Bayesian Conjugate Analysis for Transition Probabilities of Non-Homogeneous Markov Chain: A Survey

Minje Sung

“School of Business, Ajou University, Korea

Abstract

The present study surveys Bayesian modeling structure for inferences about transition probabilities of Markov chain. The motivation of the study came from the data that shows transitional behaviors of emotionally disturbed children undergoing residential treatment program. Dirichlet distribution was used as prior for the multinomial distribution. The analysis with real data was implemented in WinBUGS programming environment. The performance of the model was compared to that of alternative approaches.

Keywords: Bayesian approach, transition probability, Markov chain.

1. Introduction

We will use the Markov chain in discrete time to analyze the transient states of emotionally disturbed children and adolescents undergoing residential treatment. By understanding the nature of the transitions of the states, we will be able to predict and exercise some control over the future states of the patients. Since the patients’ states change over time with uncertainty, the future states are not deterministic, and we can handle the uncertainty of the states by assessing probabilities to the process. The statistical inferences about the patients’ transitional behavior are made under the Bayesian framework where the probability statements about the parameter are conditional on the observed data. The patients’ transitional states are described by a Markov Chain where the current patients’ states only depend on the immediate past not the whole history. At each time point, the patients’ states have multinomial distribution. Using Dirichlet as a conjugate prior, the posterior distributions are computed. The posterior distributions are represented by simulations using Markov Chain Monte Carlo (MCMC) method.

The early efforts to make inferences on the transition probabilities of a Markov chain was made in Anderson and Goodman (1957), and Billingsley (1961) where the inferences are made using maximum likelihood methods. Duncan and Lin (1972) considered models for a Markov chain having stochastic entry and exit. In other words, the authors relaxed the assumption that the number of observed individuals over time is constant and allowed for changes as a result of units entering or leaving the system. Such models, especially those allowing stochastic exits, are of interest in the analysis of data from psychiatric treatment programs where patients may leave the treatment program for different reasons. However they were non Bayesian approaches. A Bayesian analysis of the Markov chains using Dirichlet conjugate prior distribution is presented in Lee et al. (1970) but no explicit model is considered to describe time dependent transition probabilities over time. Meshkani and Billard (1992) developed empirical Bayes estimates of transition probabilities for homogeneous chains and extended
these to nonhomogeneous Markov chains by viewing the problem as a parametric Bayes problem in the sense of Morris (1983). However, the empirical Bayes approach is not a Bayesian approach. The literature does not show much of Bayesian studies for Markov chains. Sung et al. (2007) developed Bayesian models for non-homogeneous Markov chains by using logistic type of models with covariates. However, Markov chain models sometimes do not have covariate information of subjects, and just show aggregate transition patterns. In the present study, we propose a Bayesian approach for Markov chains without covariate information.

In Bayesian paradigm, inferences about unknown parameters in statistical models are made using posterior samples mostly drawn from Monte Carlo methods. For the choice of certain class of prior distributions, the posterior distribution follows the same parametric form as the prior distribution. This property is referred to as conjugacy and the prior distribution is then called a conjugate prior distribution for the particular likelihood. A formal treatment of conjugate Bayesian analysis is given in Bernardo and Smith (1994, Section 5.2).

While not many but some rigorous and novel Bayesian approaches for nonhomogeneous Markov chains were proposed in literature (for example, Sung et al., 2007), the Bayesian analysis in the present study adds a practical contribution to the literature in that the analysis of present study is straightforward but comprehensive conjugate Bayesian way, and practically the present model is much more accessible without having the barrier of having to understand the complexity of models for implementation. The performance of present model was compared to the alternative approach in the literature. Our primary concern in this study is to review a straightforward Bayesian conjugate implementation. The performance of present model was compared to the alternative approach in real data, and the model performance is compared to alternative approaches in the literature. A comprehensive description of Bayesian conjugate analysis is discussed in Section 2. Section 3 has a demonstration of model implementation using real data, and the model performance is compared to alternative approaches in the literature. Section 4 summarizes the findings of present study.

2. The Markov Probability Model

The type of stochastic process with which we will be concerned has a finite number of possible outcomes $S_j$ ($j = 1, \ldots, J$) where a discrete random variable $X_t$ ($t = 0, 1, \ldots, T$) takes at a finite number of equidistant time points. The probability distribution, $P(X_0 = S_j(0), X_1 = S_j(1), \ldots, X_T = S_j(T))$, for an individual would be $\prod_{t=0}^{T} P(X_t = S_j(t))$ if the states were independent in each stage. If we assume first order dependence where the probability distribution of a state of a given stage only on the state of the immediately preceding stage, then the process has the Markov property, and the Markov chain can be written as

$$P(X_0 = S_i, X_1 = S_j, \ldots, X_T = S_l) = P(X_0 = S_i)P(X_1 = S_j|X_0 = S_i) \cdots P(X_T = S_l|X_{T-1} = S_k).$$

(2.1)

For the nonhomogeneous Markov Chains, it is assumed that if $X_{t-1} = S_i$ and $X_t = S_j$, then $P(X_t = S_j|X_{t-1} = S_i) = p_{ij}(t)$. And, $N_{ij}$ denotes the number of individuals such that $X_{t-1} = S_i$ and $X_t = S_j$. The transition probability matrix $(I \times J)$ has the following properties.

$$0 \leq p_{ij}(t) \leq 1 \quad \text{and} \quad \sum_j p_{ij}(t) = 1, \quad \text{for} \quad i = 1, \ldots, I, \ j = 1, \ldots, J, \ t = 1, \ldots, T.$$

The number of subjects who made transitions from $S_i$ at time $t - 1$ to $S_j$ at time $t$ is denoted by the random variables, $N_{ij}(t)$, and the random vector $\hat{N}_j(t) = (N_{j1}(t), \ldots, N_{jJ}(t))$ has a multinomial