bias and mean square error fade away as the population size increases.

In Section 2.1, we briefly review the problem of non-identifiability and propose a novel restriction, in which the odds ratio of recapture probability and initial probability is a constant. In Section 2.2, we apply the hierarchical prior to the unknown parameters in the model, and put forward a objective approach to determine the hyper-parameters in the prior distributions. In Section 2.3, we derive full conditional posterior distributions of the unknown parameters for the MCMC algorithm. In Section 3, we use a real data example to illustrate the proposed Bayesian hierarchical estimation approach. In Section 4, we conduct simulation studies to examine the performance of the proposed Bayesian approaches. In Section 5, we conclude and give remarks.

2. Bayesian Analysis of Model M_{tb}

2.1. Likelihood Function

We assume that the population size is *N*, and there are *k* sampling occasions. On the jth sampling occasion, let p_{j0} denote the initial capture probability, or the probability of an unmarked animal being captured on the jth sampling occasion, and p_{j1} be the recapture probability, or the probability of a marked animal being captured. The likelihood proposed by [5] for the model M_{tb} is

$$L(N, \mathbf{p}_0, \mathbf{p}_1) \propto \frac{N!}{(N - M_{k+1})!} \prod_{j=1}^{k} \left\{ p_{j0}^{u_j} (1 - p_{j0})^{N - M_{j+1}} \right\} \prod_{j=2}^{k} \left\{ p_{j1}^{m_j} (1 - p_{j1})^{M_j - m_j} \right\} \tag{1}$$

where $\mathbf{p}_0 = (p_{10}, \cdots, p_{k0})'$, $\mathbf{p}_1 = (p_{21}, \cdots, p_{k1})'$, u_j is the number of unmarked animals being captured on the jth sampling occasion, m_j is the number of marked animals being recaptured on the jth sampling occasion, $M_{j+1} = \sum_{l=1}^{j} u_j$ is the total number of marked animals in the population when the $(j+1)^{th}$ sample is taken, and M_{k+1} is the number of distinct animals being captured during the experiment period. In this model, *N*, $\mathbf{p}_0$ and $\mathbf{p}_1$ are the 2*k* unknown parameters. The likelihood Function (1) is over-parameterized because there is a total of 2*k* unknown parameters, but the minimal sufficient statistic is $(u_1, \cdots, u_k, m_2, \cdots, m_k)$, which has dimensions of $2k - 1$. Therefore, the model is non-identifiable with more parameters than data values unless some restrictions are made to the parameters.

We propose a restriction based on the odds of initial capture and recapture, as is initially discussed in [22] (unpublished doctoral dissertation). Odds (in favor of an event) is defined as the ratio of the probability that the event will happen to the probability that the event will not happen. It is an expression of relative probabilities, which other research studies have found to be more convenient to use than probabilities. We assume that the odds of recapture bears a constant relationship to the odds of initial capture, *i.e.*,

$$\frac{p_{j1}}{1 - p_{j1}} = r \times \frac{p_{j0}}{1 - p_{j0}}, \quad j = 2, 3, \cdots, k \tag{2}$$

This assumption says that on a given sampling occasion, the odds of recapture is proportional to that of the initial capture. In a "trap happy" situation, the proportion *r* is greater than 1.0, and in a "trap shy" situation, *r* is

less than 1.0. When there is no behavioural response, $r = 1.0$.

Comparing the proposed constraint with the constant ratio constraint, which presumes that $p_{j1} = \theta p_{j0}$, we found that the proposed constant odds ratio assumption is more convenient and practical because there is no need for another layer of constraint on the constant ratio to ensure that neither initial capture nor recapture probabilities exceed 1.0. The constant odds ratio r can be allowed to range from 0 to ∞, naturally. Apply logistic transformations to (2), and let

$$c = \log(r)$$

$$\eta_{j0} = \log\left(\frac{p_{j0}}{1 - p_{j0}}\right), \qquad j = 1, \cdots, k$$

$$\eta_{j1} = \log\left(\frac{p_{j1}}{1 - p_{j1}}\right), \qquad j = 2, \cdots, k$$

The Equation (2) then becomes

$$\eta_{j1} = c + \eta_{j0}, \qquad j = 2, \cdots, k \tag{3}$$

with $\eta_{j0} \in (-\infty, \infty)$, $j = 1, 2, \cdots, k$, $\eta_{j1} \in (-\infty, \infty)$, $j = 2, 3, \cdots, k$, and $c \in (-\infty, \infty)$. In a "trap happy" situation, $c > 0$. On the contrary, in a "trap shy" situation, $c < 0$. When there is no behavioral response, $c = 0$. In addition, all parameters now range from $-\infty$ to ∞. It is worth mentioning that Equation (3) is similar to the model M_{tb} in [19] but the parameters here bear more meaningful practical interpretation.

Let $M_1 = 0$ and $m_1 = 0$ then the likelihood Function (1) can be written as

$$\frac{N!}{(N - M_{k+1})!} \prod_{j=1}^{k} \left(1 + e^{\eta_{j0}}\right)^{M_j - N} e^{\eta_{j0} n_j} \prod_{j=2}^{k} \frac{e^{c m_j}}{\left(1 + e^c e^{\eta_{j0}}\right)^{M_j}} = \frac{N!}{(N - M_{k+1})!} \prod_{j=1}^{k} \frac{e^{\eta_{j0} n_j + c m_j}}{\left(1 + e^{\eta_{j0}}\right)^{N - M_j} \left(1 + e^{c + \eta_{j0}}\right)^{M_j}} \tag{4}$$

where $n_j = m_j + u_j$ is the number of subjects caught in the jth sampling occasion. Denote $\boldsymbol{\eta} = (\eta_{10}, \cdots, \eta_{k0})$, and (4) is the likelihood function of $(N, \boldsymbol{\eta}, c)$. In this likelihood function, there are $k + 2$ unknown parameters, and the number of minimal sufficient statistics is $2k - 1$. If $2k - 1 \geq k + 2$, *i.e.*, $k \geq 3$, all unknown parameters are estimable by the capture history data.

2.2. Priors

We apply Bayesian approaches to the re-parametrized likelihood Function (4) to estimate the population size N with the hierarchical prior. For η_{j0}, a *Normal* distribution prior is applied in the first stage, *i.e.*,

$$\eta_{j0} \overset{i.i.d.}{\sim} N(\mu_0, \delta_0), \quad j = 1, \cdots, k \tag{5}$$

In the second stage, another *Normal* distribution prior is applied to μ_0, and an Inverse Gamma distribution prior is used for δ_0, *i.e.*,

$$\mu_0 \sim N\left(v_0, \tau_0^2\right) \tag{6}$$

and

$$\delta_0 \sim \text{IG}(a_0, a_0) \tag{7}$$

where v_0, τ_0, and a_0 are predetermined hyper-parameters.

The prior distribution of c is assumed to be normal, *i.e.*,

$$c \sim N(\mu_1, \delta_1) \tag{8}$$

where μ_1 and δ_1 are further assumed to have the following distributions.

$$\mu_1 \sim N\left(v_1, \tau_1^2\right) \tag{9}$$

and

$$\delta_1 \sim \mathrm{IG}(a_1, a_1) \tag{10}$$

where v_1, τ_1, and a_1 are predetermined hyper-parameters.

The hyper-parameters are chosen using vague information as follows. If the initial capture probability is assumed to be between 10^{-l} and $1 - 10^{-l}$, with $l \geq 1$, then the range of $\eta's$ will be roughly between $-l$ and l. By applying the 3σ or empirical rule, the expected value of the standard deviation of $\eta's$ is around $l/3$, namely the expected value of δ_0 is about $l^2/9$. Note that δ_0 is assumed to have a $\mathrm{IG}(a_0, a_0)$ with a mean of $a_0/(a_0 - 1)$. Let $a_0/(a_0 - 1) = l^2/9$ then we have $a_0 = l^2/(l^2 - 9)$. It is adequate in practice to presume that p_{j0} ranging between 0.0001 and 0.9999, which results in $\eta_{j0}'s$ varying from −4 to +4. Thus, it is sufficient to assume the mean μ_0 has a standard normal distribution with $v_0 = 0$ and $\tau_0 = 1$, and the expected standard deviation is about 1.33, which indicates a_0 is about 2. With similar arguments, we set $v_1 = 0$, $\tau_1 = 1$, and $a_1 = 2$.

For the prior distribution of *N*, the *Poisson-Gamma* type of hierarchical prior is a popular choice in the literature [33]-[37]. In the first stage, the unknown population size *N* is assumed to have a *Poisson* distribution with mean λ, *i.e.*,

$$N \sim \mathrm{Pois}(\lambda) \tag{11}$$

In the second stage, λ is assumed to have a Gamma distribution with hyper-parameters α_1 and β_1, *i.e.*,

$$\lambda \sim \mathrm{Gamma}(\alpha_1, \beta_1) \tag{12}$$

The marginal prior distribution of *N*, which is obtained by integrating over λ, is

$$\int_0^\infty \frac{\lambda^N}{N!} e^{-\lambda} \frac{\beta_1^{\alpha_1} \lambda^{\alpha_1 - 1}}{\Gamma(\alpha_1)} e^{-\beta_1 \lambda} \mathrm{d}\lambda \propto \frac{\Gamma(N + \alpha_1)}{(\beta_1 + 1)^{N + \alpha_1 + 1} N!}$$

In the situation where little information about λ is available, a vague prior distribution is desired. When $\alpha_1 \to 0$ and $\beta_1 \to 0$, the prior distribution of λ satisfies $\pi(\lambda) \to 1/\lambda$ and the marginal prior of *N* satisfies $\pi(N) \to 1/N$. Notice that $\pi(N) \propto 1/N$ is an improper prior. The propriety of using this prior needs to be studied. In later discussion, we choose α_1 and $\beta_1 (> 0)$ close to 0, e.g., $\alpha_1 = 10^{-2}$ and $\beta_1 = 10^{-4}$.

2.3. Bayesian Computation

The population size *N* is the parameter of interest. Under the square error loss function, the posterior mean is the Bayesian estimator. However, the closed form expression of the posterior distribution of *N* is not easy to achieve. Therefore, we will use Markov chain Monte Carlo (MCMC) methods to obtain numerical solutions. In order to implement the MCMC procedure, it is necessary to have the full conditional posterior distributions, which are the conditional distributions of one parameter given all the other unknown parameters and data.

The full conditional posterior distributions of $(N, \boldsymbol{\eta}, c, \mu_0, \delta_0, \lambda)$ given capture history **X** are as follows:

- $(N - M_{k+1} | \boldsymbol{\eta}, c, \mu_0, \delta_0, \lambda, \mathbf{X}) \sim \mathrm{Pois}\left(\lambda \prod_{j=1}^{k} \frac{1}{1 + e^{\eta_{j0}}} \right)$

- For $j = 1, \cdots, k$, $\left[\eta_{j0} | N, \boldsymbol{\eta}_{[-j]}, c, \mu_0, \delta_0, \lambda, \mathbf{X} \right] \propto \dfrac{\exp\left\{ \eta_{j0} n_j - \dfrac{(\eta_{j0} - \mu_0)^2}{2\delta_0} \right\}}{\left(1 + e^{\eta_{j0} + c}\right)^{M_j} \left(1 + e^{\eta_{j0}}\right)^{N - M_j}}$

- $\left[c | N, \boldsymbol{\eta}, \mu_0, \delta_0, \lambda, \mathbf{X} \right] \propto \dfrac{\exp\left\{ c \sum_{j=1}^{k} m_j - \dfrac{(c - \mu_1)^2}{2\sigma^2} \right\}}{\prod_{j=1}^{k} \left(1 + e^{\eta_{j0} + c}\right)^{M_j}}$

- $\left(\mu_0 \big| N, \boldsymbol{\eta}, c, \delta_0, \lambda, \mathbf{X}\right) \sim N\left(\dfrac{\tau_0^2 \sum\limits_{j=1}^{k} \eta_{j0} + \delta_0 v_0}{k\tau_0^2 + \delta_0}, \dfrac{\delta_0 \tau_0^2}{k\tau_0^2 + \delta_0}\right)$

- $\left(\delta_0 \big| N, \boldsymbol{\eta}, c, \mu_0, \lambda, \mathbf{X}\right) \sim \mathrm{IG}\left(\dfrac{k}{2} + a_0, \sum\limits_{j=1}^{k} \dfrac{\left(\eta_{j0} - \mu_0\right)^2}{2} + a_0\right)$

- $\left(\mu_1 \big| N, \boldsymbol{\eta}, c, \delta_0, \mu_0, \delta_1, \lambda, \mathbf{X}\right) \sim N\left(\dfrac{c\tau_1^2 + \delta_1 v_1}{\tau_1^2 + \delta_1}, \dfrac{\delta_1 \tau_1^2}{\tau_1^2 + \delta_1}\right)$

- $\left(\delta_1 \big| N, \boldsymbol{\eta}, c, \mu_0, \delta_0, \delta_1, \lambda, \mathbf{X}\right) \sim \mathrm{IG}\left(\dfrac{1}{2} + a_1, \dfrac{\left(c - \mu_1\right)^2}{2} + a_1\right)$

- $\left(\lambda \big| N, \boldsymbol{\eta}, c, \mu_0, \delta_0, \mathbf{X}\right) \sim \mathrm{Gamma}\left(N + \alpha_1, 1 + \beta_1\right)$.

The conditional distributions of c and $\boldsymbol{\eta}$ do not have common statistical distribution forms. Gilk [38] proposed an adaptive-reject method to sample from a log-concave density up to a normalizing constant. It can be shown that the conditional density functions $\left[\eta_{j0} \big| N, \boldsymbol{\eta}_{[-j]}, c, \mu_0, \delta_0, \lambda, \mathbf{X}\right]$, $j = 1, \cdots, k$ are all log-concave, and the conditional density function $\left[c \big| N, \boldsymbol{\eta}, \mu_0, \delta_0, \lambda, \mathbf{X}\right]$ is log-concave.

3. Real Data Examples

Otis *et al.* [5] provides an interesting data set of deer mice. The data were collected by S. Hoffman (pers. comm.) in mid-July 1974. Trapping was on 5 consecutive mornings, and 110 distinct mice were caught. It is assumed that there exists time and behavioural response of capture. The summary statistics are presented in **Table 1**. This data set is also used in Chao *et al.* [2], Lee and Chen [20], Lee *et al.* [5].

We use a Gibbs sampler in the R-package to implement the Bayesian computation. The burn-in stage is 15,000 iterations. The samples from the next 10,000 iterations are taken to construct the approximate posterior distribution of *N*. The Bayesian estimate of the population size *N* is $\hat{N}_B = 160$, with a standard error of 34.20. The three types of maximum likelihood estimates in Chao *et al.* [2] are $\hat{N}_{\mathrm{UMLE}} = 161$, $\hat{N}_{\mathrm{CMLE}} = 173$, and $\hat{N}_{\mathrm{MQLE}} = 152$, with standard errors 42.79, 45.25 and 29.94, respectively. Lee and Chen [20] and Lee *et al.* [5] obtained similar results as Chao *et al.* [2] by adjusting the values of hyper-parameters with a trial-and-error method.

The Bayesian estimate is very similar to the UMLE with a smaller standard error. It is also smaller than the CMLE, with a smaller standard error and is greater than the QMLE with a slightly greater standard error. The Bayesian estimates of initial capture probabilities are also computed as $\hat{p}_1 = 0.2406$, $\hat{p}_2 = 0.2558$, $\hat{p}_3 = 0.2119$, $\hat{p}_4 = 0.2301$, $\hat{p}_5 = 0.2118$, and $\hat{c} = 1.7917$. The initial capture probabilities are between 0.20 and 0.25, with overall catchability, which is the probability that a subject being captured at least once, around 0.73. As concluded by Chao *et al.* [2], when the overall catchability is high, the UMLE is preferable to the other two MLE estimates. Our Bayesian estimate of the deer mice population size is mostly close to the UMLE with a smaller standard error; hence, it is more desirable than the later one.

It is worth noting that the estimated behavioural effect $\hat{c} = 1.7917$ is greater than 0, which implies animals have a trap-happy response. This is consistent with the data since there are more recaptures than the initial captures on most sampling occasions.

Table 1. Summary data of deer mice example.

	Time *j*	1	2	3	4	5
Had been caught	M_j	0	37	68	77	98
Newly caught	u_j	37	31	9	21	12
Marked caught	m_j	0	23	49	44	57
Total caught	n_j	37	54	58	65	69

4. Simulation Studies

Chao *et al.* [2] reports a simulation study with the true population size of 400 and 8 different initial capture probabilities, with the recapture probabilities being calculated as $p_{j1} = 1.5 p_{j0}$. Here, we use the same true population size and the initial capture and recapture probabilities as in Chao *et al.* [2]. Together with the overall catchability, the values of the initial capture probabilities and recapture probabilities are listed in **Table 2**.

Several frequentist quantities are considered, such as bias, relative bias (which is the ratio of bias to the true population size), root mean square error (RMSE), and relative RMSE (which is the ratio of RMSE to the true population size) of Bayesian estimates of the population size. Suppose $\hat{N}_B$ is a Bayesian estimate of the population size N. The frequentist bias is the expected deviation between the estimate and the true value, *i.e.*, $E\left(\hat{N}_B - N\right)$. Though unbiasness is an attractive property, it does not take into consideration the variation associated with the estimator. The MSE is the frequentist expected squared deviation between the true value and the estimate, *i.e.*, $E\left(\hat{N}_B - N\right)^2$. It can be theoretically decomposed into a sum of the squared bias and variance of the estimator. Ideally, we would like to have an estimator with small bias and variance. However, it is also common in practice to trade off some increase in bias for a larger decrease in the variance and vice-versa so that a relatively smaller MSE can be achieved.

In the simulation, we generate capture histories 1000 times under each set of true parameters. The results of bias, relative biases (in parenthesis), RMSE, and relative RMSE (in parenthesis) are listed in **Table 3**. When the mean catchability is relatively low (cases 1 - 4), in terms of bias and relative bias, Bayesian estimates are very close to the two preferred MLE's (CMLE and MQLE) in Chao *et al.* [2]; while, Bayesian estimates consistently have smaller RMSE's than the three MLE's. In case 3, the bias of the Bayesian estimate is slightly greater than the MQLE with about 4% difference; however, there is about 10% of drop in RMSE when using the Bayesian estimate. As previously mentioned, when it comes down to practice, it is not uncommon to trade off some increase in bias for a larger decrease in the variance. Therefore, we would prefer the Bayesian estimate to the MLE's in these cases. When there is a sufficient number of captures (cases 5 - 8), according to Chao *et al.* [2], the UMLE is preferable to the other two estimators. Compared with the UMLE in these cases, the Bayesian approach yields comparable estimates in cases 5 and 8 with less RMSE. In case 7, the Bayesian estimate has a slightly larger bias than the UMLE, with a difference of less than 3%, but the RMSE dropped more than 3%. In case 6, the Bayesian estimate has a slightly larger bias with similar RMSE than the UMLE. Given the fact that the Bayesian estimates always provide a smaller RMSE than the three MLE's with comparable or slightly larger bias, we would conclude that the proposed Bayesian approach provides reasonably better estimation than the three MLE's.

It is worth mentioning that the simulated data was based on the constant ratio of initial and re-capture probabilities as in Chao *et al.* [2] rather than the constant odds ratio as we proposed. The conclusion that Bayesian estimates still outperform the MLE's, to a certain extent, indicates that the former one is robust to the constraints between the initial and recapture probabilities.

An usual difficulty of the Bayesian analysis is the need to specify prior distributions when little information is available for guidance. Here, we will examine how the proposed Bayesian inference is influenced by the choice

Table 2. Capture probabilities and behavioural effect parameter in simulation study I.

Cases	Initial Capture Probabilities					Overall Catchability
1	0.17	0.15	0.12	0.1	0.12	0.5083
2	0.12	0.16	0.2	0.12	0.16	0.5629
3	0.2	0.17	0.25	0.15	0.12	0.6275
4	0.2	0.2	0.15	0.24	0.24	0.6858
5	0.3	0.25	0.15	0.2	0.3	0.7501
6	0.23	0.28	0.35	0.31	0.25	0.8135
7	0.38	0.22	0.34	0.3	0.44	0.8748
8	0.45	0.4	0.4	0.4	0.45	0.9347

Table 3. Comparison of bayesian estimations with MLE's.

Case	Estimator	$\hat{N}$	Bias (%)	RMSE (%)	Case	Estimator	$\hat{N}$	Bias (%)	RMSE (%)
1	BAYES	332	−68 (−17.00)	87.4 (21.85)	5	BAYES	393	−7 (−1.75)	41.5 (10.38)
	UMLE	305	−95 (−23.75)	111.5 (27.88)		UMLE	400	0 (0.00)	63.9 (15.98)
	CMLE	344	−56 (−14.00)	101.2 (25.30)		CMLE	417	17 (4.25)	79.9 (19.98)
	MQLE	337	−63 (−15.75)	101.6 (25.40)		MQLE	413	13 (3.25)	75.4 (18.85)
2	BAYES	373	−27 (−6.75)	61.5 (15.38)	6	BAYES	413	13 (3.25)	40.8 (10.20)
	UMLE	354	−46 (−11.50)	90.3 (22.58)		UMLE	403	3 (0.75)	38.2 (9.55)
	CMLE	387	−13 (−3.25)	103.7 (25.93)		CMLE	409	9 (2.25)	42.4 (10.60)
	MQLE	381	−19 (−4.75)	99.4 (24.85)		MQLE	407	7 (1.75)	41 (10.25)
3	BAYES	384	−16 (−4.00)	55.5 (13.88)	7	BAYES	389	−11 (−2.75)	19.2 (4.80)
	UMLE	382	−18 (−4.50)	78.7 (19.68)		UMLE	401	1 (0.25)	29.3 (7.33)
	CMLE	407	7 (1.75)	99 (24.75)		CMLE	406	6 (1.50)	33.3 (8.33)
	MQLE	402	2 (0.50)	95.2 (23.80)		MQLE	405	5 (1.25)	32.3 (8.08)
4	BAYES	397	−3 (−0.75)	50.9 (12.73)	8	BAYES	399	−1 (−0.25)	11.3 (2.83)
	UMLE	390	−10 (−2.50)	70.5 (17.63)		UMLE	400	0 (0.00)	11.6 (2.90)
	CMLE	411	11 (2.75)	87.4 (21.85)		CMLE	402	2 (0.50)	12.4 (3.10)
	MQLE	406	6 (1.50)	83.2 (20.80)		MQLE	401	1 (0.25)	12.1 (3.03)

of different values of the hyper-parameters of the prior distribution. As described in Section 2.2, the presumed range of the capture probability is utilized to determine the value of hyper-parameters a_0 and a_1. A sensitivity study is carried out with $l = 5, 6, 7, 8$, and $a_0 = a_1 = 1.56, 1.33, 1.23, 1.16$, correspondingly and respectively. Resulting Bayesian estimates are listed in **Table 4**, which shows that the proposed Bayesian estimate is robust with the choices of the presumed range of the capture probability. When initial capture probability is low (case 1), the discrepancy of Bayesian population estimates with various hyper-parameters is less than 2.5%. The difference decreases as the initial capture probability rises. That means the impact of the prior declines as data information accumulates. When the initial capture probability is high (case 8), Bayesian inferences of population sizes under different hyper-parameters are almost identical to each other.

The third simulation study examines the performance of the Bayesian estimation with different population sizes from 100 to 1000. We use the initial capture probabilities in cases 1, 2, and 8 in **Table 2**, representing low, medium low, and high capture probabilities, respectively. The behavioural effect parameter c is set to be −1 for a trap shy situation, and 1 for a trap happy one. In each simulation case, we generate the capture histories 1000 times. The resulting relative biases and relative RMSE are shown in **Figure 1**. The plot indicates that relative biases of Bayesian population estimates declines as initial catchability and/or population size increases. When initial capture probabilities are low, the Bayesian approach produces population estimates with negative biases. When initial capture probabilities are medium low, the Bayesian approach underestimates the population, but the relative biases approach zero very quickly as the population size increases. When initial capture probabilities are high, the Bayesian estimate slightly overestimates the population size, and relative biases are very close to zero even with small population sizes. Also, the Bayesian approach produces smaller biases in trap happy situations than in trap shy ones. It is worth pointing out that, as listed in **Table 2**, the overall initial catchabilty in case 2 is just a little more than that of case 1. Therefore, we would say that, in most cases, the relative bias of the Bayesian estimate approaches zero as population size increases. **Figure 1** also shows that relative RMSE drops off as the population size and/or catchability increases.

Table 4. Sensitivity of the bayes estimation to the specification of the prior probability distributions.

l = 5, $a_0 = a_1$ = 1:56				l = 6, $a_0 = a_1$ = 1:33			
Case	$\hat{N}$	Bias	SE	RMSE	Case	$\hat{N}$	Bias	SE	RMSE
1	336.4354	−63.5646	58.7231	86.5183	1	328.7168	−71.2832	51.402	87.8682
2	384.0752	−15.9248	62.1371	64.1152	2	372.0497	−27.9503	53.8409	60.6396
3	392.1499	−7.8501	59.608	60.0931	3	381.1063	−18.8937	51.9437	55.2488
4	403.6953	3.6953	55.9828	56.0766	4	395.609	−4.391	50.1618	50.3287
5	395.9274	−4.0726	44.2326	44.3976	5	391.6046	−8.3954	40.3873	41.2309
6	415.0872	15.0872	40.3744	43.0823	6	411.8829	11.8829	38.2033	39.9904
7	389.5922	−10.4078	16.0508	19.123	7	389.195	−10.805	15.8764	19.1978
8	399.3953	−0.6047	11.2934	11.3039	8	399.1141	−0.8859	11.0794	11.1092

l = 7, $a_0 = a_1$ = 1:23				l = 8, $a_0 = a_1$ = 1:16			
Case	$\hat{N}$	Bias	SE	RMSE	Case	$\hat{N}$	Bias	SE	RMSE
1	331.5863	−68.4137	54.5113	87.4582	1	322.7735	−77.2265	47.549	90.6784
2	373.246	−26.754	55.3933	61.4909	2	369.6802	−30.3198	57.003	64.5398
3	384.2773	−15.7227	53.2808	55.5267	3	377.7821	−22.2179	54.2179	58.5686
4	397.3269	−2.6731	50.8611	50.9059	4	394.9889	−5.0111	51.897	52.1125
5	393.0034	−6.9966	40.9209	41.4945	5	390.513	−9.487	41.3206	42.3756
6	413.6375	13.6375	38.5416	40.8651	6	411.8151	11.8151	39.0769	40.8053
7	389.3894	−10.6106	16.0042	19.1954	7	388.9316	−11.0684	15.8458	19.3222
8	399.4645	−0.5355	11.3242	11.3312	8	399.0427	−0.9573	11.2079	11.2431

5. Conclusions and Remarks

In this paper, we propose a Bayesian hierarchical approach to estimate the population size of model M_{tb} with a new relationship between recapture probability and initial capture probability base on odds (in favour of capture). The capture-recapture model M_{tb} has wide applications in practical population estimation studies. However, the original model proposed by Otis *et al*. [5] is over parametrized and has an identifiability problem without further constraint on the recapture probability. One common assumption in literature is that the ratio of recapture probability and initial capture probability is a constant. Under this condition, Chao *et al*. [2] proposed three types of MLE's. The major problem with the MLEs is that no constraint is applied to the constant ratio to ensure that the resulting estimated initial capture and recapture probabilities remain between 0 and 1. Lee *et al*. [21] proposed Bayesian approach, which constrained the constant ratio in MCMC to ensure neither initial capture nor recapture probabilities exceed 1.0. However, no systematic method was discussed for the choice of prior distribution other than a trial-and-error method, and the estimate is sensitive to the value of the hyper-parameters. [19] proposed a Bayesian analysis of all eight closed population capture-recapture models within a logistic regression frame work. However, the usefulness of the method with the model M_{tb} remains unknown.

We were inspired to propose a new constraint so that the odds ratio of recapture probability and initial capture probability remains a constant and to analyse the model M_{tb} using a hierarchical Bayesian methodology with vague prior information. The benefit of the constant odds ratio assumption is that it is handier and "true-to-life", because there is no need for an extra constraint on the constant ratio to ensure that both initial capture and recapture probabilities are larger than 1.0. The constant odds ratio can naturally vary from $-\infty$ to ∞. Moreover, the simulation results show that the proposed Bayesian method provides comparatively better population estimation than the MLE's, and it is robust to the constraints on the capture probabilities. The sensitivity study shows that the proposed Bayesian estimate is robust to the choice of the hyper-parameters of the prior distribution. Also, the relative bias and RMSE of the proposed Bayesian population estimate approaches zero as the true population size grows in the usual course of application.
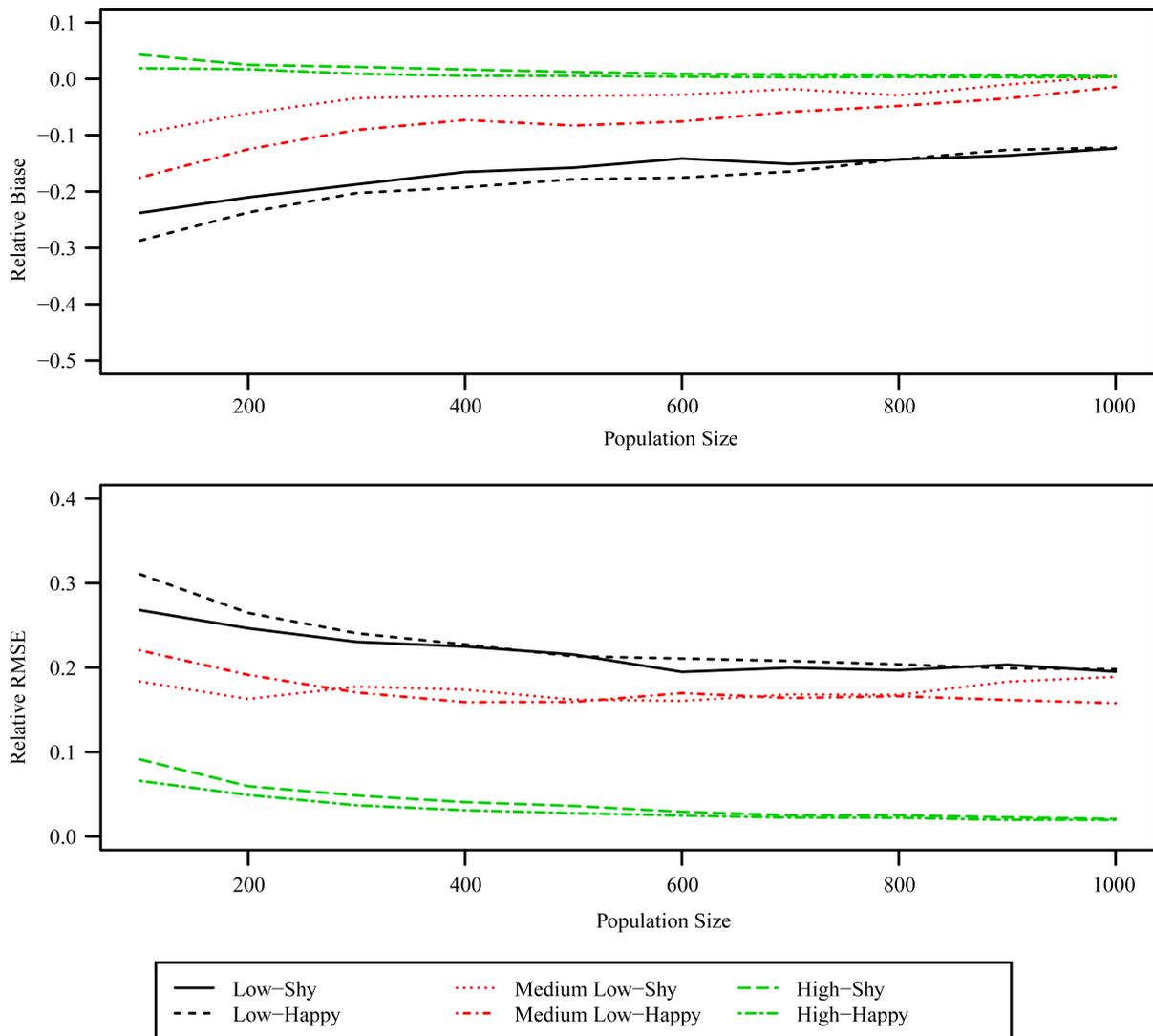
Figure 1. Relative bias and RMSE of bayesian population estimations.

Note that the findings here are mostly based on simulation results. It would be interesting to examine the theoretical bias and MSE of the Bayesian estimate under the hierarchical prior distribution studied in this paper. We'll continue studying more complex models such as M_{bh}, M_{tbh}, and open population models.

Author's Contribution

A *R*-package named *sbmtb* is created and used within the data analysis and simulation studies in the paper. The detailed help manual is in the Appendix. The *R*-package is available via e-mail request to Xiaoyin Wang at xwang@towson.edu.

References

[1] Agresti, A. (1994) Simple Capture-Recapture Models Permitting Unequal Catch Ability and Variable Sampling Effort. *Biometrics*, **50**, 494-500. http://dx.doi.org/10.2307/2533391

[2] Chao, A., Chu, W. and Hsu, C. (2000) Capture Recapture When Time and Behavioral Response Affect Capture Probabilities. *Biometrics*, **56**, 427-433. http://dx.doi.org/10.1111/j.0006-341X.2000.00427.x

[3] Norris III, J.L. and Pollock, K.H. (1995) A Capture-Recapture Model with Heterogeneity and Behavioural Response. *Environmental and Ecological Statistics*, **2**, 305-313. http://dx.doi.org/10.1007/BF00569360

[4] Norris, J.L. and Pollock, K.H. (1996) Nonparametric MLE under Two Closed Capture-Recapture Models with Heterogeneity. *Biometrics*, 52, 639-649.

[5] Otis, D.L., Burghal, K.P., White, G. and Anderson, D. (1978) Statistical Inference from Capture Data on Closed Animal Populations. *Wildlife Monographs*, **62**, 1-135.

[6] Pledger, S. (2000) Uniffed Maximum Likelihood Estimates for Closed Capture-Recapture Models Using Mixtures. *Biometrics*, **56**, 434-442. http://dx.doi.org/10.1111/j.0006-341X.2000.00434.x

[7] Chao, A., Lee, S.M. and Jeng, S.L. (1992) Estimating Population Size for the Capture-Recapture Data When Capture Probabilities Very by Time and Individual Animal. *Biometrics*, **48**, 201-216. http://dx.doi.org/10.2307/2532750

[8] Lee, S.M. (1996) Estimating Population Size for Capture-Recapture Data When Capture Probabilities Vary by Time, Behavior and Individual Animal. *Communications in Statistics-Simulation*, **25**, 431-457.

[9] Lee, S.M. and Chao, A. (1994) Estimating Population Size via Sample Coverage for Closed Capture-Recapture Model. *Biometrics*, **50**, 80-97. http://dx.doi.org/10.2307/2533199

[10] Yip, P. (1989) An Inference Procedure for a Capture and Recapture Experiment with Time-Dependent Capture Probabilities. *Biometrics*, **45**, 471-479. http://dx.doi.org/10.2307/2531490

[11] Yip, P. (1991) A Martingale Estimating Equation for a Capture-Recapture Experiment in Discrete Time. *Biometrics*, **47**, 1081-1088. http://dx.doi.org/10.2307/2532660

[12] Yip, P.S.F. and Fong, D.Y.T. (1993) Estimating Population Size by Recapture Sampling via Estimating Function. *Communications in Statistics-Stochastic Models*, **9**, 179-193. http://dx.doi.org/10.1080/15326349308807261

[13] Basu, S. and Ebrahimi, N. (2001) Bayesian Capture-Recapture Methods for Error Detection and Estimation of Population Size Heterogeneity and Dependence. *Biometrika*, **88**, 269-279. http://dx.doi.org/10.1093/biomet/88.1.269

[14] Bolfarine, H., Leite, J.G. and Rodrigues, J. (1992) On the Estimation of the Size of a Finite and Closed Population. *Biometrical Journal*, **34**, 577-593. http://dx.doi.org/10.1002/bimj.4710340507

[15] Castledine, B.J. (1981) A Bayesian Analysis of Multiple-Recapture Sampling for a Closed Population. *Biometrika*, **68**, 197-210. http://dx.doi.org/10.1093/biomet/68.1.197

[16] Chavez-Demoulin, V. (1999) Bayesian Inference for Small Sample Capture-Recapture Data. *Biometrics*, **55**, 727-731. http://dx.doi.org/10.1111/j.0006-341X.1999.00727.x

[17] Dupuis, J.A. (1995) Bayesian Estimation of Movement and Survival Probabilities from Capture-Recapture Data. *Biometrics*, **82**, 761-772.

[18] Gazey, W.J. and Staley, M.J. (1986) Population Estimation from Mark-Recapture Experiments Using a Sequential Bayes Algorithm. *Ecology*, **67**, 941-951. http://dx.doi.org/10.2307/1939816

[19] King, R. and Brooks, S. (2008) On the Bayesian Estimation of a Closed Population Size in the Presence of Heterogeneity and Model Uncertainty. *Biometrics*, **64**, 816-824. http://dx.doi.org/10.1111/j.1541-0420.2007.00938.x

[20] Lee, S.M. and Chen, C.W.S. (1998) Bayesian Inference of Population Size for Behavioral Response Models. *Statistical Sinica*, **8**, 1233-1247.

[21] Lee, S.M., Hwang, W.H. and Huang, L.H. (2003) Bayes Estimation of Population Size from Capture-Recapture Models with Time Variation and Behavior Response. *Statistical Sinica*, **13**, 477-494.

[22] Wang, X. (2002) Bayesian Analysis of Capture-Recapture Sampling for a Closed Population. Ph.D. Dissertation, University of Missouri-Columbia, Columbia.

[23] Wang, X., He, Z. and Sun, D. (2007) Bayesian Population Estimation for Small Sample Capture-Recapture Data Using Non-Informative Prior. *Journal of Statistical Planning and Inference*, **137**, 1099-1118. http://dx.doi.org/10.1016/j.jspi.2006.03.004

[24] Amstrup, S.C., McDonald, T.L. and Manly, B.F.J. (2005) Handbook of Capture Recapture Analysis. Princeton University Press, Princeton.

[25] Chao, A. (2001) An Overview of Closed Capture-Recapture Models. *Journal of Agricultural*, *Biological*, *and Environmental Statistics*, **6**, 158-175. http://dx.doi.org/10.1198/108571101750524670

[26] Pollock, K.H., Nichols, J.D., Brownie, C. and Hines, J.E. (1990) Statistical Inference for Capture-Recapture Experiments. *Wildlife Monographs*, **107**, 1-97.

[27] Schwarz, C.J. and Seber, G.A.F. (1999) Estimating Animal Abundance: Review III. *Statistical Science*, **14**, 427-456. http://dx.doi.org/10.1214/ss/1009212521

[28] Seber, G.A.F. (1992) A Review of Estimating Animal Abundance II. *International Statistical Review*, **60**, 129-166. http://dx.doi.org/10.2307/1403646

[29] Williams, B.K., Nichols, J.D. and Conroy, M.J. (2002) Analysis and Management of Animal Populations. Academic

Press, San Diego.

[30] Tanaka, R. (1951) Estimation of Vole and Mouse Population on Mount Ishizuchi and on the Uplands of Southern Shikoku. *Journal of Mammalogy*, **32**, 450-458. http://dx.doi.org/10.2307/1375793

[31] Tanaka, R. (1952) Theoretical Justification of the Mark and Release Index for Small Mammals. *Bulletin of the Kochi Women's College*, **1**, 38-47.

[32] Seber, G. (1973) The Estimation of Animal Abundance and Related Parameters. Griffin, London.

[33] George, E.I. (1992) Capture-Recapture Estimation via Gibbs Sampling. *Biometrika*, **79**, 677-683.

[34] Raftery, A.E. (1987) Inference and Prediction for General Order Statistic Model. *Journal of the American Statistical Association*, **82**, 1163-1168. http://dx.doi.org/10.1080/01621459.1987.10478554

[35] Raftery, A.E. (1988) Analysis of Simple Debugging Model. *Applied Statistics*, **37**, 12-32. http://dx.doi.org/10.2307/2347490

[36] Raftery, A.E. (1988) Inference for the Binomial *N* Parameter: A Hierarchical Bayes Approach. *Biometrika*, **75**, 223-228. http://dx.doi.org/10.1093/biomet/75.2.223

[37] Smith, P.J. (1991) Bayesian Analysis for a Multiple Capture-Recapture Model. *Biometrika*, **78**, 399-407. http://dx.doi.org/10.1093/biomet/78.2.399

[38] Gilk, W. (1992) Derivative-Free Adaptive Rejection Sampling for Gibbs Sampling. In: Bernardo, J.M., Beger, J.O., Dawi, A.P. and Smith, A.F.M., Eds., *Bayesian Statistics* 4, Oxford University Press, Oxford, 641-649.

Appendix

Package 'sbmtb'
September 18, 2014

Type   Package

Title   Bayesian Analysis of Capture-Recapture Model with Time Variation and Behavioral Response

Version   1.0

Date   2014-09-18

Author   Xiaoyin Wang

Maintainer   Xiaoyin Wang <xwang@towson.edu>

Description
   Bayesian Analysis of Capture-Recapture Model with Time Variation and Behavioral Response

License   GPL-2

Depends   ars

R topics documented:

sbmtb	*Bayesian Analysis of Capture-Recapture Model with Time Variation and Behavioral Response*

Description
   Bayesian analysis of Capture-Recapture model with time variation and behavioral response

Details

Package:	sbmtb
Type:	Package
Version:	1.0
Date:	2014-09-18
License:	GPL-2
Depends:	ars

1

2 sbmtb

Author(s)

Xiaoyin Wang <xwang@towson.edu>

References

X. Wang, Z. He and D. Sun Bayesian Estimation of Population Size via Capture-Recapture Model with Time Variation and Behavioral Response, Open Jounal of Ecology

Examples

```
n.days<-5
um.cap<-c(37,31,9,21,12)
mm.cap<-c(0,23,49,44,57)
sbmtb(n.days,um.cap, mm.cap)
```

---

sbmtb	Bayesian Analysis of Capture-Recapture Model with Time Variation and Behavioral Response

---

Description

This function provides Baysian analysis of capture-recapture model with time variation and behavioral response. It calculates the posterior mean, standard deviation and confidence interval of the population size, capture probability and behavioral effect.

Usage

```
sbmtb(n.days, u.cap, m.cap)
```

Arguments

n.days	Number of capture occasions
u.cap	a vector of number of unmarked/new captures on each capture occasion
m.cap	a vector of number of marked/recaptures on each capture occasion

Value

Population	a data frame contains the posterior mean, standard deviation, and confidence interval of the population size
Probability	a data frame contains the posterior mean, standard deviation, and confidence interval of the capture probabilities
behavior	a data frame contains the posterior mean, standard deviation, and confidence interval of the behavior effect

Author(s)

Xiaoyin Wang

sbmtb 3

Examples

```
n.days<-5
um.cap<-c(37,31,9,21,12)
mm.cap<-c(0,23,49,44,57)
sbmtb(n.days,um.cap, mm.cap)
```

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or **Online Submission Portal.**