The Korean Communications
in Statistics Vol. 7, No. 1, 2000
pp. 165-178

Bayesian Analysis for Multiple Capture-Recapture Models using Reference Priors

Younshik Chung[1] and Pongsu Park[2]

Abstract

Bayesian methods are considered for the multiple capture-recapture data. Reference priors are developed for such model and sampling-based approach through Gibbs sampler is used for inference from posterior distributions. Furthermore, approximate Bayes factors are obtained for model selection between trap and nontrap response models. Finally one methodology is implemented for a capture-recapture model in generated data and real data.

KEYWORDS : Bayes factor; Capture-recapture model; Gibbs sampler; Heterogeneity with trap response model; Heterogeneity with nontrap response model; Jeffreys' prior; Reference prior.

1. Introduction

The capture-recapture sampling method is an approach to estimate the unknown population size by a census. The multiple capture-recapture sampling method is the repetition of the single capture-recapture scheme. We take samples from a population of animals in the i^{th} stage, then there are animals which are marked before the i^{th} stage or captured first time. We count the number of marked animals and unmarked animals in the i^{th} sample, mark the previously unmarked animals, and return all the sampled animals to the population. This experiment is performed s stages. Suppose that each animal is captured with probability p_i, $i=1,\cdots,s$ in the i^{th} stage, the number of marked animals follows the hypergeometric distribution, and the sample size follows the binomial distribution with probability p_i. Then multiple capture-recapture method follows the product of hypergeometric distribution and binomial distribution. We may also consider the recapture probability c_i. The details are discussed in Section 2. Chapman(1952) and Darroch(1958) obtained the maximum likelihood estimate(MLE) for the population parameter N and variance for this estimate. Darroch(1958) found that his model did not change the maximum likelihood estimates of the populaiton size .

1) Associate Professor, Research Institute of Computer, Information and Communications, Department of Statistics, Pusan National University, Pusan, 609-735 Korea
2) Ph.D. Student, Department of Statistics, Pusan National University, Pusan, 609-735 Korea

Recently the problem of estimating the size of a closed animal population has been considered by Pollock and Otto(1983) from a classical sampling view-point. Castledine(1981)'s approach differed from previous ones by introducing prior information about the population size N and p_i. He showed that strong prior knowledge about the capture probabilities could greatly affect the inference about the population size. Pollock and Otto(1983) considered the capture probabilities which were constant over sampling times and influenced by trap response c_i. Castledine(1981) considered the capture probabilities to be same for all animals but with the probabilities of being changed over sampling times. However he did not considered the probability of trap response with noninformative priors of N and conjugate priors of p. Rodrigues et al.(1988) followed Castledine's approach but allowing for trap response but with noninformative priors of N and conjugate priors of p. In their paper, the appropriate model was chosen by using the Bayes factor. George and Robert(1992) used the Gibbs sampler to avoid the complexity of Bayesian computation. In this paper, we find the reference prior for the capture probabilities and the recapture probabilities to our model and the Bayesian model choice is investigated. And then we consider the posterior characteristics of estimated parameter N under the suitable model.

The paper is organized as follows. Section 2 derives the reference prior for capture probabilities $\underline{p}=(p_1,\cdots,p_s)$ and recapture probabilities $\underline{c}=(c_2,\cdots,c_s)$ In Section 3, two basic models are considered and the posterior analysis based on the reference prior is investigated under each model. In Section 4, Bayesian model choice is performed using the Bayes factor. In section 5, two examples are presented Markov Chain Monte Carlo (Gibbs sampler) is used to avoid the complicated Bayesian computation.

2. Modeling and Determination of the Reference Priors

2.1 Capture-Recapture Model

As in Castledine(1981), we write N for the unknown population size, $s(\geq 2)$ for the number of samples taken, p_i , $1\leq i\leq s$, for the probability of each animal to be unmarked in the i^{th} sample, c_i, $2\leq i\leq s$ for the probability of animal to be marked in the i^{th} sample, X_i for the number of unmarked animals in the i^{th} sample, Y_i for the number of marked animals in the i^{th} sample $(Y_1=0)$and N_i for the number of marked animals just before the i^{th} sample $(N_1=0)$ It follows by the definitions above that $N_{i+1}=N_i+X_i=\sum_{j=1}^{i}X_j,\ i=1,\ldots,s$ Let $n_i=X_i+Y_i$ $i=1,\ldots,s$It is assumed that the population remains closed throughout the realization of the experiment. The case where $p_i\neq c_i$, for some i, is called ''the capture-recpature model with trap response model'' by Pollock(1975), and it is hereby denoted by M_1: If $p_i = c_i$ for all i, it is called ''the capture-recapture model with nontrap response model'', denoted by M_0 which is considered by Castledine(1981).

For the capture-recapture model with trap response, let

$$X_i | p_i \sim B(N - N_i, p_i), \quad i = 1, \ldots, s \tag{2.1}$$

and

$$Y_i | c_i \sim B(N_i, c_i), \quad i = 2, \ldots, s \tag{2.2}$$

where $B(n, p)$ stands for binomial distribution with n trials and success probability p. It is also assumed that X_i is independent of Y_i (conditional on N_i). For convenience, let $Y_1 = 0$ Then it follows easily that the likelihood function is

$$L(N, \underline{p}, \underline{c} | data) = \prod_{i=1}^{s} \left\{ \binom{N - N_i}{x_i} p_i^{x_i} (1 - p_i)^{N - N_i - x_i} \right\} \prod_{i=2}^{s} \left\{ \binom{N_i}{y_i} c_i^{y_i} (1 - c_i)^{N_i - y_i} \right\}$$

$$\propto \binom{N}{r} \prod_{i=1}^{s} \{ p_i^{x_i} (1 - p_i)^{N - N_i - x_i} \} \prod_{i=2}^{s} \{ c_i^{y_i} (1 - c_i)^{N_i - y_i} \} \quad , \tag{2.3}$$

where $\underline{p} = (p_1, \cdots p_s)$ $\underline{c} = (c_2, \cdots, c_s)$ and $r = \sum_{i=1}^{s} x_i$ with the restriction that $N \geq r$. In particular if $p_i = c_i$, that is the capture-recapture model with nontrap response model. Also the following prior structure is assumed ;

$$\pi(N, \underline{p}, \underline{c}) = \pi_1(N) \pi_2(\underline{p}, \underline{c}) \ . \tag{2.4}$$

For specifying a prior distribution for N, the following choices are available ;

(1) Poisson prior on N ;

(2) Gamma-mixed Poisson prior ;

$$\pi(N) = \int \pi(N|\lambda) \pi(\lambda) d\lambda = \int \frac{e^{-\lambda} \lambda^N}{N!} \frac{\lambda^{a-1} e^{-\frac{\lambda}{b}}}{\Gamma(a) b^a} d\lambda = \frac{\Gamma(N + a)(1 + \frac{1}{b})^{-(N+a)}}{N! \Gamma(a) b^a} ;$$

(3) Jeffreys' prior $\pi(N) = N^{-1}$;

(4) Discrete uniform prior $\pi(N) \propto 1$.

Rodrigues et al.(1988) considered the case (4) above with the uniform prior of $(\underline{p}, \underline{c})$ under the same model. In the next section we want to find the reference prior of $(\underline{p}, \underline{c})$.

2.2 Reference Prior

When there are no precise information about parameters, we usually use noninformative priors. But the determination of reasonable noninformative priors in multiparameter problems is not easy ; Bernardo(1979) pointed out that if we were interested in a subset of the parameters and the rest was to be nuisance parameters, then Jeffreys' prior, common noninformative prior may be inappropriate for representing vague or little prior information. In order to overcome this problem, Bernardo(1979) proposed the reference prior approach for the development of the noninformative prior. Ye(1994) and Sun and Ye(1995) used Bayesian reference prior approach widely. Especially, Chung and Dey(1998) considered the reference prior for the estimation of intraclass correlation for balanced random effects models.

In Berger and Bernardo(1992)'s reference prior approach, the ordered group is very important. That is, the form of reference priors can be changed by the ordered grouping. Therefore the notation such as $\{p_1, p_2, \cdots, p_s, c_2, \cdots, c_s\}$ will be used to specify the group and the importance of parameters; $\{p_1, p_2, \cdots, p_s, c_2, \cdots, c_s\}$ means that p_1 is most important and c_s is least important. $\{(p_1, p_2, \cdots, p_s, c_2, \cdots, c_s)\}$ denotes that $p_1, \cdots, c_s$ are all equl likely. In general, under our model the reference prior distributions for different groups of ordering of $\{p_1, p_2, \cdots, p_s, c_2, \cdots, c_s\}$ is obtained. The following lemmas are useful to find the reference priors.

Assuming the parametric model $p(x|\boldsymbol{\theta})$, $\boldsymbol{\theta}=(\theta_1, \cdots, \theta_m)\in\Theta$ be such that the Fisher *information matrix*

$$H(\boldsymbol{\theta}) = -E_{X|\boldsymbol{\theta}} \frac{\partial^2}{\partial\theta_i\partial\theta_j} \log p(X|\boldsymbol{\theta}) .$$

Lemma 2.1 (Reference Prior : Bernardo and Smith(1994)) If H($\boldsymbol{\theta}$) is block diagonal (i.e., $\theta_1, \cdots, \theta_m$ are mutually orthogonal), with

$$H(\boldsymbol{\theta})=\begin{pmatrix} h_{11}(\boldsymbol{\theta}) & 0 & \cdots & 0 \\ 0 & h_{22}(\boldsymbol{\theta}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{mm}(\boldsymbol{\theta}) \end{pmatrix}$$

Furthermore, if

$$\{h_{jj}\}^{\frac{1}{2}} = f_j(\theta_j)g_j(\boldsymbol{\theta})$$

where $g_j(\boldsymbol{\theta})$ does not depend on θ_j, the reference prior, $\pi(\boldsymbol{\theta})$ is given by

$$\pi(\boldsymbol{\theta}) \propto \prod_{j=1}^{m} f_j(\theta_j) .$$

Lemma 2.2 For the capture-recapture model (2.3), let $p_0=0$

$$E(X_i) = N(1-p_1)(1-p_2)\cdots(1-p_{i-1})p_i , \quad i=1,\ldots,s$$

and

$$E(Y_i) = Nc_i\left\{p_1+(1-p_1)p_2+\cdots+p_{i-1}\prod_{j=1}^{i-2}(1-p_j)\right\} , \quad i=2,\ldots,s.$$

Proof. Its proof is very straightforward using the double expectations and mathematical inductions.

Lemma 2.3 For $i\geq 3$ and any p_i,

$$p_1+(1-p_1)p_2+\cdots+p_{i-1}\prod_{j=1}^{i-2}(1-p_j)+\prod_{j=1}^{i-1}(1-p_j)=1.$$

Proof. It easily follows from simple algebra.

Theorem 2.1 Under the capture-recapture model with nontrap response model M_0, the reference prior of $\underline{p}$ is

$$\pi_0 \propto \frac{1}{\sqrt{\prod_{i=1}^{s} p_i(1-p_i)}}. \tag{2.6}$$

Proof. The likelihood under the model M_0 is

$$L(N,\underline{p}|data) = \binom{N}{x_1} p_1^{x_1}(1-p_1)^{N-x_1} \prod_{i=2}^{s}\left\{\binom{N-N_i}{x_i}\binom{N_i}{y_i} p_i^{x_i+y_i}(1-p_i)^{N-(x_i+y_i)}\right\} \tag{2.7}$$

Using Lemma 2.2, the Fisher information matrix under M_0 is given by

$$H(\underline{p})=\begin{pmatrix} \frac{N}{p_1(1-p_1)} & 0 & \cdots & 0 \\ 0 & \frac{N}{p_2(1-p_2)} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \frac{N}{p_s(1-p_s)} \end{pmatrix}$$

where $\underline{p}=(p_1,p_2,\cdots,p_s)$Then set

$$\{h_{jj}(\underline{p})\}^{\frac{1}{2}}=f_j(p_j)g_j(\underline{p})$$

where

$$f_j(p_j)=\frac{1}{\sqrt{p_j(1-p_j)}}, \quad g_j(\underline{p})=N.$$

Therefore, by the equation(2.5) in Lemma 2.1, the reference prior under M_0 is given by

$$\pi_0 \propto \frac{1}{\sqrt{\prod_{i=1}^{s} p_i(1-p_i)}}.$$

Theorem 2.2 For the capture-recapture model with the trap response model M_1 in (2.3), the reference prior distribution for $(p_1,\cdots,p_s,c_2,\cdots,c_s)$is given:

$$\pi_1 \propto \frac{1}{\sqrt{p_1(1-p_1)\prod_{i=2}^{s} p_i(1-p_i)c_i(1-c_i)}}. \tag{2.8}$$

Proof. Under the model M_1 the Fisher information matrix is diagonal matrix composed by

$$H(c)=\begin{pmatrix} H_1(\underline{p},\underline{c}) & 0 \\ 0 & H_2(\underline{p},\underline{c})\end{pmatrix} \tag{2.9}$$

where

$$H_1(\underline{p},\underline{c})=\begin{pmatrix} \frac{N}{p_1(1-p_1)} & & & & 0 \\ & \frac{N(1-p_1)}{p_2(1-p_2)} & & & \\ & & \frac{N(1-p_1)(1-p_2)}{p_3(1-p_3)} & & \\ & & & \ddots & \\ 0 & & & & \frac{N\prod_{i=1}^{s-1}(1-p_i)}{p_s(1-p_s)} \end{pmatrix}$$

and

$$H_2(\underline{p},\underline{c})=\begin{pmatrix} \frac{Np_1}{c_2(1-c_2)} & & & & 0 \\ & \frac{N(p_1+(1-p_1)p_2)}{c_3(1-c_3)} & & & \\ & & \frac{N(p_1+(1-p_1)p_2+(1-p_1)(1-p_2)p_3)}{c_4(1-c_4)} & & \\ & & & \ddots & \\ 0 & & & & h_{s,s} \end{pmatrix}$$

where

$$h_{s,s}=\frac{N(p_1+(1-p_1)p_2+(1-p_1)(1-p_2)p_3+\cdots+(1-p_1)\cdots(1-p_{s-2})p_{s-1})}{c_s(1-c_s)}$$

Using Lemma 2.3, $H_2(\underline{p},\underline{c})$ can be simply expressed as

$$H_2(\underline{p},\underline{c})=\begin{pmatrix} \frac{Np_1}{c_2(1-c_2)} & & & & 0 \\ & \frac{N(1-\prod_{i=1}^{2}(1-p_i))}{c_3(1-c_3)} & & & \\ & & \frac{N(1-\prod_{i=1}^{3}(1-p_i))}{c_4(1-c_4)} & & \\ & & & \ddots & \\ 0 & & & & \frac{N(1-\prod_{i=1}^{s-1}(1-p_i))}{c_s(1-c_s)} \end{pmatrix}$$

Let

$$\{h^1_{jj}(\underline{p})\}^{\frac{1}{2}}=f_j(p_j)g_j(\underline{p}) \qquad j=1,\cdots,s$$

and

$$\{h^2_{jj}(\underline{c})\}^{\frac{1}{2}}=f_j(c_j)g_j(\underline{c}) \qquad j=2,\cdots,s$$

where

$$f_j(p_j)=\frac{1}{\sqrt{p_j(1-p_j)}} \quad , \quad g_j(\underline{p})=\sqrt{N\prod_{i=1}^{j-1}(1-p_i)} \quad , \ j=1,,\cdots,\ s$$

and

$$f_j(c_j)=\frac{1}{\sqrt{c_j(1-c_j)}} \quad , \quad g_j(\underline{c})=\sqrt{N(1-\prod_{i=1}^{j-1}(1-c_i))} \quad , \quad j=2,\ \cdots,\ s,$$

and h^k_{jj} means the diagonal element of $H_k(\underline{p},\underline{c}),\ k=1,2.$

Therefore, by the equation (2.5) in Lemma 2.1 the reference prior is given by

$$\pi_1 \ \propto \ \frac{1}{\sqrt{p_1(1-p_1)\prod_{i=2}^{s}p_i(1-p_i)c_i(1-c_i)}} \ .$$

Corollary 2.1 Under the model M_1 in (2.3), the reference prior distribution of any permutation of $\{p_1,\cdots,p_s,c_2,\cdots,c_s\}$ is given by

$$\pi_1 \ \propto \ \frac{1}{\sqrt{p_1(1-p_1)\prod_{i=2}^{s}p_i(1-p_i)c_i(1-c_i)}} \ . \tag{2.10}$$

Proof. Its proof is similar to the proof of Theorem 2.2 .

Remark 2.1 From the Fisher information matrix in (2.9), the Jeffreys' prior of $(p_1,\cdots,p_s,c_2,\cdots,c_s)$ is

$$\pi^J(p_1,\cdots,p_s,c_2,\cdots,c_s) = \sqrt{\frac{1}{p_1(1-p_1)}\prod_{i=2}^{s}\left\{\frac{\prod_{j=1}^{i-1}(1-p_j)}{p_i(1-p_i)}\times\frac{(1-\prod_{j=1}^{i-1}(1-p_j)}{c_i(1-c_i)}\right\}}$$

3. Posterior Distribution

Throughout this section we use the following notations. Brackets are used to denote densities; for example, joint, conditional and marginal forms as $[X,Y]$, $[X|Y]$ and $[X]$, respectively.

3.1 Capture-Recapture Model with nontrap response M_0

Under the reference prior of $\underline{p}=(p_1,\cdots,p_s)$ in (2.6), the joint posterior distribution is

$$[N,\underline{p}|data]\propto\binom{N}{r}\prod_{i=1}^{s}\{p_i^{x_i+y_i}(1-p_i)^{N-x_i-y_i}\}\frac{1}{\sqrt{\prod_{i=1}^{s}p_i(1-p_i)}}\ \pi(N)\ . \tag{3.1}$$

By the direct computation, the marginal posterior distribution of N given data is obtained as follows;

$$[N|data] = \int[N,\underline{p}|data]\,dp_1\cdots dp_s$$

$$\propto\binom{N}{r}\pi(N)\prod_{i=1}^{s}\left\{\frac{\Gamma(x_i+y_i+\frac{1}{2})\Gamma(N-x_i-y_i+\frac{1}{2})}{\Gamma(N+1)}\right\}\ . \tag{3.2}$$

Thus we can get the recursive relation of $[N|data]$ as follows;

$$\frac{[N+1|data]}{[N|data]}=\frac{1}{(N+1-r)(N+1)^{s-1}}\frac{\pi(N+1)}{\pi(N)}\prod_{i=1}^{s}\left\{N-x_i-y_i+\frac{1}{2}\right\}. \tag{3.3}$$

Since $\frac{[N+1|data]}{[N|data]}$ is expressed as the form in (3.3), $[N|data]$ may sometimes be approximated rather than accurately by recursion, if $\frac{\pi(N+1)}{\pi(N)}$ is readily available such as when $\pi(N)$ is Poisson. But we cannot get any analytic forms of the marginal posterior densities of $[p_i|data]$. So we apply Gibbs sampler to get the marginal posterior density $[p_i|data]$. Then the following conditional distributions are needed;

$$[N|\underline{p}, data]=\binom{N}{r}\left\{\prod_{i=1}^{s}(1-p_i)\right\}^{N}\pi(N), \tag{3.4}$$

for $i=1,\cdots,s$

$$[p_i|N,\ p_j,\ j\neq i,\ data]=\text{Beta}(x_i+y_i+\frac{1}{2},N-x_i-y_i+\frac{1}{2}), \tag{3.5}$$

where $\text{Beta}(a,b)$ means beta distribution function with parameters a and b.

For the simulation of N from $[N|\underline{p}, data]$

if $\pi(N)=P_0(\lambda)$ where P_0 denotes Poisson with mean λ, then

$$[N|\underline{p}, data]\propto\frac{N!}{r!(N-r)!}\left\{\prod_{i=1}^{s}(1-p_i)\right\}^{N}\frac{e^{-\lambda}\lambda^{N}}{N!}\propto\frac{1}{(N-r)!}\left\{\lambda\prod_{i=1}^{s}(1-p_i)\right\}^{N-r},$$

and thus

$$[N-r|\underline{p}, data]\sim P_0(\lambda\prod_{i=1}^{s}(1-p_i)).$$

If $\pi(N)=N^{-1}$, then

$$[N|\underline{p}, data]\propto\frac{N!}{r!(N-r)!}\left\{\prod_{i=1}^{s}(1-p_i)\right\}^{N}\frac{1}{N}$$

$$\propto\frac{(N-1)!}{(r-1)!(N-r)!}\left\{1-\prod_{i=1}^{s}(1-p_i)\right\}^{r}\left\{1-(1-\prod_{i=1}^{s}(1-p_i))\right\}^{N-r},$$

and thus

$$[N|\underline{p}, data]\sim NB(r,1-\prod_{i=1}^{s}(1-p_i))$$

where $NB(r,p)$ denotes negative binomial distribution function with parameters r and p.

Suppose Gamma-mixed Poisson is given. Then

$$[N|\underline{p}, data]\propto\frac{N!}{(N-r)!}\left\{\prod_{i=1}^{s}(1-p_i)\right\}^{N}\frac{\Gamma(N+a)(1+\frac{1}{b})^{-(N+a)}}{N!\Gamma(a)b^{a}}$$

$$\propto \frac{(N+a-1)!}{(a-1)!(N-r)!}\left\{1-\prod_{i=1}^{s}(1-p_i)\frac{b}{b+1}\right\}^{r}\left\{1-(1-\prod_{i=1}^{s}(1-p_i)\frac{b}{b+1})\right\}^{N-r} .$$

Therefore,

$$[N+a|\underline{p}, data] \sim NB\left(a+r, 1-\prod_{i=1}^{s}(1-p_i)\frac{b}{b+1}\right)$$

where a and b are shape and scale parameters of gamma distribution respectively.

3.2 Capture-Recapture Model with trap response M_1

Under the reference prior of $(\underline{p}, \underline{c}) = (p_1, \cdots, p_s, c_2, \cdots, c_s)$ (2.4), the joint posterior distribution is

$$[N, \underline{p}, \underline{c}|data] \propto \binom{N}{r}\prod_{i=1}^{s} p_i^{x_i}(1-p_i)^{N-N_i-x_i}\prod_{i=2}^{s} c_i^{y_i}(1-c_i)^{N_i-y_i}$$

$$\times \frac{1}{\sqrt{p_1(1-p_1)\prod_{i=2}^{s}p_i(1-p_i)c_i(1-c_i)}}\pi(N) . \qquad (3.6)$$

By the direct computation, the marginal posterior distribution of N given data is obtained as follows;

$$[N|data] = \int [N, \underline{p}, \underline{c}|data]\, dp_1 \cdots dp_s\, dc_2 \cdots dc_s$$

$$\propto \prod_{i=2}^{s}\left\{\frac{\Gamma(y_i+\frac{1}{2})\Gamma(N_i-y_i+\frac{1}{2})}{\Gamma(N_i+1)}\frac{\Gamma(x_i+\frac{1}{2})\Gamma(N-N_i-x_i+\frac{1}{2})}{\Gamma(N-N_i+1)}\right\}$$

$$\times \frac{\Gamma(x_1+\frac{1}{2})\Gamma(N-x_1+\frac{1}{2})}{\Gamma(N+1)}\binom{N}{r}\pi(N) . \qquad (3.7)$$

Since we cannot get any analytic forms of the marginal posterior densities, we apply Gibbs sampler to get the marginal posterior density $[p_i|data]$. Then the following conditional distributions are needed;

$$[N|\underline{p}, data] = \binom{N}{r}\prod_{i=1}^{s}(1-p_i)^{N}\pi(N) \qquad (3.8)$$

for $i=1,\cdots,s$

$$[p_i|N, p_j, j\neq i, \underline{c}, data] = \text{Beta}(x_i+\frac{1}{2}, N-N_i-x_i+\frac{1}{2}) \qquad (3.9)$$

and for $i=2,\cdots,s$

$$[c_i|N, \underline{p}, c_j, j\neq i, data] = \text{Beta}(y_i+\frac{1}{2}, N_i-y_i+\frac{1}{2}) . \qquad (3.10)$$

Note that by the structure of $[c_i\,|\,N, \underline{p}, c_j, j\neq i, data]$ only $[N|\underline{p}, \underline{c}, data]$ and $[p_i|N, p_j, j\neq i, \underline{c}, data]$ are needed for Gibbs sampler. The simulation of N from $[N|\underline{p}, data]$, is exactly same as the method in nontrap model.

4. Bayesian Model Choice

In this section, we test the capture-recapture model with nontrap response M_0 v.s. the capture-recapture model with trap response model M_1 using Bayes factor (BF). The formal Bayesian model choice procedure goes as follows. Let w_i be the prior probability of $M_i, i=0,1$ and let $f(y|M_i)$ be the predictive distribution for the model M_i , i.e.

$$[m]_i = f(y|M_i) = \int f(y \mid p, M_i)\pi(p|M_i)dp \ . \quad (4.1)$$

If y is the observed data, then we choose the model yielding the larger $w_i f(y|M_i)$. Often we set $w_i=\frac{1}{2}$ and compute the Bayes factor (or M_0 with respect to M_1), given by

$$BF = \frac{f(y|M_0)}{f(y|M_1)} = \frac{[m]_0}{[m]_1} \ .$$

Kass and Raftery(1995) suggested interpretive ranges for the Bayes factor. In general, M_0 is supported if $BF>1$.

More generally, assume that Y is distributed to $f(y|\beta)$ and $\pi(\beta)$ is the prior of β. Then we want to estimate $[m] = \int f(y|\beta)\pi(\beta)\,d\beta$ using the importance sampling method. Let us consider $\pi(\beta|y)$ as the importance sampling function. Then the Markov Chain Monte Carlo method, particularly Metropolis algorithm and Gibbs sampler, is used to get the sample from the posterior density $\pi(\beta|y)$. Let $\{\beta^{(g)}\}_{g=1}^{G}$ be Gibbs outputs as above where G means we repeat the Gibbs sampler G times. Then by Monte Carlo method, the approximating marginal density of Y is $\widehat{[m]}=\frac{\sum_{g=1}^{G} w_g f(y|\beta^{(g)})}{\sum_{g=1}^{G} w_g}$, where $w_g=\frac{\pi(\beta^{(g)})}{\pi(\beta^{(g)}|y)}$. Since $\pi(\beta|y) = \frac{f(y|\beta)\pi(\beta)}{[m]}$, the approximation can be expressed as

$$\widehat{[m]}=\left[\frac{1}{G}\sum_{g=1}^{G}\frac{1}{f(y|\beta^{(g)})}\right]^{-1}. \quad (4.3)$$

Also, this final form is mentioned in Kass and Raftery(1995) and Chung(1997) used the form (4.3) to choose the link function of binary regression model using Bayesian approach.

In our structure, the following BF is used to compare nontrap response model M_0 with trap response model M_1

$$BF\approx\frac{\left[\frac{1}{G_1}\sum_{g=1}^{G_1}\frac{1}{[data|N^{(g)},p^{(g)}]_0}\right]^{-1}}{\prod_{j=2}^{s}\left\{\Gamma(y_j+\frac{1}{2})\Gamma(N_j-y_j+\frac{1}{2})\,\Gamma(N_j+1)\right\}\left[\frac{1}{G_2}\sum_{g=1}^{G_2}\frac{1}{D[data|N^{(g)},p^{(g)}]_1}\right]^{-1}}\ , \quad (4.4)$$

where

$$[data|N, \underline{p}]_0 = \prod_{i=1}^{s}\left\{\binom{N-N_i}{x_i}\binom{N_i}{y_i}p_i^{x_i+y_i}(1-p_i)^{N-x_i-y_i}\right\}$$

and $\{N^{(g)}, p^{(g)}\}_{g=1}^{G_1}$ is Gibbs output using the full conditional distributions $[N|\underline{p}, data]$ and $[p_i|N, p_j, j\neq i, data]$in (3.4) and (3.5) respectively, and

$$[data|N, \underline{p}]_1 = \prod_{i=1}^{s}\left\{\binom{N-N_i}{x_i}\binom{N_i}{y_i}p_i^{x_i}(1-p_i)^{N-N_i-x_i}\right\}$$

and $\{N^{(g)}, p^{(g)}\}_{g=1}^{G_2}$ is Gibbs output using the full conditional distributions $[N|\underline{p}, data]$ and $[p_i|N, p_j, j\neq i, data]$in (3.8) and (3.9) respectively.

5. Applications

5.1 Simulation Study

We illustrate our proposed method with generated data under different priors. Data set in Table 1 is generated with different values of p_i and c_i. Therefore, the generated data is considered as that came from the capture-recapture model with trap response M_1.

Table 1 Generated Data Sets

i	1	2	3	4	5	6	7	8	9
p_i	0.10	0.12	0.13	0.15	0.09	0.05	0.10	0.12	0.20
c_i	0.00	0.09	0.08	0.10	0.09	0.152	0.08	0.10	0.13
x_i	24	27	27	27	23	9	19	21	28
y_i	0	2	4	8	10	23	11	19	17

For the reference priors in (2.6) and (2.8) and Poisson prior of N, we generate a Gibbs sequence of length 20,000 with 100 different initial values using IMSL subroutines. Table 2 lists the means, the standard deviations and a 95% credible intervals for estimated N obtained from the 2.5% and 97.5% quantiles.

Table 2. Posterior characteristics of N

	M_0			M_1		
λ	Mean of N	S.D. of N	95% credible interval	Mean of N	S.D. of N	95% credible interval
300	311	34.087	(264,365)	259	58.114	(207,370)
400	378	45.367	(309,458)	379	59.812	(237,463)
500	455	45.659	(360,539)	490	34.610	(433,541)

In table 2, S.D. denotes the standard deviation. The wide variability in posterior characteristics shows the sensitive dependence on fixed choices of λ. Since the choice of the estimate of λ is not of interest in this paper, the estimate of λ is not decided but we can determine its estimate using the empirical Bayes method or ML-II method. Next we consider the Bayesian model choice using the approximate Bayes factor in (4.4) with different value of λ . Table 3 indicates that our generated data is fitted to capture-recapture model with trap response M_1 since log(BF) is very smaller than zero.

Table 3. Bayes Factor with Poisson(λ) prior

λ	Log(BF)	Choice of model
300	-19.11	M_1
400	-5.695	M_1
500	-32.036	M_1

5.2 Gordy Lake Sunfish Data

We briefly illustrate our techniques on the famous Gordy Lake sunfish data set investigated by Castledine(1981) and George and Robert(1992). As it is shown in Table 4, it consists of s=14 capture occasions from a population of sunfish. At the i^{th} sample, n_i fish are captured out of which y_i have been previously captured. Thus, $r = \sum_{i=1}^{14} (n_i - y_i) = \sum_{i=1}^{14} x_i = 138$ is the total number of fishes which are captured differently .

Table 4. Gordy Lake Sunfish Data

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
x_i	10	27	17	7	1	5	6	15	9	18	16	5	7	19
y_i	0	0	0	0	0	0	2	1	5	5	4	2	2	3

Castledine(1981) applied the capture-recapture model with nontrap response M_0 to this data using the prior formulation $\pi(N, \underline{p}|a, b) = \pi(N)\prod\pi(p_i|a, b)$ with Jeffreys' prior $\pi(N) = N^{-1}$ and $\pi(p_i|a, b) = \text{Beta}(a, b)$ for various fixed (a, b). But instead of treating (a,b) as fixed, George and Robert(1992) pursued the hierarchical approach of putting five different priors on (a,b). They also used the Gibbs sampler to avoid the complicated Bayesian computation. Here, we consider the reference priors in (2.6) and (2.8) and Poisson prior N into the models M_0 and M_1.

Setting initial values by maximum likelihood, we simulated the values of $p^{(k)}$, $c^{(k)}$ and $N^{(k)}$ in the Gibbs sequence using IMSL subroutine. Then with different prior parameters, Table 5 show the posterior characteristics.

Table 5. Posterior characteristics of N Poisson(λ) prior

	M_0			M_1		
λ	Mean of N	S.D. of N	95% credible interval	Mean of N	S.D. of N	95% credible interval
300	308	41.17	(248,355)	264	60.67	(138,331)
400	394	49.25	(325,454)	386	37.16	(314,439)
500	490	54.80	(412,556)	489	42.34	(420,539)

Next we compare the nontrap model M_0 with the trap model M_1 using the Bayes factor given in (4.4) . Table 6 says that the trap model M_1 is strongly supportive since $BF \ll 10^{-2}$ regardless of the values of λ.

Table 6. Bayes Factor with Poisson(λ) prior

λ	Log(BF)	Choice of model
300	-10.2473	M_1
400	-8.1746	M_1
500	-3.8278	M_1

References

[1] Berger, J. and Bernardo, J.(1992), On the development of the reference prior method. In: J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith. *Bayesian Statistics 4*, Oxford University Press, London.

[2] Bernardo, J.(1979), Reference posterior distributions for Bayesian inference(with discussion). *Journal of the Royal Statistical Society*, B. Vol. 41, 113-147.

[3] Bernardo, J. and Smith, A.F.M(1994), *Bayesian Theory* , John Wiley & Sons, New York.

[4] Castledine, B.(1981), A Bayesian analysis of the multiple-recapture sampling for a close population. *Biometrika*, Vol. 67, 1,197-210.

[5] Chapman, D.G.(1952), Inverse, multiple and sequential sample censuses. *Biometrics*, Vol. 8, 286-306.

[6] Chung, Y.(1997), Sampling based approach to Bayesian analysis of binary regression model with incomplete data *Journal of the Korean Statistical Society*, Vol. 26, 493-506.

[7] Chung, Y. and Dey, D.K.(1998), Bayesian approach to estimation of intraclass correlation using reference prior. *Communications in Statistics-Thoery and Methods,* Vol 27, 2241-2255.

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

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

[10] Kass, R.E. and Raftery, A.E.(1995), Bayes factors. *Journal of the American Statistical Association,* Vol. 90, 773-795.

[11] Pollock, K.(1975), Building models of capture-recapture experiments. *The Statistician,* Vol. 25, 253-259.

[12] Polloack, K. and Otto, M.(1983), Robust estimation of population size in closed animal population from capture-recapture experiments. *Biometrika* Vol. 39, 1035-1049.

[13] Rodrigues, J. Bolfrine, H. and Leite, J.G.(1988), A Bayesian analysis in closed animal populations from capture recapture experiments with trap response. *Communication in Statistics - Simulation and Computations,* Vol. 17, 407-430 .

[14] Sun, D. and Ye, K.(1995), Reference prior Bayesian analysis for normal mean products. *Journal of the American Statistical Association,* Vol. 90, 589-597.

[15] Ye, K.(1994), Bayesian reference prior analysis on the ratio of variances for the balanced one-way random effect model. *Journal of the Statistical Planning and Inference,* Vol. 41, 267-280.