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 $r$th process has $K_r$ states. Let $X_r(t)$ be the state occupied at time $t$ for the $r$th process by an individual, then $X(t) = (X_1(t), \ldots, X_R(t))'$ is a realization at time $t$ for an individual. The $\{X(t); t \geq 0\}$ is a vector-valued Markov process if

\[
Pr\{X(t) = v|X(s) = u, X(\tau), \ 0 \leq \tau < s\} = \Pr\{X(t) = v|X(s) = u\} = p_{uv}(s, t)
\]

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

Let $f_{uv}(t; \Delta t_1, \ldots, \Delta t_R)$ 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 $u$ at time $t$, the probability that the patient’s state is changed to the state $v$ in a very short time interval is very small in most cases. Therefore, it is reasonable to assume that the probability $f_{uv}(t, \Delta t_1, \ldots, \Delta t_R)$ is of order $\Delta t_1 \ldots \Delta t_R$. Then we define

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

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

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

If $u = v$ then we define $q_{uu}(t) = -\sum_{v \neq u} q_{uv}(t)$. Let $\Omega = \{u = (u_1, \ldots, u_R)'|u_r \in \{1, \ldots, K_r\}, \text{ for } r = 1, \ldots, R\}$, then order the elements of $\Omega$ in some way, say, in the following way: $\Omega = \{w_1, w_2, \ldots, w_\sigma\}$.
where $\sigma = K_1 \times K_2 \times \cdots \times K_R$. Therefore, for any $u = (u_1, \ldots, u_R)'$ there must exist a $1 \leq i \leq \sigma$ such that $u = (u_1, u_2, \ldots, u_R)' = w_i$ and $w_i \in \Omega$, then we define: $\psi_{ij} = q_{w_i,w_j}$ if $\delta_{w_i,w_j} = 1$ or 0; $\psi_{ij} = 0$, otherwise. Thus it can be proved that $\dot{\tilde{p}}_{ij}(t) = \sum_{k=1}^{\sigma} \psi_{ik}\tilde{p}_{kj}(t)$, where $\tilde{p}_{ij}(t) = p_{w_i,w_j}(t)$.

That is $P'(t) = \Psi P(t)$ where $\Psi = (\psi_{ij})_{\sigma \times \sigma}$. It can be shown that the solution of the matrix differential equations is given by $P(t) = e^{\Psi t}$.

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

$$Y_{l0} = (y_{l,0}^1, y_{l,0}^2, \ldots, y_{l,0}^R)', \ldots, Y_{l,m_l} = (y_{l,m_l}^1, \ldots, y_{l,m_l}^R)', \quad l = 1, \ldots, n.$$ 

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

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

(4)

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

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

REFERENCES


