The Korean Communications
in Statistics Vol. 7, No. 1, 2000
pp. 211-224

Markov Chain Monte Carlo Estimation in Two Successive Occasion Sampling with Randomized Response Model

Kay-O Lee[1)]

Abstract

The Bayes estimation of the proportion in successive occasions sampling with randomized response model is discussed by means of Acceptance Rejection sampling. Bayesian estimation of transition probabilities in two successive occasions is suggested via Markov Chain Monte Carlo algorithm and its applicability is represented in a numerical example.

1. Introduction

In social public surveys, an unknown amount of error could have arisen from non-response or incorrect response when survey items are sensitive or stigmatizing. Using the randomized response technique reduced this bias. Even in the instances when the survey questions are sensitive, bayesian approach was successful in obtaining an estimate using information on an auxiliary variable.

In recent times Markov Chain Monte Carlo method has been used in Bayesian models to obtain either a random sample from the posterior distribution or some functionals of the posterior distribution as the posterior distribution which often becomes mathematically or computationally intractable by traditional methods. To our knowledge, it has not been used in the context of randomized response model, and such a study could be very useful in application of bayesian inference to sensitive surveyed data.

In many sample surveys, successive occasion sampling scheme is used to obtain time series information as well as more efficient estimates. When a successive occasion sampling scheme is implemented in the survey of sensitive subjects or public opinion survey on election, the main interest is not only in getting precise estimate on current occasion but also in estimating the change of the parameter over two successive occasions.

In particular, when the study variable is dichotomous, the change in the characteristics over successive occasions can be parameterized in terms of transition probabilities.

The objective of this article is to obtain the bayesian estimate of the characteristics of dichotomous variable in randomized response model via Monte Carlo method and to implement Monte Carlo simulation for estimating the transition probability.

---

1) Professor, Department of Computer Science and Statistics, Korea Air Force Academy, Chungbuk, 363-849

The bayesian estimation of randomized response model using Monte Carlo method is dealt with in section 2, and estimation of transition probability using Markov Chain Monte Carlo method is discussed in section 3.

2. Bayes Estimation of Warner's Model

2.1 Background

In Warner's randomized response model, the respondent uses a randomization device to choose one of two questions; " Do you belong to Group A?" or " Do you belong to Group A^c?" with known probabilities P and $1-P$, and then replies " yes" or "no" to the randomly chosen question.

We assume that n respondents are selected by simple random sampling and the parameter of interest is θ_A, the proportion of sensitive Group A of whole population.

Let X_i be the response of the ith individual.

$$\text{where} \quad X_i = \begin{cases} 1 & \text{if i-th respondent replies "yes"} \\ 0 & \text{if i-th respondent replies "no"} \end{cases}$$

Then the probability of getting a "yes" response, defined by λ as follows,

$$\lambda = \Pr(X_i = 1) = (2P-1)\theta_A + (1-P)$$

If the number of "yes" respondents in the sample is denoted by $Y = \sum_{i=1}^{n} X_i$, then the likelihood function of for given observed data is expressed by

$$l(\theta_{A;}y, n) \approx [(2P-1)\theta_A + (1-P)]^y [P-(2P-1)\theta_a]^{n-y}, \ 0 < \theta_A < 1 \qquad (1)$$

Suppose that the prior distribution of θ_A is a beta distribution with parameters α and β.

Hence the posterior distribution of θ_A for the given data is given by

$$f(\theta_A | y, n) = \frac{l(\theta_A; y, n) p(\theta_A)}{\int_0^1 l(\theta_A; y, n) p(\theta_A) d\theta_A} \qquad (2)$$

Often the normalizing constant can not be analytically computed, so the posterior distribution may be written as.

$$f(\theta_A | y, n) \approx [(2P-1)\theta_A + (1-P)]^y [P-(2P-1)\theta_A]^{n-y} \theta_A^{\alpha-1} (1-\theta_A)^{\beta-1} \qquad (3)$$

where the normalizing constant is a function of y and n.

From (3) it is apparent that the posterior density is a mixture of beta distributions with a large number of terms for even moderate n. Mean and variance of the posterior distribution are extremely cumbersome to compute.

In this paper we shall discuss the utility of Monte Carlo method to obtain Bayes estimate of θ_A. The Monte Carlo method using the Acceptance-Rejection method of sampling is the main focus of this paper.

2.2. Acceptance-Rejection Sampling

(1) Algorithm

An appropriate procedure of implementation of Acceptance-Rejection(A-R) sampling technique are described as follows:

(1) Choose a M such that $M = l(\widehat{\theta_A}; y, n)$ where $\widehat{\theta_A}$ maximizes the likelihood function $l(\theta_A; y, n)$ given by equation (1).

(2) Generate θ_A^i from $p(\theta_A)$

(3) Generate U_i from uniform (0,1)

(4) If $U_i \le f_y(\theta_A^i)/(M \cdot p(\theta_A^i))$, then accept θ_A^i as a sample; otherwise, repeat steps (2), (3), and (4).

(5) Repeat steps (2), (3), and (4) until the desired numbers of samples are obtained.

(6) Compute the sample mean,

$$\overline{\theta_A} = \frac{1}{m} \sum_{i=1}^{m} \theta_A^i \tag{4}$$

It can be easily proved that a generated variable θ_A^i in step (4) of the preceding procedure is a random sample from the posterior distribution $f(\theta_A | y, n)$ of θ_A for the given data. The Acceptance-Rejection sampling method used here may be described as a method of drawing a sample from a p.d.f. where the kernel of the density $f_y(\theta_A)$ is known, but not be the density function itself. The proof is straightforward.

Let's define two sets as follows:

$A_0 = \{(\theta_A, U): \theta_A \le \theta_0,\ U \le f_y(\theta_A)/M \cdot p(\theta_A)\}$, and

$A = \{(\theta_A, U): 0 < \theta_A < 1,\ U \le f_y(\theta_A)/M \cdot p(\theta_A)\}$

Then the distribution function of accepted θ_A is computed,

$$\Pr(\theta_A \le \theta_0 | \theta_A \text{ accepted}) = \frac{\mathrm{P_r}(\theta_A \le \theta_0, \theta_A \text{ accepted})}{\mathrm{P_r}(\theta_A \text{ accepted})}$$

$$= \frac{\int\int 1_{A_0}(\theta_A, U)P(\theta_A)dUd\theta_A}{\int\int 1_A(\theta_A, U)p(\theta_A)dUd\theta_A}$$

$$= \frac{\int_0^{\theta_0}\int_0^{c(\theta_A)} p(\theta_A)dUd\theta_A}{\int_0^{1}\int_0^{c(\theta_A)} p(\theta_A)dUd\theta_A}, \quad \text{where } c(\theta_A) = \frac{f_y(\theta_A)}{M \cdot p(\theta_A)}$$

$$= \frac{\int_0^{\theta_0} f_y(\theta_A)d\theta_A}{\int_0^{1} f_y(\theta_A)d\theta_A}$$

It follows that an accepted θ_A has the probability density $f(\theta_A|y,n) \propto f_y(\theta_A)$.

(2) Numerical Example

To illustrate the A-R sampling technique proposed in the previous subsection, let's consider the following data as the observations of a randomized response model ; n=20, y=8 (number of "yes" respondent), P=0.85.

The maximum likelihood estimator of can be computed by θ_A and its variance is obtained by $\widehat{\theta_A} = (y/n - 1 + P)/(2P - 1)$

$$Var(\widehat{\theta_A}) = \frac{\theta_A(1-\theta_A)}{n} + \frac{P(1-P)}{n(2P-1)^2}$$

The computed estimate of $\widehat{\theta_A}$ is 0.357 and estimate of variance of $\widehat{\theta_A}$ is 0.0245. These estimates could be used in analyzing the results of numerical example.

To implement the algorithm, the values of parameters of beta prior density function have to be appropriately chosen. In this subsection, we consider two cases so that the influence of choice of prior density function might be examined. In one case uniform density function is considered as a prior beta function, $\alpha = \beta = 1$. In the other case, the estimates of parameters of beta density function are obtained via the moment matching method.

In the present case we apply an empirical approach to choice of the parameters of beta prior density function. When the information of prior density function can be obtained from the previous surveys or the judgement of expert in the study topic, the mean and variance of prior distribution can be computed approximately. The predetermined mean and variance are equated to the corresponding to them of prior distribution. This may be considered as being variant of empirical bayes approach, but posterior distribution of given by (3) can be expanded into a mixture of binomial terms. Hence, this method can be used in the choice of prior parameters α and β without violation of the spirit of empirical bayes approach.

$$\alpha_0 = \theta_A\left\{\frac{\theta_A(1-\theta_A)}{Var(\theta_A)} - 1\right\}$$

$$\beta_0 = (1-\theta_A)\left\{\frac{\theta_A(1-\theta_A)}{Var(\theta_A)} - 1\right\} \tag{5}$$

For choosing the appropriate range of α and β by using the previous information, values of α and β are decided on the basis of the rejection rate of samples in practical simulation studies. The computed values of α and β are α=3.0 and β=5.5.

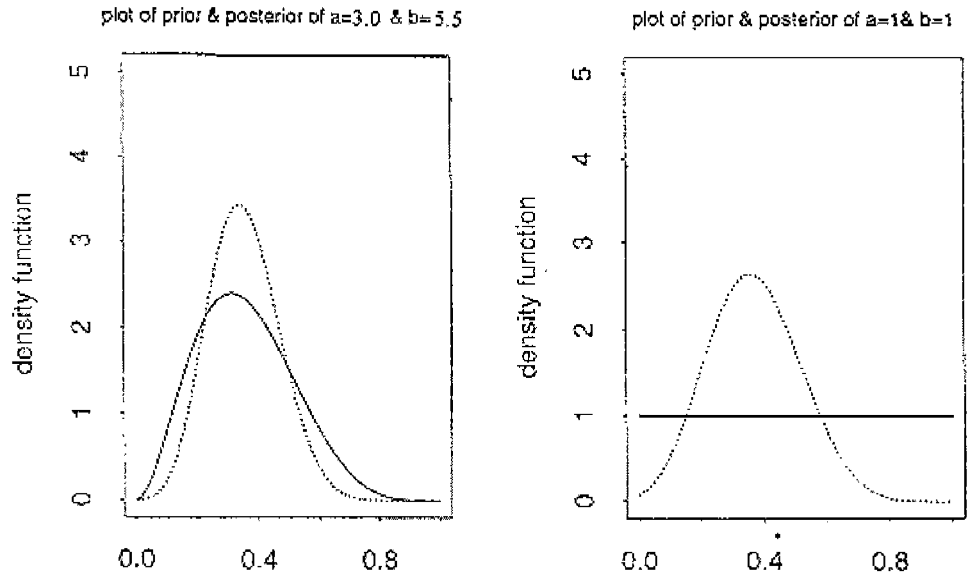

Figure 1. Plot of Prior and Posterior Functions

The prior and posterior density functions are illustrated in figure 1. In the beta prior density given by moment method, it can be easily observed that the posterior density function is very similar to the prior density.

Before making any judgements on the results of implementation of A-R sampling, the properties of Monte Carlo method should be explained. If the size of sample is not sufficiently large, then the characteristics of posterior distribution computed by A-R sampling can not be stable.

The results of implementation of A-R sampling for several sizes of samples are given in Table 1.

The illustrations of Table 1 are summarized as follows: the effect of two prior distributions on the estimate is indicated to be accounted much of determination of prior parameters. The uniform prior distribution provides a positive shift in the estimate toward the prior mean of θ_A, 0.5. As the size of sample increases the estimate of posterior mean becomes more stable and the estimated standard deviation becomes small. If the sample size is less than 750, then estimates of mean and standard deviation can fluctuate. This means the size of sample has to be larger than 750.

Table 1. Estimates of Mean and Standard Deviation

Beta Prior	$\alpha=3.0,\ \beta=5.5$		$\alpha=1.0,\ \beta=1.0$	
Parameter	mean	standard deviation	mean	standard deviation
n=100	0.3689	0.12515	0.3807	0.14347
n=250	0.3469	0.11586	0.3626	0.14560
n=500	0.3490	0.11391	0.3624	0.14914
n=750	0.3512	0.11143	0.3694	0.14859
n=1000	0.3520	0.11140	0.3689	0.14869
n=2000	0.3531	0.11260	0.3669	0.14679

3. Bayesian Estimation of Transition Probability Via Markov Chain Monte Carlo Method

Under the two successive occasion sampling scheme the estimation of the trend of characteristic change over two occasions is derived from the property of Markov chain with stationary distribution. When the study variable is dichotomous, a system of stationary distribution of Markov chain can be expressed in terms of transition matrix as follows:

$$(\pi_t(A), \pi_t(A^c)) \begin{pmatrix} \pi_{11} & \pi_{10} \\ \pi_{01} & \pi_{00} \end{pmatrix} = (\pi_{t+1}(A), \pi_{t+1}(A^c)) \tag{6}$$

where $\pi_t(A)$ is the probability that a unit belongs to sensitive group (A) on the t occasion.

π_{11} is the transition probability that the unit belonged to sensitive group on the t and (t+1) occasions.

π_{01} is the transition probability that the unit did not belong to sensitive group on the t occasion and belongs to sensitive group on the (t+1) occasion.

$$\pi_{10} = 1 - \pi_{11}$$

$$\pi_{00} = 1 - \pi_{01}$$

If we can estimate the transition matrix, then $\pi_{t+1}(A)$ and $\pi_{t+1}(A^c)$ can be computed. In this section the estimation of transition matrix is mainly discussed.

3.1 Likelihood function of π_{11} and π_{01}

A combination of successive occasions sampling with randomized response technique enables it to estimate not only the proportions of sensitive group on the two occasions but also the trend over the time.

The sampling scheme that provides the estimates of π_{11} and π_{01} is summarized as follows:

On the first occasion, the sample of n units is randomly selected and is investigated by a randomized response technique. On the second occasion, the observed sample on the previous occasion is divided into two partitions according to its reply ; one group is composed of " yes" replied units, the other consists of units that replied "no". Then, two independently matched samples are chosen from each group, respectively, and the investigation is made by randomized response technique.

On the other hand, unmatched sample of u units are chosen independently from the N-n remaining units which were not observed on the first occasion and are investigated by the randomized response method.

N otations to be used in this section are defined as follows:

n_1 is the number of units consisting of "yes" group on the first occasion.

m_1 is the size of subsample drawn from the n_1 units.

m_0 is the size of subsample chosen from the n_0 units.

m_{11} is the number of units that replied "yes" on the second occasion and belongs to m_1

$m_{10} = m_1 - m_{11}$

m_{01} is the number of units that replied " yes" on the second occasion and belongs to m_0

$m_{00} = m_0 - m_{01}$

Assume that the above sampling scheme has some properties of conditional distribution and randomized response model. The likelihood function of transition probabilities π_{11} and π_{01} for given observed data can be written as follows:

$$
\begin{aligned}
&lik((\pi_{11},\ \pi_{01});(m_{11}, m_{01})) = \\
&\qquad k \cdot \{P \cdot \pi_1(A) \cdot [(2P-1)\pi_{11} + (1-P)] + (1-P) \cdot (1-\pi_1(A)) \\
&\qquad \cdot [(2P-1)\pi_{01} + (1-P)]\}^{m_{11}} \{P \cdot \pi_1(A)[P-(2P-1)\pi_{11}]P \cdot \pi_1(A) \qquad (7) \\
&\qquad \cdot [P-(2P-1)\pi_{11}] + (1-P)(1-\pi_1(A))[P-(2P-1)\pi_{01}]\}^{m_1 - m_{11}}
\end{aligned}
$$

$$\cdot \{(1-P)\cdot \pi_1(A)[(2P-1)\pi_{11}+(1-P)]+P(1-\pi_1(A))$$
$$\cdot [(2P-1)\pi_{11}+(1-P)]\}^{m_{01}}\{(1-P)\cdot \pi_1(A)[P-(2P-1)\pi_{11}]$$
$$+P\cdot (1-\pi_1(A))\cdot [P-(2P-1)\pi_{01}]\}^{m_0-m_{01}}$$

where P is the probability choosing sensitive question in randomizing device.

$\pi_1(A)$ is the proportion of units belonging to sensitive group at the first occasion.

π_{11} is the transition probability of which unit belongs to sensitive group on the first and second occasions.

π_{01} is the transition probability of which unit belongs to sensitive group on the first occasion but belongs to non-sensitive group on the second occasion.

Bayes estimators of π_{11} and π_{01} are studied in the following subsections.

3.2 Bayes Estimation of π_{11} and π_{01}

Since observed variable of the mentioned sampling scheme is a kind of dichotomous variable, the likelihood function given in equation (7) is considered as a product of two binomial density functions. Hence, the beta function can be regarded as conjugate prior function of each transition probability.

In this sampling scheme, we can assume that the prior density functions of π_{11} and π_{01} are correlated without loss of generality. For instance, if the value of π_{11} goes up, then π_{01} also increases. On the contrary, if π_{11} decreases, then π_{01} goes down.

Therefore it is reasonable assumption that the prior joint distribution of π_{11} and π_{01} has a positive correlation structure.

The property of beta density function can be used for choosing the positive correlated prior joint distribution as follows : the product of beta random variables has a beta distribution, that is, if $X\sim beta(a,b)$ and $Y\sim beta(a,b+c)$ then $XY\sim beta(a,b+c)$ whenever X and Y are independent.

The prior joint density function of π_{11} and π_{01} may be assumed to have the following form (the detailed derivation is given in appendix).

$$f(\pi_{11},\pi_{01})=k_0\,\pi_{11}^{a_1-1}\pi_{01}^{a_2-1}\int_B(z-\pi_{11})^{b_1-1}(z-\pi_{01})^{b_2-1}z^{1-a}(1-z)^{b-1}dz \qquad (8)$$

where k_0 is the normalizing constant.

π_{11} has beta distribution with parameters a_1 and a_1+b

π_{01} has beta distribution with parameters a_2 and a_2+b

$B=\{(\pi_{11},\pi_{01},z):\ 0<\pi_{11},\pi_{01}<z<1\}$

It can be easily proved that the joint density function of π_{11} and π_{01} has a positive correlation structure, and its correlation coefficient is obtained by

$$Corr(\pi_{11}, \pi_{01}) = \frac{b}{a}\sqrt{\frac{\pi_{11}}{1-\pi_{11}} \cdot \frac{\pi_{01}}{1-\pi_{01}}}$$

The derivation of the correlated prior joint density function is given in the appendix.

Hence the posterior distribution function of π_{11} and π_{01} for given data is expressed by the product of likelihood function and prior joint density function.

$$\begin{aligned} f(\pi_{11}, \ \pi_{01}|m_{11}, m_{01}) \propto & \{P\pi_1(A)[(2P-1)\pi_{11}+(1-P)]+(1-P)(1-\pi_1(A)) \\ & [(2P-1)\pi_{01}+(1-P)]\}^{m_{11}}\{P\pi_1(A)[P-(2P-1)\pi_{11}] \\ & +(1-P)(1-\pi_1(A))[P-(2P-1)\pi_{01}]\}^{m_1-m_{11}}\{(1-P)\pi_1(A) \\ & [(2P-1)\pi_{11}+(1-P)]+P(1-\pi_1(A))[(2P-1)\pi_{01}+(1-P)]\}^{m_{01}} \\ & \{(1-P)\pi_1(A)[P-(2P-1)\pi_{11}]+P(1-\pi_1(A)) \\ & [P-(2P-1)\pi_{01}]\}^{m_0-m_{01}}\pi_{11}^{a_1-1}\pi_{01}^{a_2-1} \\ & \cdot \int_B (z-\pi_{11})^{b_1-1}(z-\pi_{01})^{b_2-1}z^{1-a}(1-z)^{b-1}dz \end{aligned} \tag{9}$$

It is not easy to compute the Bayes estimators of π_{11} and π_{01} with respect to the squared error loss function for given the posterior joint distribution function. Let's consider a numerical method of Bayes estimates of π_{11} and π_{01} via Gibbs sampler.

Since the prior joint density function of π_{11} and π_{01} is defined in terms of the product of two beta random variables for introducing the positive correlation structure, it is not simple to compute numerical integration given in (9).

When Markov Chain Monte Carlo(MCMC) sampler is applied to computation of Bayes estimates of π_{11} and π_{01}, we can easily solve the problems of computing normalizing constant and numerical integration. MCMC method via Gibbs sampler is applied to computation of Bayes estimates and is studied in the following subsection.

The parameters of prior density function should be determined before studying the MCMC method. It is a reasonable method to estimate the parameters of prior density function using the information obtained at the first occasion's survey.

3.3 Markov Chain Monte Carlo Algorithm

For the given joint posterior distribution, the Bayesian estimation of and can be obtained from each marginal posterior distribution. Each mean of marginal posterior distribution can be computed on the basis of Gibbs samples which is a kind of MCMC sampler. The joint posterior given in (9) is a trivariate case, and hence, three conditional posterior distributions are derived so as to apply Gibbs sampling to estimate of posterior means.

Gibbs sampling procedure is performed on the basis of A-R sampling method, and its algorithm is given as follows:

(1) Selection of sample from the conditional posterior density $f_{1.23}(\pi_{11}|\pi_{01}^0, \pi^0)$

(1-1) Compute the Conditional distribution of π_{11} for given π_{10}^0 and Z^0.

$f_{1.23}(\pi_{11}|\ \pi_{01}^0, Z^0) \propto lik((\pi_{11}, \pi_{01}^0);(m_{11,} m_{01})) \cdot f(\pi_{11}, \pi_{01}^0, Z^0)$

$$= c_1(\pi_{01}^0, Z^0) \cdot lik((\pi_{11}, \pi_{01}^0);(m_{11}, m_{01})) \cdot \left(\frac{\pi_{11}}{Z^0}\right)^{a_1-1}\left(1-\frac{\pi_{11}}{Z^0}\right)^{b_1-1} Z^0$$

where $lik((\pi_{11}, \pi_{01}^0);(m_{11,} m_{01}))$ is likelihood function given in (6)

$f(\pi_{11}, \pi_{01}^0, Z^0) = k\pi_{11}^{a_1-1}(Z^0-\pi_{11})^{b_1-1}(\pi_{01}^0)^{a_2-1}(Z^0-\pi_{01}^0)^{b_2-1}(Z^0)^{1-a}(1-Z^0)^{b-1}$

$c_1(\pi_{01}^0, Z^0) = \int_0^1 lik((Z^0\xi, \pi_{01}^0 0;(m_{11}, m_{01}))\xi^{a_1-1}(1-\xi)^{b_1-1}d\xi$

(1-2) Evaluate the maximum value of likelihood function.

$m_1 = \max\{lik((Z^0\xi, \pi_{01}^0);(m_{11}, m_{01}))\}$

(1-3) Draw a sample ξ^1 from beta density function with the parameters a_1 and b_1

(1-4) Draw a uniform sample u_1^1

(1-5) Accept ξ^1 as a sample whenever $u_1^1 \le \dfrac{lik((Z^0\xi, \pi_{01}^0);(m_{11}, m_{01}))}{m_1}$

If reject ξ^1, then repeat (1-3), (1-4) and (1-5) steps until a sample ξ_*^1 is accepted.

Let $\pi_{11}^1 = Z^0\xi_*^1$ and sample π_{01} can be selected from $f_{2.13}(\pi_{01}|\pi_{11}^1, Z^0)$ by the similar process to above. Also, assume π_{01}^1 is a selected sample from $f_{2.13}(\pi_{01}|\pi_{11}^1, Z^0)$ and sample Z^1 is drawn from $f_{3.12}(Z|\pi_{11}^1, \pi_{01}^1)$ with the similar gibbs sampler as the above.

After implementing the 3 steps gibbs sampler,a triplet sample $(\pi_{11}^1, \pi_{01}^1, Z^1)$ is obtained. Repeat whole sample selection processes until the required size of sample is obtained.

After collecting Gibbs samples from implementing the above algorithm the output data are analyzed and diagnosed for the following quantities; the number of iterations required to give valid result, the number of iterations needed for the burn in, and dependence measure. Thus quantities are computed via Gibbsit program developed by Raftery, A.E. and S. M. Lewis(1992 b).

The estimates of π_{11} and π_{01} are obtained as follows:

$$\widetilde{\pi_{11}} = \frac{1}{M_1}\sum_{i=N_1}^{N_1+M_1}\pi_{11}^{(i)}$$

$$\widetilde{\pi_{01}} = \frac{1}{M_2} \sum_{i=N_2}^{n_2+M_2} \pi_{01}^{(i)} \tag{10}$$

where N_1 and N_2 denote the number of iterations need for the burn-in for each component.

M_1 and M_2 denote the number of samples.

And its variance is computed by

$$Var(\widetilde{\pi_{11}}) = \frac{1}{m_1 - 1} \sum_{i=N_1}^{n_1+m_1} (\pi_{11}^{(i)} - \widetilde{\pi_{11}})^2$$

$$Var(\widetilde{\pi_{01}}) = \frac{1}{m_2 - 1} \sum_{i=N_2}^{n_2+m_2} (\pi_{01}^{(i)} - \widetilde{\pi_{01}})^2 \tag{11}$$

3.4 Numerical Example

The applicability of MCMC algorithm given in the previous subsection is investigated by a numerical simulation. For simulating a numerical example, a given data are assumed to be collected from successive occasions sampling : $n=40$, $n_1=22$, $n_0=18$, $m_1=11$, $m_0=9$, $m_{11}=6$, and $m_{01}=4$.

Table 3 The simulated Data Diagnostics via Gibbsit Program

parameters	Kthin	Nburn	Nprec	Nmin	I-RL	Kind
π_{11}	1	3	730	600	1.22	2
π_{01}	1	2	570	600	0.95	1
Z	1	2	570	600	0.95	1

where Kthin is the thinning parameter required to make the chain first order Markov.

Nburn is the number of iterations needed for the burn-in

Nperc is the number of iterations required to achieve the specified precision.

Nmin is the number of iterations required if they were independent

I-RL is the ratio of Nburn+Nprec to Nmin.

Kind is thinning parameter required to make the chain into an independence chain.

The parameters of prior density functions are computed on the basis of moment matching methods for observed data on the first occasion's survey and the obtained quantities are given by $a_1=4.2$, $b_1=1.4$, $b_2=2.8$, and $b=1.4$.

The simulation of Gibbs Sampler is implemented by a FORTRAN Program and the number of iterations is 1000 samples for each component. The generated samples are diagnosed via Gibbsit Software and their results are summarized into Table 3.

The generated Gibbs Samples indicate a low level of dependence between sequences of samples and Kthin=1 means that all simulated data after burning in Nburn can be used to estimate π_{11} and π_{01}. The small values of Nburn show that the starting values of Gibbs Sampler are chosen well.

In this numerical computation, the rates of rejections are relatively small; 0.2187 for π_{11}, 0.1409 for π_{01} and 0.3993 for Z. The small rejection rates mean that the choices of prior parameters and starting values of full conditional parameters are well given.

The Bayes estimates of π_{11} and π_{01} computed on the basis of the equation (10) are $\widetilde{\pi_{11}}$ =0.6033 and $\widetilde{\pi_{01}}$=0.3904. From the equation (11) the variances of $\widetilde{\pi_{11}}$ and $\widetilde{\pi_{01}}$ is obtained and their estimates are $Var(\widetilde{\pi_{11}})$=0.0249 and $Var(\widetilde{\pi_{01}})$=0.0225.

The run of MCMC Sampler is monitored by graphical method. The results of monitoring generations of Gibbs samples show no stickiness and indicate the first order Markov Chain. Gibbs samples' means and estimated variances are very stable when the number of iterations are greater than 600 for each parameter.

From the system of stationary distributions of Markov Chain given in (6), the current occasion characteristics $\pi_{t+1}(A)$ is computed by $\tilde{\pi}_{t+1}(A) = \tilde{\pi}_{t+1}(A)\pi_{11} + \tilde{\pi}_t(A^c)\pi_{01}$. The estimates of $\pi_t(A)$ and $\pi_t(A^c)$ are obtained on the basis of observed data, that is, $\tilde{\pi}_t(A) = ((n_1/n) - 1 + P)/(2P-1)$= 0.5714. The estimates of $\pi_{t+1}(A)$ and $\pi_{t+1}(A^c)$ are computed as $\tilde{\pi}_{t+1}(A)$=0.515 and $\tilde{\pi}_{t+1}(A^c)$=0.485.

4. Conclusion and Remarks

The Markov Chain Monte Carlo proportion estimation in randomized response model has studied by using theoretical and numerical method. Since this model does not have a complete form of beta conjugate family, Bayes estimate of population mean can not be computed by an analytic method.

Acceptance Rejection Sampling is used to compute Bayes estimate of mean and is simulated for several large sample sizes. When sample size is over 250, the estimates mean and standard deviation are shown to be very stable.

The rate of rejection sample is mainly dependent on the values of prior density parameters. In the simulation study of Acceptance Rejection Sampling, uniform prior density with parameters a=1.0 and b=1.0 and a beta prior density with parameters a=3.0 and b=5.5 are discussed. The rate of rejection sample in uniform prior is twice higher than that of beta prior. It is shown that the moment matching method is very useful to determine the beta prior parameters.

Because estimation of prior density parameters is a kind of robust estimation problem in Bayesian inferences, we leave this topic for further research.

The Bayesian estimation of transition matrix is studied on the basis of simple two successive occasion sampling for the randomized response technique. The distribution of the correlated transition probability is derived in terms of the distribution of product of two beta random variables. The Gibbs sampler is implemented for estimating the two transition probabilities and auxiliary variable and the diagnosis of Gibbs sampler illustrates that the generated data have no stickiness and satisfy the property of the first order Markov Chain.

Further research is needed to consider the estimation of prior density parameters of transition probabilities. Further research is also useful to apply the proposed MCMC methods to a practical field Survey data, and to study the sampling scheme that provides the estimates of transition probabilities as well as the proportion on the current occasion.

Appendix

Let $X_1 \sim beta(a_1, b_1)$, $X_2 \sim beta(a_2, b_2)$ and $Z \sim beta(a, b)$ be independent random variables and $a = a_1 + b_1 = a_2 + b_2$. The joint prior density function of $\pi_{11}(= X_1 \cdot Z)$ and $\pi_{01}(= X_2 \cdot Z)$ is given by

$$f(\pi_{11}, \pi_{01}) = k_0 \pi_{11}^{a_1-1} \pi_{01}^{a_2-1} \int_B (Z - \pi_{11})^{b_1-1} (Z - \pi_{01})^{b_2-1} Z^{1-a} (1-Z)^{b-1} dZ,$$

where k_0 is normalizing constant and $B = \{(\pi_{11}, \pi_{01}, z): 0 < \pi_{11}, \pi_{01} < z < 1\}$.

(Proof) From the assumption, the joint pdf of X_1, X_2 and Z is expressed in the set A as.

$$f(X_1, X_2, Z) = k X_1^{a_1-1} (1 - X_1)^{b_2-1} X_2^{a_2-1} (1 - X_2)^{b_2-1} Z^{a-1} (1-Z)^{b-1}$$

where k is normalizing constant and $A = \{(X_1, X_2, Z) : 0 < X_1, X_2, Z < 1\}$.

Consider the transformation $\pi_{11} = X_1 \cdot Z$, $\pi_{01} = X_2 \cdot Z$ and $Z = Z$. This transformation maps the domain A onto the domain B.

The Jacobian of the transformation is Z^{-2}. And the joint p.d.f. of π_{11}, π_{01} and Z is obtained as

$$\begin{aligned} f(\pi_{11}, \pi_{01}, Z) &= k\left(\frac{\pi_{11}}{Z}\right)^{a_1-1} \left(1 - \frac{\pi_{11}}{Z}\right)^{b_1-1} \left(\frac{\pi_{01}}{Z}\right)^{a_2-1} \left(1 - \frac{\pi_{01}}{Z}\right)^{b_2-1} Z^{a-3} (1-Z)^{b-1} \\ &= k \pi_{11}^{a_1-1} (Z - \pi_{11})^{b_1-1} \pi_{11}^{a_1-1} (Z - \pi_{11})^{b_1-1} Z^{1-a} (1-Z)^{b-1} \end{aligned}$$

The marginal density function of π_{11} and π_{01} is obtained by integrating with respect to Z.

$$f(\pi_{11}, \pi_{01}) = k_0\ \pi_{11}^{a_1-1}\pi_{01}^{a_2-1}\int_B (Z-\pi_{11})^{b_1-1}(Z-\pi_{01})^{b_2-1}Z^{1-a}(1-Z)^{b-1}dZ$$

q.e.d.

Reference

[1] Brooks, S. P. (1998), *Markov Chain Monte Carlo and its Application.* The Statistician 47, 69-100

[2] Diebolt, J. and C. Robert (1994), Estimation of Finite Mixture Distributions through Bayesian Sampling. Journal of the Royal Statistical Society, Series B 56, 589-590

[3] Gelfand, A. E. and A. F. M. Smith (1990), Sampling Based Approaches to Calculating Marginal Densities. Journal of the American Statistical Association 85, 398-409

[4] Geyer, C. J. (1992), Practical Markov Chain Monte Carlo. Statistical Science 7, 473-511

[5] Gilks, W. R. and P. Wild (1992), Adaptive Rejection Sampling for Gibbs Sampling. Applied Statistics 41, 337-348

[6] Hasting, W. K. (1970), Monte Carlo Sampling Methods Using Markov Chain and their Application. Biometrika 57, 97-107

[7] Pesku, P. H. (1973), Optimum Monte Carlo Sampling Using Markov Chain. Biometrika 60, 607-612

[8] Raftery, A. E. and S. M Lewis (1992 a), How Many Iterations in the Gibbs Sampler? In J. M. Bernardo, A.F.M. Smith, A. P. Dawid and J.O. Berger(eds.), Bayesian Statistics 4, Oxford University Press

[9] Raftery, A. E. and S. M. Lewis (1992 b), One Long Run Diagnostics : Implementation Strategies for Markov Chain Monte Carlo. Statist. Sci. 7, 493-497

[10] Ripley, B. D. (1987), *Stochastic Simulation* . New York : Wiley

[11] Tierney, L. (1994), Markov Chains for Exploring Posterior Distributions. Annals of Statistics 22, 1701-1762