Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online): 2229-5666
Vol. 9, Issue 2 (2016): 81-90

COMPARISON OF BAYESIAN APPROACH WITH CLASSICAL APPROACH FOR ESTIMATING THE PARAMETER OF MARKOV MODEL

***Janardan Mahnata and Soma Chowdhury Biswas**
Department of Statistics, University of Chittagong, Chittagong, Bangladesh.
E Mail: *johnstat@cu.ac.bd

Received December 25, 2015
Modified December 02, 2016
Accepted December 16, 2016

Abstract

This paper has introduced Bayesian analysis and its application to estimate the parameter of the Markov model. To use Markov model, comparison between Bayesian approach and method of maximum likelihood have been done. Bayesian approach gives better result than classical approach. Jeffery's non-informative prior and squared error loss function have been used in Bayesian inference. Tierney-Kadnae (T.K.) algorithm has been used to solve the Bayesian integral.

Key Words: Credible Interval, Markov Model, Squared Error (SE), Tierney-Kadnae (T.K.).

1. Introduction

In follow-up studies Markov chain based logistic models are used. A logistic regression model to analyze the transitional probabilities from one state to another was applied by Muenz-Rubinstein (1985). Azzalini (1994) examined the influence of time dependent covariate on the marginal distribution of the binary outcome variables in serially correlated data. Raftery and Tavare (1994) suggested a Markov chain model of higher order than one that involves only one parameter for each extra lag variable. Islam and Chowdhury (2006) reviewed the first order model of Muenz-Rubinstein (1985) and developed a general procedure based on the Chapman-Kolmogorov equation for transition. Cook and Ng (1997) applied a logistic bivariate normal mixture model for a two state Markov chain. Recently, Islam at el. (2012) have used classical approach for the estimation of parameters for analyzing polytomous outcome data using logistic link function. Noorian and Ganjali (2012) applied Bayesian analysis of transitional model for longitudinal ordinal response data. However, in their study they applied Markov Chain Monte Carlo (MCMC) algorithm. Although MCMC is easiest and widely used but there is no clear idea about estimating procedure of the model to use Markov Chain Monte Carlo integration for solving Bayesian integral. This is a programming based operation. Mahanta et al. (2015) applied Bayesian approach for estimating the parameters of Muenz-Rubinstein model. This paper has analyzed Azzalini's two state Markov model and the theoretical idea about the estimator of the parameter of Azzalini's model to estimate the parameters by Bayesian approach as well as method of maximum likelihood.

For estimating the parameter of model using pregnancy complication data, the data were collected from Bangladesh Institute of Research for Promotion of

Essential and Reproductive Health and Technologies (BIRPERHT) survey during the period from November 1992 to December 1993 on maternal morbidity in the rural areas of Bangladesh.

2. Model

Let $(y_1, y_2, ..., y_t)$ be the binary response data observed on n objects at time $t = 1,2,....,T$ whose values are 0 or 1. $(x_1, x_2, ..., x_t)$ are associated covariates recorded for each subject at each occasion. Our aim is to obtain estimates for the regression of y_t on x_t using binary Markov chain of first order. In practice, we often concern with the non-stationary case, in which $\theta_t = E(Y_t) = Pr(Y_{it} / Y_{it-1})$ varies with t via some link function such as,

$$logit(\theta_{it}) = log\frac{\theta_t}{1-\theta_t} = X_t'\beta \tag{2.1}$$

where, $\theta_t = Pr(Y_{it} = 1 / Y_{it-1}) = \dfrac{e^{X_t\beta}}{1+e^{X_t\beta}}$

A first order two state Markov model represented by

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

assumes that the current response variable is dependent on the history only through the immediate preceding response,

i.e., $Pr(Y_{it} / Y_{ij}; j > t) = Pr(Y_{it} / Y_{it-1})$

the transition probabilities

$$p_0 = p_{it,0} = Pr(Y_{it} = 1 / Y_{it-1} = 0)$$

and $1 - p_0 = Pr(Y_{it} = 0 / Y_{it-1} = 0)$

$$p_1 = p_{it,1} = Pr(Y_{it} = 1 / Y_{it-1} = 1)$$

and $1 - p_1 = Pr(Y_{it} = 0 / Y_{it-1} = 1)$

define the Markov process but do not directly parameterize the marginal mean. Azzalini (1994) parameterize the transition probabilities through two assumptions. First, a marginal mean regression model is adopted that constrains the transition probabilities to satisfy

$$\theta_{it} = p_{it,1}\,\theta_{it-1} + p_{it,0}\,(1-\theta_{it-1}). \tag{2.2}$$

Second, the transition probabilities are structured through assumptions on the pair wise odds ratio.

$$\psi_{it} = \frac{p_{it,1} / (1-p_{it,1})}{p_{it,0} / (1-p_{it,0})}. \tag{2.3}$$

This quantifies the strength of serial correlation. The simplest dependence model assumes a time homogeneous association, $\psi_{it} = \psi_0$. However, models that allow ψ_{it} to depend on covariates or to depend on time are possible.

Solving (2.2) and (2.3) for p_0 and p_1 for any $t > 1$ and by mathematical induction for any $t > 1$, we can finally represent the transition probabilities p_j as,

$$p_j = \begin{cases} \theta_t, for\ \psi = 1 \\ \dfrac{\delta - 1 + \{\psi - 1)(\theta_t - \theta_{t-1})}{2(\psi - 1)(1 - \theta_{t-1}\ )} + j\dfrac{1 - \delta + (\psi - 1)(\theta_t + \theta_{t-1} - 2\theta_t\theta_{t-1})}{2(\psi - 1)\theta_{t-1}(1 - \theta_{t-1})}; \psi \neq 1 \\ \\ \end{cases} \quad (2.4)$$

$$for\, t = 2,3,4$$

where, $\delta^2 = 1 + (\psi - 1)\{(\theta_t - \theta_{t-1})^2 \psi - (\theta_t - \theta_{t-1})^2 + 2(\theta_t + \theta_{t-1})\}$.

It can be shown that the $p_j's$ always lie in $(0,1)$ and $log\,\psi = \lambda\ or, \psi = e^{\lambda}$.

The above relationships generate a process having the desired properties. On taking $Pr(Y_t = 1) = \theta_1$ and then generating $y_1, y_2,..., y_t$ via a non-homogeneous Markov chain with transition probabilities p_j we obtain a sequence such that $E(Y_t) = \theta_t$ for $t = 1,2,....,T$ and the odds ratios for (y_{t-1}, y_t) are equal to ψ .

Define $Y_{it} = \begin{cases} 1, if\ the\,ith\,individual\,succeed\,at\,time\,t \\ 0, if\ the\,ith\,individual\ failed\,at\,time\,t \end{cases}$

and $X_{it} = (X_{i1}, X_{i2},...., X_{it})$ be the covariate matrix θ_{it} be denoted by the excepted value of Y_{it} and $log\,it(\theta_{it}) = X_t'\beta$.

For the binary random variable Y_t with covariate X_t the marginal distribution is given by

$$\begin{aligned} f(y_t / x_t) &= p_{y_{t-1}}^{y_t}\left(1 - p_{y_{t-1}}\right)^{1 - y_{t-1}} \\ &= \left(\frac{p_{y_{t-1}}}{1 - p_{y_{t-1}}}\right)^{y_t}\left(1 - p_{y_{t-1}}\right) \end{aligned} \quad (2.5)$$

3. Prior and Posterior Distribution

Selection of a prior distribution is an important part in Bayesian approach. When proper information is not available, use of non-informative has an extensive tradition in statistics. Mahanta et al. (2015) used non-informative prior along with the uniform prior and defined as $g(\beta, \lambda) = I$.

Where, $g(\beta, \lambda)$ is the joint prior density of parameter β and λ. That is $g(\beta, \lambda) = g(\beta)g(\lambda)$, since β & λ are independent and I represent identity vector.

Then the posterior density of β and λ for the given observation is

$$f(\beta, \lambda / X) = \frac{\prod_{i=1}^{n} f(y_t / x_t) g(\beta, \lambda)}{\int \prod_{i=1}^{n} f(y_t / x_t) g(\beta, \lambda) d(\beta, \lambda)}. \quad (3.1)$$

4. Bayes Estimators

Loss function is the important ingredient for Bayesian approach. Squared error loss function of parameter β is defined as

$$L\left(\hat{\beta};\beta\right)=\left(\hat{\beta}-\beta\right)^2 . \tag{4.1}$$

Bayes estimators (Podder & Roy, 2003) are the mean of the posterior density under squared error loss function

$$\hat{\beta}_{BSE}=\frac{\int\beta\prod_{i=1}^{n}f(y_t/x_t)g(\beta)}{\int\prod_{i=1}^{n}f(y_t/x_t)g(\beta)d(\beta)} . \tag{4.2}$$

The Bayesian integral of Markov model cannot be solved to have a closed form. Tierney-Kadane (1986) approximation is the approximation that solves this type of integral.

If the form of the Bayes integral is

$$I(X)=E\left(u(\beta)/X\right)=\frac{\int u(\beta)e^{l_t(\beta)+p(\beta)}d\beta}{\int e^{l_t(\beta)+p(\beta)}d\beta} \tag{4.3}$$

where, $I(X)$ represent the form of the integral, l_t is the log-likelihood, $p(\beta)$ is the log of prior and $u(\beta)$ is the functional form of the parameter β that is expected with respect to posterior density.

Then according to Tierney-Kadane (1986), the integral can be approximately be evaluated as

$$I(X)=\frac{\hat{\sigma}^*}{\hat{\sigma}}\exp\left[n\left\{\xi^*\left(\hat{\beta}^*\right)-\xi\left(\hat{\beta}\right)\right\}\right]$$

where, $\xi(\beta)=\frac{1}{n}\{l_t(\beta)+p(\beta)\}$ (4.4)

and $\xi^*(\beta^*)=\frac{\log\ u(\beta)}{n}+\xi(\beta)$ (4.5)

$\hat{\beta}^*$ maximizes $\xi^*(\beta^*)$ and β is the posterior mode and therefore maximizes $\xi(\beta)$ and

$$\hat{\sigma}^{-2}=\left(-\frac{\delta^2\xi(\beta)}{\delta\beta^2}\right)_{\beta=\hat{\beta}}\quad and\quad \hat{\sigma}^{*-2}=\left(-\frac{\delta^2\xi^*(\beta^*)}{\delta\beta^{*2}}\right)_{\beta^*=\hat{\beta}^*}$$

where, and $\hat{\beta}$ and $\hat{\beta}^*$ are the maximum likelihood estimators of parameter β and β^* respectively.

For estimating the parameter β by Bayesian approach using chain rule of differentiation, and considering the t^{th} term of the log-likelihood function, Azzalini (1994) applied its derivatives computed via Chain rule of differentiation

$$\frac{\partial \xi_t}{\partial \beta} = \frac{\partial \xi_t}{\partial p_{y_{t-1}}}\left(\frac{\partial p_{y_{t-1}}}{\partial \theta_t}.\frac{\partial \theta_t}{\partial \beta} + \frac{\partial p_{y_{t-1}}}{\partial \theta_{t-1}}.\frac{\partial \theta_{t-1}}{\partial \beta}\right); \quad t = 1, 2,3,4 \tag{4.6}$$

Following the first order differentiation of Azzalini (1994), the second order differentiation is

$$\frac{\partial^2 \xi_t}{\partial \beta^2} = \frac{\partial^2 \xi_t}{\partial p^2{}_{y_{t-1}}}\left(\frac{\partial^2 p_{y_{t-1}}}{\partial \theta^2{}_t}.\frac{\partial^2 \theta_t}{\partial \beta^2} + \frac{\partial^2 p_{y_{t-1}}}{\partial \theta^2{}_{t-1}}.\frac{\partial^2 \theta_{t-1}}{\partial \beta^2}\right). \tag{4.7}$$

Taking log on both sides in equation (2.5), the log-likelihood function is

$$\xi_t = l_t = \sum_{t=1}^{T}\left\{y_t \, logit\left(p_{y_{t-1}}\right) + log\left(1 - p_{y_{t-1}}\right)\right\}. \tag{4.8}$$

Now differentiating successively both sides with respect to $p_{y_{t-1}}$

$$\frac{\partial \xi_t}{\partial p_{y_{t-1}}} = \sum\left\{\frac{y_t - p_{y_{t-1}}}{p_{y_{t-1}}\left(1 - p_{y_{t-1}}\right)}\right\}$$

and, $$\frac{\partial^2 \xi_t}{\partial p^2{}_{y_{t-1}}} = \sum\left\{\frac{y_t\left(2p_{y_{t-1}} - 1\right) - p^2{}_{y_{t-1}}}{p^2{}_{y_{t-1}}\left(1 - p_{y_{t-1}}\right)^2}\right\}$$

again, successive differentiation of equation (2.4) with respect to θ_t gives

$$\frac{\partial p_{y_{t-1}}}{\partial \theta_t} = \frac{1}{A}(-(2y_{t-1} - 1)\frac{\partial \delta}{\partial \theta_t} + \psi - 1)$$

and, $$\frac{\partial^2 p_{y_{t-1}}}{\partial \theta^2{}_t} = \frac{1}{A}\left[-\left(2y_{t-1} - 1\right)\frac{\partial^2 \delta}{\delta \theta_t^2}\right]$$

where, $A = 2(\psi - 1)\{1 - y_{t-1} + (2y_{t-1} - 1)\theta_{t-1}\}$

and, $\delta^2 = 1 + (\psi - 1)\left\{(\theta_t - \theta_{t-1})^2\psi - (\theta_t - \theta_{t-1})^2 + 2(\theta_t + \theta_{t-1})\right\}$.

Again, from equation (2.1) we have,

$$\theta_t = \frac{exp(X_t'\beta)}{1 + exp(X_t'\beta)}$$

now differentiating successively both sides with respect to β

$$\frac{\partial \theta_t}{\partial \beta} = \theta_t(1 - \theta_t)X_t$$

$$= \frac{X_t\theta_t}{1 + exp(X_t\beta)}$$

and, $$\frac{\partial^2 \theta_t}{\partial \beta^2} = \frac{\{1 + exp(X_t\beta)\}X_t\frac{\partial \theta_t}{\partial \beta} - X_t^2\theta_t\, exp(X_t\beta)}{\{1 + exp(X_t\beta)\}^2}$$

also, we know that,

$$\delta^2 = 1 + (\psi - 1)\left\{(\theta_t - \theta_{t-1})^2\psi - (\theta_t - \theta_{t-1})^2 + 2(\theta_t + \theta_{t-1})\right\}.$$

Now applying partially and successively differentiating δ^2 with respect to θ_t and θ_{t-1}

$$\frac{\partial\delta}{\partial\theta_t} = \frac{1}{\delta}\left[(\psi-1)\{\psi(\theta_t-\theta_{t-1})-(\theta_t-\theta_{t-1})+1\}\right]$$

and, $$\frac{\partial^2\delta}{\partial\theta_t^{\,2}} = \frac{(\psi-1)}{\delta^2}\left[\delta(\psi-1)-\{\psi(\theta_t-\theta_{t-1})-(\theta_t-\theta_{t-1})+1\}\frac{\partial\delta}{\partial\theta_t}\right]$$

also, $$\frac{\partial\delta}{\partial\theta_{t-1}} = \frac{1}{\delta}\left[(\psi-1)\{\psi(\theta_{t-1}-\theta_t)-(\theta_{t-1}-\theta_t)+1\}\right]$$

and, $$\frac{\partial^2\delta}{\partial\theta_{t-1}^{\,2}} = \frac{(\psi-1)}{\delta^2}\left[\delta(\psi-1)-\{\psi(\theta_{t-1}-\theta_t)-(\theta_{t-1}-\theta_t)+1\}\frac{\partial\delta}{\partial\theta_{t-1}}\right].$$

Moreover differentiating successively with respect to ψ

$$\frac{\partial\delta}{\partial\psi} = \frac{1}{\delta}\left\{\psi(\theta_t-\theta_{t-1})^2+(\theta_t+\theta_{t-1})\right\}$$

and, $$\frac{\partial^2\delta}{\partial\psi^2} = \frac{1}{\delta^2}\left\{(\theta_t-\theta_{t-1})^2\left(\delta-\psi\frac{\partial\delta}{\partial\psi}\right)-(\theta_t+\theta_{t-1})\frac{\partial\delta}{\partial\psi}\right\}.$$

Again, differentiating equation (2.4) with respect to θ_{t-1}

$$\frac{\partial p_{y_{t-1}}}{\partial\theta_{t-1}} = \frac{1}{A^2}\left[\{(2y_{t-1}-1)(-\frac{\partial\delta}{\partial\theta_{t-1}}+\theta_{t-1})+\theta_t\}A-2B\{(1-y_{t-1}+2y_{t-1}\theta_t\}\right]$$

where, $B=(2y_{t-1})\{(1-\delta+(\psi-1)\theta_{t-1}\}+(\psi-1)\theta_t$

differentiating successively B with respect to θ_{t-1}

$$\frac{\partial B}{\partial\theta_{t-1}} = (2y_{t-1}-1)\left\{-\frac{\partial\delta}{\partial\theta_{t-1}}-(\psi-1)\right\}$$

and, $$\frac{\partial^2 B}{\partial\theta^2{}_{t-1}} = (2y_{t-1}-1)\left(-\frac{\partial^2\delta}{\partial\theta_{t-1}^2}\right).$$

We know, $A=2(\psi-1)\{1-y_{t-1}+(2y_{t-1}-1)\theta_{t-1}\}$

now differentiation both sides with respect to θ_{t-1}

$$\frac{\partial A}{\partial\theta_{t-1}} = 2(\psi-1)(2y_{t-1}-1)$$

Therefore,

$$\frac{\partial^2 p_{y_{t-1}}}{\partial\theta^2{}_{t-1}} = \frac{1}{A^4}\left[A^2\frac{\partial E}{\partial\theta_{t-1}}-E\times 2\times A\frac{\partial A}{\partial\theta_{t-1}}\right]$$

where, $$E=\left[\{(2y_{t-1}-1)(-\frac{\partial\delta}{\partial\theta_{t-1}}+\theta_{t-1})+\theta_t\}A-2B(1-y_{t-1}+2y_{t-1}\theta_t)\right]$$

now differentiating E with respect to θ_{t-1}

$$\frac{\partial E}{\partial\theta_{t-1}} = \left[(2y_{t-1}-1)\left\{\frac{\partial A}{\partial\theta_{t-1}}\left(\theta_{t-1}-\frac{\partial\delta}{\partial\theta_{t-1}}\right)+A\left(1-\frac{\partial^2\delta}{\partial\theta^2{}_{t-1}}\right)\right\}+\theta_t\frac{\partial A}{\partial\theta_{t-1}}-2\times\frac{\partial B}{\partial\theta_{t-1}}\times(1-y_{t-1}+2y_{t-1}\theta_t)\right]$$

From equation (4.5),
we have, $\xi^*(\beta^*) = \ln u(\beta) + \xi(\beta)$. (4.9)

Therefore, $\xi^*(\beta^*) = \sum_{t=1}^{T}\{y_t \, logit(p_{y_{t-1}}) + log(1 - p_{y_{t-1}})\} + \ln u(\beta)$.

For squared error loss function, $u(\beta) = \beta$.

Therefore, $\xi^*(\beta^*) = \sum_{t=1}^{T}\{y_t \, logit(p_{y_{t-1}}) + log(1 - p_{y_{t-1}})\} + \ln \beta^*$. (4.10)

Now differentiating successively equation (4.10) with respect to β^*.

$$\frac{\partial \xi^*(\beta^*)}{\partial \beta^*} = \frac{\partial \xi^*(\beta^*)}{\partial p^*_{y_{t-1}}}\left(\frac{\partial p^*_{y_{t-1}}}{\partial \theta^*_t}.\frac{\partial \theta^*_t}{\partial \beta^*} + \frac{\partial p^*_{y_{t-1}}}{\partial \theta^*_{t-1}}.\frac{\partial \theta^*_{t-1}}{\partial \beta^*}\right) + \frac{1}{\beta^*} \quad (4.11)$$

and, $$\frac{\partial^2 \xi^*(\beta^*)}{\partial \beta^{*2}} = \frac{\partial^2 \xi^*(\beta^*)}{\partial p^{*2}_{y_{t-1}}}\left(\frac{\partial^2 p^*_{y_{t-1}}}{\partial \theta^{*2}_t}.\frac{\partial^2 \theta^*_t}{\partial \beta^{*2}} + \frac{\partial^2 p^*_{y_{t-1}}}{\partial \theta^{*2}_{t-1}}.\frac{\partial^2 \theta^*_{t-1}}{\partial \beta^{*2}}\right) - \frac{1}{\beta^{*2}}. \quad (4.12)$$

Now we use equations (4.6), (4.7) and (4.11), (4.12) to estimate the maximum likelihood estimator of parameter β and β^* respectively.
Then according to Tierney-Kadnae, the integral can be approximately be evaluated as

$$I(X) = \frac{\hat{\sigma}^*}{\hat{\sigma}} exp\left[\xi^*\left(\hat{\beta}^*\right) - \xi\left(\hat{\beta}\right)\right].$$

Therefore, The Bayes estimator of β under squared error loss function is

$$\hat{\beta}_{BSE} = \frac{\hat{\sigma}^*}{\hat{\sigma}} exp\left[\xi^*\left(\hat{\beta}^*\right) - \xi\left(\hat{\beta}\right)\right] \quad (4.13)$$

where, $\hat{\beta}_{BSE}$ represents that Bayesian estimator under squared error loss function, $\xi\left(\hat{\beta}\right)$ & $\xi^*\left(\hat{\beta}^*\right)$ have been obtained from equation (4.4) & (4.9) respectively, $\hat{\sigma}$ & $\hat{\sigma}^*$ have been estimated by equation (4.6) & (4.12) respectively and $\hat{\beta}$ & $\hat{\beta}^*$ are the maximum likelihood (Azzalini, 1994) estimators of β & β^* respectively. λ can be estimated similarly.

5. Bayesian Credible Interval

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

$$P(\beta \in (\beta_1, \beta_2) / X) = \int_{\beta_1}^{\beta_2} f(\beta / X) d\beta = 1 - \alpha \quad (5.1)$$

Mahanta *et al.* (2015) used ($1 - \alpha$) 100% Bayesian credible interval of β.

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 $P(\beta \in (\beta_1, \beta_2)/x) \geq 1-\alpha$ and is completely determined from the current observed data x and the prior distribution.

To estimate the parameter of Azzalini model by pregnancy complication data, three covariates have been utilized in this study because of complexity to fit the model. Three highly significant covariates viz. any miscarriage, socio economic status and age at marriage are used. Maximum likelihood approach and Bayesian approach have been applied for estimating the parameters of the model.

6. Numerical Result

In this paper, we have used pregnancy complication data collected from Bangladesh Institute of Research for Promotion of Essential & Reproductive Health and Technologies (BIRPERHT) for the period November 1992 to December 1993. The data were collected using both cross-sectional and prospective study designs. A total of 1059 pregnant women were interviewed in the follow-up component of the study.

Confidence intervals for maximum likelihood estimators and credible intervals of Bayesian estimators under squared error have been used to calculate the parameter of Azzalini model and have been presented in table 1 and table 2 respectively.

Covariates	Point Estimate	Odds ratio	Interval Estimate (95%)		
			Lower	Upper	Length
Constant	0.0994	-	0.0869	0.1119	0.0250
Any miscarriage	0.1001	1.1053	0.0973	0.1030	0.0058
Economic Status	0.1003	1.1055	0.0919	0.1087	0.0168
Age at Marriage	0.1000	1.1052	0.0973	0.1027	0.0054
λ	0.1000	1.1052	0.0996	0.1004	0.0008

Table 1: Confidence interval for Maximum Likelihood Estimate

Covariates	Point Estimate	Odds ratio	Credible Interval (95%)		
			Lower	Upper	Length
Constant	0.0993	-	0.0981	0.1006	0.0025
Any miscarriage	0.1000	1.1052	0.0988	0.1012	0.0025
Economic Status	0.1001	1.1053	0.0989	0.1013	0.0025
Age at Marriage	0.0999	1.1050	0.0987	0.1011	0.0025
λ	0.0999	1.1050	0.0986	0.1011	0.0025

Table 2: Credible interval for Bayesian estimate under squared error loss function

All covariates that are positively associated with pregnancy complication have been presented in Table 1 and Table 2. Lengths of all Baysian credible intervals are found to be smaller than the corresponding lengths of maximum likelihood confidence intervals. Therefore, Bayes estimator under squared error loss function is more preferable than method of maximum likelihood estimator to use estimate the parameters of Azzalini model. All the calculations were performed by using R-Software (Version-2.10.0).

1. Conclusions

Longitudinal data are widely used in various sectors such as medical science, social science, biological science etc. For this data, Markov model such as Azzalini model has been applied. Two approaches of estimation method viz. Bayesian approach and method of maximum likelihood have been employed for estimating the parameters of Azzalini model. Comparisons between Bayesian approach and maximum likelihood method have been made and the results show that the length of Bayesian credible interval is smaller than the corresponding length of confidence interval. According to decision rule, estimate having smaller length of interval is preferable. Thus, Bayesian approach under squared error loss function gives better estimate. Estimation is important part for decision-making. If Markov model is applied in follow-up data of medical science, social science, biological science, it is better to use Bayesian approach to estimate the parameters to get more accurate predictions.

References

1. Azzalini, A. (1994). Logistic regression for autocorrelated data with application to repeated measure, Biometrika, 81(4), p. 767-775.
2. Biswas, S. C., Islam, M. A. and Islam, J. N. (2011). Higher order markov structure-based logistic model and likelihood inference fo ordinal data, Journal of Modern Applied Statistical Methods , 10, p. 528-538.
3. Cook, R. J. and Ng, E. M. (1997). A Logistic-bivariate normal model for overdispersed two-state markov processes, Biometrics, 53, p. 358-364.
4. Islam, M. A. and Chowdhury, R. I. (2006). A higher order Markov model of analyzing covariate dependence, Applied Mathematical Modeling, 30, p .477-488.
5. Islam, M. A., Chowdhury, R. I. and Singh, K. P. (2012). A Markov model for analyzing polytomous outcome data, Pakistan Journal of Statistics and Operation Research, 8(3), p. 593-603.
6. Jaynes, E.T. (2003). Probability Theory: The Logic of Science, Cambridge, UK, Cambridge University Press.
7. Mahanta, J., Biswas, S. C., Roy, M. K. and Islam, M. A. (2015). A comparison of bayesian and classical approach for estimating Markov based logistic model, American Journal of Mathematics and Statistics, 5(4), p. 178-183.
8. Muenz, K.R., & Rubinstein, L.V. (1985). Markov models for covariance dependence of binary sequences. Biometrics, 41, p. 91-101.
9. Noorian, S. and Ganjali, M. (2012). Bayesian analysis of transition model for longitudinal ordinal response data: Application to insomnia data, International Journal of Statistics in Medical Research, 1, p. 148-161.

10. Podder, C. K. and Roy, M. K. (2003). Bayesian estimation of the parameter of Maxwell distribution under MLINEX loss function, Journal of Statistical Studies, 23, p. 11-16.
11. Raftery, A. and Tavare, S. (1994). Estimating and modeling repeated patterns in higher order Markov chains with the mixture transition distribution model, Applied Statistics, 43(1), p. 179-199.
12. Tierney, L. and Kadane, J. B. (1986). Accurate approximation for posterior moments and marginal densiies, Journal of American Statistical Association , 81(393), p. 82-86.