

SIS models for recurrent infections

SISMID/July 13–15, 2015

Instructors: Kari Auranen, Elizabeth Halloran, Vladimir Minin

Outline

- ▶ Background: recurrent infections
- ▶ Binary Markov processes and their generalizations
- ▶ Counting process likelihood
- ▶ Incomplete observations
 - ▶ discrete-time transition models
 - ▶ Bayesian data augmentation and reversible jump MCMC
- ▶ A computer class exercise

Background

- ▶ Many infections can be considered recurrent, i.e., occurring as an alternating series of presence and absence of infection
 - ▶ Nasopharyngeal carriage of *Streptococcus pneumoniae* (Auranen et al.; Cauchemez et al.; Melegaro et al.)
 - ▶ Nasopharyngeal carriage of *Neisseria meningitidis*
 - ▶ multi-resistant *Staphylococcus aureus* (Cooper et al.)
 - ▶ some parasitic infections (e.g. Nagelkerke et al.)
- ▶ Observation of these processes requires active sampling of the underlying epidemiological states
- ▶ Acquisition and clearance times often remain unobserved ⇒ incompletely observed data

A binary Markov process

A simple model for a recurrent infection is the binary Markov process:

- ▶ The state of the individual alternates between “susceptible” (state 0) and “infected” (state 1)
- ▶ The hazard of acquiring infection is λ :

$$P(\text{acq. in } [t, t + dt] | \text{ susceptible at time } t-) = \lambda dt$$

- ▶ The hazard of clearing infection is μ :

$$P(\text{clearance in } [t, t + dt] | \text{ infected at time } t-) = \mu dt$$

The complete data

- ▶ For each individual i , the complete data include the times of acquisition and clearance during the observation period $[0, T]$:
 - ▶ Denote the ordered acquisition times for individual i during $]0, T[$ by $\mathbf{t}^{(i)} = (t_{i1}, \dots, t_{iN_{01}^{(i)}})$
 - ▶ Denote the ordered clearance times for individual i during $]0, T[$ by $\mathbf{r}^{(i)} = (r_{i1}, \dots, r_{iN_{10}^{(i)}})$
 - ▶ Denote the ordered acquisition and clearance times together as $u_{i1} = 0, u_{i2}, \dots, u_{i,N^{(i)}} = T$
 - ▶ Note: these include times 0 and T
(so that $N^{(i)} = N_{01}^{(i)} + N_{10}^{(i)} + 2$)

Keeping track who is susceptible

- ▶ The indicators for individual i to be susceptible or infected at time t are denoted by $S_i(t)$ and $I_i(t)$, respectively
 - ▶ Both indicators are taken to be *predictable*, i.e., their values at time t are determined by the initial value $S_i(0)$ and the complete data observed up to time t —
 - ▶ Note that $I_i(t) = 1 - S_i(t)$ for all times $t \geq 0$

The process of acquisitions

- ▶ In each individual, acquisitions occur with intensity $\lambda S_i(t)$
 - ▶ The intensity is λ when the individual is in state 0 (susceptible) and 0 when the individual is in state 1 (infected)
- ▶ The probability density of the acquisition events is proportional to

$$\prod_{k=1}^{N^{(i)}} \left[\beta \mathbf{1}_{(u_k \text{ is time of acq.})} \exp \underbrace{-\beta S_i(u_k)(u_k - u_{k-1})}_{\text{total time susceptible}} \right]$$
$$\propto \beta^{N_{01}^{(i)}} \times \exp \left[-\beta \sum_{k=1}^{N^{(i)}} S_i(u_k)(u_k - u_{k-1}) \right]$$

The process of clearances

- ▶ In each individual, the clearances occur with intensity $\mu l_i(t)$
 - ▶ The intensity is μ when the individual is in state 1 (infected) and 0 when the individual is in state 0 (susceptible)
- ▶ The probability density of the clearance events is proportional to

$$\prod_{k=1}^{N^{(i)}} \left[\mu^{1(u_k \text{ is time of clearance})} \exp^{-\mu l_i(u_k)(u_k - u_{k-1})} \right]$$
$$= \mu^{N_{10}^{(i)}} \times \exp^{-\mu \overbrace{\sum_{k=1}^{N^{(i)}} l_i(u_k)(u_k - u_{k-1})}^{\text{total time infected}}}$$

The complete data likelihood

- ▶ The likelihood function of parameters β and μ , based on the complete data from individual i :

$$\begin{aligned} & \overbrace{f(\mathbf{t}^{(i)}, \mathbf{r}^{(i)} | \beta, \mu)} \\ & L_i(\beta, \mu; \mathbf{t}^{(i)}, \mathbf{r}^{(i)}) \\ &= \beta^{N_{01}^{(i)}} \mu^{N_{10}^{(i)}} \times \exp \left(- \sum_{k=1}^{M^{(i)}} (\beta S_i(u_k) + \mu I_i(u_k))(u_k - u_{k-1}) \right) \\ &= \beta^{N_{01}^{(i)}} \mu^{N_{10}^{(i)}} \times \exp \left(- \int_0^T \{ \beta S_i(u) + \mu I_i(u) \} du \right) \end{aligned}$$

- ▶ Likelihood for all M individuals is $\prod_{i=1}^M L_i(\beta, \mu; \mathbf{t}^{(i)}, \mathbf{r}^{(i)})$

More complex models

- ▶ In the following six slides, the binary model is formulated as a process of counting transitions “ $0 \rightarrow 1$ ” (acquisitions) and “ $1 \rightarrow 0$ ” (clearances)
- ▶ More complex models can then be defined, allowing e.g.
 - ▶ different (sero)types/strains of infection
 - ▶ taking into account exposure from other individuals in the relevant mixing groups, e.g., modelling transmission in households

A counting process formulation

- ▶ For individual i , the binary process can be described in terms of two counting processes (jump processes):
 - ▶ $N_{01}^{(i)}(t)$ counts the number of acquisitions for individual i from time 0 up to time t
 - ▶ $N_{10}^{(i)}(t)$ counts the number of clearances for individual i from time 0 up to time t
- ▶ Specify the initial state: (e.g.) $N_{01}^{(i)}(0) = N_{10}^{(i)}(0) = 0$
- ▶ Denote $H_t^{(i)}$ the history of the processes up to time t :
 $H_t^{(i)} = \{N_{01}^{(i)}(s), N_{10}^{(i)}(s); 0 \leq s \leq t\}$

Stochastic intensities

- ▶ The two counting processes can be specified in terms of their stochastic intensities:

$$P(dN_{01}^{(i)}(t) = 1 | H_{t-}^{(i)}) = \alpha_{01}^{(i)}(t) Y_0^{(i)}(t) dt$$

$$P(dN_{10}^{(i)}(t) = 1 | H_{t-}^{(i)}) = \alpha_{10}^{(i)}(t) Y_1^{(i)}(t) dt$$

- ▶ Here, $Y_j^{(i)}(t)$ is indicator for individual i being in state j at time t —
- ▶ In the simple Markov model, $\alpha_{01}^{(i)}(t) = \lambda$, $\alpha_{10}^{(i)}(t) = \mu$, $Y_0^{(i)}(t) = S_i(t)$, and $Y_1^{(i)}(t) = I_i(t)$

Several types of infection

- ▶ The infection can involve a “mark”, e.g. the serotype of the infection
 - ▶ $N_{0j}^{(i)}(t)$ counts the number of times that individual i has acquired infection of type j from time 0 up to time t
 - ▶ $N_{j0}^{(i)}(t)$ counts the number of times that individual i has cleared infection of type j from time 0 up to time t
 - ▶ Stochastic intensities can be defined accordingly for all possible transitions between the states. For example, for K serotypes, $\alpha_{rs}^{(i)}(t) Y_r^{(i)}(t)$, $r, s = 0, \dots, K$

Modelling transmission

- ▶ The hazard of infection may depend on the presence of infected individuals in the family, day care group, school class etc.
- ▶ The statistical unit is the relevant mixing group
- ▶ Denote $H_t^{(i, \text{fam})}$ the joint history of all members in the mixing group (e.g. family) of individual i :

$$P(dN^{(i)}(t) = 1 | H_{t-}^{(i, \text{fam})}) = \alpha_{01}^{(i)}(t) S_i(t) dt \equiv \frac{\beta C^{(i)}(t)}{M_{\text{fam}}^{(i)} - 1} S_i(t) dt$$

where $C^{(i)}(t) = \sum_{j=1}^{M_{\text{fam}}^{(i)}} I_j^{(i)}(t)$ is the number of infected individuals in the family of individual i at time $t-$

The counting process likelihood

- ▶ For M individuals followed from time 0 to time T , the *complete data* record all transitions between states 0 and 1 (equivalent to observing all jumps in the counting processes):

$$y_{\text{complete}} = \{T_{rs}^{(ik)}; r, s = 0, 1 (r \neq s), k = 1, \dots, N_{rs}^{(i)}(T), i = 1, \dots, M\}$$

- ▶ The likelihood of the rate parameters θ , based on the complete (event-history) data

$$\underbrace{L(\theta; y_{\text{complete}})}^{f(y_{\text{complete}}|\theta)} = \prod_i^N \prod_{r \neq s} \prod_k^{N_{rs}^{(i)}(T)} \left[\alpha_{rs}^{(i)}(T_{rs}^{(ik)}) \times \exp \left(- \int_0^T \alpha_{rs}^{(i)}(u) Y_r^{(i)}(u) du \right) \right]$$

Remarks

- ▶ The likelihood is valid even when the individual processes are dependent on the histories of *other* individuals, e.g. in the case of modelling transmission (cf. Andersen et al)
- ▶ The likelihood is correctly normalized with respect to any number of events occurring between times 0 and T (cf. Andersen et al)
 - ▶ This is crucial when performing MCMC computations through data augmentation with an unknown number of events

Incomplete observations

- ▶ Usually, we do not observe complete data
- ▶ Instead, the status $y_j^{(i)}$ of each individual is observed at pre-defined times $t_j^{(i)}$
 - ▶ This creates *incomplete data*: the process is only observed at discrete times (panel data)
 - ▶ The observed data likelihood is now a complicated function of the model parameters
- ▶ How to estimate the underlying continuous process from discrete observations?
 - ▶ a discrete-time Markov transition model
 - ▶ Bayesian data augmentation

Markov transition models

- ▶ Treat the problem as a discrete-time Markov transition model
- ▶ This is parameterized in terms of transition probabilities $P(X^{(i)}(t) = s | X^{(i)}(u) = r)$ for all r, s in the state space χ , and for all times $t \geq u \geq 0$
- ▶ In a time-homogeneous model the transition probabilities depend only on the time difference:

$$p_{rs}(t) = P(X^{(i)}(t) = s | X^{(i)}(0) = r)$$

- ▶ This defines a transition probability matrix P_t with entries $[P_t]_{rs} = p_{rs}(t)$, where $\sum_s p_{rs}(t) = 1$ for all r and all $t \geq 0$

The likelihood

- ▶ When observations $y_j^{(i)}$ are made at equal time intervals (Δ), the likelihood is particularly simple

$$L(P_\Delta) = \prod_{r,s} [p_{rs}(\Delta)]^{N_{rs}(T)} = \prod_{r,s} [P_\Delta]_{rs}^{N_{rs}(T)}$$

- ▶ When observation are actually made at intervals $k\Delta$ only (e.g. $\Delta = \text{day}$ and $k = 28$), the likelihood is

$$L(P_\Delta) = \prod_{r,s} [P_\Delta^k]_{rs}^{N_{rs}(T)}$$

Modeling transmission

- ▶ In a mixing group of size M , the state space is $\chi_1 \times \chi_2 \times \dots \times \chi_M$
 - ▶ For example, in a family of three the states then are: (0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), (1,1,1)
 - ▶ For M individuals, the dimension of the state space is 2^M
- ▶ Application to pneumococcal carriage in families (Melegaro et al.)
 - ▶ The transition probability matrix in a family of 3 (next page), assuming the same probabilities (per day) for each family member
 - ▶ Notation: $q_{ii} = 1$ - the sum of the i th row

Transition probability matrix

$$P_{\Delta} = \begin{matrix} & \begin{matrix} (0,0,0) & (1,0,0) & (0,1,0) & (0,0,1) & (1,1,0) & (1,0,1) & (0,1,1) & (1,1,1) \end{matrix} \\ \begin{pmatrix} q_{11} & \kappa & \kappa & \kappa & 0 & 0 & 0 & 0 \\ \mu & q_{22} & 0 & 0 & \lambda/2 + \kappa & \lambda/2 + \kappa & 0 & 0 \\ \mu & 0 & q_{33} & 0 & \beta/2 + \kappa & 0 & \beta/2 + \kappa & 0 \\ \mu & 0 & 0 & q_{44} & 0 & \lambda/2 + \kappa & \lambda/2 + \kappa & 0 \\ 0 & \mu & \mu & 0 & q_{55} & 0 & 0 & \lambda + \kappa \\ 0 & \mu & 0 & \mu & 0 & q_{66} & 0 & \lambda + \kappa \\ 0 & 0 & \mu & \mu & 0 & 0 & q_{77} & \lambda + \kappa \\ 0 & 0 & 0 & 0 & \mu & \mu & \mu & q_{88} \end{pmatrix} \end{matrix}$$

Potential problems

- ▶ The dimension of the state space
 - ▶ With M individuals and $K + 1$ types of infection, the dimension of the state space is $(K + 1)^M$
 - ▶ With 13 serotypes and 25 individuals (see Hoti et al.), the dimension is $\sim 4.5 \times 10^{28}$
- ▶ Non-Markovian sojourn times
 - ▶ e.g. a Weibull duration of infection may be more realistic than the exponential one
- ▶ Handling of varying observation intervals and individuals with completely missing data are still cumbersome

Bayesian data augmentation

- ▶ Retaining the continuous-time model formulation, the unknown event times are taken as additional model unknowns (parameters)
- ▶ Statistical inference on all model unknowns (θ and y_{complete})

$$\overbrace{f(y_{\text{observed}} | y_{\text{complete}})}^{\text{observation model}} \quad \overbrace{f(y_{\text{complete}} | \theta)}^{\text{complete data likelihood}} \quad \overbrace{f(\theta)}^{\text{prior}}$$

- ▶ The observation model often only ensures agreement with the observed data (as an indicator function)
- ▶ The computational problem:
how to sample from $f(y_{\text{complete}} | y_{\text{observed}}, \theta)$?

The sampling algorithm

- ▶ Initialize the model parameters and the latent processes
- ▶ For each individual, update the latent processes
 - ▶ Update the event times using standard MH
 - ▶ Add/delete episodes using reversible jump MH
 - ▶ with 0.5 probability propose to add a new episode
 - ▶ with 0.5 probability propose to delete an existing episode
- ▶ Update the model parameters using single-step MH
- ▶ Iterate the updating steps for a given number of MCMC iterations
 - ▶ See the computer class exercise

Adding/deleting episodes

- ▶ Choose one interval at random from among the K sampling intervals (see page+2)
- ▶ Choose to add an episode (delete an existing episode) within the chosen interval with probability $\pi_{\text{add}} = 0.5$ ($\pi_{\text{delete}} = 0.5$)
 - ▶ If 'add', choose random event times $\bar{t}_1 < \bar{t}_2$ uniformly from Δ (= the length of the sampling interval). These define the new episode.
 - ▶ If 'delete', delete the two event times
- ▶ The 'add' move is accepted with probability ("acceptance ratio")

$$\min \left(\frac{f(y_{\text{observed}} | y_{\text{complete}}^*) f(y_{\text{complete}}^* | \theta) q(y_{\text{complete}} | y_{\text{complete}}^*)}{f(y_{\text{observed}} | y_{\text{complete}}) f(y_{\text{complete}} | \theta) q(y_{\text{complete}}^* | y_{\text{complete}})}, 1 \right)$$

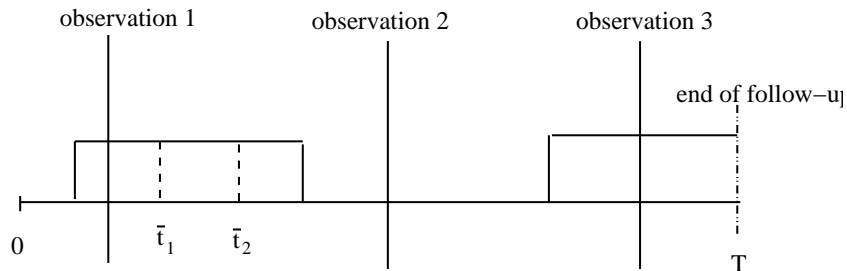
Adding/deleting episodes cont.

- ▶ The ratio of the proposal densities is

$$\frac{q(y_{\text{complete}} | y_{\text{complete}}^*)}{q(y_{\text{complete}}^* | y_{\text{complete}})} = \frac{\pi_{\text{delete}} \frac{1}{K} \frac{1}{L}}{\pi_{\text{add}} \frac{1}{K} \frac{1}{L} \frac{2}{\Delta^2}} = \frac{\Delta^2}{2}$$

- ▶ The ratio of the proposal densities in the 'delete' move is the inverse of the expression above
- ▶ Technically, the add/delete step relies on so called reversible jump MCMC (see page+2)
- ▶ Reversible jump types should be devised to assure irreducibility of the Markov chain
- ▶ For a more complex example, see Hoti et al.

Adding/deleting latent processes cont.



The number of sampling intervals $K=4$

The number of 'sub-episodes' within the second interval $L=2$

Reversible jump MCMC

- ▶ “When the number of things you don’t know is one of the things you don’t know”
- ▶ For example, under incomplete observation of the previous (Markov) processes, the exact number of events is not observed
- ▶ This requires a joint model over ‘sub-spaces’ of different dimensions
- ▶ And a method to do numerical integration (MCMC sampling) in the joint state space

References

- [1] Andersen et al. "Statistical models based on counting processes", Springer, 1993
- [2] Auranen et al. "Transmission of pneumococcal carriage in families – a latent Markov process model for binary data. J Am Stat Assoc 2000; 95:1044-1053.
- [3] Melegaro et al. Estimating the transmission parameters of pneumococcal carriage in families. Epidemiol Infect 2004; 132:433-441.
- [4] Cauchemez et al. Streptococcus pneumoniae transmission according to inclusion in conjugate vaccines: Bayesian analysis of a longitudinal follow-up in schools. BMC Infectious Diseases 2006, 6:14.
- [5] Nakelkerke et al. Estimation of parasitic infection dynamics when detectability is imperfect. Stat Med 1990; 9:1211-1219.
- [6] Cooper et al. "An augmented data method for the analysis of nosocomial infection data. Am J Epidemiol 2004; 168:548-557.
- [7] Bladt et al. "Statistical inference for discretely observed Markov jump processes. J R Statist Soc B 2005; 67:395-410.
- [8] Andersen et al. Multi-state models for event history analysis. Stat Meth Med Res 2002; 11:91-115.
- [9] Hoti et al. Outbreaks of Streptococcus pneumoniae carriage in day care cohorts in Finland – implications to elimination of carriage. BMC Infectious Diseases, 2009 (in press)
- [10] Green P. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 1995; 82:711-732.