

Bayesian Hierarchical Mixed Effects Analysis of Time Non-Homogeneous Markov Chains

Minje Sung^{a,1}

^aSchool of Business, Ajou University

(Received December 21, 2013; Revised February 3, 2014; Accepted February 19, 2014)

Abstract

The present study used a hierarchical Bayesian approach was used to develop a mixed effect model to describe the transitional behavior of subjects in time nonhomogeneous Markov chains. The posterior distributions of model parameters were not in analytically tractable forms; subsequently, a Gibbs sampling method was used to draw samples from full conditional posterior distributions. The proposed model was implemented with real data.

Keywords: Markov chain, Bayesian analysis, mixed effect model.

1. 서론

어떤 개체의 상태가 이산 시점에 따라 변화하는 상황을 나타내는 경우에 마코프 체인(Markov chain)이 널리 사용된다. Lee 등 (1970)의 연구에 마코프 전이 행렬에 대한 베이지안 추론이 처음 설명되었고, 비교적 최근 Sung 등 (2007)의 연구에서는 공변량(covariate)을 사용한 비동차 마코프 체인(nonhomogeneous Markov chain)에 대한 베이지안 추론이 소개되어 전이 확률에 대한 개개인의 차이가 공변량이 서로 다르므로 설명되었다. 그런데 어떤 선행연구에서도 공변량이 설명하지 못하는 개체 각각의 특이성을 설명하는 모델이 제시된 적이 없고, Erkanli 등 (2001)의 연구에서 서술되었듯이, 대단히 많은 경우에 전이확률에 대해 공변량이 설명하지 못하고 설명되지 않은 채로 남아있는 상이성(heterogeneity)이 존재하게 된다. 이러한 개개인의 특이한 상이성은 임의효과(random effect)를 나타내는 항목을 모델에 포함함으로써 설명할 수 있다. 본 연구에서는 Sung 등 (2007)의 마코프 체인에 대한 베이지안 연구에서 제시된 로짓(logit) 연결함수 모형을 확장시켜 고정효과(fixed effect)와 임의효과(random effect) 항목을 포함하는 혼합효과(mixed effect)모형을 소개하고 실제 자료를 이용하여 모델의 유용성을 보이고자 한다.

2. 비동차 마코프 체인(Nonhomogeneous Markov Chain) 모델

이산 확률과정 변수 s_{mt} 는 m 번째 ($m = 1, \dots, M$)의 사람이 t 시점에 ($t = 1, \dots, T$) 취하게 되는 상태를 나타낸다고 하자. 그러면 s_{mt} 의 상태가 여러 시점에 걸쳐 변하는 과정을 마코프 체인(Markov

¹School of Business, Ajou University, San 5 Woncheon-Dong Yeongtong-Gu, Suwon, 443-749, Korea.

E-mail: sungmj@ajou.ac.kr

chain)으로 다음과 같이 나타낼 수 있다.

$$p(s_{mt} = j | s_{m,t-1} = i, \dots, s_{m0} = k) = p(s_{mt} = j | s_{m,t-1} = i) = \pi_{mijt}, \quad (2.1)$$

여기에서 $i, j \in \varepsilon = \{1, \dots, J\}$ 이고 π_{mijt} 은 마코프 체인의 전이확률을 나타낸다. 식 (2.1)은 상태공간 $\varepsilon = \{1, \dots, J\}$ 을 가지는 마코프 성질을 나타낸다. 만약 여기에서 $t \neq u$ 인 경우에 $\pi_{mijt} \neq \pi_{miju}$ 라면 이러한 마코프 체인은 비동차(non-homogeneous) 마코프 체인이라 불린다 (Chatfield, 1973). 비동차 마코프 체인의 전이확률 행렬은 다음과 같이 나타낼 수 있다.

$$\Pi_m(t) = \begin{bmatrix} \pi_{m11t} & \cdots & \pi_{m1Jt} \\ \vdots & \ddots & \vdots \\ \pi_{mJ1t} & \cdots & \pi_{mJJt} \end{bmatrix}. \quad (2.2)$$

이항변수 y_{mijt} 가 m 번째 개체가 시점 $t-1$ 에 i 번째 상태에 있다가 시점 t 에 j 번째 시점으로 바뀌었는가의 여부를 나타낸다고 하면,

$$y_{mijt} = 1(s_{mt} = j | s_{m,t-1} = i),$$

여기서 만약 사건 A 가 발생했다면 $1(A) = 1$ 이고 아니면 $1(A) = 0$ 이 된다. 마코프 체인의 i 상태에서 전이를 나타내는 벡터 $\tilde{y}_{mit} = (y_{mi1t}, \dots, y_{miJt})$ 로 정의하고 전이확률 벡터는 $\tilde{\pi}_{mit} = (\pi_{mi1t}, \dots, \pi_{miJt})$ 로 정의하면 각각의 개체가 한 시점에 상태를 옮기는 전이는 다음과 같은 다항분포(multinomial)를 따르게 된다.

$$\tilde{y}_{mit} | \tilde{\pi}_{mit}, 1 \sim \text{multinomial}(\tilde{\pi}_{mit}, 1). \quad (2.3)$$

그러므로 식 (2.2)의 전이확률 행렬의 i 번째 행은 식 (2.3)의 다항분포의 모수행렬이 된다.

본 연구는 J 번째 상태를 기저상태(baseline)로 하는 다항 로짓(multinomial logit)연결함수를 사용하여 공변량(covariate)이 전이확률에 미치는 영향을 판단하고자 한다. 다항 로짓은 다음과 같이 정의된다.

$$\eta_{mijt} = \text{logit}(\pi_{mijt}) = \log \left(\frac{\pi_{mijt}}{\pi_{miJt}} \right), \quad (2.4)$$

여기서 $m = 1, \dots, M$, $i = 1, \dots, J$, $j = 1, \dots, J-1$, 그리고 $t = 1, \dots, T$ 이다. 공변량을 사용해서 전이 행태를 나타내는데 있어 개체간 상이성은 공변량이 설명하지 못하는 어떤 알려지지 않은 다른 요소가 있을 수 있다 (Erkanli 등, 2001). 이러한 개체간 상이성을 각각의 개체에 대한 임의효과(random effect)를 나타내는 항을 모델에 포함시킴으로써 설명될 수 있다.

그러면 공변량과 임의효과를 포함하는 로짓 연결 함수는 다음과 같다.

$$\eta_{mijt} = \gamma_j + \gamma_{ij} + \beta_{mijt} z_{mt}, \quad (2.5)$$

여기서 γ_j 는 행렬의 모든 행에 공통으로 포함되는 항목으로, 행렬 (2.2)에서 다른 행들의 공통점을 나타내는 요인을 말한다. γ_{ij} 는 각 행 i 에 따라 다른 항목이고, β_{mijt} 는 시간에 따라 변하는 공변량 z_{mt} 가 전이확률에 미치는 영향을 나타내는 임의효과계수이다. 임의효과 β_{mijt} 는 인덱스 m, i, j, t 모두에 대해 다른 값을 가지게 된다. 그러므로 전이 행태의 시점에 따라 상이한 비동차성(time nonhomogeneity)은 시점에 따라 변하는 β_{mijt} 에 의해서 설명된다. 식 (2.5)는 벡터 $\tilde{\eta}_{mit} = (\eta_{mi1t}, \dots, \eta_{miJ-1t})'$ 에 대해 다음과 같이 나타낼 수 있다.

$$\tilde{\eta}_{mit} = \tilde{\gamma} + \tilde{\gamma}_i + z_{mt} \tilde{\beta}_{mit}, \quad (2.6)$$

또는 다음과 같이 벡터의 식으로 나타낼 수 있다.

$$\begin{pmatrix} \eta_{mi1t} \\ \vdots \\ \eta_{miJ-1t} \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_{J-1} \end{pmatrix} + \begin{pmatrix} \gamma_{i1} \\ \vdots \\ \gamma_{iJ-1} \end{pmatrix} + z_{mt} \begin{pmatrix} \beta_{mi1t} \\ \vdots \\ \beta_{miJ-1t} \end{pmatrix}.$$

식 (2.6)은 공변량의 수 $q > 1$ 를 가지는 일반적 형태로 다음과 같이 나타낼 수 있다.

$$\tilde{\eta}_{mit} = \tilde{\gamma} + \tilde{\gamma}_i + z_{mt1}\tilde{\beta}_{mit1} + \cdots + z_{mtq}\tilde{\beta}_{mitq}.$$

계층적 베이지안(hierarchical Bayesian) 모델링을 사용하여, $\tilde{\gamma}$ 항의 사전분포로서 다변량 정규분포(multivariate normal distribution)를 가정하면,

$$\tilde{\gamma} \sim \text{MVN}(\tilde{\mu}_\gamma, W_1), \tag{2.7}$$

위에서 평균벡터 $\tilde{\mu}_\gamma$ 는 사전적으로 고정(specified)되었다고 가정하고, 알려지지 않은 공분산 행렬의 역행렬은 다음의 위샤트(Wishart)분포를 따른다고 가정한다.

$$W_1^{-1} \sim \text{Wish}(\tilde{R}, k), \tag{2.8}$$

위에서 두 모수 \tilde{R} 과 k 는 고정된 값이라고 가정한다. 위샤트 분포는 다변량 정규분포의 역 공분산 행렬의 켈레 사전확률(conjugate prior) 분포이다. 식 (2.8)에서 \tilde{R} 는 J 번째 상태가 기저상태로 사용되었기 때문에 $(J - 1) \times (J - 1)$ 차원의 행렬이다.

각 행의 고정효과를 나타내는 모수 $\tilde{\gamma}_i$ 는 다음과 같이 다변량 정규분포를 따른다고 가정한다.

$$\tilde{\gamma}_i \sim \text{MVN}(\tilde{\mu}_{\gamma_i}, W_2), \tag{2.9}$$

여기서 평균 벡터 $\tilde{\mu}_{\gamma_i}$ 에 대해서는 사전적으로 값이 주어질 것이고 알려지지 않은 역 공분산 행렬은 다음과 같이 위샤트 분포를 따른다고 가정한다.

$$W_2^{-1} \sim \text{Wish}(\tilde{R}, k), \tag{2.10}$$

여기서 \tilde{R} 과 k 값은 사전적으로 주어질 것이다.

임의효과를 나타내는 모수 $\tilde{\beta}_{mit}$ 는 개체를 나타내는 인덱스 m 에 대해 교환가능성(exchangeability)을 가정한다. 사전분포의 모수(hyperparameter)는 시간 인덱스인 t 에 대해서 일차 마코프 의존성(first order Markov dependence)을 가진다고 가정한다. 그러면 $\tilde{\beta}_{mit}$ 는 다음과 같은 다변량 정규분포를 가지는 데,

$$\tilde{\beta}_{mit} = \text{MVN}(\tilde{\mu}_{\beta_{it}}, W_\beta). \tag{2.11}$$

알려지지 않은 평균 벡터 $\tilde{\mu}_{\beta_{it}}$ 에 대한 일차 마코프 의존성은 다음과 같이 나타낸다.

$$\tilde{\mu}_{\beta_{it}} \sim \begin{cases} \text{MVN}(\tilde{\mu}_{\beta_{it-1}}, W_{\mu_\beta}), & \text{if } t > 0, \\ \text{MVN}(\tilde{0}, W_{\mu_\beta}), & \text{if } t = 0, \end{cases} \tag{2.12}$$

여기서 $t = 0, \dots, T$ 이다. 식 (2.11)에 나타난 교환가능성은 $\tilde{\beta}_{1it}, \dots, \tilde{\beta}_{Mit}$ 들 사이의 완전 대칭성(complete symmetry)을 가지게 되고, $\{\tilde{\beta}_{1it}, \dots, \tilde{\beta}_{Mit}\}$ 들의 모수가 알려지지 않은 평균 벡터 $\tilde{\mu}_{\beta_{it}}$ 와

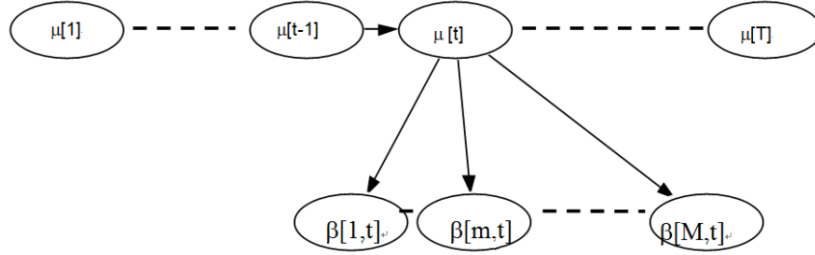


Figure 2.1. Assumptions for exchangeability and first order Markov property of random effects parameter $\tilde{\beta}_{mit}$

공분산 행렬 W_β 을 가지는 다변량 정규분포를 따르는 임의의 변량이라는 것을 의미한다. W_β 의 역행렬은 식 (2.10)과 유사한 위사트 분포를 따른다고 가정한다.

Figure 2.1에서 임의 효과 모수인 $\tilde{\beta}_{mit}$ 에 대한 계층적 구조와 조건부 독립성 가정이 그림으로 나타나 있다. Figure 2.1의 그림에서 인덱스 i 와 j 는 그림을 단순화 하기 위해 생략되었다. 그림의 아래 부분에 나타나 있는 조건부 독립성은 식 (2.11)에서의 인덱스 m 에 대한 $\tilde{\beta}_{mit}$ 의 교환가능성(exchangeability)을 나타낸다. 그림의 위 부분은 식 (2.12)에서 사전 모수들(hyperparameters)이 t 에 대해 일차 의존성을 가지는 구조를 나타낸다. 다시 말하면 $\tilde{\mu}_{\beta_{ti}}$ 가 주어졌다면, $\tilde{\beta}_{mit}$ 는 $m \neq n$ 에 대해 조건부 독립적이고, $\tilde{\beta}_{mit}$ 는 $t \neq u$ 에 대해서 $\tilde{\mu}_{\beta_{tu}}$ 와 독립적이다.

3. 사후 분석(Posterior Analysis)

이번 절에서는 2절에 소개된 베이지안 모델에 포함된 알려지지 않은 모수에 대한 사후 확률분포(posterior distribution)를 소개한다. 결합 사후(joint posterior) 확률분포가 구해지면 주변분포(marginal distribution), 신용구간(credible interval), 그리고 결합 사후 확률분포에 대한 확률 구문(probability statement)들을 사용하여 다양한 통계적 추론을 내릴 수 있게 된다. 그러나 아래에 소개 되는 바와 같이 결합 혹은 주변 사후 분포들은 분석적으로 알려진 분포의 형태로 구해지지 않게 된다. 따라서 마코프 체인 몬테칼로(Markov chain Monte Carlo; MCMC) 시뮬레이션 방법을 사용하여 통계적 추론을 내리게 된다. 본 연구에서는 MCMC의 한 형태인 깁스(Gibbs) 샘플링 (Gelfand와 Smith, 1990) 방법이 사용된다.

M 개의 개체와 T 시간 동안에 걸친 전이 행태에 대한 모든 자료들을 가지고 베이지안 분석을 위한 결합 사후확률은 다음과 같다.

$$p\left(\Pi_1(1), \dots, \Pi_M(T), \tilde{\gamma}, \tilde{\Gamma}, \tilde{B}_1, \dots, \tilde{B}_T, \tilde{\mu}_B, W_1, W_2, W_\beta \mid \tilde{S}_1, \dots, \tilde{S}_M\right), \quad (3.1)$$

여기서 $\Pi_m(t) = (\tilde{\pi}_{m0t}, \dots, \tilde{\pi}_{m1t})$, $\tilde{S}_m = (s_{m0}, \dots, s_{mT})$, $\tilde{\gamma} = (\gamma_1, \dots, \gamma_J)$, $\tilde{\Gamma} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_J)$, $\tilde{B}_t = ((\tilde{\beta}_{11t}, \dots, \tilde{\beta}_{M1t})', \dots, (\tilde{\beta}_{1Jt}, \dots, \tilde{\beta}_{MJt})')$ 로 정의되고 $m = 1, \dots, M$, $i = 1, \dots, J$, $j = 1, \dots, J$, $t = 1, \dots, T$ 이다. 식 (3.1)의 결합 사후확률에 조건부 독립성을 적용하면 다음의 비례식이 된다.

$$\begin{aligned} & \propto \prod_{m=1}^M \prod_{i=1}^J \prod_{t=1}^T p\left(\tilde{y}_{mit} \mid \tilde{\gamma}, \tilde{\gamma}_i, \tilde{\beta}_{mit}\right) p\left(\tilde{\gamma} \mid \tilde{\mu}_\gamma, W_1\right) p\left(\tilde{\gamma}_i \mid \tilde{\mu}_{\gamma_i}, W_2\right) \\ & \times p\left(\tilde{\beta}_{mit} \mid \tilde{\mu}_{\beta_{it}}, W_\beta\right) p\left(W_1^{-1}\right) p\left(W_2^{-1}\right) p\left(\tilde{\mu}_{\beta_{it}} \mid \tilde{\mu}_{\beta_{it-1}}, W_{\mu_\beta}\right) p\left(W_\beta^{-1}\right) p\left(W_{\mu_\beta}^{-1}\right). \end{aligned} \quad (3.2)$$

위 식 (3.2)의 결합 사후 확률 분포는 분석적인 형태로 구해질 수 없는 분포이다. 따라서 본 연구에서는 식 (3.2)의 분포로부터 샘플을 추출하기 위해 깁스(Gibbs) 샘플러를 이용하여 시물레이션을 통해 샘플을 구하게 된다. 깁스 샘플링을 수행하려면 알려지지 않은 모든 변량들에 대해 다음과 같은 조건부 사후 분포(full conditional posterior distribution)들이 필요하게 된다.

$$\begin{aligned} & \left(\tilde{\gamma} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\gamma}^{(-)} \right), \quad \left(\tilde{\gamma}_i \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\gamma}_i^{(-)} \right), \quad \left(\tilde{\beta}_{mit} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\beta}_{mit}^{(-)} \right), \\ & \left(\tilde{\mu}_{\beta_{it}} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\mu}_{\beta_{it}}^{(-)} \right), \quad \left(W_1^{-1} \mid \tilde{S}_1, \dots, \tilde{S}_M, W_1^{(-)} \right), \quad \left(W_2^{-1} \mid \tilde{S}_1, \dots, \tilde{S}_M, W_2^{(-)} \right), \\ & \left(W_\beta^{-1} \mid \tilde{S}_1, \dots, \tilde{S}_M, W_\beta^{(-)} \right), \quad \left(W_{\mu_\beta}^{-1} \mid \tilde{S}_1, \dots, \tilde{S}_M, W_{\mu_\beta}^{(-)} \right). \end{aligned}$$

식 (3.2)에서 $\tilde{\pi}_{mit}$ 에 대해서는 조건부 사후분포가 필요하지 않는데 그 이유는 식 (2.6)에서 로짓 연결 함수에 의해서 $\tilde{\pi}_{mit}$ 값들이 구해질 수 있기 때문이다. 고정효과 항목 $\tilde{\gamma}$ 는 전이 행렬의 모든 행에 대해서 공통점을 나타내는데 다음의 조건부 사후확률분포로부터 샘플링이 될 것이다.

$$\begin{aligned} & p \left(\tilde{\gamma} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\gamma}^{(-)} \right) \\ & \propto \prod_{m=1}^M \prod_{i=1}^J \left[\prod_{t=1}^T p \left(\tilde{y}_{mit} \mid \tilde{\gamma}, \tilde{\gamma}_i, \tilde{\beta}_{mit} \right) \right] p \left(\tilde{\gamma} \mid \tilde{\mu}_\gamma, W_1 \right) \\ & \propto \prod_{m,i,j,t} \left(\frac{\exp(\gamma_j + \gamma_{ij} + \beta_{mij} z_{mt})}{\sum_{j=1}^J \exp(\gamma_j + \gamma_{ij} + \beta_{mij} z_{mt})} \right)^{y_{mijt}} \exp \left\{ -\frac{1}{2} (\tilde{\gamma} - \tilde{\mu}_\gamma)' W_1^{-1} (\tilde{\gamma} - \tilde{\mu}_\gamma) \right\}. \quad (3.3) \end{aligned}$$

전이 행렬의 행마다 다른 고정효과를 나타내는 $\tilde{\gamma}_i$ 는 다음의 분포로부터 샘플이 추출 될 것이다.

$$\begin{aligned} & p \left(\tilde{\gamma}_i \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\gamma}_i^{(-)} \right) \\ & \propto \prod_{m=1}^M \left[\prod_{t=1}^T p \left(\tilde{y}_{mit} \mid \tilde{\gamma}, \tilde{\gamma}_i, \tilde{\beta}_{mit} \right) \right] p \left(\tilde{\gamma}_i \mid \tilde{\mu}_{\gamma_i}, W_2 \right) \\ & \propto \prod_{m,j,t} \left(\frac{\exp(\gamma_j + \gamma_{ij} + \beta_{mij} z_{mt})}{\sum_{j=1}^J \exp(\gamma_j + \gamma_{ij} + \beta_{mij} z_{mt})} \right)^{y_{mijt}} \exp \left\{ -\frac{1}{2} (\tilde{\gamma}_i - \tilde{\mu}_{\gamma_i})' W_2^{-1} (\tilde{\gamma}_i - \tilde{\mu}_{\gamma_i}) \right\}. \quad (3.4) \end{aligned}$$

위에서 식 (3.3)과 (3.4)의 조건부 사후분포들은 전이행렬 행을 나타내는 인덱스 i 에 대한 곱의 항목을 제외하면 비슷한 형태의 분포를 가지게 된다. 그런데 식 (3.4)에서 인덱스 i 에 대한 곱의 항목은 생략되었는데, 그 이유는 서로 다른 행들간, 즉 $i \neq j$ 에 대해서 $\tilde{\gamma}_i$ 들은 독립적이기 때문이다.

$(\tilde{\beta}_{mit} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\beta}_{mit}^{(-)})$ 에 대한 사후 시물레이션은 다음의 조건부 사후분포로부터 이루어진다.

$$\begin{aligned} & p \left(\tilde{\beta}_{mit} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\beta}_{mit}^{(-)} \right) \\ & \propto p \left(\tilde{y}_{mit} \mid \tilde{\gamma}, \tilde{\gamma}_i, \tilde{\beta}_{mit} \right) p \left(\tilde{\beta}_{mit} \mid \tilde{\mu}_{\beta_{it}}, W_\beta \right) \\ & \propto \prod_j \left(\frac{\exp(\gamma_j + \gamma_{ij} + \beta_{mij} z_{mt})}{\sum_{j=1}^J \exp(\gamma_j + \gamma_{ij} + \beta_{mij} z_{mt})} \right)^{y_{mijt}} \exp \left\{ -\frac{1}{2} (\tilde{\beta}_{mit} - \tilde{\mu}_{\beta_{it}})' W_\beta^{-1} (\tilde{\beta}_{mit} - \tilde{\mu}_{\beta_{it}}) \right\}. \quad (3.5) \end{aligned}$$

위의 식 (3.5)에서 인덱스 m, i, t 에 대한 곱의 항목은 생략되었는데 각각의 다른 m, i, t 에 대해 독립적인 체인으로부터 $\tilde{\beta}_{mit}$ 의 값이 시뮬레이션이 되기 때문이다.

역 공분산 행렬 W_1^{-1} 은 다음과 같은 조건부 사후분포를 따르게 된다.

$$\begin{aligned} & p\left(W_1^{-1} \mid \tilde{S}_1, \dots, \tilde{S}_M, W_1^{(-)}\right) \\ & \propto p\left(\tilde{\gamma} \mid \tilde{\mu}_\gamma, W_1^{-1}\right) p\left(W_1^{-1} \mid \tilde{R}, k\right) \\ & \propto \left|\tilde{W}_1\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\tilde{\gamma}-\tilde{\mu}_\gamma\right)' W_1^{-1}\left(\tilde{\gamma}-\tilde{\mu}_\gamma\right)\right\}\left|W_1^{-1}\right|^{\frac{k-p-1}{2}} \exp\left\{-\frac{1}{2} \operatorname{tr}\left(\tilde{R} W_1^{-1}\right)\right\} \\ & \propto\left|W_1^{-1}\right|^{\frac{k-p}{2}} \exp\left\{-\frac{1}{2} \operatorname{tr}\left[\left(\tilde{R}+\left(\tilde{\gamma}-\tilde{\mu}_\gamma\right)\left(\tilde{\gamma}-\tilde{\mu}_\gamma\right)'\right) W_1^{-1}\right]\right\}, \end{aligned} \quad (3.6)$$

여기에서 $E\left[W_1^{-1}\right]=\left(\tilde{R} / k\right)^{-1}$ 이기 때문에 \tilde{R} / k 는 k 개의 변수들의 공분산에 대한 사전 추정값으로 해석될 수 있다. 식 (3.6)의 조건부 사후분포는 자유도 $k+1$ 과 척도(scale) 행렬 $1/2\left(\tilde{R}+\left(\tilde{\gamma}-\tilde{\mu}_\gamma\right)\left(\tilde{\gamma}-\tilde{\mu}_\gamma\right)'\right)$ 을 가지는 위샤프 분포이다.

$\tilde{\mu}_{\beta_{it}}$ 에 대한 조건부 사후분포는 다음과 같다.

$$\begin{aligned} & p\left(\tilde{\mu}_{\beta_{it}} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\mu}_{\beta_{it}}^{(-)}\right) \\ & \propto \prod_{m=1}^M p\left(\tilde{\beta}_{mit} \mid \tilde{\mu}_{\beta_{it}}, W_\beta\right) p\left(\tilde{\mu}_{\beta_{it}} \mid \tilde{\mu}_{\beta_{it-1}}, W_{\mu_\beta}\right) p\left(\tilde{\mu}_{\beta_{it+1}} \mid \tilde{\mu}_{\beta_{it}}, W_{\mu_\beta}\right) \\ & \propto \prod_{m=1}^M \exp\left\{-\frac{1}{2}\left(\tilde{\beta}_{mit}-\tilde{\mu}_{\beta_{it}}\right)' W_2^{-1}\left(\tilde{\beta}_{mit}-\tilde{\mu}_{\beta_{it}}\right)\right\} \exp\left\{-\frac{1}{2}\left(\tilde{\mu}_{\beta_{it}}-\tilde{\mu}_{\beta_{it-1}}\right)' W_{\mu_\beta}^{-1}\left(\tilde{\mu}_{\beta_{it}}-\tilde{\mu}_{\beta_{it-1}}\right)\right\} \\ & \quad \times \exp\left\{-\frac{1}{2}\left(\tilde{\mu}_{\beta_{it+1}}-\tilde{\mu}_{\beta_{it}}\right)' W_{\mu_\beta}^{-1}\left(\tilde{\mu}_{\beta_{it+1}}-\tilde{\mu}_{\beta_{it}}\right)\right\} \\ & \propto \exp\left[-\frac{1}{2}\left\{\tilde{\mu}_{\beta_{it}}'\left(M W_2^{-1}+2 W_{\mu_\beta}^{-1}\right) \tilde{\mu}_{\beta_{it}}-2 \tilde{\mu}_{\beta_{it}}'\left(W_2^{-1} \sum_{m=1}^M \beta_{mit}+W_{\mu_\beta}^{-1}\left(\tilde{\mu}_{\beta_{it-1}}-\tilde{\mu}_{\beta_{it+1}}\right)\right)+h\right\}\right], \end{aligned}$$

여기서 h 는 상수항으로서 변량인 $\tilde{\mu}_{\beta_{it}}$ 을 포함하지 않는 항목이다. 그러면 위의 이차식을 정리하면 다음과 같이 나타낼 수 있는데

$$\left(\tilde{\mu}_{\beta_{it}}-\tilde{B}\right)' \tilde{C}^{-1}\left(\tilde{\mu}_{\beta_{it}}-\tilde{B}\right),$$

여기서 $\tilde{C}^{-1}=M W_2^{-1}+2 W_{\mu_\beta}^{-1}$, 그리고 $\tilde{B}=\tilde{C}\left(W_2^{-1} \sum_{m=1}^M \beta_{mit}+W_{\mu_\beta}^{-1}\left(\tilde{\mu}_{\beta_{it-1}}+\tilde{\mu}_{\beta_{it+1}}\right)\right)$ 로 정의된다. 그러므로 다음과 같은 간편 비례식으로 나타낼 수 있다.

$$p\left(\tilde{\mu}_{\beta_{it}} \mid \tilde{S}_1, \dots, \tilde{S}_M, \tilde{\mu}_{\beta_{it}}^{(-)}\right) \propto \exp\left\{-\frac{1}{2}\left(\tilde{\mu}_{\beta_{it}}-\tilde{B}\right)' \tilde{C}^{-1}\left(\tilde{\mu}_{\beta_{it}}-\tilde{B}\right)\right\}.$$

위의 형태는 평균 벡터 \tilde{B} 와 공분산 행렬 \tilde{C}^{-1} 를 가지는 다변량 정규분포이다.

W_2^{-1} 에 대한 조건부 사후분포는 식 (3.6)와 비슷한 위샤프 분포를 가지는데, 인덱스 i 에 대한 곱을 해야 하기 때문에 조건부 사후분포는 자유도 $k+J$ 와 척도(scale) 행렬 $1/2\left(\tilde{R}+\sum_i\left(\tilde{\gamma}_i-\tilde{\mu}_{\gamma_i}\right)\left(\tilde{\gamma}_i-\tilde{\mu}_{\gamma_i}\right)'\right)$ 를 가지는 위샤프 분포이다. W_β^{-1} 의 조건부 사후분포는 자유도 $k+M+J+T$ 와 척도 행렬 $1/2\left(\tilde{R}+\sum_{m,i,t}\left(\tilde{\beta}_{mit}-\tilde{\mu}_{\beta_{it}}\right)\left(\tilde{\beta}_{mit}-\tilde{\mu}_{\beta_{it}}\right)'\right)$ 을 가지는 위샤프 분포이다. 비슷한 방식으로 $W_{\mu_\beta}^{-1}$ 는 자유도 $k+J+T$ 와 척도 행렬 $1/2\left(\tilde{R}+\sum_{i,t}\left(\tilde{\mu}_{\beta_{it}}-\tilde{\mu}_{\beta_{it-1}}\right)\left(\tilde{\mu}_{\beta_{it}}-\tilde{\mu}_{\beta_{it-1}}\right)'\right)$ 을 가지는 위샤프 분포를 따른다.

4. 자료분석

2절에서 소개된 모델은 Nhan (1998)의 연구에서 수집된 실제 자료들에 대해 적용이 되어, 본 절에서 개발된 모델의 적용 예가 소개된다. 자료는 나이 8세부터 17세까지의 아동과 청소년기의 $N = 348$ 명에 대해 미국 버지니아주의 한 정신건강 치료 병원에서 환자의 상태변화를 나타낸다. 1의 상태가 가장 심각한 상태이고 4가 가장 가벼운 상태이다. 상태 5는 환자가 병원으로부터 퇴원을 하여 전이행렬 시스템으로부터 ‘이탈(exit)’하게 되는 상태를 나타낸다. 환자들의 상태에 대한 관찰은 병원에 입원 후 30일 이후부터 시작되었으며 그 뒤로 매 3달 간격으로 환자들의 상태가 관찰되었다. 본 연구에서는 $t = 1, \dots, 7$ 총 7시점의 마코프 전이 행렬 자료를 사용하여 모델의 적용 예를 소개하고자 한다. 본 연구에서 공변량은 나이가 고려되었다.

4.1. 사전분포 설정

전이확률에 대한 추론을 사후분포로부터 내리기 위해서 베이저안 방법론을 사용하려면, 모수들에 대한 사전 분포 값을 설정해 주어야 하는데, 본 연구에서는 모델의 모수들에 대한 불확실성을 나타내기 위해 무정보 사전분포(noninformative prior)값들을 사용하였다. 즉, 식 (2.8)과 (2.9)의 $\tilde{\gamma}$ 과 $\tilde{\gamma}_i$ 에 대해서는 평균 벡터값 $(0, 0, 0, 0)$ 을 사용하였고, 미지의 공분산 행렬에 대해서는 특정값을 설정하는 대신 역 공분산 행렬 W_1^{-1} 과 W_2^{-1} 들이 자유도 $k = 4$ 와 척도 행렬(scale matrix) $\tilde{R} = \text{diag}(.01, .01, .01, .01)$ 를 가지는 위사트 분포를 따른다고 가정하였다. 식 (2.11)과 (2.12)에서의 역 공분산 행렬 W_β^{-1} 과 $W_{\mu_\beta}^{-1}$ 들에 대해서도 같은 값의 모수들을 가지는 위사트 분포를 가정하였다.

4.2. 사후분포 분석

개발된 모델을 사용한 자료분석은 WinBUGS (Spiegelhalter 등, 1996) 프로그램을 이용하였으며, 처음 500,000의 샘플은 버리고, 혹시 발생할 수 있는 샘플간 연관성을 줄이기 위해 매 50번째의 샘플을 저장하여 총 5,000개의 샘플을 사용하여 사후분포에 대한 추정을 하게 되었다.

식 (2.5)에 나타난 로지스틱 전이 모델에서 모수 γ_j 과 γ_{ij} 는 고정효과를 나타내고 β_{mijt} 는 공변량(covariate)에 대한 임의효과를 나타낸다. 그러면 전이행렬의 i 행으로부터 j 열로 상태가 전이하는 상대확률 비(odds ratio)는 다음과 같다.

$$\frac{\pi_{mijt}}{\pi_{mi5t}} = \exp(\gamma_j + \gamma_{ij} + \beta_{mijt}z_{mt}). \quad (4.1)$$

위의 식 (4.1)은 시스템으로부터의 이탈(exit) 상태, 즉 상태 5에 대한 상대 확률의 비를 나내는 식으로, 확률의 로그값들의 차이의 형태로 다음과 같이 나타낼 수 있다.

$$\log(\pi_{mijt}) - \log(\pi_{mi5t}) = \gamma_j + \gamma_{ij} + \beta_{mijt}z_{mt},$$

여기서 $(\gamma_j + \gamma_{ij})$ 값은 공변량에 의해 설명되지 않은 확률의 로그값 변동 기대치로 해석 될 수 있다.

Gibbs 샘플링 방법을 사용하여 얻은 모수들에 대한 표본추출(sampling)이 안정되게 진행되었는가를 검증하기 위해 trace plot을 Figure 4.1에 나타난 것처럼 구해보니 큰 문제점은 발견되지 않았다. Figure 4.1에서 첫 행은 γ_j ($j = 1, 2$)에 대해, 둘째 행은 γ_{ij} ($i = 1, j = 1, 2$), 셋째 행은 β_{mijt} ($m = 1, i = 1, j = 1, 2, t = 1$)에 대한 trace plot들이다. 고정효과 모수들인 γ_j 와 γ_{ij} 는 plot에 보여지는 것처럼 수렴이 빠르게 진행되었고 임의효과 모수인 β_{mijt} 는 상대적으로 느리게 진행되었으나 사전분포와는 확실히 다른 값들을 가지는 사후분포를 나타내고 있어 데이터로부터의 배움(learning)이 발생했다는 것을 나타내며 Gibbs 샘플링의 큰 문제점은 나타나지 않았다.

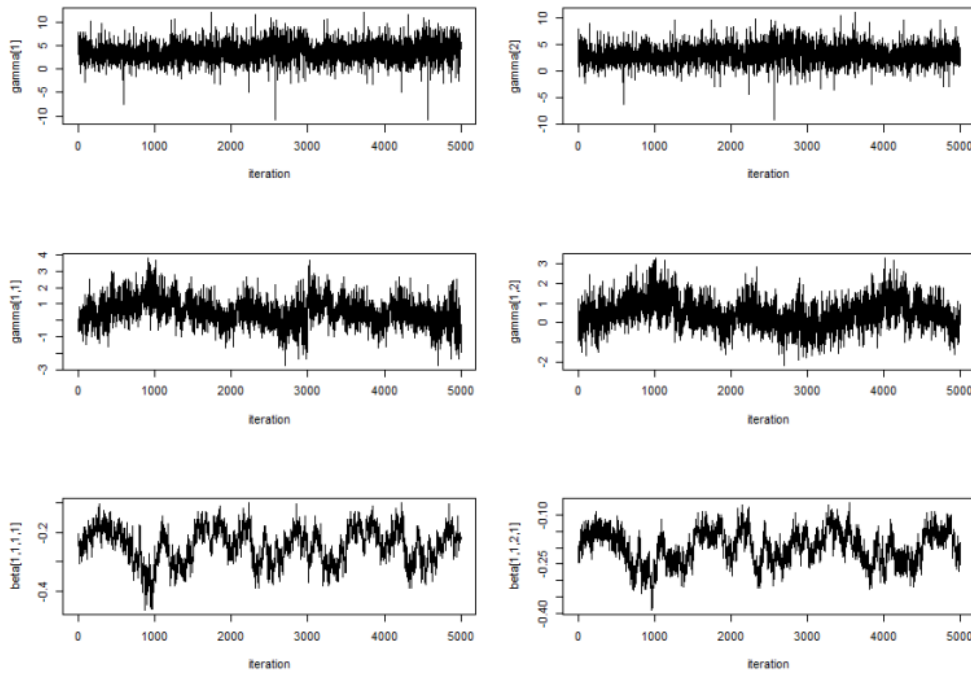


Figure 4.1. Trace plots of $\gamma_j, \gamma_{ij}, \beta_{mijt}$ for convergence

Figure 4.2는 인덱스 $i, j = 1, 2, 3, 4$ 에 대한 고정효과 모수들의 합 ($\gamma_j + \gamma_{ij}$)의 사후분포를 보여주는데, 각각의 사후분포는 이탈 상태(상태 5)에 대한 상대확률(odds) 값 분포를 나타낸다. 즉, 각각의 $i = 1, 2, 3, 4$ 상태들로부터 각각의 $j = 1, 2, 3, 4$ 상태로의 전이에 대해 상태 5를 기준으로 상대확률 비를 나타낸다. Figure 4.2의 한 패널에서 오른쪽에서 왼쪽으로 옮겨갈수록, 즉 주어진 상태에서 더 좋은 상태로 옮겨갈 때, 사후분포의 박스 도표(box plot)들은 점차적으로 아래로 내려가는 경향을 보이는데, 이는 상대확률의 로그(log odds)값들이 점차로 적어진다는 것을 의미한다. 다시 말하면, 공변량에 대한 임의효과가 통제되었을 때, 환자들이 더 좋은 상태로 바뀌는 가능성이 이탈상태로 바뀌는 가능성과 비교하여 점차적으로 줄어든다는 것을 말한다. 이러한 현상은 상태 1, 2, 3으로부터의 전이에서는 공통적으로 관찰되지만, 최상의 상태인 상태 4로부터의 전이의 경우에는 다른 상황이 관찰되었다.

전이확률에 대한 임의효과는 식 (2.11)에 표현된 대로 다변량 정규분포를 따르는데 개체를 나타내는 인덱스 m 에 대해 교환가능성(exchangeability)을 가정하고 있다. 공변량인 나이에 대한 한계효과(marginal effect)는 식 (2.12)의 미지의 평균벡터로 표현되었는데, 시점에 대해 마코프 특성을 가지고 있다고 가정되었다. Figure 4.3의 각 패널은 $\mu_{\beta_{tij}}$ 의 사후분포가 시점 $t = 1, \dots, 7$ 에 따라 어떻게 변화하는지 나타내고 있다. 여기서 주목할 것은 각각의 패널에서 $\mu_{\beta_{tij}}$ 사후분포가 시간이 지남에 따라 감소하는 경향을 보이는데, 이는 나이가 전이확률에 미치는 한계 효과가 시간에 따라 적어짐을 의미한다. 오른쪽 패널에서 왼쪽 패널로 옮겨가면 낮은 상태에서 높은 상태로의 전이에 대한 모수의 효과를 나타내는데, 사후분포가 높은 값을 가지는 경향이 관찰되어 시간이 갈수록 더 나은 상태로 옮겨갈 확률이 높아짐을 의미한다.

임의효과 모수 β_{mijt} 의 사후분포도 각각의 인덱스 m, i, j, t 들 모두에 대해 구할 수 있는데, 모든 사후분포를 본 논문에서 보이는 것은 의미가 없고, 예로서 $m = 1$ 인 특정 환자가 시점에 따라 $i = 1$ 상태

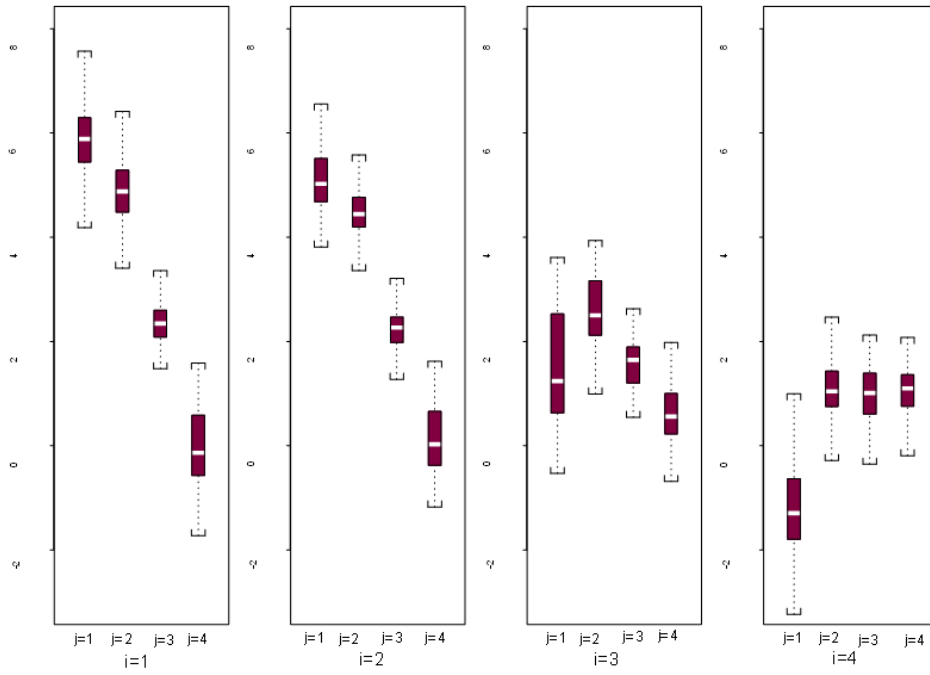


Figure 4.2. Posterior distributions of fixed effects parameters $(\gamma_j + \gamma_{ij})$ for $i, j = 1, 2, 3, 4$

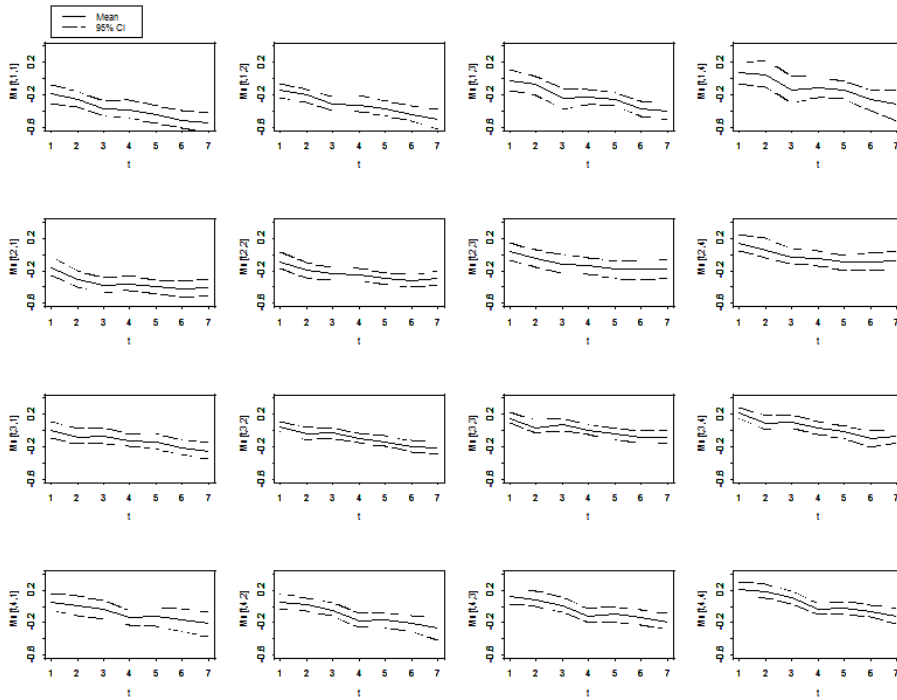
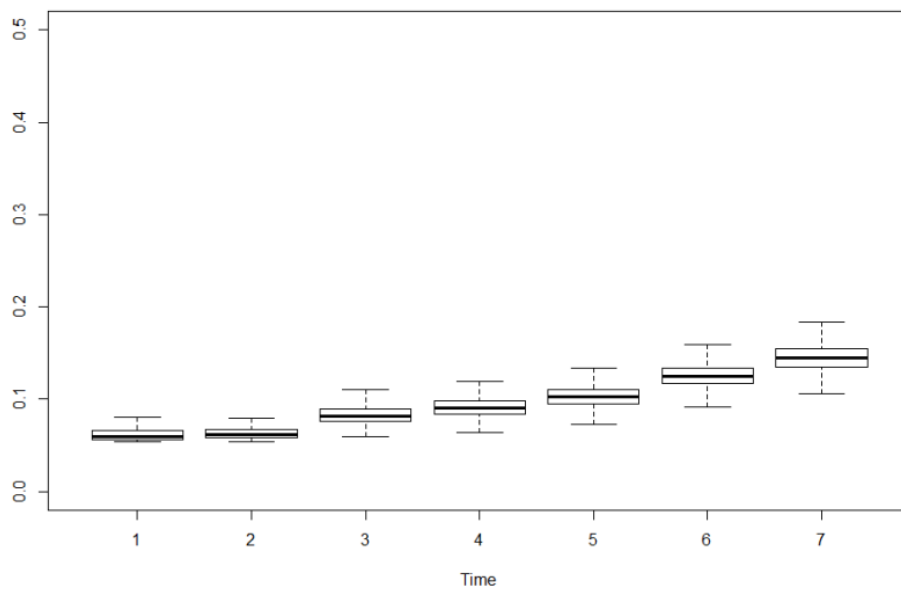


Figure 4.3. Posterior distributions of random effects parameters $\mu_{\beta_{tij}}$

Table 4.1. Posterior means and 95% credible intervals for posterior distributions of β_{11jt}

	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$
$i, j = 1, 1$	-.07 (-.28, .13)	-.15 (-.35, .11)	-.26 (-.43, -.06)	-.27 (-.46, -.06)	-.32 (-.52, -.07)	-.39 (-.58, -.17)	-.42 (-.62, -.22)
$i, j = 1, 2$	-.05 (-.23, .15)	-.11 (-.30, .12)	-.23 (-.39, -.03)	-.22 (-.40, -.04)	-.27 (-.44, -.04)	-.34 (-.51, -.15)	-.39 (-.59, -.21)
$i, j = 1, 3$.06 (-.15, .28)	.01 (-.22, .24)	-.18 (-.44, .03)	-.14 (-.35, .08)	-.16 (-.39, .09)	-.28 (-.50, -.04)	-.31 (-.58, -.06)
$i, j = 1, 4$.11 (-.15, .36)	.07 (-.20, .31)	-.14 (-.41, .11)	-.08 (-.34, .16)	-.12 (-.38, .16)	-.12 (-.38, .16)	-.29 (-.58, .00)

**Figure 4.4.** Posterior distributions of transition probabilities of moving from state $i = 1$ to state $J = 5$ for $m = 1$

에서 각각의 $j = 1, 2, 3, 4$ 로의 전이에 대한 임의효과가 시간에 따라 변하는 것을 Table 4.1에 요약하였다. 임의효과 β_{11jt} 는 모든 환자들에게 공통적인 고정효과가 나타내지 못하는 환자 개인들간의 상이성(heterogeneity)을 특정개인 $m = 1$ 에 대해 나타낸 것이다. 다른 개인 환자들에 대해서도 각각의 특이한 효과에 대한 추론을 비슷한 사후분포를 통해서 구할 수 있다는 것을 이 예를 통해 알 수 있다. 이 특정 개인의 경우 Table 4.1의 열에서 위에서 아래로 내려갈수록 값이 커지는 것을 관찰 할 수 있는데, 이는 이 환자는 주어진 시점에 더 나은 상태로 옮겨갈 가능성이 점차 높음을 알 수 있다. 또한 주어진 열에서 오른쪽에서 왼쪽으로 갈수록 값이 적어지는데, 주어진 시점에서는 더 나은 상태로 옮겨갈 가능성이 높지만 시간이 지날수록 이러한 가능성이 점점 적어짐을 나타낸다.

공변량의 효과가 식 (2.4)와 (2.5)의 다항 로짓 함수로 나타내어졌는데, 자료분석에서 편의상 $J = 5$, 즉 이탈(exit) (혹은 치료시설로부터의 퇴원) 상태가 기저상태(baseline)로 사용되어 식 (2.5)의 모수들이 기저상태인 ‘이탈’ 상태에 대해서는 정의되지 않아 모수를 통해서 ‘이탈’ 상태로의 전이에 대한 추론을

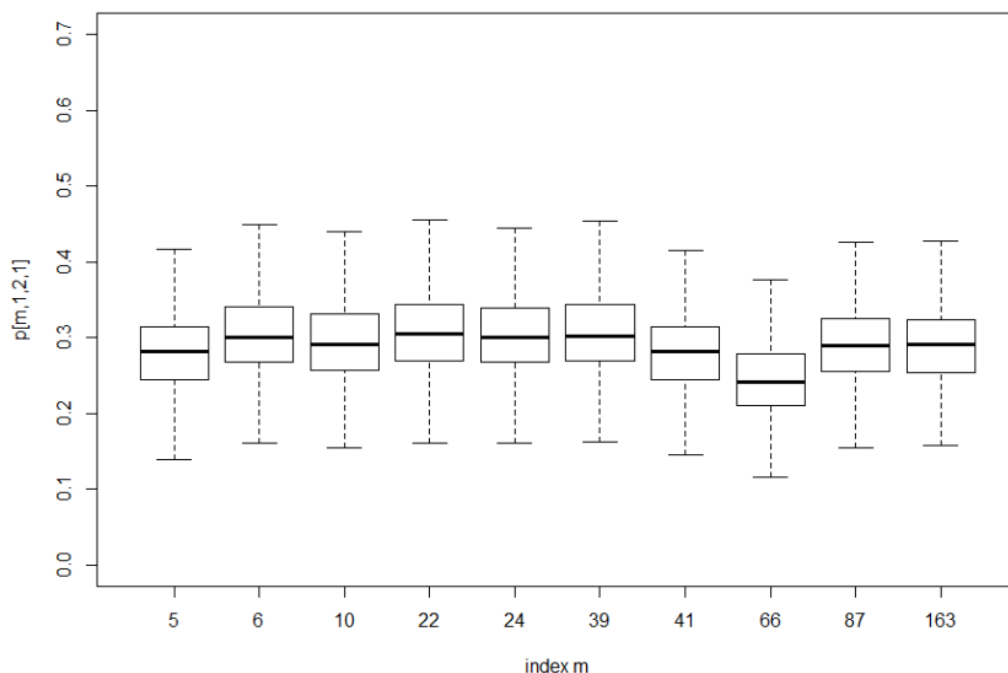


Figure 4.5. Comparisons of posterior probability distributions of moving from $i = 1$ to $j = 2$ at $t = 1$ for age = 10

직접적으로 내릴 수는 없다. 그러나 본 논문의 리뷰위들이 지적한대로 ‘이탈’ 혹은 퇴원은 환자의 상태가 호전되는 경우 혹은 기대만큼 상태가 호전되지 않아 퇴원하는 경우 등 여러 가능성이 있어 연구자들의 관심이 될 만한 중요한 상태이다. 모수를 통한 직접적 영향 평가는 내릴 수 없지만 궁극적으로 연구자들의 관심사인 이탈 (혹은 퇴원) 상태로의 전이확률 $\pi_{miJt} = 1 - \sum_{j=1}^{J-1} \pi_{mijt}$ 로써 구할 수 있어 환자들의 이탈 상태로의 추론을 내릴 수 있다. 예를 들면 Figure 4.4는 특정 환자 ($m = 1$)가 가장 낮은 상태인 $i = 1$ 의 상태에서 퇴원으로 전이하는 사후 확률 분포를 시점에 따라 나타내고 있는데, 시간이 지날수록 퇴원의 가능성이 점차 높아짐을 알 수 있다.

통상적으로 고정효과 모델은 같은 값을 가지는 공변량에 대해 모든 개체들이 같은 평균적 효과를 가지는 것으로 표현하는데 반하여, 임의효과 모델의 장점은 모델에 포함 된 공변량이 설명하지 못하는 상이성(heterogeneity)을 개체간에 따라 다르게 나타낼 수 있다는 것이다. 본 연구에서는 편의상 환자들의 나이만을 공변량으로 포함하였는데 나이 외에 성별, 사회 경제적 수준 등 다른 요소들도 환자들의 상태에 영향을 미칠 수 있다. 이렇게 본 연구에서 고려되지 않은 요소들 혹은 다른 개개인의 특성에 기인한 환자들의 전이 행태의 상이함이 임의 효과 모수에 반영되어 같은 공변량 값을 가진다 하더라도 서로 다른 상태의 변화를 가져 올 수 있다. Figure 4.5는 나이가 모두 10세인 환자들이 $t = 1$ 시점에 $i = 1$ 에서 $j = 2$ 상태로 호전되는 전이확률들의 사후분포를 box plot으로 비교해 나타내 주고 있다. 대체적으로 비슷한 사후확률 분포를 나타내고 있는데 $m = 5$ 와 $m = 66$ 번째 환자들 이 다른 환자들에 비해서 상대적으로 낮은 전이 확률을 나타내고 있다. 이렇게 다른 전이 행태를 보이는 환자들을 판별해 내어 모델에 포함되지 않은 어떤 요소들의 특징을 가지고 있는가를 관찰 하여 본다면 질병의 상태가 호전되는 혹은 악화되는 제 3의 요인들에 대한 시사점을 찾을 수도 있을 것이다.

5. 결론

본 연구에서는 베이지안 방법론을 사용하여 이산 시점에 따라 변하는 마코프 전이 확률에 대한 추론을 내리는데 있어 공변량이 설명하지 못하는 개인의 특이성을 설명할 수 있는 임의효과와 고정효과를 포함하는 혼합효과 모델을 제시하였다. 베이지안 방법론을 사용하면 사후분포를 통하여 모수에 대한 통계적 추론을 내리게 되는데 3절에서 보여진 것처럼 본 모델에서의 모수들의 사후분포는 분석적으로 알려진 분포를 가지지 않는다. 따라서 MCMC 시뮬레이션 기법의 일종인 깁스(Gibbs) 샘플링 방법론을 사용하여 모수에 대한 사후분포를 구하였고 고정효과와 임의효과를 나타내는 모수의 사후분포에 대한 분석을 통하여 공통적으로 환자의 전이 행태에 대한 이해와 특정 개인이 다르게 행동하는 전이 행태에 대한 해석이 가능함을 보였다.

본 연구에서는 의학 자료를 사용하여 예시하였지만, 예를 들면 마케팅 분야에서 특정 제품에 대한 소비자의 선호가 시점에 따라 변하는 자료 등 다양한 다른 분야에 적용될 수 있다.

References

- Chatfield, C. (1973). Statistical inference regarding Markov chain models, *Applied Statistics*, **22**, 7–20.
- 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.
- Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, *Journal of American Statistical Association*, **85**, 972–985.
- Lee, T. C., Judge, G. G. and Zellner, A. (1970). *Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data*, North-Holland and Pub. Co., Amsterdam.
- Nhan, N. (1998). *Assessing Change Among Patients in Residential Treatment*, Technical Report, Graydon Manor Research Department, Virginia.
- Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1996). *Bayesian Inference Using Gibbs Sampling Manual (version ii)*, MRC Biostatistics Unit, Cambridge University.
- Sung, M., Soyer, R. and Nhan, N. (2007). Bayesian analysis of non-homogenous Markov chains: Application to mental health data, *Statistics in Medicine*, **26**, 3000–3017.

계층적 베이지안 혼합 효과 모델을 사용한 비동차 마코프 체인의 분석

성민제^{a,1}

^a아주대학교 경영학과

(2013년 12월 21일 접수, 2014년 2월 3일 수정, 2014년 2월 19일 채택)

요약

본 연구에서는 비동차 마코프 체인에서 개체들의 전이 행태를 분석하기 위한 계층적 베이지안 방법론을 사용하여 혼합 효과 모델을 소개 하였다. 모델의 모수들에 대한 사후분포가 분석적으로 구해질 수 없는 형태를 가지기 때문에 깁스(Gibbs) 샘플링 시뮬레이션 방법을 사용하여 조건부 사후확률로부터 샘플이 추출되었고, 실제 자료분석을 예를 사용하였다.

주요용어: 마코프 체인, 베이지안 분석, 혼합 효과 모델.

¹(443-749) 수원시 원천동 산 5, 아주대학교 경영학과. E-mail: sungmj@ajou.ac.kr