# Bayesian Analysis of Nonhomogeneous Markov Chains Having Random Exit

Minje Sung
*The George Washington University, Dept. of Management Science*
*Washington, DC 20052, USA*
*sungmj@gwu.edu*

Refik Soyer
*The George Washington University, Dept. of Management Science*
*Washington, DC 20052, USA*
*soyer@gwu.edu*

## 1. INTRODUCTION

We present a hierarchical Bayesian approach for analysis of time nonhomogeneous Markov chains having random exit. The Markov chain is used for describing random exit behavior of emotionally disturbed children in a treatment program. We consider models where the transition probabilities are functions of covariates as well as time. Here, the categorical measurements of interest describe the functioning level of the patients in the program at different points in time. Our primary interest lies in assessing the effectiveness of the treatment program and the random exit behavior of the patients at different points in time.

Bayesian approaches for analyzing Markov chains have been considered by few authors. Lee, Judge and Zellner (1968) presented conjugate Bayesian methods to estimate transition probabilities of a homogeneous chain. Recently, Erkanli, Soyer and Angold (2001) considered a binary Markov chain using a Markov logistic regression setup and presented a Bayesian analysis of the model. But the previous approaches did not provide a formal treatment of nonhomogeneous Markov chains. In this paper, we present a formal treatment of nonhomogeneous Markov chains by introducing models to describe the time evolution of transition probabilities. In so doing, we take the logistic regression representation of a Markov chain, and assume that the parameters evolve over time according to a first-order Markov process. To describe the random exit behavior, we develop a Markov chain model which has a single absorbing state representing exits as suggested by Duncan and Lin (1972).

## 2. MARKOV MODEL FOR NONHOMOGENEOUS MARKOV CHAINS

We consider a first order Markov process on the parameters of the logistic regression model to describe the transition probabilities of the nonhomogeneous Markov chain having random exit. The model for nonhomogeneous Markov chains using the logistic regression setup is an extension of the Bayesian models of Erkanli et al. (2001). The categorical variable has $J$ states where the $J$th state represents the exit state.

Let's define $x_{mijt}$ as a binary variable representing the transition of the $m$th individual from state $i$ at time $(t\text{-}1)$ to state $j$ at time $t$. Then, the vector $\tilde{x}_{mit} = (x_{mi1t},\ldots,x_{miJt})$ is a multinomial random variable with probability vector $\tilde{\pi}_{mit} = (\pi_{mi1t},\ldots,\pi_{miJt})$ where $\sum_{j=1} \pi_{mijt} = 1$ for $i$=1, ..,$J$-1, $\pi_{mJjt} = 0$ for $j$=1,..,$J$-1, and $\pi_{mJJt} = 1$ as the $J$th state is an absorbing state.

The multinomial model for the transitions from the $i$th state of the chain is given by

$$(\tilde{x}_{mit} \mid \tilde{\pi}_{mit}) \sim Multinomial(\tilde{\pi}_{mit}, 1),$$

for $m$=1, ..,$M$, $i$,$j$=1, ..,$J$, $t$=1, ..,$T$. The multinomial logit transform for the elements of $\tilde{\pi}_{mit}$ is defined as

$$\eta_{mijt} = logit(\pi_{mijt}) = log(\frac{\pi_{mijt}}{\pi_{miJt}}),$$

for $m$=1,..,$M$, $i$=1,..,$J$ ,$j$=1,..,$J$-1, $t$=1,..,$T$. The logit vector $\tilde{\eta}_{mjt} = (\eta_{mi1t},\ldots,\eta_{mi,J-1,t})$ is given by

$$\tilde{\eta}_{mit} = \tilde{\gamma}_t + \tilde{\gamma}_{it} + z_m \beta_i .$$

(1)

In (1) vector $\tilde{g}_t = (g_{1t},...,g_{J-1,t})$ represents common factors across the rows whereas vector $\tilde{g}_{it} = (g_{i1t},...,g_{i,J-1,t})$ is row specific and thus describes the row effects on transition probabilities. The vectors $\tilde{g}_t$ and $\tilde{g}_{it}$ represent the fixed effects for the *i*th row whereas the vector $\tilde{\boldsymbol{b}}_i = (\boldsymbol{b}_1,...,\boldsymbol{b}_{J-1})$ represents the covariate effect in the model and $z_m$ is a covariate. As *J*th state is used as baseline category, $g_{Jt} = g_{iJt} = \boldsymbol{b}_J = 0$ for all *i*'s.

Following the Markov structure used by Grunwald, Raftery and Guttorp (1993), and Cargnoni, Muller and West (1997) to describe a first order dependence of the time evolving parameters, we assume that the parameters in (1) follow a state space model of the form

$$\tilde{g}_t = \tilde{g}_{t-1} + \tilde{w}_t,$$
$$\tilde{g}_{it} = \tilde{g}_{i,t-1} + \tilde{w}_{it}.$$

where $\tilde{w}_t$ and $\tilde{w}_{it}$ are independent of each other following multivariate normal distributions $MVN(\tilde{0},\tilde{W}_1)$ and $MVN(\tilde{0},\tilde{W}_2)$, respectively, for all $t$. Furthermore, $\tilde{w}_t$ and $\tilde{w}_{it}$ are independent sequences over time. It follows from the above that

$$(\tilde{g}_t \mid \tilde{g}_{t-1},\tilde{W}_1) \sim \begin{cases} MVN(\tilde{g}_{t-1},\tilde{W}_1) & if \quad t > 0 \\ MVN(\tilde{0},\tilde{W}_1) & if \quad t = 0 \end{cases},$$

where $\tilde{W}_1^{-1} \sim Wishart(\tilde{R},k)$ with $\tilde{R}$ and $k$ specified. Similarly, for the row specific component $\tilde{g}_{it}$ we have

$$(\tilde{g}_{it} \mid \tilde{g}_{i,t-1},\tilde{W}_2) \sim \begin{cases} MVN(\tilde{g}_{i,t-1},\tilde{W}_2) & if \quad t > 0 \\ MVN(\tilde{0},\tilde{W}_2) & if \quad t = 0 \end{cases},$$

where $\tilde{W}_2^{-1} \sim Wishart(\tilde{R},k)$. As before $\tilde{g}_{it}$ is independent of $\tilde{g}_{jt}$ for all $i \neq j$. The covariate effect $\tilde{\boldsymbol{b}}_i$ in a given row *i* is independently distributed of the other rows with a multivariate normal $MVN(\tilde{m}_{b_i},\tilde{W}_b)$ with known mean vector $\tilde{m}_{b_i}$ and the unknown inverse covariance matrix following a Wishart distribution $\tilde{W}_b^{-1} \sim Wishart(\tilde{R},k)$. We note that covariance matrix $\tilde{W}_b$ is common for all the rows. Furthermore, apriori it is assumed that the vectors $\tilde{g}_t$, $\tilde{g}_{it}$ and $\tilde{\boldsymbol{b}}_i$ are independent of each other.

The Bayesian inference for parameters of the proposed model can not be developed analytically. But posterior inference can be obtained via Markov chain Monte Carlo methods using the programming environment WinBugs (Spiegelhalter et al., 1996). The proposed methodology is applied to some real data from a psychiatric treatment program.

**REFERENCES**


Cargnoni, C., Muller, P. and West, M. (1997). Bayesian forecasting of multinomial time series through conditionally Gaussian dynamic models. *J. Amer. Statist. Assoc.* **92**, 640-647.

Duncan, G., and Lin, L. (1972). Inference for Markov chains having stochastic entry and exit. *J. Amer. Statist. Assoc.* **67**, 761-767.

Erkanli, A., Soyer, R. and Angold, A. (2001). Bayesian analyses of longitudinal binary data using Markov regression models of unknown order. *Statistics in Medicine* **20**, 755-770.

Grunwald, G., Raftery, A. and Guttorp, P. (1993). Time series for continuous proportions. *J. Roy. Statist. Soc. B* **55**, 103-116.

Lee, T.C., Judge, G.G. and Zellner, A. (1968). Maximum likelihood and Bayesian estimation of transition probabilities. *J. Amer. Statist. Assoc.* **63**, 1162-1179.

Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1996). Bayesian Inference Using Gibbs Sampling Manual (version ii). MRC Biostatistics Unit, Cambridge University.


**RESUME**


Nous présentons l'approche Bayesian hiérarchique afin d'analyser les chaî nes Markov du temp nonhomogenes. La chaî ne Markov est utilisée pour décrire le mode de transition chez les enfants émotionnellement perturbés dans le cadre d'un programme de traitement. Les probabilités de transition sont fonction des facteurs et du temps dans les modèles utilisés.