Modelling Non-homogeneous Markov Processes via Time Transformation

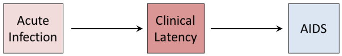
Biostat 572

Jon Fintzi fintzij@uw.edu

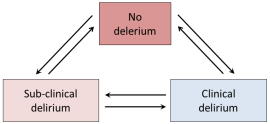
Motivation - Scientific Objectives

Goal: Describe a disease process in terms of its transitions through discrete states.

- Progressive disease: subjects traverse disease states in one direction.
  - Stages of HIV infection



- Non-progressive disease: subjects may visit some or all states repeatedly, just once, or not at all.
  - Delirium



Motivation - Methodological Objectives

Goal: Obtain estimates and covariance matrices for transition intensity parameters in non-homogeneous Markov process models.
Panel Data: Subjects are observed at a sequence of discrete times, observations consist of the states occupied by the subjects at those times.

- The exact transition times are not observed.
- The complete sequence of states visited by a subject may not be known.

Key Assumption: The process is *non-homogeneous* only through a time-varying multiplicative change in the state transition intensities. (*Spoiler alert: this will allow us to model the process on an operational time scale on which we get homogeneity "for free".*)

A Crash Course in Stochastic Processes

Notation, Definitions, and Equations

- $X(t) =$ state occupied at time t, $X(t) \in \mathcal{S} = \{1, \ldots, m\}$
- $p_{ij}(s, t) = \Pr\{X(t) = j | X(s) = i\}$
  $\mathbf{P}(t, t+s) = [p_{ij}(t, t+s)]$ for $s, t \geq 0$
- $q_{ij}(t) = \lim_{s \to 0} \frac{p_{ij}(t, t+s)}{s}$, $q_{ii} = -\sum_{j \neq i} q_{ij}(t)$
  $\mathbf{Q}(t) = [q_i j(t)]$ for $t \geq 0$

A Crash Course in Stochastic Processes

- **definition:** The process is said to be *Markov* if the future state of the process depends only on its current state. Formally,$\forall s, t \geq 0$ and $\forall i, j \in \mathcal{S}$, the process satisfies

$$P\{X(t+s) = j | X(t) = i, X(u) = x(u), 0 \leq u < s\} = p_{ij}(t, t+s)$$

- **definition:** The process is said to be *homogeneous* if the transition probabilities do not depend on the chronological time t. That is,

$$p_{ij}(t, t+s) = p_{ij}(s), \text{ and } q_{ij}(t) = q_{ij}$$

A Crash Course in Stochastic Processes

We would suspect that having a process be both homogeneous and Markov would be a good thing.

- More data available to estimate fewer parameters (Homogeneity)
- Simpler form for the likelihood (Markov)

A Crash Course in Stochastic Processes

If the process is a homogeneous Markov process, it satisfies the *Chapman-Kolmogorov* equation, which describes the transition probabilities in terms of the possible paths the process can take:

$$\begin{aligned} p_{ij}(s) &= \sum_{k \in \mathcal{S}} p_{ik}(u) p_k j(s-u),\ 0 < u < s \\ \text{i.e. } \mathbf{P}(s) &= \mathbf{P}(u)\mathbf{P}(s-u) \end{aligned}$$

From this, we get the *forward* and *backward equations*

$$\begin{aligned} \frac{d}{ds}\mathbf{P}(s) &= \mathbf{QP}(s) = \mathbf{P}(S)\mathbf{Q} \\ \text{so that: } \mathbf{P}(s) &= exp(\mathbf{Q}s) = \sum_{n=0}^{\infty} \frac{(\mathbf{Q}s)^n}{n!} \end{aligned}$$

A Crash Course in Stochastic Processes

Therefore for a homogeneous Markov process, we can equivalently characterize the process by its transition intensity matrix and by its matrix of transition probabilities.

Homogeneity is good. Yay homogeneity!

Key Proposition: Time Transformation

Let u denote the original time scale of the observations. If there exists an invertible transformation of the time scale such that the process is homogeneous on t=*h(u)*, with transition intensity matrix Q_0, then

$$\begin{aligned} \mathbf{P}(u_1, u_2) &= \mathbf{P}(h(u_1), h(u_2)) = \mathbf{P}(t_2 - t_1) \\ &= exp[Q_0(t_2 - t_1)] = exp[Q_0(h(u_2) - h(u_1))] \end{aligned}$$

N.B. $h(u)$ is a time scale, so we require it to be non-negative and have non-negative first derivative.

Likelihood Function

The likelihood for $\mathbf{X} = (X(u_1), \ldots, X(u_n))^T$ is:

$$L(Q_0) = P(X(u_1) = x_1) \prod_{i=2}^{n} p_{x_{i-1}, x_i}(u_{i-1}, u_i)$$

Since the transformed process is homogeneous, we have

$$\begin{aligned} L(Q_0, \theta) &= P(X(u_1) = x_1) \prod_{i=2}^{n} p_{x_{i-1}, x_i}(h(u_i; \theta) - h(u_{i-1}; \theta)) \\ &= P(X(u_1) = x_1) \prod_{i=2}^{n} \left[e^{Q_0(h(u_i;\theta) - h(u_{i-1};\theta))} \right]_{x_{i-1} x_i} \end{aligned}$$

Seminal Paper: *Kalbfleish & Lawless (1985)*

- Derived procedures for obtaining MLEs and covariance matrices for transition intensity parameters in modeling panel data with continuous time homogeneous Markov processes.
- Used Fisher scoring algorithm based on the score function and its first derivatives.

Seminal Paper: *Kalbfleish & Lawless (1985)*

- Assume independence across subjects. Letting ψ be the vector of functionally independent elements of Q_0 and θ, we have that the MLEs are the solutions to $\frac{\partial}{\partial\psi} logL_m(\psi) = 0$ and

$$\hat{\psi} \sim \mathcal{N}(0, \mathcal{I}_m^{-1}(\psi_0)), \text{ where } \mathcal{I}_m = E\left[\left(\frac{\partial}{\partial\psi} logL_m(\psi)\right)\left(\frac{\partial}{\partial\psi} logL_m(\psi)\right)^T\right]$$

- Given initial estimates $\psi^(0)$, the $(k+1)^{st}$ step for the MLEs is

$$\hat{\psi}^{(k+1)} = \hat{\psi}^{(k)} + \hat{\mathcal{I}}_m(\hat{\psi}^{(k)})^{-1} \frac{\partial}{\partial\psi} logL_m(\psi)$$

Next Steps

- More pictures of Cesar.
- Simulation results under exponential and kernel time transformation functions, and comparisons with functions in the *msm* package in R.
- Model results based on real-world data (dataset yet to be selected).
- More pictures of Cesar!