

# A Comparison of Bayesian and Classical Approach for Estimating Markov Based Logistic Model

Janardan Mahanta<sup>1,\*</sup>, Soma Chowdhury Biswas<sup>1</sup>, Manindra Kumar Roy<sup>2</sup>, M. Ataharul Islam<sup>3</sup>

<sup>1</sup>Department of Statistics, University of Chittagong

<sup>2</sup>Department of Statistics, Mawlana Bhasani Science and Technology University, Tangail

<sup>3</sup>Department of Applied Statistics, East West University, Dhaka, Bangladesh

**Abstract** Inferential Statistics is the main subject of statistics. Classical and non-classical are the two types of approaches of estimation. In Bayesian estimation selection of appropriate loss function and prior density are very important. Jeffery's non-informative prior and squared error loss function were used. Lindley approximation was used to solve the Bayesian integral. The proposed procedure is applied to a longitudinal data on pregnancy complication in rural Bangladesh collected from the Bangladesh Institute of Research for Promotion of Essential and Reproductive Health and Technologies (BIRPERHT). In this study, we have conducted Markov model by maximum likelihood and Bayesian approach and compare them above the approaches. Bayesian approach of estimation found to be better than the classical approach in this particular case.

**Keywords** BSE, Credible interval, MLE, T.K.

## 1. Introduction

Now-a-days Bayesian approach is widely used for decision-making. In this paper, we have applied Bayesian approach in Markov chain based logistic model. Markov chain models can be used in analyzing longitudinal data. There is several discrete time Markov chain models proposed for analyzing repeated categorical data over decades. To analyze the binary sequence of presence or absence of diseases, Muenz and Rubinstein [10] introduced a discrete time Markov chain for expressing the transition probabilities in terms of covariates. The technique suggests by them is applicable for first order Markov model. For analyzing sequences of ordinal data from relapsing and remitting of a disease, Albert [1] developed a finite Markov chain model. In addition, Albert and Waclawiw [2] developed a class of quasilielihood models for a two state Markov chain with stationary transition probabilities for heterogeneous transitional data. Raftery [14], Raftery and Tavaré [13] proposed a higher order Markov chain model with dependence on contribution of the past transitions. Islam and Chowdhury [8] applied a three state Markov model for analyzing covariate dependence, also Islam and Chowdhury [7] presented a higher order version of the covariate dependent Markov model. For analyzing repeated observations, there is a renewed interest in the development

of multivariate models based on Markov chains. These models can be employed for analyzing data generated from meteorology, epidemiology and survival analysis, reliability, econometric analysis, biological concerns, etc. Muenz and Rubinstein [10] employed logistic regression models to analyze the transition probabilities from one state to another. The estimation for first-order Markov models is quite straight forward, but still there is serious lack of generalization in estimation and testing for models applicability for higher order Markov chains. Islam and Chowdhury [7] provided a further generalization for covariate dependent higher order models. Following Azzalini [3], Heagerty and Zeger [5] presented a class of marginalized transition models (MTM) and Heagerty [4] proposed a class of generalized MTMs to allow serial dependence of first or higher order. These models are computationally tedious and the form of serial dependence is quite restricted. Heagerty [4] provided derivatives for score and information computations. Muenz and Rubinstein [10] introduced a discrete time Markov chain for expressing the transition probabilities in terms of function of covariates for a binary sequence of presence or absence of a disease. Islam, Chowdhury and Singh [6] suggest covariate dependence in a higher order Markov models is examined. The proposed model and inference procedures are simple and the covariate dependence of the transition probabilities of any order can be examined without making the underlying model complex. Another advantage of the model lies in the fact that the estimation and test procedures for both the specific parameter of interest and the overall model remain easy for practical applications for any longitudinal

\* Corresponding author:

johnstatcu@gmail.com (Janardan Mahanta)

Published online at <http://journal.sapub.org/ajms>

Copyright © 2015 Scientific & Academic Publishing. All Rights Reserved

data. A simple alternative is also proposed. Applications are illustrated using maternal morbidity during pregnancy.

In this paper, an attempt is made to estimate the probability of being pregnancy complications of women according to the characteristics of the women and estimate the parameters by Bayesian approach and method of maximum likelihood approach. This characteristics or variables may often be related to each other. Higher difficulties occur to any miscarriage, economic status and age at marriage in this study.

## 2. Model

Muenz and Rubinstein [10] in 1985 proposed a two state Markov chain to model a discrete time-binary sequence. The transition matrix of the chain is:

$$M = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

where  $p_{00}$  denotes the transition probabilities from 0 to 0 and  $p_{10}$  is the transition probability from 1 to 0. At each time point, a vector of length two contains the probability of outcome of interest and its complement.

The first such vector (row) is  $P(1) = (p(1), 1 - p(1))$ , with the first element equal to the probability of the complement of the event of interest.

The matrix  $M$  and initial vector  $P(1)$  suffice to characterize the Markov chain. After time point 1, the vector of occupation probabilities is:

$$P(j) = (p(j), 1 - p(j)) = P(1)M^{j-1}; j > 1$$

Muenz and Rubinstein [10] proposed model the transition probabilities  $p_{00}$  and  $p_{10}$  by logistic regressions.

$$p_{00} = \frac{\exp(\beta'X)}{1 + \exp(\beta'X)} \text{ and}$$

$$p_{10} = \frac{\exp(\alpha'X)}{1 + \exp(\alpha'X)}$$

The vector  $X$  contains covariates and for the  $q$ th person in the study is equal to,  $X_q = (1, X_{q1}, \dots, X_{qp})$ . There are two logistic regressions, one having parameter vector  $\beta = (\beta_0, \dots, \beta_p)'$  and the other having parameter vector  $\alpha = (\alpha_0, \dots, \alpha_p)'$ . Large positive (negative) values of  $\beta'X$  and  $\alpha'X$  yield large (small) transition probabilities. Since the transition matrix is a function of covariates,  $P(j)$  relates the occupation probabilities to the initial state for each pattern of covariates.

### For transition 0 to 0

For binary random variable  $Y_i$  with covariate  $X_i$  the joint density function is as

$$f(x/\beta) = \prod_{i=0}^n \{p_{00}\}^{n_{00i}} \{1 - p_{00}\}^{n_{01i}}$$

where,  $n_{00i}$  and  $n_{01i}$  are the number of transitions.

## 3. Bayesian Approach

In this study, we use the Bayesian paradigm to make inferences about parameters of the model [10] for pregnancy complication data.

### 3.1. Prior and Posterior Distribution

In Bayesian analysis prior and posterior distribution are important. Since the parameters of Markov model of  $\beta$  lies between  $-\infty$  to  $+\infty$ , then according to Jeffrey's non-informative prior [15] of  $\beta$  are  $g(\beta) = I$ . Where,  $I$  represent unit vector.

Then the posterior density of  $\beta$  for the given sample is

$$f(\beta/X) = \frac{\prod_{i=0}^n \{p_{00}\}^{n_{00i}} \{1 - p_{00}\}^{n_{01i}}}{\int \prod_{i=0}^n \{p_{00}\}^{n_{00i}} \{1 - p_{00}\}^{n_{01i}} d\beta}$$

$$f(\beta/X) = \frac{\prod_{i=0}^n \left[ \left\{ \frac{\exp(\beta'X)}{1 + \exp(\beta'X)} \right\}^{n_{00i}} \left\{ \frac{1}{1 + \exp(\beta'X)} \right\}^{n_{01i}} \right]}{\int_{-\infty}^{\infty} \prod_{i=0}^n \left[ \left\{ \frac{\exp(\beta'X)}{1 + \exp(\beta'X)} \right\}^{n_{00i}} \left\{ \frac{1}{1 + \exp(\beta'X)} \right\}^{n_{01i}} \right] d\beta}$$

### 3.2. Bayes Estimators under Squared Error Loss Function

The squared error loss function is defined as

$$L\left(\hat{\beta}; \beta\right)=\left(\hat{\beta}-\beta\right)^2$$

For squared error loss function, Bayes estimators are the mean of the posterior density [11]

$$\hat{\beta}_{BSE} = \frac{\int_{-\infty}^{\infty} \beta \prod_{i=0}^n \left[ \left\{ \frac{\exp(\beta' X)}{1 + \exp(\beta' X)} \right\}^{n_{00i}} \left\{ \frac{1}{1 + \exp(\beta' X)} \right\}^{n_{01i}} \right] d\beta}{\int_{-\infty}^{\infty} \prod_{i=0}^n \left[ \left\{ \frac{\exp(\beta' X)}{1 + \exp(\beta' X)} \right\}^{n_{00i}} \left\{ \frac{1}{1 + \exp(\beta' X)} \right\}^{n_{01i}} \right] d\beta}$$

The two integrals appear in the ratio cannot be solved to have a closed form. For evaluating them, we use the Lindley [12] approximation.

We have,

$$I(X) = E(u(\beta) / X) = \frac{\int u(\beta) e^{L_0(\beta) + p(\beta)} d\beta}{\int e^{L_0(\beta) + p(\beta)} d\beta}$$

where,  $L_0(\beta)$  is the log likelihood and  $p(\beta)$  is the logarithm prior distribution.

Since,  $g(\beta) = I$  then  $p(\beta) = \log g(\beta)$

or,  $p(\beta) = \log I = 0$

Then according to Lindley,  $I(X)$  can be approximately evaluated as

$$I(X) = u(\hat{\beta}) + \frac{1}{2} \left[ u''(\hat{\beta}) + 2u'(\hat{\beta})p'(\hat{\beta}) + L_0'''(\hat{\beta})u(\hat{\beta})\sigma^2 \right] \hat{\sigma}^2$$

where,  $u(\beta)$  is the functional form of the parameter  $\beta$ , which is expected to posterior density and  $\hat{\beta}$  is the maximum likelihood [10] estimator of  $\beta$ .

$$\begin{aligned} u'(\beta) &= \frac{\partial u(\beta)}{\partial \beta}, \quad u''(\beta) = \frac{\partial^2 u(\beta)}{\partial^2 \beta} \\ p'(\beta) &= \frac{\partial p(\beta)}{\partial \beta}, \quad L_0'''(\beta) = \frac{\partial^3 L_0(\beta)}{\partial^3 \beta} \\ \text{and } \hat{\sigma}^2 &= -\frac{1}{L_0''(\hat{\beta})} \end{aligned}$$

Now we know that the likelihood function of binary dependence Markov model is

$$L = \prod_{i=1}^n \left[ \{p_{00}\}^{n_{00i}} \{1 - p_{00}\}^{n_{01i}} \right] \left[ \{p_{10}\}^{n_{01i}} \{1 - p_{10}\}^{n_{11i}} \right]$$

Therefore,  $L = \prod_{i=1}^n l_0 l_1$

Taking log on both sides

$$\begin{aligned}\log L &= \sum_{i=1}^n (\log l_0 + \log l_1) \\ &= L_0 + L_1\end{aligned}$$

where,  $L_0$  and  $L_1$  are the likelihood function of transition 0 to 0 and 1 to 0 respectively.

Now considering

$$\begin{aligned}L_0 &= \sum_{i=1}^n [n_{00i} \log p_{00} + n_{01i} \log(1 - p_{00})] \\ &= \sum_{i=1}^n \left[ n_{00i} \log \left\{ \frac{\exp(\beta' X_i)}{1 + \exp(\beta' X_i)} \right\} + n_{01i} \log \left\{ \frac{1}{1 + \exp(\beta' X_i)} \right\} \right] \\ &= \sum_{i=1}^n \left[ n_{00i} \beta X_i - (n_{00i} + n_{01i}) \log \{1 + \exp(\beta X_i)\} \right]\end{aligned}$$

Differentiating both sides with respect to  $\beta$ , we have

$$\frac{\partial L_0}{\partial \beta} = \sum_{i=1}^n \left[ n_{00i} X_i - (n_{00i} + n_{01i}) \frac{\exp(\beta' X_i) X_i}{1 + \exp(\beta' X_i)} \right]$$

$$\frac{\partial^2 L_0}{\partial^2 \beta} = - \sum_{i=1}^n (n_{00i} + n_{01i}) X_i^2 p_{00} p_{01}$$

$$\frac{\partial^3 L_0}{\partial^3 \beta} = 2 \sum_{i=1}^n (n_{00i} + n_{01i}) X_i^3 p_{00} p_{01} (p_{00} - p_{01})$$

$$\hat{\sigma}^2 = - \frac{1}{L_0''} = \frac{1}{\sum_{i=1}^n (n_{00i} + n_{01i}) X_i^2 \hat{p}_{00} \hat{p}_{01}}$$

Therefore, Bayes estimator under squared error loss function is

$$\hat{\beta}_{BSE} = \hat{\beta} + \hat{\beta} \left[ \frac{\sum_{i=1}^n (n_{00i} + n_{01i}) X_i^3 \hat{p}_{00} \hat{p}_{01} (\hat{p}_{00} - \hat{p}_{01})}{\left\{ \sum_{i=1}^n (n_{00i} + n_{01i}) X_i^2 \hat{p}_{00} \hat{p}_{01} \right\}^2} \right]$$

where  $\hat{\beta}$  is the maximum likelihood [10] estimate of  $\beta$ .

If  $\sum_{i=1}^n (n_{00i} + n_{01i})$  is large then Bayes estimator tends to maximum likelihood estimator.

In addition, the same way we estimate the parameters of the transition 1 to 0.

## 4. Bayesian Credible Interval

If  $f(\theta/X)$  is the posterior distribution given the sample, we may be interested in finding an interval such that

$$P(\theta \in (\theta_1, \theta_2) / x) = \int_{\theta_1}^{\theta_2} f(\theta / X) d\theta = 1 - \alpha$$

This interval is called  $(1 - \alpha)$  100% Bayesian credible [9] interval of  $\theta$ .

In Bayesian analysis, credible interval becomes the counterpart of the classical confidence interval, also credible interval may be unique for all models. The Bayesian credible interval, on the other hand, has a direct probability interpretation  $P(\theta \in (\theta_1, \theta_2) / x) \geq 1 - \alpha$  and is completely determined from the current observed data  $x$  and the prior distribution.

## 5. Numerical Analysis

The covariate dependent Markov model proposed in this paper is applied to the pregnancy complication data conducted from BIRPERHT data for the period November 1992 to December 1993. The data were collected using both cross-sectional and prospective study designs. A multistage sampling design was used for collecting the data for this study. Districts were selected randomly in the first stage, one district from each Division. Then Thanass were selected randomly in the second stage, one Thana from each of the selected Districts. At the third stage, two Unions were selected randomly from each selected Thana. The subjects comprised of pregnant women with less than 6 months duration in the selected Unions. All the selected pregnant women from the selected Unions were followed on regular basis (roughly at an interval of 1 month) throughout the pregnancy. Again, the subjects were followed at the time of delivery for a full-term pregnancy and 90 days after delivery or 90 days after any other pregnancy outcome. A total of 1059 pregnant women were interviewed in the follow-up component of the study. The following pregnancy complications are considered under the complications in this study: hemorrhage, fits, convulsion, edema, excessive vomiting, and cough or fever for more

than three days. If hemorrhage, fits, convulsion, edema, excessive vomiting, and cough or fever for more than three days occurred to the respondents, we considered complications and was coded as 1, if no complications then coded as 0. The explanatory variables are: age at marriage (15 years or lower, more than 15 years), economic status (lower or upper), any miscarriage (yes, no). The number of transitions for the two-state Markov chain of first order is displayed in Table 1.

**Table 1.** Transition counts of Markov chain for first orders of pregnancy complication of different flow-up

Transition Counts			
States	0	1	Total
0	1338	250	1588
1	250	354	604

Estimate the Muenz-Rubinstein model using pregnancy complication data. In this study, we have used three covariates because of complexity to fit the model. Three highly significant covariates are used in our study. Bayesian approaches have been applied for estimating the parameter of the above model.

### 5.1. Comparison between Credible Interval and Confidence Interval

Confidence intervals for maximum likelihood estimators and credible intervals of Bayesian estimators under squared error loss function were calculated for Muenz-Rubinstein model and two types of transition are presented in the following table.

#### For 0 to 0 transition

**Table 2.** Confidence interval for maximum likelihood estimators

Covariates	Point Estimate	Interval Estimate (95%)		
		Lower	Upper	Length
Constant	1.775	1.453	2.097	0.644
Any miscarriage	0.160	0.047	0.274	0.226
Economic Status	-0.306	-0.459	-0.154	0.305
Age at Marriage	0.024	-0.098	0.146	0.244

**Table 3.** Credible interval for Bayesian estimator under squared error loss function

Covariates	Point Estimate	Credible Interval (95%)		
		Lower	Upper	Length
Constant	1.776	1.697	1.855	0.158
Any miscarriage	0.160	0.130	0.191	0.061
Economic Status	-0.307	-0.339	-0.274	0.065
Age at Marriage	0.024	-0.005	0.054	0.060

From the above results, we have found that the length of Bayesian credible is lower than the length of confidence interval for all covariates.

Moreover, 1 to 0 transitions

**Table 4.** Confidence interval for maximum likelihood estimators

Covariates	Point Estimate	Interval Estimate (95%)		
		Lower	Upper	Length
Constant	-0.702	-1.296	-0.108	1.187
Any miscarriage	0.285	0.060	0.510	0.450
Economic Status	0.235	-0.034	0.505	0.539
Age at Marriage	-0.163	-0.397	0.070	0.466

**Table 5.** Credible interval for Bayesian estimator under squared error loss function

Covariates	Point Estimate	Credible Interval (95%)		
		Lower	Upper	Length
Constant	-0.703	-0.777	-0.628	0.149
Any miscarriage	0.286	0.226	0.346	0.120
Economic Status	0.236	0.177	0.294	0.117
Age at Marriage	-0.164	-0.221	-0.106	0.115

Therefore, from the above tables it is also seen that, Bayesian credible intervals are smaller length than confidence interval in 1 to 0 transitions. Therefore, we can say that Bayesian approach provides better result than usual method maximum approach in Muenz-Rubinstein model. In addition, economic status is negatively associated, any miscarriage and age at marriage are positively associated with pregnancy complication.

All the numerical analysis was performed by R-Language (Version-2.10.0).

## 6. Conclusions

Markov chain is very important part in real world situation. Its application increasing day by day, in this study we applied pregnancy complication data and only three covariates were used because complexity of the model. Meanwhile different types of approach were used for estimating the parameters in this model. In Bayesian approach under squared error of loss function were used and compare with method of maximum likelihood and we have found that length of Bayesian credible interval is smaller than length of confidence interval for all covariates in two types of transitions.

From the above analysis, we conclude that Bayesian estimators under squared error loss function give better result in Muenz Rubinstein model. An application is included in this paper to illustrate the use of the proposed model for real life problems.

## REFERENCES

- [1] Albert, P. S. (1994). A Markov model for sequence of ordinal data from a relapsing remitting disease. *Biometrics*, 50, 51-60.

- [2] Albert, P. S., & Waclawiw, M. A. (1998). A two state Markov chain for heterogeneous transitional data: A quasilielihood approach. *Statistics in Medicine*, 17, 1481-1493.
- [3] Azzalini, A. (1994). Logistic regression for autocorrelated data with application to repeated measures. *Biometrika*, 81, 767-775.
- [4] Heagerty, P. J. (2002). Marginalized transition models and likelihood inference for longitudinal categorical data. *Biometrics*, 58, 342-351.
- [5] Heagerty, P. J., & Zeger, S. L. (2000). Marginalized multi-level models and likelihood inference (with Discussion). *Statistical Science*, 15, 1-26.
- [6] Islam, M. A., Chowdhury, R.I., & Singh, K. P. (2008). Covariate Dependent Markov Models for Analysis of Repeated Binary Outcomes. *Journal of Modern Applied Statistical Methods*, 6(2), 561-572.
- [7] Islam, M. A., & Chowdhury, R. I. (2006). A higher-order Markov model for analyzing covariate dependence. *Applied Mathematical Modelling*, 30, 477-488.
- [8] Islam, M. A., & Chowdhury, R. I. (2004). A three state Markov model for analyzing covariate dependence. *International Journal of Statistical Sciences*, 3, 241-249.
- [9] Kazmi, S. M., Aslam, M., & Ali, S. (2012). Preference of prior for the class of life-time distributions under different loss functions. *Pak. J. Statist*, 28(4), 467-484.
- [10] Muenz, L. R., & Rubinstein, L. V. (1985). Markov models for covariate dependence of binary sequences. *Biometrics*, 41, 91-101.
- [11] Podder, C. K., & Roy, M. K. (2003). Bayesian estimation of the parameter of Maxwell distribution under MLINEX loss function. *Journal of Statistical Studies*, 23, 11-16.
- [12] Press, S. J. (1989). *Bayesian Statistics: Principles, Models, and Applications*. New York: John Wiley & Sons.
- [13] Raftery, A., & Tavare, S. (1994). Estimating and modeling repeated patterns in higher order Markov chains with the mixture transition distribution model. *Applied Statistics*, 43(1), 179-199.
- [14] Raftery, A. E. (1985). A model for higher order Markov chains. *Journal of Royal Statistical Society B*, 47, 1528-39.
- [15] Smith, G. P. (n.d.). *Expressing Prior Ignorance of a Probability Parameter*. Notes, University of Missouri <http://www.stats.org.uk/priors/noninformative/Smith.pdf> on informative priors.