Journal of Data Science 567-592 , DOI: 10.6339/JDS.201807_16(3).0007

Estimating the parameters of Azzalini model by Bayesian approach under symmetric and asymmetric loss functions

Janardan Mahanta[1], Soma Chowdhury Biswas[2], Manindra Kumar Roy[3]

[1]Department of Statistics,University of Chittagong,Chittagong, Bangladesh.Email: johnstat@cu.ac.bd, [2]Department of Statistics,University of Chittagong,Chittagong, Bangladesh.Email: soma_chow@hotmail.com, [3] Ranada Prasad Shaha University, Narayanganj, Bangladesh.Email: manindraroy@yahoo.com

Abstract:This paper has been proposed to estimate the parameters of Markov based logistic model by Bayesian approach for analyzing longitudinal binary data. In Bayesian estimation selection of appropriate loss function and prior density are most important ingredient. Symmetric and asymmetric loss functions have been used for estimating parameters of two state Markov model and better performance has been observed by Bayesian estimate under squared error loss function.

Key words:Modified linear exponential (MLINEX) loss function, Markov model, Linear exponential (LINEX) loss function, Squared error loss function (SE).

1. Introduction

Bayesian estimation is the important part in real world situation. It application is increasing day by day. This paper has been applied Bayesian estimation to estimate the parameter of Markov model. For analyzing longitudinal data, Markov model are used. The longitudinal studies or follow-up studies are repeated over an extended period of time in order to measure the rate and degree of change occurring in patterns of response. For detailed study in Markov model were offered to Cox and Miller (1965), Kemeny and Snell (1976), Chiang (1980), and Karlin and Taylor (1981). The discrete time Markov models appear to be restricted due to over-parameterization and several attempts have been made to simplify the application of the models. Rafter (1985), Raftery and Tavare (1994), Berchtold, and Raftery (2002) aimed one such area of problems in estimating transition probabilities. Markov chain model for analyzing sequences of ordinal data from a relapsing remitting of disease was developed by Albert (1994) and quasi-likelihood models for a two state Markov chain with stationary transition probabilities for heterogeneous transitional data was developed by Albert and Waclawiw (1998). There were several studies on discrete time Markov chain models proposed analyzing repeated categorical data over decades. Kalbfleisch and Lawless (1985) proposed a continuous time Markov process model. They presented procedures for obtaining estimates for transition intensity parameters in homogenous models. Azzalini (1994) introduced a first order Markov model, to examine influence of time dependent covariate on the marginal distribution of the binary outcome variables in serially correlated binary data. Heagerty and Zeger (2000) presented a class of marginalized transition models (MTM) and Heagerty (2002) proposed a class of generalized MTMs to allow serial dependence of first order. Ware, Lipsitz and Speizer (1988) made a first broad distinction between transitional and marginal models depending on whether the covariates determine the conditional distribution given the past or the marginal distribution of any nominated observation. Muenz and Rubinstein (1985) employed logistic regression models to analyze the transition probabilities from one state to another. A logistic-bivariate normal mixture model for a two-state Markov chain was applied by Cook and Ng (1997). Liu (2010) showed the application of Markov chain in time series data. Recently, Islam *et al.* (2012) analyzed polytomous outcome data; they were used logistic link functions to analyze the transitional probability. Islam *et al.* (2008) examined the covariate dependence in a higher order Markov models. Chan (2012) analyzed of correlation structures using generalized estimating equation approach for longitudinal binary data. A selection model for longitudinal data was employed by Gad (2011), the proposed model applied to an example from the international breast cancer study group. Dey and Islam (2017) applied GEE approach in conditional count model for repeated data. Sirdari and Islam (2018) employed higher order binary Markov chain model using Health and Retirement Study (HRS) data. Following Azzalini (1994) model, Biswas *et al.* (2011) introduced a second order Markov model. The present study focuses on a logistic model with Markovian dependence proposed by Azzalini (1994) for the study of the influence of time dependent covariates on the binary response in serially correlated binary data. All of them applied method of maximum likelihood approach for decision-making. However, the maximum likelihood approach has some limitations that can be resolved by adoption of the more flexible Bayesian approach. Bhatta and Nandram (2013) apllied nonlinear regression model using Bayesian adjustment of the Heligman-Pollard (HP) law. Noorian and Ganjali (2012) was applied Bayesian analysis of transitional model for longitudinal ordinal response data, but in their study, they were applied Markov Chain Monte Carlo (MCMC). Although MCMC method is well known and popular but this method is to be solved by programming. There is no theoretical idea about the estimating procedure of parameter. Mahanta *et al.* (2015) employed Bayesian approach in Muenz and Rubinstein model. Mahanta and Biswas (2016) applied Bayesian approach under squared error loss function to estimate the parameter of Azzalini model. In Bayesian analysis, squared error loss function is the symmetric loss function but in most of the case loss is asymmetric in nature.

To get proper idea of the estimating procedure. This paper has been applied Bayesian approach under asymmetric loss function such as linear exponential (LINEX) and modified linear exponential (MLINEX) functions as well as squared error loss function to estimate the parameters of Azzalini model theoretically and the numerical findings were obtained using R programming.

2. Model

Let the binary response data $\left(y_1, y_{2,} ....y_t\right)$ be observed on n subjects at time $t = 1,2,...,T$ which is assumed to be generated by a Markov chain taking value 0 and 1. We assumed that associated covariates $x_t = (x_1, x_2, ..x_t)$ are recorded for each subject at each occasions. Our statistical objective is to obtain estimates for the regression of y_t on x_t using binary Markov chain of first order.

In practice, we often concerned 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,

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

where, $\theta_t = \Pr\left(Y_{it} = 1 / Y_{it-1} = 0\right) = \dfrac{e^{X_t \beta}}{1 + e^{X_t \beta}}$

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 previous response,

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

the transition probabilities

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

$$p_1 = p_{it,1} = \Pr(Y_{it} = 1 / Y_{it-1} = 1) \text{ and } 1 - p_1 = \Pr(Y_{it} = 0 / Y_{it-1} = 1)$$

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

$$\theta_{it} = p_{it,1}\, \theta_{it-1} + p_{it,0}\, (1 - \theta_{it-1}) \tag{2}$$

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

$$\psi_{it} = \frac{p_{it,1} / (1 - p_{it,1})}{p_{it,0} / (1 - p_{it,0})} \tag{3}$$

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) and (3) for p_o 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,

$$p_j = \begin{cases} \theta_t, \; for \;\psi = 1 \\ \dfrac{\delta - 1 + \{\psi - 1)(\theta_t - \theta_{t-1})}{2(\psi - 1)(1 - \theta_{t-1})} + j \dfrac{1 - \delta + (\psi - 1)(\theta_t + \theta_{t-1} - 2\theta_t \theta_{t-1})}{2(\psi - 1)\theta_{t-1}(1 - \theta_{t-1})}; \psi \neq 1 \end{cases} \tag{4}$$

for t = 2, 3, 4

where,

$$\delta^2 = 1 + (\psi - 1)\left\{(\theta_t - \theta_{t-1})^2 \psi - (\theta_t - \theta_{t-1})^2 + 2(\theta_t + \theta_{t-1})\right\} \tag{5}$$

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

and $X_{it} = (X_{i1}, X_{i2}, ..., X_{it})$ be the covariate matrix θ_{it} be denoted by the expected 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

$$\begin{aligned} f(y_t / x_t) &= p_{y_{t-1}}^{y_t}\left(1 - p_{y_{t-1}}\right)^{1 - y_{t-1}} \\ &= \left(\frac{p_{y_{t-1}}}{1 - p_{y_{t-1}}}\right)^{y_t}\left(1 - p_{y_{t-1}}\right) \end{aligned} \tag{6}$$

3. Prior and posterior distribution

Specification of a prior distribution is an important issue in Bayesian analysis. When a subjective prior is not available, use of non-informative has an extensive tradition in statistics. A commonly used non-informative prior is Jeffrey's prior along with the uniform prior.

Since the parameters of Markov model of β and λ lies between $-\infty$ to $+\infty$, then according to Jeffrey's non-informative prior (Mahanta and Biswas, 2016) of β and λ are $g(\beta,\lambda)=I$. Where, I represent the identity vector.

Then the posterior density function of β and λ for the given sample is

$$f(\beta,\lambda / X)=\frac{\prod_{i=1}^{n} f(y_t / x_t)g(\beta,\lambda)}{\int\prod_{i=1}^{n} f(y_t / x_t)g(\beta,\lambda)d(\beta,\lambda)} \tag{7}$$

where, $g(\beta,\lambda)$ is the joint prior density of β and λ .

$$\therefore g(\beta,\lambda)=g(\beta)g(\lambda).$$

Since, β and λ are independent.

4. Bayes estimators and their posterior risk

This paper have been used the Bayesian paradigm to make inferences about parameters of the Azzalini model using different loss functions namely squared error, linear exponential (LINEX) and modified linear exponential (MLINEX) loss function, also comparing their estimate to use posterior risk of respective estimates.

4.1. Bayes estimator under squared error loss function

The squared error loss function is defined as

$$L(\hat{\beta};\beta)=(\hat{\beta}-\beta)^2 \tag{8}$$

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

$$\hat{\beta}_{BSE}=\frac{\int\beta\prod_{i=1}^{n} f(y_t / x_t)g(\beta)}{\int\prod_{i=1}^{n} f(y_t / x_t)g(\beta)d(\beta)} \tag{9}$$

The Bayesian integration of Markov based logistic model is complex one. The Bayesian integral appeared in the computation cannot be solvable. Now we use the approximation to evaluate the Bayesian integral. By Lindley's approximation, we solve this type of Bayesian integral.

Lindley (1980) suggest that, if the form of the integrals is

$$I(X)=E(u(\beta)/X)=\frac{\int u(\beta)e^{l_t(\beta)+p(\beta)}d\beta}{\int e^{l_t(\beta)+p(\beta)}d\beta} \tag{10}$$

where, $I(X)$ represent the form of the integral, l_t is the log likelihood and $p(\beta)$is the log of prior. Then according to Lindley (Press, 1989), the integral can be approximately be evaluated as

$$I(X)=u(\hat{\beta})+\frac{1}{2}\left[u''(\beta)+2u'(\hat{\beta})p'(\hat{\beta})+l_t'''(\hat{\beta})u(\hat{\beta})\hat{\sigma}^2\right]\hat{\sigma}^2 \tag{11}$$

where, $u(\beta)$ is the functional form of the parameter β, that is used in expectation of posterior density and $\hat{\beta}$ is the maximum likelihood estimators of β.

where, $u'(\beta)=\frac{\partial u(\beta)}{\partial\beta}$, $\quad u''(\beta)=\frac{\partial^2 u(\beta)}{\partial^2\beta}$

$$p'(\beta)=\frac{\partial p(\beta)}{\partial\beta}, \qquad l_t'''(\beta)=\frac{\partial^3 l_t(\beta)}{\partial^3\beta} \qquad and\,\hat{\delta}^2=-\frac{1}{l_t''(\beta)}$$

The parameters β is estimated by Bayesian approach and using chain rule of differentiation. Now considering the t^{th} term of the log-likelihood function, its derivatives are computed via Chain rule of differentiation (Azzalini, 1994).

$$\frac{\partial l_t}{\partial\beta}=\frac{\partial l_t}{\partial p_{y_{t-1}}}\left(\frac{\partial p_{y_{t-1}}}{\partial\theta_t}.\frac{\partial\theta_t}{\partial\beta}+\frac{\partial p_{y_{t-1}}}{\partial\theta_{t-1}}.\frac{\partial\theta_{t-1}}{\partial\beta}\right);t=1,2,3,4 \tag{12}$$

Following the first order differentiation of Azzalini (1994), we want to differentiate second and third order differentiation.

Therefore,

$$\frac{\partial^2 l_t}{\partial\beta^2}=\frac{\partial^2 l_t}{\partial p^2{}_{y_{t-1}}}\left(\frac{\partial^2 p_{y_{t-1}}}{\partial\theta^2{}_t}.\frac{\partial^2\theta_t}{\partial\beta^2}+\frac{\partial^2 p_{y_{t-1}}}{\partial\theta^2{}_{t-1}}.\frac{\partial^2\theta_{t-1}}{\partial\beta^2}\right) \tag{13}$$

and

$$\frac{\partial^3 l_t}{\partial\beta^3}=\frac{\partial^3 l_t}{\partial p^3{}_{y_{t-1}}}\left(\frac{\partial^3 p_{y_{t-1}}}{\partial\theta^3{}_t}.\frac{\partial^3\theta_t}{\partial\beta^3}+\frac{\partial^3 p_{y_{t-1}}}{\partial\theta^3{}_{t-1}}.\frac{\partial^3\theta_{t-1}}{\partial\beta^3}\right) \tag{14}$$

The log-likelihood function from equation (6)

$$l_t=\log L(\beta)=\sum_{t=1}^{T}\left\{y_t\log it\left(p_{y_{t-1}}\right)+\log\left(1-p_{y_{t-1}}\right)\right\}$$

Now, three times differentiation both sides with respect to $p_{y_{t-1}}$

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

again

$$\frac{\partial^2 l_t}{\partial p^2{}_{y_{t-1}}} = \Sigma\left\{\frac{y_t\left(2p_{y_{t-1}}-1\right)-p^2{}_{y_{t-1}}}{p^2{}_{y_{t-1}}\left(1-p_{y_{t-1}}\right)^2}\right\} \tag{15}$$

and

$$\frac{\partial^3 l_t}{\partial p^3{}_{y_{t-1}}} = \sum\left\{\frac{2\left[p^5{}_{y_{t-1}} - p^4_{y_{t-1}} + y_t\left\{p_{y_{t-1}} - 4p^2_{y_{t-1}} + 6p^3_{y_{t-1}} - 3p^4_{y_{t-1}}\right\}\right]}{p^4{}_{y_{t-1}}\left(1-p_{y_{t-1}}\right)^4}\right\} \tag{16}$$

Again, from equation (4) we differentiate successively $p_{y_{t-1}}$ with respect to θ_t

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

again

$$\frac{\partial^2 p_{y_{t-1}}}{\partial\theta^2{}_t} = \frac{1}{A}\left[-(2y_{t-1}-1)\frac{\partial^2\delta}{\delta\theta_t^2}\right] \tag{17}$$

and

$$\frac{\partial^3 p_{y_{t-1}}}{\partial\theta^3{}_t} = \frac{1}{A}\left[-(2y_{t-1}-1)\frac{\partial^3\delta}{\delta\theta_t^3}\right] \tag{18}$$

other terms are vanished, since these are independent of θ_t

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

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

Again, we know that,

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

differentiating θ_t successively by β

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

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

again

$$\frac{\partial^2\theta_t}{\partial\beta^2} = \frac{\{1+exp(X_t\beta)\}X_t\frac{\partial\theta_t}{\partial\beta} - X_t^2\theta_t\, exp(X_t\beta)}{\{1+exp(X_t\beta)\}^2} \tag{19}$$

and

$$\frac{\partial^3\theta_t}{\partial\beta^3} = \frac{\{1+exp(X_t\beta)\}^2\frac{\partial D}{\partial\beta} - D.2\{1+exp(X_t\beta)\}X_t\, exp(X_t\beta)}{\{1+exp(X_t\beta)\}^4} \tag{20}$$

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

D have been differentiated by β

$$\frac{\partial D}{\partial\beta} = \{1+\exp(X_t\beta)\}X_t\frac{\partial^2\theta_t}{\partial\beta^2} - X_t^3\theta_t\exp(X_t\beta).$$

From equation (5)

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

δ has been successively differentiate by θ_t

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

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 δ has been successively differentiate by θ_{t-1}

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

again

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

and

$$\frac{\partial^3\delta}{\partial\theta_{t-1}^{\,3}}=\frac{(\psi-1)}{\delta^4}\left[\delta^2\{(\theta_{t-1}-\theta_t)(1-\psi)-1\}\times\frac{\partial^2\delta}{\partial\theta_{t-1}^2}-2\times\delta\times T\right]$$

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

now successive differentiation δ with respect to ψ

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

again

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

and

$$\frac{\partial^3\delta}{\partial\psi^3}=\frac{1}{\delta^3}\left[\left\{-(\theta_t-\theta_{t-1})^2\psi\frac{\partial^2\delta}{\partial\psi^2}-(\theta_t+\theta_{t-1})\frac{\partial^2\delta}{\partial\psi^2}\right\}-2\frac{\partial\delta}{\partial\psi}\left\{(\theta_t-\theta_{t-1})^2\left(\delta-\psi\frac{\partial\delta}{\partial\psi}\right)-(\theta_t+\theta_{t-1})\frac{\partial\delta}{\partial\psi}\right\}\right].$$

Again, from equation (4) $p_{y_{t-1}}$ is differentiated successively by θ_{t-1}

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

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

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

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

again

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

and

$$\frac{\partial^3 B}{\partial\theta^3_{t-1}}=(2y_{t-1}-1)\left(-\frac{\partial^3\delta}{\partial\theta_{t-1}^3}\right).$$

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

differentiating both sides with respect to θ_{t-1}

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

therefore,

$$\frac{\partial^2 p_{y_{t-1}}}{\partial \theta^2_{t-1}} = \frac{1}{A^4}\left[A^2\frac{\partial E}{\partial \theta_{t-1}} - E\times 2\times A\frac{\partial A}{\partial \theta_{t-1}}\right] \quad (21)$$

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

now applying successive differentiation by θ_{t-1}

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

ain

$$\frac{\partial^2 E}{\partial \theta^2_{t-1}} = \left[(2y_{t-1}-1)\left\{2\frac{\partial A}{\partial \theta_{t-1}}\left(1-\frac{\partial \delta}{\partial \theta_{t-1}}\right) + A\left(-\frac{\partial^3 \delta}{\partial \theta^3_{t-1}}\right)\right\} - 2\times\frac{\partial^2 B}{\partial \theta^2_{t-1}}\times(1-y_{t-1}+2y_{t-1}\theta_t)\right]$$ a

nd

$$\frac{\partial^3 p_{y_{t-1}}}{\partial \theta^3_{t-1}} = \frac{1}{A^8}\left[A^4\left\{A^2\frac{\partial^2 E}{\partial \theta^2_{t-1}} - 2\times E\left(\frac{\partial A}{\partial \theta_{t-1}}\right)^2\right\} - \left\{A\frac{\partial E}{\partial \theta_{t-1}} - 2EA\frac{\partial A}{\partial \theta_{t-1}}\right\}\times 4A^3\frac{\partial A}{\partial \theta_{t-1}}\right] \quad (22)$$

Again, we know that,

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

Now differentiating successively both sides with respect to ψ

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

again

$$\frac{\partial^2 p_{y_{t-1}}}{\partial \psi^2} = \frac{(\psi-1)^2\frac{\partial^2 \delta}{\partial \psi^2} - 2(\psi-1)\frac{\partial \delta}{\partial \psi} + 2(\delta-1)}{2(\psi-1)^3(1-\theta_{t-1})}\left\{1-\frac{y_{t-1}}{\theta_{t-1}}\right\}$$

and

$$\frac{\partial^3 p_{y_{t-1}}}{\partial \psi^3} = \frac{(\psi-1)\frac{\partial^3 \delta}{\partial \psi^3} - 3(\psi-1)^2\frac{\partial^2 \delta}{\partial \psi^2} - 6(\psi-1)\frac{\partial \delta}{\partial \psi} - 6(\delta-1)}{2(\psi-1)^4(1-\theta_{t-1})}\left\{1-\frac{y_{t-1}}{\theta_{t-1}}\right\}$$

From integral (9) and (11) we have,

$$u(\hat{\beta}) = \hat{\beta}, \qquad \frac{\partial u(\hat{\beta})}{\partial \beta} = 1 \qquad \text{and} \qquad \frac{\partial^2 u(\hat{\beta})}{\partial \beta^2} = 0.$$

Therefore, the Bayes estimator under squared error loss function is

$$\hat{\beta}_{BSE} = \hat{\beta} + \frac{1}{2}\left[l_t'''(\beta)\times\hat{\beta}\times\hat{\sigma}^2\right]\times\hat{\sigma}^2. \quad (23)$$

Where, $l'''(t)=\frac{\partial^3 l_t}{\partial\beta^3}$ has been obtained from equation (16), (18), (20) and (22). $\hat{\sigma}^2=-\frac{1}{l_t''}$ has been estimated by using equation (15), (17), (19), and (21). $\hat{\beta}_{BSE}$ is the Bayesian estimator under squared error loss function and $\hat{\beta}$ is the maximum likelihood (Azzalini, 1994) estimators of β.

4.2. Posterior risk function under squared error loss function

Posterior risk function is the expected value of loss function with respect to posterior density

$$\begin{aligned} R_p(\beta) &= \int\left(\hat{\beta}-\beta\right)^2 f(\beta/X)d\beta \\ &= \int\left(\hat{\beta}^2-2\hat{\beta}\beta+\beta^2\right)f(\beta/X)d\beta \\ &= -\hat{\beta}_{BSE}+\int\beta^2 f(\beta/X)d\beta \end{aligned} \qquad (24)$$

4.3. Bayes Estimator under LINEX Loss Function

We know that, Bayes estimator under LINEX loss function (Zellner, 1971) is

$$\hat{\beta}_{BL}=-\frac{1}{c}\log E\left(e^{-c\beta}\right)$$

From integral I(X) in equation (11) we have,

$$u(\beta)=e^{-c\beta}, \quad u'(\beta)=-ce^{-c\beta} \text{ and, } \quad u''(\beta)=c^2e^{-c\beta}$$

Therefore, the Bayes estimator of LINEX loss function is

$$\hat{\beta}_{BL}=\hat{\beta}-\frac{1}{c}\log\left[1+\frac{1}{2}\left\{c^2+l_t'''(\hat{\beta})\hat{\sigma}^2\right\}\hat{\sigma}^2\right] \qquad (25)$$

where, $\hat{\beta}_{BL}$ is Bayes estimator under LINEX loss function and $\hat{\beta}$ is the maximum likelihood (Azzalini, 1994) estimates of β.

4.4. Posterior risk function under LINEX loss function

Posterior risk under LINEX loss function evaluated as

$$\begin{aligned} R_p(\beta) &= \int\left[e^{c\left(\hat{\beta}-\beta\right)}-c\left(\hat{\beta}-\beta\right)-1\right]f(\beta/X)d\beta \\ &= e^{\hat{\beta}_{BL}}\int e^{-c\beta}f(\beta/X)d\beta-c\hat{\beta}_{BL}+c\hat{\beta}_{BSE}-1 \end{aligned} \qquad (26)$$

4.5. Bayes estimator under MLINEX loss function

We know that Bayes estimator under MLINEX loss function (Wahed & Uddin, 1998) is

$$\hat{\beta}_{BML}=\left[E\left(\beta^{-c}\right)\right]^{-\frac{1}{c}}$$

where, $E\left(\beta^{-c}\right)=\frac{\int\beta^{-c}e^{L_0(\beta)+p(\beta)}d\beta}{\int e^{L_0(\beta)+p(\beta)}d\beta}$

From integral I(X), in equation (11) we have,

$u(\beta)=\beta^{-c}$, $u'(\beta)=-c\beta^{-(c+1)}$ and $u''(\beta)=c(c+1)\beta^{-(c+2)}$

Therefore, Bayes estimator of β under MLINEX loss function

$$\hat{\beta}_{BML} = \hat{\beta}\left[1+\frac{1}{2}\left\{c(c+1)\hat{\beta}^{-2}+l_t'''(\hat{\beta})\hat{\sigma}^2\right\}\hat{\sigma}^2\right]^{-1/c} \tag{27}$$

where, $\hat{\beta}_{BML}$ is the Bayes estimator under MLINEX loss function and $\hat{\beta}$ is the maximum likelihood (Azzalini, 1994) estimators of β.

4.6. Posterior risk function under MLINEX loss function

Using MLINEX loss function posterior risk is

$$R_p(\beta)=\int\left[\left(\frac{\hat{\beta}}{\beta}\right)^c - c\,log\left(\frac{\hat{\beta}}{\beta}\right)-1\right]f(\beta/X)d\beta \tag{28}$$

$$= \hat{\beta}_{BML}\int\beta^{-c} f(\beta/X)d\beta - c\,log\,\hat{\beta}_{BML} + c\int log\beta\, f(\beta/X)d\beta - 1$$

Estimate the parameter λ and their posterior risk in similar way.

5. Numerical results

The covariate dependent Markov model proposed in this paper is applied to the pregnancy complication data conducted 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. Three highly significant covariates viz. age at marriage, economic status, any miscarriage have been used.

To estimate the parameter of Azzalini model by Bayesian approach under squared error loss function, LINEX loss function and MLINEX loss function have been obtained from equation (23), (25) and (27) respectively and the result of these estimates have been presented in table 1, table 2 and table 3 respectively.

Table 1: Bayes estimates under squared error loss function

Covariates	Estimate	Odds ratio	Posterior risk
Constant	0.09938	-	0.00004
Any miscarriage	0.10014	1.10533	0.00000
Economic Status	0.10027	1.10548	0.00002
Age at Marriage	0.10004	1.10521	0.00000
λ	0.10000	1.09842	0.00000

Table 2: Bayes estimates and posterior risk under LINEX loss function

Covariates	Estimate	Odds ratio	Posterior risk
Constant	0.10503	-	0.00564
Any miscarriage	0.10057	1.10580	0.00042
Economic Status	0.10038	1.10559	0.00011
Age at Marriage	0.10046	1.10568	0.00042
λ	0.17196	1.18763	0.07196

Table 3: Bayes estimates and posterior risk under MLINEX loss function

Covariates	Estimate	Odds ratio	Posterior risk
Constant	0.71244	-	3.91307
Any miscarriage	0.10939	1.11560	0.08410
Economic Status	0.10243	1.10787	0.02025
Age at Marriage	0.10930	1.11549	0.08430
λ	0.14481	1.15582	2.14771

On comparing the posterior risk of Bayes estimator under different loss functions, it has been observed that Bayes estimator under squared error loss function has smallest posterior risk than other Bayes estimator under LINEX and Bayes estimator under MLINEX loss functions for fixed value c=1. Where, c is the shape parameter of the LINEX and MLINEX loss functions. According to decision rule, the less Bayes posterior risk under squared error loss function is more preferable estimate. In addition, three covariates viz. any miscarriage, economic status and age at marriage are positively associated with pregnancy complication and result have been shown by table 1, table 2 and table 3.

Now we want to observe that the performance of Bayesian estimate under LINEX and MLINEX loss function under different values of c.

Table 4: Estimates the parameters of Bayesian approach under LINEX and MLINEX loss function under different values of c

Value of c	Covariates	LINEX		MLINEX	
		Estimate	Posterior risk	Estimate	Posterior risk
2	Constant	0.11077	0.02277	0.06391	1.00361
	Any miscarriage	0.10099	0.00170	0.11592	0.28407
	Economic Status	0.10049	0.00042	0.10361	0.06322
	Age at Marriage	0.10089	0.00170	0.11584	0.28480
	λ	0.14929	0.09858	0.05032	0.18137
5	Constant	0.12971	0.15163	0.05702	1.06147
	Any miscarriage	0.10228	0.01066	0.13030	3.29500
	Economic Status	0.10081	0.00265	0.10820	0.37483
	Age at Marriage	0.10217	0.01066	0.12989	3.28472
	λ	0.16156	0.30780	0.05318	2.26991
8	Constant	0.15520	0.44655	0.06267	2.14242
	Any miscarriage	0.10358	0.02751	0.09159	1.25143
	Economic Status	0.10113	0.00681	0.11979	1.41420
	Age at Marriage	0.10347	0.02751	0.09145	1.24815
	λ	0.20922	0.87377	0.06023	3.83617

Value of c	Covariates	LINEX	MLINEX		
		Estimate	Posterior risk	Estimate	Posterior risk
10	Constant	0.18211	0.82727	0.06581	2.68808
	Any miscarriage	0.10448	0.04332	0.08798	0.66296
	Economic Status	0.10134	0.01066	0.12048	3.82532
	Age at Marriage	0.10437	0.04332	0.08786	0.66014
	λ	0.30865	2.08649	0.06386	4.71075
-2	Constant	0.08800	0.02277	0.03712	4.08295
	Any miscarriage	0.09930	0.00170	0.09582	0.09680
	Economic Status	0.10006	0.00042	0.09921	0.02342
	Age at Marriage	0.09919	0.00170	0.09570	0.09701
	λ	0.05071	0.09858	0.08310	2.81546
-5	Constant	0.06906	0.15163	0.15874	0.05820
	Any miscarriage	0.09801	0.01066	0.06894	1.88835
	Economic Status	0.09974	0.00265	0.09564	0.24204
	Age at Marriage	0.09790	0.01066	0.06870	1.90037
	λ	0.03844	0.30780	0.17268	0.38156
-8	Constant	0.04357	0.44655	0.15261	0.97796
	Any miscarriage	0.09671	0.02751	0.10414	1.72108
	Economic Status	0.09942	0.00681	0.08969	0.90073
	Age at Marriage	0.09660	0.02751	0.10407	1.71735
	λ	-0.00922	0.87377	0.16079	0.01886
-10	Constant	0.01666	0.82727	0.14705	1.35131
	Any miscarriage	0.09581	0.04332	0.11102	1.01111
	Economic Status	0.09921	0.01066	0.07451	2.97983
	Age at Marriage	0.09570	0.04332	0.11094	1.00821
	λ	-0.10865	2.08649	0.15345	0.05606

Table 4 has been presented that, smaller posterior risk of Bayesian estimator under LINEX loss function than MLINEX loss function for all cases. That means, Bayesian estimate under LINEX loss function gives better estimate than MLINEX loss function. The patterns of their risk have been shown in the following graph.

Now we want to observe the performance of the estimates of Bayesian approach under LINEX and MLINEX loss function using their posterior risk in the following graph.

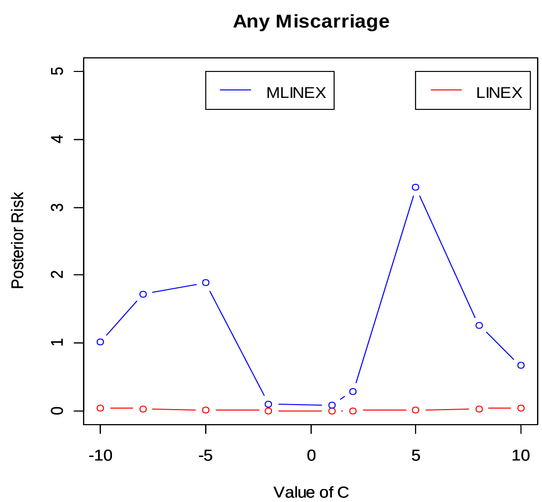

Figure 1: Graph of Posterior risk of Bayesian estimates of LINEX and MLINEX loss function for covariate Any Miscarriage

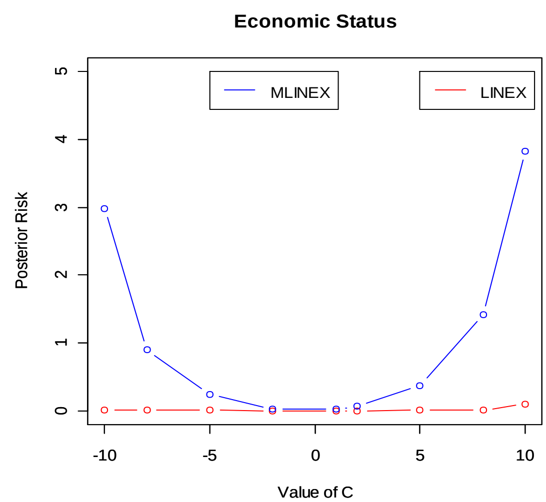

Figure 2**:** Graph of Posterior risk of Bayesian estimates of LINEX and MLINEX loss function for covariate Economic Status

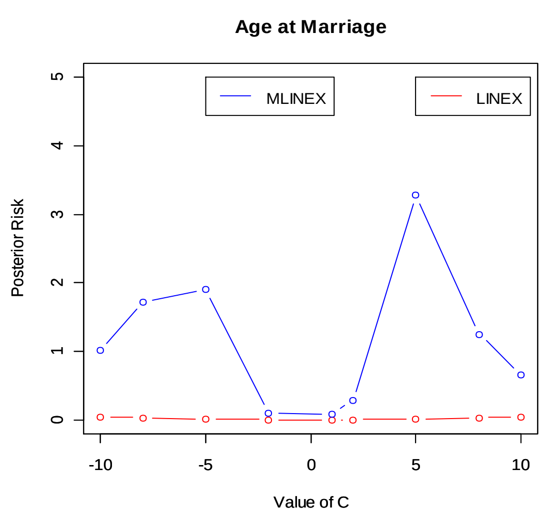

Figure 3: Graph of Posterior risk of Bayesian estimates of LINEX and MLINEX loss function for covariate age at marriage

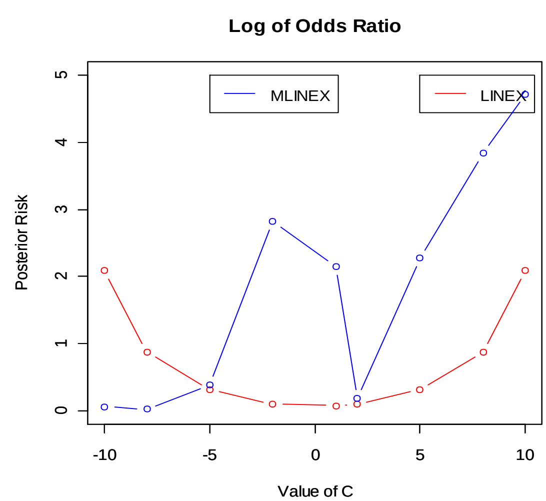

Figure 4: Graph of Posterior risk of Bayesian estimates of LINEX and MLINEX loss function for log of odds ratio

From the above graphical study, have been concluded that for three covariates Bayes estimator under LINEX loss function has lower posterior risk than MLINEX loss function for all values of c. All the calculations were performed by R-Language (Version-2.10.0). The program has been shown in appendix section.

6. Conclusions

Azzalini model is used in the case of transitional patterns from one state to another such as present status of patient's complication depend on the directly past history, future sugar level of a diabetic patient depends on present sugar level etc.. Azzalini (1994) estimated the parameter of the model by method of maximum likelihood but maximum likelihood method has some limitations. This paper has been used Bayesian approach to estimate the parameters of the model. This article illustrates theoretical elaborations about Bayesian approach on two state Markov model. In Bayesian approach symmetric and asymmetric loss functions have been used namely squared error, LINEX and MLINEX loss functions.

Performances of Bayesian estimate under LINEX loss function and MLINEX loss function for different value of c have been observed. Bayesian estimate under LINEX loss function has smaller posterior risk than MLIEX loss function for all value of c. In addition, squared error loss function has smallest posterior risk than LINEX and MLINEX loss function for all cases.

Therefore, Bayesian approach under squared error loss function can be suggested to estimate the parameters of Azzalini model.

References

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

[2] Albert, P. S., & Waclawiw, M. A. (1998). A two state Markov chain for heterogenous transitional data: a quasi-likelihood approach. *Statistics in Medicine, 17,* 1481-1493.

[3] Azzalini, A. (1994). Logistic Regression for Autocorrelated data with Application to Repeated Measure. *Biometrika, 81, 4*, 767-775.

[4] Bhatta, D., & Nandram, B. (2013). A Bayesian Adjustment of the HP Law via a Switching Nonlinear Regression Model. *Journal of Data Science , 11*, 85-108.

[5] Berchtold, A., & Raftery, A. E. (2002). The mixture transition distribution model for high-order Markov chains and non-Gaussian time series. *Statistical Science, 17,* 328-356.

[6] Biswas, S. C., Islam, M. A., & Islam, J. N. (2011). Higher Order Markov Structure-Based Logistic Model and Likelihood Inference fo Ordinal Data. *Journal of Modern Applied Statistical Methods , 10*, 528-538.

[7] Chan, J. S. (2012). Analysis of Correlation Structures using Generalized. *Journal of Data Science , 12*, 293-305.

[8] Chiang, C. L. (1980). *An Intrduction to Stochastic Processes and Their Applications.* New York: Wiley.

[9] Cook, R. J., & Ng, E. M. (1997). A Logistic-Bivariate Normal Model for Overdispersed Two-State Markov Processes. *Biometrics, 53,* 358-364.

[10] Cox, D. R., & Miller, H. D. (1965). *The Theory of Stochastic Processes.* London: Methuen.

[11] Dey, R., & Islam, M. A. (2017). A conditional count model for repeated count data and its application to GEE approach. (504, Ed.) *Statistical Papers , 58* (2), 485.

[12] Gad, A. M. (2011). A Selection Model for Longitudinal Data with Non-Ignorable Non-Monotone Missing Values. *Journal of Data Science , 9*, 171-180.

[13] Heagerty, P. J. (2002). Marginalized transition models and likelihood inference for longitudinal categorical data. *Biometrics, 58,* 342-351.

[14] Heagerty, P. J., & Zeger, S. L. (2000). Marginalized multi-level models and likelihood inference (with Discussion). *Statistical Science, 15,* 1-26.

[15] Islam, M. A., Chowdhury, R. I., & Singh, K. P. (2012). A Markov Model for Analyzing Polytomous Outcome Data. *Pakistan Journal of Statistics and Operation Research, 8 (3),* 593-603.

[16] Health and Retirement Study (HRS). (2009). *Produced and distributed by the University of Michigan with funding from the National Institute on Aging* (Grant No. NIAU01AG09740). Waves [1-8], Year [1992–2006] [Online], [Accessed 2009]. Retrieved from World Wide Web: http://hrsonline.isr.umich.edu/data/index.html.

[17] Islam, M. A., Chowdhury, R. I., & Singh, K. P. (2008). Covariate Dependent Markov Models foe Analysis of Repeated Binary Outcomes. *Journal of Modern Applied Statistical Methods, 6 (2),* 561-572.

[18] Kalbfleisch, J. D., & Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of American Statistical Association, 88,* 863-871.

[19] Karlin, S., & Taylor, H. M. (1981). *A Second Course in Stochastic Processes.* New York: Academic Press.

[20] Kemeny, J. G., & Snell, J. L. (1976). *Finite Markov Chains.* New York: Springer.

[21] Lindley, D. V. (1980). *Approximate Bayes methods. Bayesian Statistics.* Valencia, Spain: University Press.

[22] Liu, T. (2010). Application of Markov Chains to Analyze and Predict the Time Series. *Modern Applied Science , 4* (5), 162-166.

[23] Mahanta, J., & Biswas, S. C. (2016). Comparison of Bayesian approach with classical approach for estimating the Markov Model. *Journal of reliability and Statistical studies*, 9(2), p. 81-90.

[24] Mahanta, J., Biswas, S. C., Roy, M. K., & 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.

[25] Muenz, K.R., & Rubinstein, L.V. (1985). Markov chain for covariance dependence of binary sequences. *Biometrics*, 41, 91-101.

[26] Noorian, S., & 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*, 148-161.

[27] Podder, C. K., & Roy, M. K. (2003). Bayesian estimation of the parameter of Maxwell distribution under MLINEX loss function. *Journal of Statistical Studies,* 11-16.

[28] Press, S. J. (1989). *Bayesian Statistics: Principles, Models, and Applications.* New York: John Wiley & Sons.

[29] Raftery, A. E. (1985). A model for high-order Markov chains. *Journal of the Royal Statistical Society, Series B 47*, 528-539.

[30] 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.

[31] Sirdari, M. Z., & Islam, M. A. (2018). Goodness of fit test for higher order binary Markov chain models. *Cogent Mathematics & Statistics , 5*, 1-15.

[32] Wahed, A. S., & Uddin, B. (1998). Bayes estimation under asymetric loss function. *Dhaka University Journal of Science, 46 (6),* 355-361.

[33] Ware, J. H., Lipsitz, S., & Speizer, F. E. (1988). Issues in the analysis of repeated categorical outcome. *Statistics in Medicine, 7*, 95-107.

[34] Zellner, A. (1971). *An introduction to Bayesian inference in econometrics.* New York: Wiley

Appendix

```
library(foreign)
data=read.table("D:/dat.txt", header=TRUE)
initial<-c(.1,.1,.1,.1,.1)
c<-(1)
bin1<-function(data,initial)
{
id <- data[,1]
fup <- data[,2]
y <- data[,4]
A<-data[,3]
 B <- data[,5]
C <-data[,6]
count <- 0
k <- 0
repeat {
b0 <- initial[1]
b1 <- initial[2]
b2 <- initial[3]
 b3 <- initial[4]
 l<- initial[5]
b<-as.vector(c(b0,b1,b2,b3,l))
score <- c(0,0,0,0,0)
inf <- matrix(c(0,0,0,0,0,0,0,0, 0, 0, 0, 0, 0, 0, 0, 0,0,0,0,0,0,0,0,0,0),
ncol =5,byrow = T)
for(i in 1:(length(fup)-3))
        {
if(fup[i] == 1 && fup[i + 1] == 2 && fup[i + 2] == 3 && fup[i + 3] == 4)
        {
        k <- k + 1
#cal of theta
m1<-
exp(b0+b1*A[i]+b2*B[i]+b3*C[i])/(1+exp(b0+b1*A[i]+b2*B[i]+b3*C[i]))
m2<-
exp(b0+b1*A[i+1]+b2*B[i+1]+b3*C[i+1])/(1+exp(b0+b1*A[i+1]+b2*B[i+
1]+b3*C[i+1]))
m3<-
exp(b0+b1*A[i+2]+b2*B[i+2]+b3*C[i+2])/(1+exp(b0+b1*A[i+2]+b2*B[i+
2]+b3*C[i+2]))
m4<-
exp(b0+b1*A[i+3]+b2*B[i+3]+b3*C[i+3])/(1+exp(b0+b1*A[i+3]+b2*B[i+
3]+b3*C[i+3]))
v<-exp(l)
#cal of delta
d2<-sqrt(1+(v-1)*((m2-m1)^2)*v-(m2-m1)^2+2*(m2+m1))
d3<-sqrt(1+(v-1)*((m3-m2)^2)*v-(m3-m2)^2+2*(m3+m2))
d4<-sqrt(1+(v-1)*((m4-m3)^2)*v-(m4-m3)^2+2*(m4+m3))
#cal A&B
A2<-2*(v-1)*(1-y[i]+(2*y[i]-1)*m1)
A3<-2*(v-1)*(1-y[i+1]+(2*y[i+1]-1)*m2)
A4<-2*(v-1)*(1-y[i+2]+(2*y[i+2]-1)*m3)

B2<-  (2*y[i]-1)*(1-d2+(v-1))*m1 +(v-1)*m2
B3<-(2*y[i+1]-1)*(1-d3+(v-1))*m2 +(v-1)*m3
B4<-(2*y[i+2]-1)*(1-d4+(v-1))*m3 +(v-1)*m4
#calculation of del(l) by del(p)
#calculation of del(l) by del(p)
#calcuation of pyt-1
p1<-((d2-1+(v-1)*(m2-m1))/(2*(v-1)*(1-m1)))+(y[i]*(1-d2+(v-
1)*(m2+m1-2*m1*m2)))/(2*(v-1)*m1*(1-m1))
p2<-((d3-1+(v-1)*(m3-m2))/(2*(v-1)*(1-m2)))+(y[i+1]*(1-d3+(v-
1)*(m3+m2-2*m2*m3)))/(2*(v-1)*m2*(1-m2))
p3<-((d4-1+(v-1)*(m4-m3))/(2*(v-1)*(1-m3)))+(y[i+2]*(1-d4+(v-
1)*(m4+m3-2*m3*m4)))/(2*(v-1)*m3*(1-m3))
q2<-(y[i+1]-p1)/(p1*(1-p1))
q3<-(y[i+2]-p2)/(p2*(1-p2))
q4<-(y[i+3]-p3)/(p3*(1-p3))
#del of delta by theta(t-1)
dd21<-(v-1)*(-v*(m2-m1)+(m2-m1)+1)/d2
dd31<-(v-1)*(-v*(m3-m2)+(m3-m2)+1)/d3
dd41<-(v-1)*(-v*(m4-m3)+(m4-m3)+1)/d4
#del of delta by theta(t)
dd22<-(v-1)*(v*(m2-m1)-(m2-m1)+1)/d2
dd32<-(v-1)*(v*(m3-m2)-(m3-m2)+1)/d3
dd42<-(v-1)*(v*(m4-m3)-(m4-m3)+1)/d4
#del(p) by del(theta(t))
p21<-(1/A2)*(-(2*y[i]-1)*dd22+v-1)
p31<-(1/A3)*(-(2*y[i+1]-1)*dd32+v-1)
p41<-(1/A4)*(-(2*y[i+2]-1)*dd42+v-1)
#del(p) by del(theta(t-1))
p23<-(((2*y[i]-1)*(-dd21+m1)+m2)*A2-2*B2*(1-y[i]+2*y[i]*m2))/(A2)^2
p33<-(((2*y[i+1]-1)*(-dd31+m2)+m3)*A3-2*B3*(1-
y[i+1]+2*y[i+1]*m3))/(A3)^2
p43<-(((2*y[i+2]-1)*(-dd41+m3)+m4)*A4-2*B4*(1-
y[i+2]+2*y[i+2]*m4))/(A4)^2
#del theta(t) by del beta
p022<-m2*(1-m2)*1
p032<-m3*(1-m3)*1
p042<-m4*(1-m4)*1

p122<-m2*(1-m2)*B[i+1]
p132<-m3*(1-m3)*B[i+2]
p142<-m4*(1-m4)*B[i+3]

p222<-m2*(1-m2)*A[i+1]
p232<-m3*(1-m3)*A[i+2]
p242<-m4*(1-m4)*A[i+3]

p322<-m2*(1-m2)*C[i+1]
p332<-m3*(1-m3)*C[i+2]
p342<-m4*(1-m4)*C[i+3]

#del theta(t-1) by del beta
p024<-m1*(1-m1)*1
p034<-m2*(1-m2)*1
p044<-m3*(1-m3)*1

p124<-m1*(1-m1)*B[i]
p134<-m2*(1-m2)*B[i+1]
p144<-m3*(1-m3)*B[i+2]

p224<-m1*(1-m1)*A[i]
p234<-m2*(1-m2)*A[i+1]
p244<-m3*(1-m3)*A[i+2]

p324<-m1*(1-m1)*C[i]
p334<-m2*(1-m2)*C[i+1]
p344<-m3*(1-m3)*C[i+2]

#likelihood & score function
l01<-(y[i]-m1)
l02<-q2*(p21*p022+p23*p024)
l03<-q3*(p31*p032+p33*p034)
l04<-q4*(p41*p042+p43*p044)
sc0<-(l01+l02+l03+l04)

l11<-(y[i]-m1)*B[i]
l12<-q2*(p21*p122+p23*p124)
l13<-q3*(p31*p132+p33*p134)
l14<-q4*(p41*p142+p43*p144)
sc1<-(l11+l12+l13+l14)

l21<-(y[i]-m1)*A[i]
l22<-q2*(p21*p222+p23*p224)
l23<-q3*(p31*p232+p33*p234)
l24<-q4*(p41*p242+p43*p244)
sc2<-(l21+l22+l23+l24)

l31<-(y[i]-m1)*C[i]
l32<-q2*(p21*p322+p23*p324)
l33<-q3*(p31*p332+p33*p334)
l34<-q4*(p41*p342+p43*p344)
sc3<-(l31+l32+l33+l34)
#del(l)---del(p) #del(shi)--del(lamda)
q2<-q2
q3<-q3
q4<-q4
r<-v
#del(delta)--del(shi)
dk2<-(v*(m2-m1)^2+(m2+m1))/d2
dk3<-(v*(m3-m2)^2+(m3-m2))/d3
```

```
dk4<-(v*(m4-m3)^2+(m4-m3))/d4
#del(p)--del(shi)
p201<-(((v-1)^2)*dk2-(d2-1))*(1-y[i]/m1)/(2*((v-1)^2)*(1-m1))
p301<-(((v-1)^2)*dk3-(d3-1))*(1-y[i+1]/m2)/(2*((v-1)^2)*(1-m2))
p401<-(((v-1)^2)*dk4-(d4-1))*(1-y[i+2]/m3)/(2*((v-1)^2)*(1-m3))
#del(l)--del(lamda)
llamda1<-q2*p201*r
llamda2<-q3*p301*r
llamda3<-q4*p401*r
sc4<-(llamda1+llamda2+llamda3)

score<-score+c(sc0,sc1,sc2,sc3,sc4)
inf<-inf + score %*% t(score)

#cal del srq l by del pyt-1 srq
q22<-(y[i+1]*(2*p1-1)-p1*p1)/((p1^2)*(1-p1)^2)
q33<-(y[i+2]*(2*p2-1)-p2*p2)/((p2^2) *(1-p2)^2)
q44<-(y[i+3]*(2*p3-1)-p3*p3)/((p3^2)*(1-p3)^2)
ddl<-q22+q33+q44
#cal of del sq delta by del theta(t) sq
sdd22<-((v-1)/d2^2)*(d2*(v-1)-((m4-m1)-(m4-m1)+1)*dd22)
sdd32<-((v-1)/d3^2)*(d3*(v-1)-((m4-m4)-(m4-m4)+1)*dd32)
sdd42<- ((v-1)/d4^2)*(d4*(v-1)-((m4-m4)-(m4-m4)+1)*dd42)
sp2<-(1/A2)*(-(2*y[i]-1)*sdd22)
sp3<-(1/A3)*(-(2*y[i+1]-1)*sdd32)
sp4<-(1/A4)*(-(2*y[i+2]-1)*sdd42)
ddp<-sp2+sp3+sp4
#cal of del sq delta by del theta(t-1) sq
sdd21<-((v-1)/d2^2)*(d2*(v-1)-((m1-m4)+(m4-m1)+1)*dd21)
sdd31<-((v-1)/d3^2)*(d3*(v-1)-((m4-m4)+(m4-m4)+1)*dd31)
sdd41<- ((v-1)/d4^2)*(d4*(v-1)-((m4-m4)+(m4-m4)+1)*dd41)
#cal of dl sq pyt by del sq theta(t-1)
#cal of E
E2<-((2*y[i]-1)*(-dd21+m1)+m4)*A2-2*B2*(1-y[i]+2*y[i]*m4)
E3<-((2*y[i+1]-1)*(-dd31+m4)+m4)*A3-2*B3*(1-y[i+1]+2*y[i+1]*m4)
E4<-((2*y[i+2]-1)*(-dd41+m4)+m4)*A4-2*B4*(1-y[i+2]+2*y[i+2]*m4)
#cal of del A by del theta(t-1)
t2<-2*(v-1)*(2*y[i]-1)
t3<-2*(v-1)*(2*y[i+1]-1)
t4<-2*(v-1)*(2*y[i+2]-1)
#cal of delB by del theta(t-1)
z2<-(2*y[i]-1)*(-dd21-(v-1))
z3<-(2*y[i+1]-1)*(-dd31-(v-1))
z4<-(2*y[i+2]-1)*(-dd41-(v-1))
#cal of delE by del thete (t-1)
de2<-(2*y[i]-1)*(t2*(m1-dd21)+A2*(1-sdd21))+m4*t2-2*z2*(1-
y[i]+2*y[i]*m4)
de3<-(2*y[i+1]-1)*(t3*(m4-dd31)+A3*(1-sdd31))+m4*t3-2*z3*(1-
y[i+1]+2*y[i+1]*m4)
de4<-(2*y[i+2]-1)*(t4*(m4-dd41)+A4*(1-sdd41))+m4*t4-2*z4*(1-
y[i+2]+2*y[i+2]*m4)
sq2<-(1/A2^4)*((A2^2)*de2-E2*2*A2*t2)
sq3<-(1/A3^4)*((A3^2)*de3-E3*2*A3*t3)
sq4<-(1/A4^4)*((A4^2)*de4-E4*2*A4*t4)
ddp1<-sq2+sq3+sq4
#cal of del sq theta by del beta sq(t)
dr022<-((1/(1-m2))*p022*l-((m2^2)/1-m2)*l^2)/((1/1-m2)^2)
dr032<-((1/(1-m3))*p032*l-((m3^2)/1-m3)*l^2)/((1/1-m3)^2)
dr042<-((1/(1-m4))*p042*l-((m4^2)/1-m4)*l^2)/((1/1-m4)^2)
dr02<-dr022+dr032+dr042

dr122<-((1/(1-m2))*p222*A[i]-((m2^2)/1-m2)*A[i]^2)/((1/1-m2)^2)
dr132<-((1/(1-m3))*p232*A[i+1]-((m3^2)/1-m3)*A[i+1]^2)/((1/1-
m3)^2)
dr142<-((1/(1-m4))*p242*A[i+2]-((m4^2)/1-m4)*A[i+2]^2)/((1/1-
m4)^2)
dr12<-dr122+dr132+dr142

dr222<-((1/(1-m2))*p122*B[i]-((m2^2)/1-m2)*B[i]^2)/((1/1-m2)^2)
dr232<-((1/(1-m3))*p132*B[i+1]-((m3^2)/1-m3)*B[i+1]^2)/((1/1-
m3)^2)
dr242<-((1/(1-m4))*p142*B[i+2]-((m4^2)/1-m4)*B[i+2]^2)/((1/1-
m4)^2)
dr22<-dr222+dr232+dr242

dr322<-((1/(1-m2))*p322*C[i]-((m2^2)/1-m2)*C[i]^2)/((1/1-m2)^2)
dr332<-((1/(1-m3))*p332*C[i+1]-((m3^2)/1-m3)*C[i+1]^2)/((1/1-
m3)^2)
dr342<-((1/(1-m4))*p342*C[i+2]-((m4^2)/1-m4)*C[i+2]^2)/((1/1-
m4)^2)
dr32<-dr322+dr332+dr342
#cal of del sq theta by del beta sq(t-1)
dr024<-((1/(1-m1))*p024*l-((m1^2)/1-m1)*l^2)/((1/1-m1)^2)
dr034<-((1/(1-m2))*p034*l-((m2^2)/1-m2)*l^2)/((1/1-m2)^2)
dr044<-((1/(1-m3))*p044*l-((m3^2)/1-m3)*l^2)/((1/1-m3)^2)
dr04<-dr024+dr034+dr044

dr124<-((1/(1-m1))*p224*A[i]-((m1^2)/1-m1)*A[i]^2)/((1/1-m1)^2)
dr134<-((1/(1-m2))*p234*A[i+1]-((m2^2)/1-m2)*A[i+1]^2)/((1/1-
m2)^2)
dr144<-((1/(1-m3))*p244*A[i+2]-((m3^2)/1-m3)*A[i+2]^2)/((1/1-
m3)^2)
dr14<-dr124+dr134+dr144

dr224<-((1/(1-m1))*p124*B[i]-((m1^2)/1-m1)*B[i]^2)/((1/1-m1)^2)
dr234<-((1/(1-m2))*p134*B[i+1]-((m2^2)/1-m2)*B[i+1]^2)/((1/1-
m2)^2)
dr244<-((1/(1-m3))*p144*B[i+2]-((m3^2)/1-m3)*B[i+2]^2)/((1/1-
m3)^2)
dr24<-dr224+dr234+dr244

dr324<-((1/(1-m1))*p324*C[i]-((m1^2)/1-m1)*C[i]^2)/((1/1-m1)^2)
dr334<-((1/(1-m2))*p334*C[i+1]-((m2^2)/1-m2)*C[i+1]^2)/((1/1-
m2)^2)
dr344<-((1/(1-m3))*p344*C[i+2]-((m3^2)/1-m3)*C[i+2]^2)/((1/1-
m3)^2)
dr34<-dr324+dr134+dr344
dr1<-c(dr04,dr14,dr24,dr34)
dr<-c(dr02,dr12,dr22,dr32)
#del l by bel beta
sec1<-ddl*(ddp*dr+ddp1*dr1)

#calculation del sq del by del shi sq
dk202<-(1/d2^2)*(((m2-m1)^2)*(d2-v*dk2)-(m2+m1)*dk2)
dk302<-(1/d3^2)*(((m3-m2)^2)*(d3-v*dk3)-(m3+m2)*dk3)
dk402<-(1/d4^2)*(((m4-m3)^2)*(d4-v*dk4)-(m4+m3)*dk4)
#cal del sq pyt-1 by del shi sq
p202<-(((v-1)^2)*dk202-2*(v-1)*dk2+2*(d2-1))*(1-y[i]/m1)/(2*((v-
1)^3)*(1-m1))
p302<-(((v-1)^2)*dk302-2*(v-1)*dk3+2*(d3-1))*(1-y[i+1]/m2)/(2*((v-
1)^3)*(1-m2))
p402<-(((v-1)^2)*dk402-2*(v-1)*dk4+2*(d4-1))*(1-y[i+2]/m3)/(2*((v-
1)^3)*(1-m3))
f421<-q22*p202*r
f422<-q33*p302*r
f423<-q44*p402*r
sec2<-f421+f422+f423
sec0<-c(sec1,sec2)
sec<-sum(sec0)
sig<-(sec)
score1<-score+(1/b)
inf1<-inf + score1 %*% t(score1)
}
}
count<-count+1
print(count)
score<-score
infinv<-solve(inf)
score1<-score1
infinv1<-solve(inf1)
est<-initial+infinv %*% score
conv<-abs(initial-est)
if(conv[1]<=0.001 &&  conv[2]<=0.001 &&  conv[3]<=0.001 &&
conv[4]<=0.001 &&  conv[5]<=0.001)
        break
initial<-est
}
est<-initial+infinv %*% score
cat("The likelihood estimate is\n")
print(est)
est1<-initial+infinv1 %*% score1
cat("The another estimator  is\n")
print(est1)
#cal of theta
m1<-
exp(est[1]+est[2]*A[i]+est[3]*B[i]+est[4]*C[i])/(1+exp(est[1]+est[2]*A[i]+
est[3]*B[i]+est[4]*C[i]))
m2<-
exp(est[1]+est[2]*A[i+1]+est[3]*B[i+1]+est[4]*C[i+1])/(1+exp(est[1]+est[
2]*A[i+1]+est[3]*B[i+1]+est[4]*C[i+1]))
```

```
m3<-
exp(est[1]+est[2]*A[i+2]+est[3]*B[i+2]+est[4]*C[i+2])/(1+exp(est[1]+est[
2]*A[i+2]+est[3]*B[i+2]+est[4]*C[i+2]))
m4<-
exp(est[1]+est[2]*A[i+3]+est[3]*B[i+3]+est[4]*C[i+3])/(1+exp(est[1]+est[
2]*A[i+3]+est[3]*B[i+3]+est[4]*C[i+3]))
v<-exp(est[5])

#calculation of I(Beta)
Ib1<-log(m1/(1-m1))+log(1-m1)
Ib2<-log(m2/(1-m2))+log(1-m2)
Ib3<-log(m3/(1-m3))+log(1-m3)
Ib4<-log(m4/(1-m4))+log(1-m4)
Ib<-Ib1+ Ib2+ Ib3+ Ib4

#cal of delta
d2<-sqrt(1+(v-1)*((m2-m1)^2)*v-(m2-m1)^2+2*(m2+m1))
d3<-sqrt(1+(v-1)*((m3-m2)^2)*v-(m3-m2)^2+2*(m3+m2))
d4<-sqrt(1+(v-1)*((m4-m3)^2)*v-(m4-m3)^2+2*(m4+m3))
#cal A&B
A2<-2*(v-1)*(1-y[i]+(2*y[i]-1)*m1)
A3<-2*(v-1)*(1-y[i+1]+(2*y[i+1]-1)*m2)
A4<-2*(v-1)*(1-y[i+2]+(2*y[i+2]-1)*m3)

B2<-  (2*y[i]-1)*(1-d2+(v-1))*m1 +(v-1)*m2
B3<-(2*y[i+1]-1)*(1-d3+(v-1))*m2 +(v-1)*m3
B4<-(2*y[i+2]-1)*(1-d4+(v-1))*m3 +(v-1)*m4
#calculation of del(l) by del(p)
#calculation of del(l) by del(p)
#calcuation of pyt-1
p1<-((d2-1+(v-1)*(m2-m1))/(2*(v-1)*(1-m1)))+(y[i]*(1-d2+(v-
1)*(m2+m1-2*m1*m2)))/(2*(v-1)*m1*(1-m1))
p2<-((d3-1+(v-1)*(m3-m2))/(2*(v-1)*(1-m2)))+(y[i+1]*(1-d3+(v-
1)*(m3+m2-2*m2*m3)))/(2*(v-1)*m2*(1-m2))
p3<-((d4-1+(v-1)*(m4-m3))/(2*(v-1)*(1-m3)))+(y[i+2]*(1-d4+(v-
1)*(m4+m3-2*m3*m4)))/(2*(v-1)*m3*(1-m3))
q2<-(y[i+1]-p1)/(p1*(1-p1))
q3<-(y[i+2]-p2)/(p2*(1-p2))
q4<-(y[i+3]-p3)/(p3*(1-p3))
#del of delta by theta(t-1)
dd21<-(v-1)*(-v*(m2-m1)+(m2-m1)+1)/d2
dd31<-(v-1)*(-v*(m3-m2)+(m3-m2)+1)/d3
dd41<-(v-1)*(-v*(m4-m3)+(m4-m3)+1)/d4
#del of delta by theta(t)
dd22<-(v-1)*(v*(m2-m1)-(m2-m1)+1)/d2
dd32<-(v-1)*(v*(m3-m2)-(m3-m2)+1)/d3
dd42<-(v-1)*(v*(m4-m3)-(m4-m3)+1)/d4
#del(p) by del(theta(t))
p21<-(1/A2)*(-(2*y[i]-1)*dd22+v-1)
p31<-(1/A3)*(-(2*y[i+1]-1)*dd32+v-1)
p41<-(1/A4)*(-(2*y[i+2]-1)*dd42+v-1)
#del(p) by del(theta(t-1))
p23<-(((2*y[i]-1)*(-dd21+m1)+m2)*A2-2*B2*(1-y[i]+2*y[i]*m2))/(A2)^2
p33<-(((2*y[i+1]-1)*(-dd31+m2)+m3)*A3-2*B3*(1-
y[i+1]+2*y[i+1]*m3))/(A3)^2
p43<-(((2*y[i+2]-1)*(-dd41+m3)+m4)*A4-2*B4*(1-
y[i+2]+2*y[i+2]*m4))/(A4)^2
#del theta(t) by del beta
p022<-m2*(1-m2)*1
p032<-m3*(1-m3)*1
p042<-m4*(1-m4)*1

p122<-m2*(1-m2)*B[i+1]
p132<-m3*(1-m3)*B[i+2]
p142<-m4*(1-m4)*B[i+3]

p222<-m2*(1-m2)*A[i+1]
p232<-m3*(1-m3)*A[i+2]
p242<-m4*(1-m4)*A[i+3]

p322<-m2*(1-m2)*C[i+1]
p332<-m3*(1-m3)*C[i+2]
p342<-m4*(1-m4)*C[i+3]

#del theta(t-1) by del beta
p024<-m1*(1-m1)*1
p034<-m2*(1-m2)*1
p044<-m3*(1-m3)*1

p124<-m1*(1-m1)*B[i]
p134<-m2*(1-m2)*B[i+1]
p144<-m3*(1-m3)*B[i+2]

p224<-m1*(1-m1)*A[i]
p234<-m2*(1-m2)*A[i+1]
p244<-m3*(1-m3)*A[i+2]

p324<-m1*(1-m1)*C[i]
p334<-m2*(1-m2)*C[i+1]
p344<-m3*(1-m3)*C[i+2]

#likelihood & score function
l01<-(y[i]-m1)
l02<-q2*(p21*p022+p23*p024)
l03<-q3*(p31*p032+p33*p034)
l04<-q4*(p41*p042+p43*p044)
sc0<-(l01+l02+l03+l04)

l11<-(y[i]-m1)*B[i]
l12<-q2*(p21*p122+p23*p124)
l13<-q3*(p31*p132+p33*p134)
l14<-q4*(p41*p142+p43*p144)
sc1<-(l11+l12+l13+l14)

l21<-(y[i]-m1)*A[i]
l22<-q2*(p21*p222+p23*p224)
l23<-q3*(p31*p232+p33*p234)
l24<-q4*(p41*p242+p43*p244)
sc2<-(l21+l22+l23+l24)

l31<-(y[i]-m1)*C[i]
l32<-q2*(p21*p322+p23*p324)
l33<-q3*(p31*p332+p33*p334)
l34<-q4*(p41*p342+p43*p344)
sc3<-(l31+l32+l33+l34)
#del(l)---del(p) #del(shi)--del(lamda)
q2<-q2
q3<-q3
q4<-q4
r<-v
#del(delta)--del(shi)
dk2<-(v*(m2-m1)^2+(m2+m1))/d2
dk3<-(v*(m3-m2)^2+(m3-m2))/d3
dk4<-(v*(m4-m3)^2+(m4-m3))/d4
#del(p)--del(shi)
p201<-(((v-1)^2)*dk2-(d2-1))*(1-y[i]/m1)/(2*((v-1)^2)*(1-m1))
p301<-(((v-1)^2)*dk3-(d3-1))*(1-y[i+1]/m2)/(2*((v-1)^2)*(1-m2))
p401<-(((v-1)^2)*dk4-(d4-1))*(1-y[i+2]/m3)/(2*((v-1)^2)*(1-m3))
#del(l)--del(lamda)
llamda1<-q2*p201*r
llamda2<-q3*p301*r
llamda3<-q4*p401*r
sc4<-(llamda1+llamda2+llamda3)

score<-score+c(sc0,sc1,sc2,sc3,sc4)
inf<-inf + score %*% t(score)

#cal del srq l by del pyt-1 srq
q22<-(y[i+1]*(2*p1-1)-p1*p1)/((p1^2)*(1-p1)^2)
q33<-(y[i+2]*(2*p2-1)-p2*p2)/((p2^2) *(1-p2)^2)
q44<-(y[i+3]*(2*p3-1)-p3*p3)/((p3^2)*(1-p3)^2)
ddl<-q22+q33+q44
#cal of del sq delta by del theta(t) sq
sdd22<-((v-1)/d2^2)*(d2*(v-1)-((m4-m1)-(m4-m1)+1)*dd22)
sdd32<-((v-1)/d3^2)*(d3*(v-1)-((m4-m4)-(m4-m4)+1)*dd32)
sdd42<- ((v-1)/d4^2)*(d4*(v-1)-((m4-m4)-(m4-m4)+1)*dd42)
sp2<-(1/A2)*(-(2*y[i]-1)*sdd22)
sp3<-(1/A3)*(-(2*y[i+1]-1)*sdd32)
sp4<-(1/A4)*(-(2*y[i+2]-1)*sdd42)
ddp<-sp2+sp3+sp4
#cal of del sq delta by del theta(t-1) sq
sdd21<-((v-1)/d2^2)*(d2*(v-1)-((m1-m4)+(m4-m1)+1)*dd21)
sdd31<-((v-1)/d3^2)*(d3*(v-1)-((m4-m4)+(m4-m4)+1)*dd31)
sdd41<- ((v-1)/d4^2)*(d4*(v-1)-((m4-m4)+(m4-m4)+1)*dd41)
#cal of dl sq pyt by del sq theta(t-1)
#cal of E
E2<-((2*y[i]-1)*(-dd21+m1)+m4)*A2-2*B2*(1-y[i]+2*y[i]*m4)
E3<-((2*y[i+1]-1)*(-dd31+m4)+m4)*A3-2*B3*(1-y[i+1]+2*y[i+1]*m4)
E4<-((2*y[i+2]-1)*(-dd41+m4)+m4)*A4-2*B4*(1-y[i+2]+2*y[i+2]*m4)
#cal of del A by del theta(t-1)
```

```
t2<-2*(v-1)*(2*y[i]-1)
t3<-2*(v-1)*(2*y[i+1]-1)
t4<-2*(v-1)*(2*y[i+2]-1)
#cal of delB by del theta(t-1)
z2<-(2*y[i]-1)*(-dd21-(v-1))
z3<-(2*y[i+1]-1)*(-dd31-(v-1))
z4<-(2*y[i+2]-1)*(-dd41-(v-1))
#cal of delE by del thete (t-1)
de2<-(2*y[i]-1)*(t2*(m1-dd21)+A2*(1-sdd21))+m4*t2-2*z2*(1-
y[i]+2*y[i]*m4)
de3<-(2*y[i+1]-1)*(t3*(m4-dd31)+A3*(1-sdd31))+m4*t3-2*z3*(1-
y[i+1]+2*y[i+1]*m4)
de4<-(2*y[i+2]-1)*(t4*(m4-dd41)+A4*(1-sdd41))+m4*t4-2*z4*(1-
y[i+2]+2*y[i+2]*m4)
sq2<-(1/A2^4)*((A2^2)*de2-E2*2*A2*t2)
sq3<-(1/A3^4)*((A3^2)*de3-E3*2*A3*t3)
sq4<-(1/A4^4)*((A4^2)*de4-E4*2*A4*t4)
ddp1<-sq2+sq3+sq4
#cal of del sq theta by del beta sq(t)
dr022<-((1/(1-m2))*p022*l-((m2^2)/1-m2)*l^2)/((1/1-m2)^2)
dr032<-((1/(1-m3))*p032*l-((m3^2)/1-m3)*l^2)/((1/1-m3)^2)
dr042<-((1/(1-m4))*p042*l-((m4^2)/1-m4)*l^2)/((1/1-m4)^2)
dr02<-dr022+dr032+dr042

dr122<-((1/(1-m2))*p222*A[i]-((m2^2)/1-m2)*A[i]^2)/((1/1-m2)^2)
dr132<-((1/(1-m3))*p232*A[i+1]-((m3^2)/1-m3)*A[i+1]^2)/((1/1-
m3)^2)
dr142<-((1/(1-m4))*p242*A[i+2]-((m4^2)/1-m4)*A[i+2]^2)/((1/1-
m4)^2)
dr12<-dr122+dr132+dr142

dr222<-((1/(1-m2))*p122*B[i]-((m2^2)/1-m2)*B[i]^2)/((1/1-m2)^2)
dr232<-((1/(1-m3))*p132*B[i+1]-((m3^2)/1-m3)*B[i+1]^2)/((1/1-
m3)^2)
dr242<-((1/(1-m4))*p142*B[i+2]-((m4^2)/1-m4)*B[i+2]^2)/((1/1-
m4)^2)
dr22<-dr222+dr232+dr242

dr322<-((1/(1-m2))*p322*C[i]-((m2^2)/1-m2)*C[i]^2)/((1/1-m2)^2)
dr332<-((1/(1-m3))*p332*C[i+1]-((m3^2)/1-m3)*C[i+1]^2)/((1/1-
m3)^2)
dr342<-((1/(1-m4))*p342*C[i+2]-((m4^2)/1-m4)*C[i+2]^2)/((1/1-
m4)^2)
dr32<-dr322+dr332+dr342
#cal of del sq theta by del beta sq(t-1)
dr024<-((1/(1-m1))*p024*l-((m1^2)/1-m1)*l^2)/((1/1-m1)^2)
dr034<-((1/(1-m2))*p034*l-((m2^2)/1-m2)*l^2)/((1/1-m2)^2)
dr044<-((1/(1-m3))*p044*l-((m3^2)/1-m3)*l^2)/((1/1-m3)^2)
dr04<-dr024+dr034+dr044

dr124<-((1/(1-m1))*p224*A[i]-((m1^2)/1-m1)*A[i]^2)/((1/1-m1)^2)
dr134<-((1/(1-m2))*p234*A[i+1]-((m2^2)/1-m2)*A[i+1]^2)/((1/1-
m2)^2)
dr144<-((1/(1-m3))*p244*A[i+2]-((m3^2)/1-m3)*A[i+2]^2)/((1/1-
m3)^2)
dr14<-dr124+dr134+dr144

dr224<-((1/(1-m1))*p124*B[i]-((m1^2)/1-m1)*B[i]^2)/((1/1-m1)^2)
dr234<-((1/(1-m2))*p134*B[i+1]-((m2^2)/1-m2)*B[i+1]^2)/((1/1-
m2)^2)
dr244<-((1/(1-m3))*p144*B[i+2]-((m3^2)/1-m3)*B[i+2]^2)/((1/1-
m3)^2)
dr24<-dr224+dr234+dr244

dr324<-((1/(1-m1))*p324*C[i]-((m1^2)/1-m1)*C[i]^2)/((1/1-m1)^2)
dr334<-((1/(1-m2))*p334*C[i+1]-((m2^2)/1-m2)*C[i+1]^2)/((1/1-
m2)^2)
dr344<-((1/(1-m3))*p344*C[i+2]-((m3^2)/1-m3)*C[i+2]^2)/((1/1-
m3)^2)
dr34<-dr324+dr134+dr344
dr1<-c(dr04,dr14,dr24,dr34)
dr<-c(dr02,dr12,dr22,dr32)
#del l by bel beta
sec1<-ddl*(ddp*dr+ddp1*dr1)

#calculation del sq del by del shi sq
dk202<-(1/d2^2)*(((m2-m1)^2)*(d2-v*dk2)-(m2+m1)*dk2)
dk302<-(1/d3^2)*(((m3-m2)^2)*(d3-v*dk3)-(m3+m2)*dk3)
dk402<-(1/d4^2)*(((m4-m3)^2)*(d4-v*dk4)-(m4+m3)*dk4)
#cal del sq pyt-1 by del shi sq
p202<-(((v-1)^2)*dk202-2*(v-1)*dk2+2*(d2-1))*(1-y[i]/m1)/(2*((v-
1)^3)*(1-m1))
p302<-(((v-1)^2)*dk302-2*(v-1)*dk3+2*(d3-1))*(1-y[i+1]/m2)/(2*((v-
1)^3)*(1-m2))
p402<-(((v-1)^2)*dk402-2*(v-1)*dk4+2*(d4-1))*(1-y[i+2]/m3)/(2*((v-
1)^3)*(1-m3))
f421<-q22*p202*r
f422<-q33*p302*r
f423<-q44*p402*r
sec2<-f421+f422+f423
sec0<-c(sec1,sec2)
sec<-sum(sec0)
den<-(sec)

#cal of theta
m1<-
exp(est1[1]+est1[2]*A[i]+est1[3]*B[i]+est1[4]*C[i])/(1+exp(est1[1]+est1[
2]*A[i]+est1[3]*B[i]+est1[4]*C[i]))
m2<-
exp(est1[1]+est1[2]*A[i+1]+est1[3]*B[i+1]+est1[4]*C[i+1])/(1+exp(est1[
1]+est1[2]*A[i+1]+est1[3]*B[i+1]+est1[4]*C[i+1]))
m3<-
exp(est1[1]+est1[2]*A[i+2]+est1[3]*B[i+2]+est1[4]*C[i+2])/(1+exp(est1[
1]+est1[2]*A[i+2]+est1[3]*B[i+2]+est1[4]*C[i+2]))
m4<-
exp(est1[1]+est1[2]*A[i+3]+est1[3]*B[i+3]+est1[4]*C[i+3])/(1+exp(est1[
1]+est1[2]*A[i+3]+est1[3]*B[i+3]+est1[4]*C[i+3]))
v<-exp(est1[5])

#calculation of I(Beta*)
Ib11<-log(m1/(1-m1))+log(1-m1)
Ib12<-log(m2/(1-m2))+log(1-m2)
Ib13<-log(m3/(1-m3))+log(1-m3)
Ib14<-log(m4/(1-m4))+log(1-m4)
Ib1<-Ib11+ Ib12+ Ib13+ Ib14

#cal of delta
d2<-sqrt(1+(v-1)*((m2-m1)^2)*v-(m2-m1)^2+2*(m2+m1))
d3<-sqrt(1+(v-1)*((m3-m2)^2)*v-(m3-m2)^2+2*(m3+m2))
d4<-sqrt(1+(v-1)*((m4-m3)^2)*v-(m4-m3)^2+2*(m4+m3))
#cal A&B
A2<-2*(v-1)*(1-y[i]+(2*y[i]-1)*m1)
A3<-2*(v-1)*(1-y[i+1]+(2*y[i+1]-1)*m2)
A4<-2*(v-1)*(1-y[i+2]+(2*y[i+2]-1)*m3)

B2<-  (2*y[i]-1)*(1-d2+(v-1))*m1 +(v-1)*m2
B3<-(2*y[i+1]-1)*(1-d3+(v-1))*m2 +(v-1)*m3
B4<-(2*y[i+2]-1)*(1-d4+(v-1))*m3 +(v-1)*m4
#calculation of del(l) by del(p)
#calculation of del(l) by del(p)
#calcuation of pyt-1
p1<-((d2-1+(v-1)*(m2-m1))/(2*(v-1)*(1-m1)))+(y[i]*(1-d2+(v-
1)*(m2+m1-2*m1*m2)))/(2*(v-1)*m1*(1-m1))
p2<-((d3-1+(v-1)*(m3-m2))/(2*(v-1)*(1-m2)))+(y[i+1]*(1-d3+(v-
1)*(m3+m2-2*m2*m3)))/(2*(v-1)*m2*(1-m2))
p3<-((d4-1+(v-1)*(m4-m3))/(2*(v-1)*(1-m3)))+(y[i+2]*(1-d4+(v-
1)*(m4+m3-2*m3*m4)))/(2*(v-1)*m3*(1-m3))
q2<-(y[i+1]-p1)/(p1*(1-p1))
q3<-(y[i+2]-p2)/(p2*(1-p2))
q4<-(y[i+3]-p3)/(p3*(1-p3))
#del of delta by theta(t-1)
dd21<-(v-1)*(-v*(m2-m1)+(m2-m1)+1)/d2
dd31<-(v-1)*(-v*(m3-m2)+(m3-m2)+1)/d3
dd41<-(v-1)*(-v*(m4-m3)+(m4-m3)+1)/d4
#del of delta by theta(t)
dd22<-(v-1)*(v*(m2-m1)-(m2-m1)+1)/d2
dd32<-(v-1)*(v*(m3-m2)-(m3-m2)+1)/d3
dd42<-(v-1)*(v*(m4-m3)-(m4-m3)+1)/d4
#del(p) by del(theta(t))
p21<-(1/A2)*(-(2*y[i]-1)*dd22+v-1)
p31<-(1/A3)*(-(2*y[i+1]-1)*dd32+v-1)
p41<-(1/A4)*(-(2*y[i+2]-1)*dd42+v-1)
#del(p) by del(theta(t-1))
p23<-(((2*y[i]-1)*(-dd21+m1)+m2)*A2-2*B2*(1-y[i]+2*y[i]*m2))/(A2)^2
p33<-(((2*y[i+1]-1)*(-dd31+m2)+m3)*A3-2*B3*(1-
y[i+1]+2*y[i+1]*m3))/(A3)^2
p43<-(((2*y[i+2]-1)*(-dd41+m3)+m4)*A4-2*B4*(1-
y[i+2]+2*y[i+2]*m4))/(A4)^2
#del theta(t) by del beta
p022<-m2*(1-m2)*1
```

```
p032<-m3*(1-m3)*1
p042<-m4*(1-m4)*1

p122<-m2*(1-m2)*B[i+1]
p132<-m3*(1-m3)*B[i+2]
p142<-m4*(1-m4)*B[i+3]

p222<-m2*(1-m2)*A[i+1]
p232<-m3*(1-m3)*A[i+2]
p242<-m4*(1-m4)*A[i+3]

p322<-m2*(1-m2)*C[i+1]
p332<-m3*(1-m3)*C[i+2]
p342<-m4*(1-m4)*C[i+3]

#del theta(t-1) by del beta
p024<-m1*(1-m1)*1
p034<-m2*(1-m2)*1
p044<-m3*(1-m3)*1

p124<-m1*(1-m1)*B[i]
p134<-m2*(1-m2)*B[i+1]
p144<-m3*(1-m3)*B[i+2]

p224<-m1*(1-m1)*A[i]
p234<-m2*(1-m2)*A[i+1]
p244<-m3*(1-m3)*A[i+2]

p324<-m1*(1-m1)*C[i]
p334<-m2*(1-m2)*C[i+1]
p344<-m3*(1-m3)*C[i+2]

#likelihood & score function
l01<-(y[i]-m1)
l02<-q2*(p21*p022+p23*p024)
l03<-q3*(p31*p032+p33*p034)
l04<-q4*(p41*p042+p43*p044)
sc0<-(l01+l02+l03+l04)

l11<-(y[i]-m1)*B[i]
l12<-q2*(p21*p122+p23*p124)
l13<-q3*(p31*p132+p33*p134)
l14<-q4*(p41*p142+p43*p144)
sc1<-(l11+l12+l13+l14)

l21<-(y[i]-m1)*A[i]
l22<-q2*(p21*p222+p23*p224)
l23<-q3*(p31*p232+p33*p234)
l24<-q4*(p41*p242+p43*p244)
sc2<-(l21+l22+l23+l24)

l31<-(y[i]-m1)*C[i]
l32<-q2*(p21*p322+p23*p324)
l33<-q3*(p31*p332+p33*p334)
l34<-q4*(p41*p342+p43*p344)
sc3<-(l31+l32+l33+l34)
#del(l)---del(p) #del(shi)--del(lamda)
q2<-q2
q3<-q3
q4<-q4
r<-v
#del(delta)--del(shi)
dk2<-(v*(m2-m1)^2+(m2+m1))/d2
dk3<-(v*(m3-m2)^2+(m3-m2))/d3
dk4<-(v*(m4-m3)^2+(m4-m3))/d4
#del(p)--del(shi)
p201<-(((v-1)^2)*dk2-(d2-1))*(1-y[i]/m1)/(2*((v-1)^2)*(1-m1))
p301<-(((v-1)^2)*dk3-(d3-1))*(1-y[i+1]/m2)/(2*((v-1)^2)*(1-m2))
p401<-(((v-1)^2)*dk4-(d4-1))*(1-y[i+2]/m3)/(2*((v-1)^2)*(1-m3))
#del(l)--del(lamda)
llamda1<-q2*p201*r
llamda2<-q3*p301*r
llamda3<-q4*p401*r
sc4<-(llamda1+llamda2+llamda3)

score<-score+c(sc0,sc1,sc2,sc3,sc4)
inf<-inf + score %*% t(score)

#cal del srq l by del pyt-1 srq
q22<-(y[i+1]*(2*p1-1)-p1*p1)/((p1^2)*(1-p1)^2)
q33<-(y[i+2]*(2*p2-1)-p2*p2)/((p2^2) *(1-p2)^2)
q44<-(y[i+3]*(2*p3-1)-p3*p3)/((p3^2)*(1-p3)^2)
ddl<-q22+q33+q44
#cal of del sq delta by del theta(t) sq
sdd22<-((v-1)/d2^2)*(d2*(v-1)-((m4-m1)-(m4-m1)+1)*dd22)
sdd32<-((v-1)/d3^2)*(d3*(v-1)-((m4-m4)-(m4-m4)+1)*dd32)
sdd42<- ((v-1)/d4^2)*(d4*(v-1)-((m4-m4)-(m4-m4)+1)*dd42)
sp2<-(1/A2)*(-(2*y[i]-1)*sdd22)
sp3<-(1/A3)*(-(2*y[i+1]-1)*sdd32)
sp4<-(1/A4)*(-(2*y[i+2]-1)*sdd42)
ddp<-sp2+sp3+sp4
#cal of del sq delta by del theta(t-1) sq
sdd21<-((v-1)/d2^2)*(d2*(v-1)-((m1-m4)+(m4-m1)+1)*dd21)
sdd31<-((v-1)/d3^2)*(d3*(v-1)-((m4-m4)+(m4-m4)+1)*dd31)
sdd41<- ((v-1)/d4^2)*(d4*(v-1)-((m4-m4)+(m4-m4)+1)*dd41)
#cal of dl sq pyt by del sq theta(t-1)
#cal of E
E2<-((2*y[i]-1)*(-dd21+m1)+m4)*A2-2*B2*(1-y[i]+2*y[i]*m4)
E3<-((2*y[i+1]-1)*(-dd31+m4)+m4)*A3-2*B3*(1-y[i+1]+2*y[i+1]*m4)
E4<-((2*y[i+2]-1)*(-dd41+m4)+m4)*A4-2*B4*(1-y[i+2]+2*y[i+2]*m4)
#cal of del A by del theta(t-1)
t2<-2*(v-1)*(2*y[i]-1)
t3<-2*(v-1)*(2*y[i+1]-1)
t4<-2*(v-1)*(2*y[i+2]-1)
#cal of delB by del theta(t-1)
z2<-(2*y[i]-1)*(-dd21-(v-1))
z3<-(2*y[i+1]-1)*(-dd31-(v-1))
z4<-(2*y[i+2]-1)*(-dd41-(v-1))
#cal of delE by del thete (t-1)
de2<-(2*y[i]-1)*(t2*(m1-dd21)+A2*(1-sdd21))+m4*t2-2*z2*(1-
y[i]+2*y[i]*m4)
de3<-(2*y[i+1]-1)*(t3*(m4-dd31)+A3*(1-sdd31))+m4*t3-2*z3*(1-
y[i+1]+2*y[i+1]*m4)
de4<-(2*y[i+2]-1)*(t4*(m4-dd41)+A4*(1-sdd41))+m4*t4-2*z4*(1-
y[i+2]+2*y[i+2]*m4)
sq2<-(1/A2^4)*((A2^2)*de2-E2*2*A2*t2)
sq3<-(1/A3^4)*((A3^2)*de3-E3*2*A3*t3)
sq4<-(1/A4^4)*((A4^2)*de4-E4*2*A4*t4)
ddp1<-sq2+sq3+sq4
#cal of del sq theta by del beta sq(t)
dr022<-((1/(1-m2))*p022*l-((m2^2)/1-m2)*l^2)/((1/1-m2)^2)
dr032<-((1/(1-m3))*p032*l-((m3^2)/1-m3)*l^2)/((1/1-m3)^2)
dr042<-((1/(1-m4))*p042*l-((m4^2)/1-m4)*l^2)/((1/1-m4)^2)
dr02<-dr022+dr032+dr042

dr122<-((1/(1-m2))*p222*A[i]-((m2^2)/1-m2)*A[i]^2)/((1/1-m2)^2)
dr132<-((1/(1-m3))*p232*A[i+1]-((m3^2)/1-m3)*A[i+1]^2)/((1/1-
m3)^2)
dr142<-((1/(1-m4))*p242*A[i+2]-((m4^2)/1-m4)*A[i+2]^2)/((1/1-
m4)^2)
dr12<-dr122+dr132+dr142

dr222<-((1/(1-m2))*p122*B[i]-((m2^2)/1-m2)*B[i]^2)/((1/1-m2)^2)
dr232<-((1/(1-m3))*p132*B[i+1]-((m3^2)/1-m3)*B[i+1]^2)/((1/1-
m3)^2)
dr242<-((1/(1-m4))*p142*B[i+2]-((m4^2)/1-m4)*B[i+2]^2)/((1/1-
m4)^2)
dr22<-dr222+dr232+dr242
dr322<-((1/(1-m2))*p322*C[i]-((m2^2)/1-m2)*C[i]^2)/((1/1-m2)^2)
dr332<-((1/(1-m3))*p332*C[i+1]-((m3^2)/1-m3)*C[i+1]^2)/((1/1-
m3)^2)
dr342<-((1/(1-m4))*p342*C[i+2]-((m4^2)/1-m4)*C[i+2]^2)/((1/1-
m4)^2)
dr32<-dr322+dr332+dr342
#cal of del sq theta by del beta sq(t-1)
dr024<-((1/(1-m1))*p024*l-((m1^2)/1-m1)*l^2)/((1/1-m1)^2)
dr034<-((1/(1-m2))*p034*l-((m2^2)/1-m2)*l^2)/((1/1-m2)^2)
dr044<-((1/(1-m3))*p044*l-((m3^2)/1-m3)*l^2)/((1/1-m3)^2)
dr04<-dr024+dr034+dr044
dr124<-((1/(1-m1))*p224*A[i]-((m1^2)/1-m1)*A[i]^2)/((1/1-m1)^2)
dr134<-((1/(1-m2))*p234*A[i+1]-((m2^2)/1-m2)*A[i+1]^2)/((1/1-
m2)^2)
dr144<-((1/(1-m3))*p244*A[i+2]-((m3^2)/1-m3)*A[i+2]^2)/((1/1-
m3)^2)
dr14<-dr124+dr134+dr144
dr224<-((1/(1-m1))*p124*B[i]-((m1^2)/1-m1)*B[i]^2)/((1/1-m1)^2)
dr234<-((1/(1-m2))*p134*B[i+1]-((m2^2)/1-m2)*B[i+1]^2)/((1/1-
m2)^2)
dr244<-((1/(1-m3))*p144*B[i+2]-((m3^2)/1-m3)*B[i+2]^2)/((1/1-
m3)^2)
```

```
dr24<-dr224+dr234+dr244
dr324<-((1/(1-m1))*p324*C[i]-((m1^2)/1-m1)*C[i]^2)/((1/1-m1)^2)
dr334<-((1/(1-m2))*p334*C[i+1]-((m2^2)/1-m2)*C[i+1]^2)/((1/1-
m2)^2)
dr344<-((1/(1-m3))*p344*C[i+2]-((m3^2)/1-m3)*C[i+2]^2)/((1/1-
m3)^2)
dr34<-dr324+dr134+dr344
dr1<-c(dr04,dr14,dr24,dr34)
dr<-c(dr02,dr12,dr22,dr32)
#del l by bel beta
sec1<-ddl*(ddp*dr+ddp1*dr1)
#calculation del sq del by del shi sq
dk202<-(1/d2^2)*(((m2-m1)^2)*(d2-v*dk2)-(m2+m1)*dk2)
dk302<-(1/d3^2)*(((m3-m2)^2)*(d3-v*dk3)-(m3+m2)*dk3)
dk402<-(1/d4^2)*(((m4-m3)^2)*(d4-v*dk4)-(m4+m3)*dk4)
#cal del sq pyt-1 by del shi sq
p202<-(((v-1)^2)*dk202-2*(v-1)*dk2+2*(d2-1))*(1-y[i]/m1)/(2*((v-
1)^3)*(1-m1))
p302<-(((v-1)^2)*dk302-2*(v-1)*dk3+2*(d3-1))*(1-y[i+1]/m2)/(2*((v-
1)^3)*(1-m2))
p402<-(((v-1)^2)*dk402-2*(v-1)*dk4+2*(d4-1))*(1-y[i+2]/m3)/(2*((v-
1)^3)*(1-m3))
f421<-q22*p202*r
f422<-q33*p302*r
f423<-q44*p402*r
sec2<-f421+f422+f423
sec0<-c(sec1,sec2)
sec<-sum(sec0)
neo<-(sec)
term<-exp(Ib1-Ib)
neo1<-neo+ (1/(t(est1)%*%est1))
bse=est1%*%((neo1/den)^(-1/2))*term
cat("The Bayes estimator  is\n")
print(bse)
#calculation of posterior risk
neo2<- neo+ (2/(t(est1)%*%est1))
bse2=est1*est1%*%((neo2/den)^(-1/2))*term
rbse<-(-bse^2)+bse2
cat("The Risk of squared  error is\n")
print(rbse)
bbl<-est1+(1/(2*c))*log(neo/den)-(1/c)*term
cat("The Bayes estimate under LINEX  is\n")
print(bbl)
#calculation of posterior risk
xtra<-exp(-c*est1)*((neo/den)^(-.5))*exp(term)
rbbl<-exp(c*bbl)*xtra-c*bbl+c*bse-1
cat("The Risk of LINEX is\n")
print(abs(rbbl))
bbml<-est1%*%(((((-c/t(est1)%*%est1)+neo)/den)^(-1/2))*term)^(-1/c)
cat("The Bayes estimate for MLINEX  is\n")
print(bbml)
#calculation of posterior risk
aest1<-abs(est1)
ghj<-(est1+log(aest1))/(est1*est1*log(aest1)*log(aest1))
xtra2<-log(aest1)*(((ghj+neo)/den)^(-.5))*term
xtra1<- (est1^-c)%*%(((((-c/(t(est1)%*%est1))+neo)/den)^(-1/2))*term)
rbbml<-bbml*xtra1-c*log(abs(bbml))+c*xtra2-1
cat("The Risk of MLINEX is\n")
print(abs(rbbml))
        }
bin1(data, initial)
```