Multi-state Markov Models for Analyzing Disease History Data

Shenghai Zhang
Department of Statistics and Actuarial Sciences
University of Waterloo
Waterloo, Canada N2L3G1
E-Mail shzhang@math.uwaterloo.ca

A number of dependent multi-state processes arise when several response variables are required to measure the outcome of interest.This type of data is quite common in clinical and observational studies. Suppose that R processes are periodically observed on each individual and rth process has K_r states. Let $X_r(t)$ be the state occupied at time t for the rth process by a individual, then $\boldsymbol{X(t)} = (X_1(t), \cdots, X_R(t))'$ is a realization at time t for a individual. The $\{\boldsymbol{X}(t); t \geq 0\}$ is a vector-valued Markov process if

$$\begin{aligned} &Pr\{\boldsymbol{X}(t) = \boldsymbol{v} | \boldsymbol{X}(s) = \boldsymbol{u}, \boldsymbol{X}(\tau), \quad 0 \leq \tau < s\} \\ &= Pr\{\boldsymbol{X}(t) = \boldsymbol{v} | \boldsymbol{X}(s) = \boldsymbol{u}\} = p_{\boldsymbol{uv}}(s,t) \end{aligned} \tag{1}$$

for $s < t$, where $\boldsymbol{u} = (u_1, u_2, \ldots, u_R)'$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_R)'$, u_r and v_r take values of states of the rth process, $r = 1, \ldots, R$. We try to define intensities for a vector-valued Markov process such that the process $\{\boldsymbol{X}(t); t \geq 0\}$ can be specified in terms of the transition intensities.

Let $f_{\boldsymbol{uv}}(t; \Delta t_1, \ldots, \Delta t_R)$ be the probability of the event $\{X_1(t + \Delta t_1) = v_1, \ldots, X_R(t + \Delta t_R) = v_R\}$ given $\{X_1(t) = u_1, \ldots, X_R(t) = u_R\}$. Suppose a patient is in the state $\boldsymbol{u}$ at time t, the probability that the patient's state is changed to the state $\boldsymbol{v}$ in a very short time interval is very small in most cases. Therefore, it is reasonable to assume that the probability $f_{\boldsymbol{u}v}(t, \Delta t_1, \ldots, \Delta t_R)$ is of order $\Delta t_1 \ldots \Delta t_R$. Then we define

$$q_{\boldsymbol{uv}}(t) = \lim_{\max\{\Delta t_r\} \to 0} \frac{f_{\boldsymbol{uv}}(t; \Delta t_1, \ldots, \Delta t_R)}{\Delta t_1 \cdots \Delta t_R} \qquad u_r \neq v_r \text{ for } r = 1, \ldots, R. \tag{2}$$

Now, we consider the case that a patient is in the stat $\boldsymbol{u}$ at time t, after a short time interval, some components of his sate is changed, others are not. *i.e.*: $\boldsymbol{X}(t) = \boldsymbol{u}$ and $\boldsymbol{X}(t + \boldsymbol{\Delta_t}) = \boldsymbol{v}$, where there exist $m_1, \ldots, m_a$ such that the states $u_{m_l} = v_{m_l}$ for $l = 1, \ldots, a$ $(1 \leq a < R)$; $u_i \neq v_i$ otherwise, then we let $g_{\boldsymbol{uv}}(t; \Delta t_{k_1}, \ldots, \Delta t_{k_{R-a}})$ is the probability of $\{X_{k_r}(t + \Delta t_{k_r}) = v_{k_r} \quad | \quad r = 1, \ldots R - a\}$ given $\{X_1(t) = u_1, \ldots, X_R(t) = u_R\}$. We define

$$q_{\boldsymbol{uv}}(t) = \lim_{\max\{\Delta t_r\} \to 0} \frac{g_{\boldsymbol{uv}}(t; \Delta t_{k_1}, \ldots, \Delta t_{k_{R-a}})}{\prod_{r=1}^{R-a} \Delta t_{k_r}}. \tag{3}$$

If $\boldsymbol{u} = \boldsymbol{v}$ then we define $q_{\boldsymbol{uu}}(t) = -\sum_{\boldsymbol{v} \neq \boldsymbol{u}} q_{\boldsymbol{uv}}(t)$. Let $\boldsymbol{\Omega} = \{\boldsymbol{u} = (u_1, \ldots, u_R)' | u_r \in \{1, \ldots, K_r\}, \text{ for } r = 1, \ldots, R\}$, then order the elements of $\boldsymbol{\Omega}$ in some way, say, in the following way: $\boldsymbol{\Omega} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_\sigma\}$

where $\sigma = K_1 \times K_2 \times \cdots \times K_R$. Therefore, for any $\boldsymbol{u} = (u_1, \ldots, u_R)'$ there must exist a $1 \leq i \leq \sigma$ such that $\boldsymbol{u} = (u_1, u_2, \ldots, u_R)' = \boldsymbol{w}_i$ and $\boldsymbol{w}_i \in \boldsymbol{\Omega}$, then we define: $\psi_{ij} = q_{\boldsymbol{w_i w_j}}$ if $\delta_{\boldsymbol{w_i w_j}} = 1$ or 0; $\psi_{ij} = 0$, otherwise. Thus it can be proved that $\frac{\tilde{p}_{ij}(t)}{dt} = \sum_{k=1}^{\sigma} \psi_{ik} \tilde{p}_{kj}(t)$, where $\tilde{p}_{ij}(t) = p_{\boldsymbol{w_i w_j}}(t)$. That is $\boldsymbol{P}'(t) = \boldsymbol{\Psi P}(t)$ where $\boldsymbol{\Psi} = (\psi_{ij})_{\sigma \times \sigma}$. It can be shown that the solution of the matrix differential equations is given by $\boldsymbol{P}(t) = e^{\boldsymbol{\Psi} t}$.

Suppose that lth individual is observed at times $t_{l,0} < t_{l,1} < \cdots < t_{l,m_l}$, and the states observed at these times are

$$\boldsymbol{Y}_{l0} = (y_{l,0}^1, y_{l,0}^2, \ldots, y_{l,0}^R)', \ldots, \boldsymbol{Y}_{l,m_l} = (y_{l,m_l}^1, \ldots, y_{l,m_l}^R)', \qquad l = 1, \ldots, n.$$

So, we get the ith processes: $\{X_l^i(t); t \geq 0\}$ for lth individual ($i = 1, \ldots, R; l = 1 \ldots, n$). Suppose the corresponding states are $1, \ldots, K_i$ for $i = 1, \ldots, R$ and $\{\boldsymbol{X}_l(t); t \geq 0\}$ is a time-homogeneous Markov process, where $\boldsymbol{X}_l(t) = (X_l^1(t), \ldots, X_l^R(t))'$, for $l = 1, \ldots, n$. First of all, we want to get Likelihood function for estimating the parameters $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is of interest and the transition probabilities depend on it. The likelihood function should be as the following

$$L(\boldsymbol{\theta}) = \prod_{l=1}^{n} \prod_{h=1}^{m_l} \prod_{i=1}^{\sigma} \prod_{j=1}^{\sigma} \tilde{p}_{ij}(\Delta t_{lh}; \boldsymbol{\theta})^{\delta_{lhi}} \tag{4}$$

where $\tilde{p}_{ij}(t; \boldsymbol{\theta}) = p_{\boldsymbol{w_i w_j}}(t; \boldsymbol{\theta})$ and $\delta_{lhi} = I\{\boldsymbol{X}(t_{l,h-1}) = \boldsymbol{w_i}\}$.

The MLE $\hat{\boldsymbol{\theta}}$ is obtained by maximizing the $\log L(\boldsymbol{\theta})$. However, whatever the estimate functions we use, we have to compute $p_{ij}^{(r)}(t; \boldsymbol{\theta})$ and $\frac{\partial p_{ij}^{(r)}(t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ or $\tilde{p}_{ij}(t; \boldsymbol{\theta})$ and $\frac{\partial \tilde{p}_{ij}(t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$. It is generally the case that $\boldsymbol{P}(t; \boldsymbol{\theta})$ and $\frac{\partial \boldsymbol{P}(t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is a complicated function of $\boldsymbol{\theta}$ even for very simply situation. Without adopting efficient algorithm it is very difficult to get the estimates. An algorithm provided by Kalbfleish and Lawless (1985) could be used to solve this problem.

REFERENCES

Kalbfleisch, J. D. and Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association* , **80**, 863-871.

Lee, E. W and Kim, Y. M. (1998). The analysis of correlated pannel data using a continuous time Markov Model. *Biometrics*, **54**, 1638-1644.

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

Young, P. J., Weeden, S. and Kirwan, J. R. (1999). The analysis of a bivariate multi-state Markov transition model for rheumatoid arthritis with an incomplete disease history. *Statistics in Medicine*, **18**, 1677-1690.

Gentleman, R. C., Lawless, J. F., Lindsey, J. C. and Yan, P. (1994). Multi-state Markov models for analysing incomplete disease history data with illustrations for HIV disease. *Statistics in Medicine*, **13**, 805-821.