Multi-state Models for Intermittently Observed Data

Chris Jackson
MRC Biostatistics Unit, University of Cambridge, UK
chris.jackson@mrc-bsu.cam.ac.uk

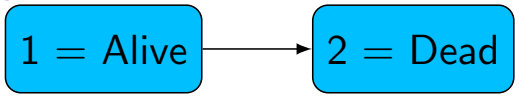
May 26, 2021
INSERM Workshop
Recent advances in statistical analysis of survival data

Contents

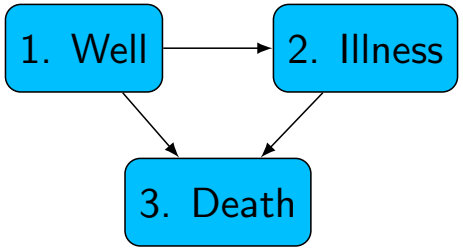
Examples of multi-state models

General discrete(finite)-state, continuous-time stochastic process

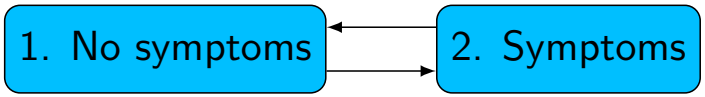
Survival model



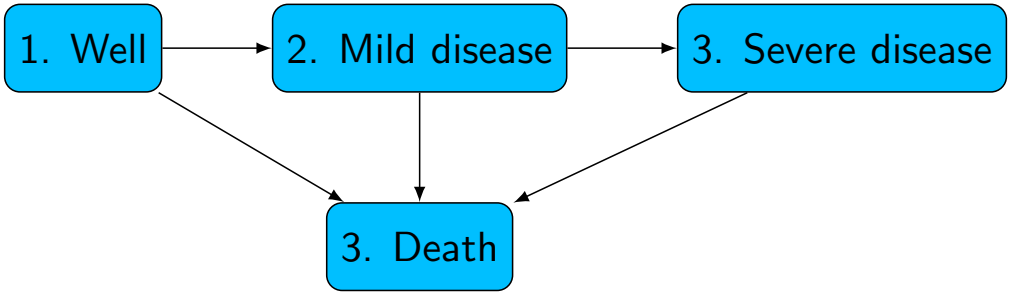
Illness-death model



Relapsing and remitting non-fatal condition



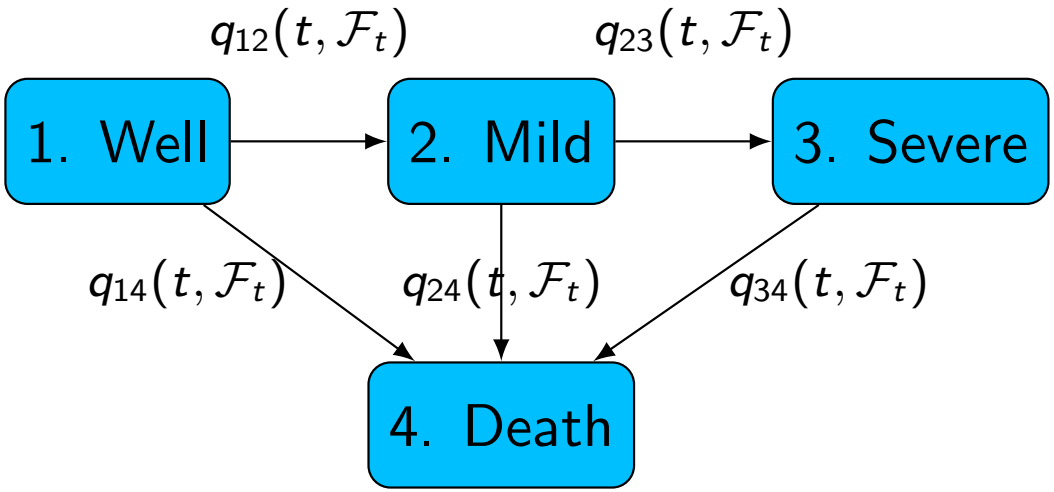
Staged disease progression model



- Many situations where we represent a disease as a multi-state process
- States chosen from clinical convention, or corresponding to measured data, though disease severity may be continuous

Contents

Transition intensities

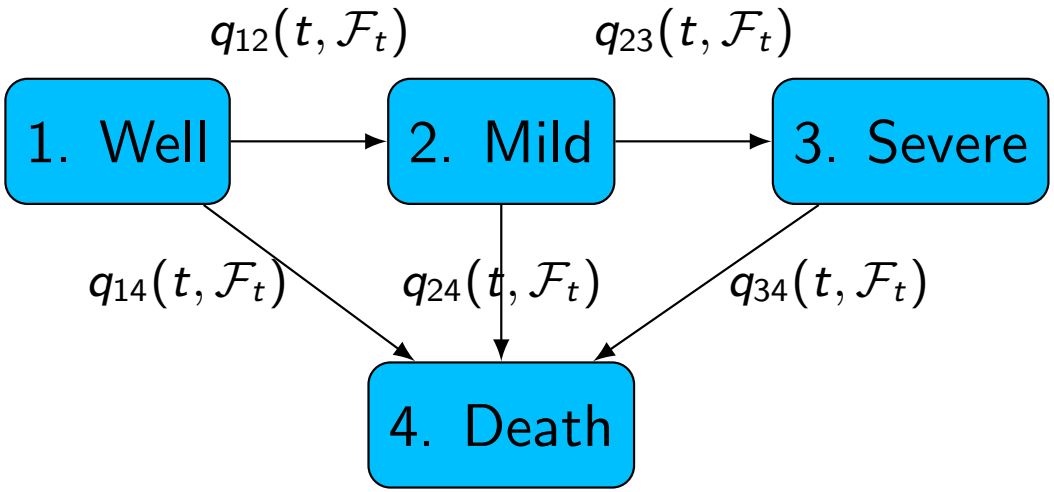


Multi-state models are fully specified by their transition intensities $q_{rs}(t, \mathcal{F}_t) : r \neq s, r, s = 1, ..., R$.
Rate of transition to state s for someone currently in state r.

$$q_{rs}(t; \mathcal{F}_t) = \lim_{\delta t \to 0} \frac{P(X(t + \delta t) = s | X(t) = r, \mathcal{F}_t)}{\delta t}, r \neq s, r, s = 1, ..., R$$

- Not a probability but a rate
- May depend on current time t and also the history $\mathcal{F}_t$

Transition intensities



Multi-state models are fully specified by their transition intensities $q_{rs}(t, \mathcal{F}_t) : r \neq s, r, s = 1, ..., R$.
Rate of transition to state s for someone currently in state r.

$$q_{rs}(t; \mathcal{F}_t) = \lim_{\delta t \to 0} \frac{P(X(t + \delta t) = s | X(t) = r, \mathcal{F}_t)}{\delta t}, r \neq s, r, s = 1, ..., R$$

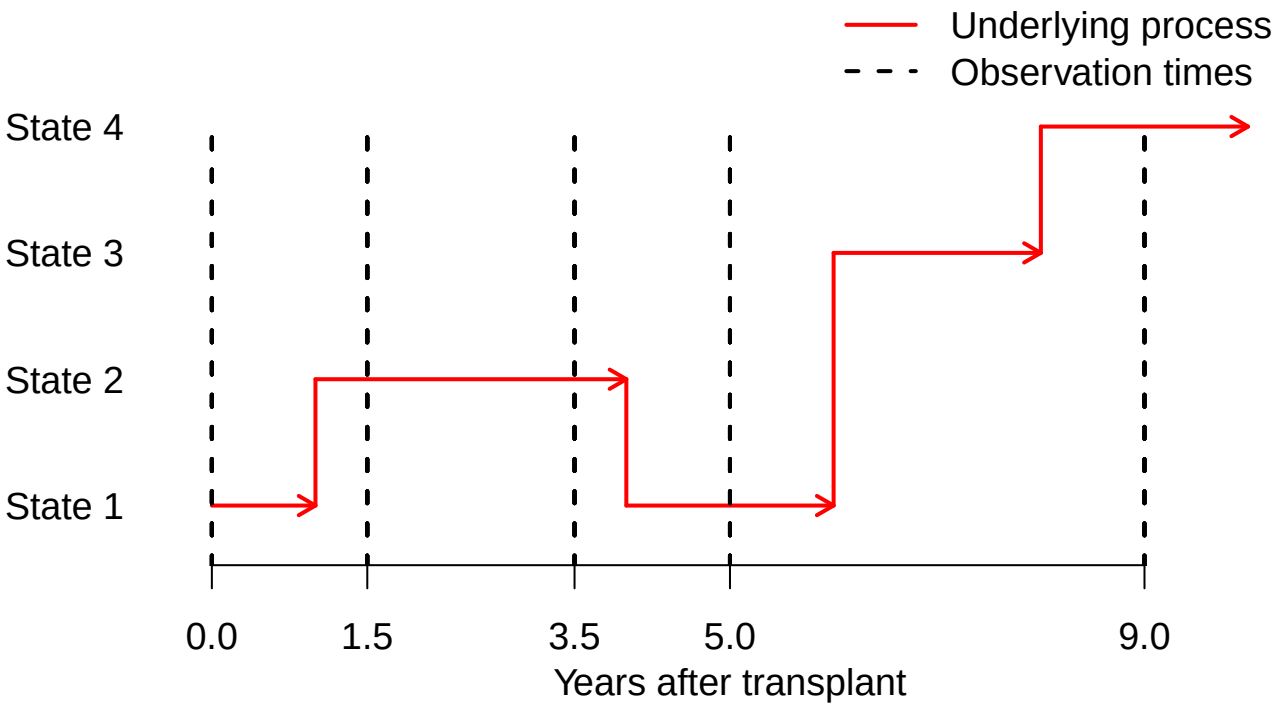
- Not a probability but a rate
- May depend on current time t and also the history $\mathcal{F}_t$

Context for this talk (previous sessions)

Previous talks dealt with fitting multi-state models to data where:

- we know the state at all times
- that is, the process is continuously observed
- Data can be represented as times to events
  - or times from one event to another
  - potentially with censoring
- Multi-state modelling methods can be constructed as generalisations of survival analysis.

Context for this talk



Instead, this talk will mainly deal with intermittently observed data

- Only know the state at a finite series of times

To be able to fit the models to data, first we'll need some more probability theory...

Contents

Markov processes

If $q_{rs}(t; \mathcal{F}_t) = q_{rs}(t)$ where t is time since start of process
i.e. no dependence on history $\mathcal{F}_t$:

- time spent in the current state
- states visited previously by the individual and time spent in them

then the process is a (continuous-time) Markov process, or Markov model

Extensive theory developed for continuous-time Markov models, e.g. textbooks

- Cox & Miller *Theory of Stochastic Processes*
- Norris *Markov Chains*
- Kulkarni *Modeling and Analysis of Stochastic Systems*

Other special cases and variants

Semi-Markov model $q_{rs}(t; \mathcal{F}_t)$ depends only on $\mathcal{F}_t$ through time since entered current state,

- Most general non-Markov model considered in this talk.

Time-homogeneous Markov model $q_{rs}(t; \mathcal{F}_t) = q_{rs}$

- transition rate is constant (over some time period)
- piecewise-constant models are important with intermittent observation...

Dependence on covariates $\mathbf{z}(t)$: $q_{rs}(\mathbf{z}(t); \mathcal{F}_t)$

- covariates could be constant or time-varying
- proportional hazards common
  $q_{rs}(\mathbf{z}(t); \mathcal{F}_t) = q_{rs}^{(0)}(t, \mathcal{F}_t) \exp(\beta' \mathbf{z}(t))$

Transition intensity matrix in a time-homogeneous Markov model

Transition intensity matrix $\mathbf{Q}$: r, s entry equals the intensity q_{rs}

$$\begin{bmatrix} \cdot & q_{12} & q_{13} & \cdots & q_{1n} \\ q_{21} & \cdot & q_{23} & \cdots & q_{2n} \\ q_{31} & q_{32} & \cdot & \cdots & q_{3n} \\ \vdots & & & \ddots & \vdots \end{bmatrix}$$

Additionally define the diagonal entries $q_{rr} = -\sum_{s \neq r} q_{rs}$, so that rows of Q sum to zero. Then we have:

- Sojourn time T_r (spent in state r before moving) has exponential distribution
  - rate $-q_{rr}$, i.e. mean $-1/q_{rr}$.
  - Pr(still in state r in t units' time) $P(T > t)$? $\exp(q_{rr} t)$
- Pr(next state is s | in r now)? $-q_{rs}/q_{rr} = q_{rs} / \sum_{j \neq r} q_{rj}$

Transition intensity matrix in a time-homogeneous Markov model

Transition intensity matrix $\mathbf{Q}$: r, s entry equals the intensity q_{rs}

$$\begin{bmatrix} q_{11} = -\sum_{s \neq 1} q_{1s} & q_{12} & q_{13} & \cdots & q_{1n} \\ q_{21} & q_{22} = -\sum_{s \neq 2} q_{2s} & q_{23} & \cdots & q_{2n} \\ \vdots & q_{32} & \ddots & \cdots & q_{3n} \end{bmatrix}$$

Additionally define the diagonal entries $q_{rr} = -\sum_{s \neq r} q_{rs}$, so that rows of Q sum to zero. Then we have:

- Sojourn time T_r (spent in state r before moving) has exponential distribution
  - rate $-q_{rr}$, i.e. mean $-1/q_{rr}$.
  - Pr(still in state r in t units' time) $P(T > t)$? $\exp(q_{rr} t)$
- Pr(next state is s | in r now)? $-q_{rs}/q_{rr} = q_{rs} / \sum_{j \neq r} q_{rj}$

Transition intensity matrix in a time-homogeneous Markov model

Transition intensity matrix $\mathbf{Q}$: r, s entry equals the intensity q_{rs}

$$\begin{bmatrix} q_{11} = -\sum_{s \neq 1} q_{1s} & q_{12} & q_{13} & \cdots & q_{1n} \\ q_{21} & q_{22} = -\sum_{s \neq 2} q_{2s} & q_{23} & \cdots & q_{2n} \\ \vdots & q_{32} & \ddots & \cdots & q_{3n} \end{bmatrix}$$

Additionally define the diagonal entries $q_{rr} = -\sum_{s \neq r} q_{rs}$, so that rows of Q sum to zero. Then we have:

- Sojourn time T_r (spent in state r before moving) has exponential distribution
  - rate $-q_{rr}$, i.e. mean $-1/q_{rr}$.
  - Pr(still in state r in t units' time) $P(T > t)$? $\exp(q_{rr} t)$
- Pr(next state is s | in r now)? $-q_{rs}/q_{rr} = q_{rs}/\sum_{j!=r} q_{rj}$

Transition intensity matrix in a time-homogeneous Markov model

Transition intensity matrix $\mathbf{Q}$: r, s entry equals the intensity q_{rs}

$$\begin{bmatrix} q_{11} = -\sum_{s \neq 1} q_{1s} & q_{12} & q_{13} & \cdots & q_{1n} \\ q_{21} & q_{22} = -\sum_{s \neq 2} q_{2s} & q_{23} & \cdots & q_{2n} \\ \vdots & q_{32} & \ddots & \cdots & q_{3n} \end{bmatrix}$$

Additionally define the diagonal entries $q_{rr} = -\sum_{s \neq r} q_{rs}$, so that rows of Q sum to zero. Then we have:

- Sojourn time T_r (spent in state r before moving) has exponential distribution
  - rate $-q_{rr}$, i.e. mean $-1/q_{rr}$.
  - Pr(still in state r in t units' time) $P(T > t)$? $\exp(q_{rr} t)$
- Pr(next state is s | in r now)? $-q_{rs}/q_{rr} = q_{rs}/\sum_{j!=r} q_{rj}$

Transition probability matrix $P(t)$

Probability of being in some state at a specific time in the future

$$p_{rs}(t_0, t) = P(\text{state } s \text{ at time } t_0 + t \mid \text{state } r \text{ at time } t_0)$$

$P(t_0, t)$: matrix with r, s entry $p_{rs}(t_0, t)$, solves the Kolmogorov forward equation in terms of intensities $Q(t)$

$$\frac{dP(t_0, t)}{dt} = P(t_0, t)Q(t_0 + t) \quad \text{with} \quad P(t_0, t_0) = I$$

If Q is time-homogeneous over desired time interval

- where $Q(t) = Q$ and $p_{rs}(t_0, t) = p_{rs}(t)$ is independent of t_0
- explicit solution using the matrix exponential

$$P(t) = Exp(tQ) = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n$$

(can be computed with specialised numerical methods, or element-wise solution for each $p_{rs}(t)$ for some simpler models)

Transition probability matrix $P(t)$

Probability of being in some state at a specific time in the future

$$p_{rs}(t_0, t) = P(\text{state } s \text{ at time } t_0 + t \mid \text{state } r \text{ at time } t_0)$$

$P(t_0, t)$: matrix with r, s entry $p_{rs}(t_0, t)$, solves the Kolmogorov forward equation in terms of intensities $Q(t)$

$$\frac{dP(t_0, t)}{dt} = P(t_0, t)Q(t_0 + t) \quad \text{with} \quad P(t_0, t_0) = I$$

If Q is time-homogeneous over desired time interval

- where $Q(t) = Q$ and $p_{rs}(t_0, t) = p_{rs}(t)$ is independent of t_0
- explicit solution using the matrix exponential

$$P(t) = Exp(tQ) = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n$$

(can be computed with specialised numerical methods, or element-wise solution for each $p_{rs}(t)$ for some simpler models)

Transition probability matrix $P(t)$

Probability of being in some state at a specific time in the future

$$p_{rs}(t_0, t) = P(\text{state } s \text{ at time } t_0 + t \mid \text{state } r \text{ at time } t_0)$$

$P(t_0, t)$: matrix with r, s entry $p_{rs}(t_0, t)$, solves the Kolmogorov forward equation in terms of intensities $Q(t)$

$$\frac{dP(t_0, t)}{dt} = P(t_0, t)Q(t_0 + t) \quad \text{with} \quad P(t_0, t_0) = I$$

If Q is time-homogeneous over desired time interval

- where $Q(t) = Q$ and $p_{rs}(t_0, t) = p_{rs}(t)$ is independent of t_0
- explicit solution using the matrix exponential

$$P(t) = Exp(tQ) = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n$$

(can be computed with specialised numerical methods, or element-wise solution for each $p_{rs}(t)$ for some simpler models)

Time-varying transition intensities

Transition intensities $Q(t)$ commonly not constant with t .

If they can be modelled as piecewise constant

- Compute $P(t_0, t) = P(t_0, t_1)P(t_1, t_2) \dots P(t_{n-1}, t_0 + t)$ over n intervals where Q is constant

More generally if $Q(t)$ not constant, compute $P(t_0, t)$ by solving the Kolmogorov equation numerically

$$\frac{dP(t_0, t)}{dt} = P(t_0, t)Q(t + t_0) \quad \text{with} \quad P(t_0, t_0) = I$$

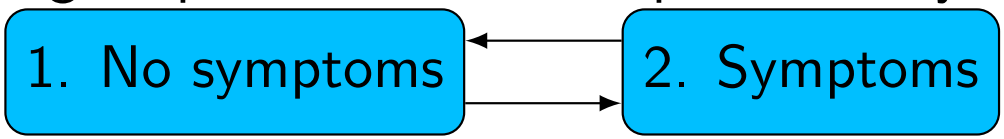
- e.g. Titman, Biometrics 67:780–7, 2011
- `deSolve` R package

Quantities of interest in a multi-state model

e.g. for person in state r at current time $t = 0$, total time they are expected to spend in state s before time t is

$$E\left\{\int_0^t I_{X(u)=s}\,du\right\} = \int_0^t p_{rs}(u)\,du$$

e.g. expected total time spent with symptoms over 10 years in



Other quantities of interest, e.g.

- expected first passage time (time to reach particular state)
- expected number of visits to a state

straightforward functions of transition probabilities or intensities.

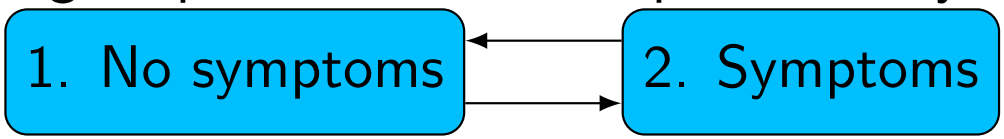
van Loan (1978) IEEE Trans Automatic Control 23(3)395–404
van Rosmalen et al. (2013) Med. Decis. Making 33:767-779
`msm` package documentation

Quantities of interest in a multi-state model

e.g. for person in state r at current time $t = 0$, total time they are expected to spend in state s before time t is

$$E\left\{\int_0^t I_{X(u)=s}\,du\right\} = \int_0^t p_{rs}(u)\,du$$

e.g. expected total time spent with symptoms over 10 years in



Other quantities of interest, e.g.

- expected first passage time (time to reach particular state)
- expected number of visits to a state

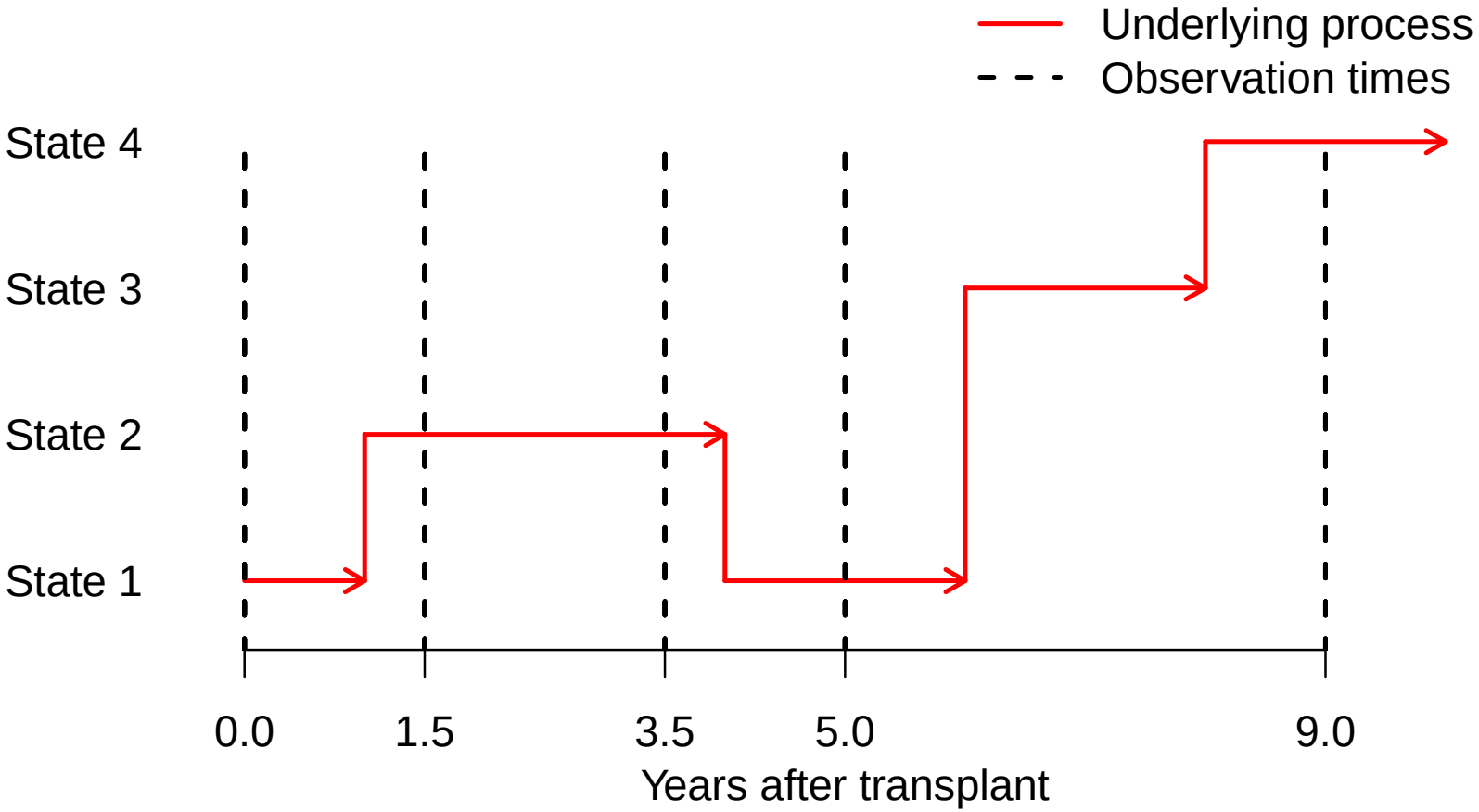
straightforward functions of transition probabilities or intensities.

van Loan (1978) IEEE Trans Automatic Control 23(3)395–404
van Rosmalen et al. (2013) Med. Decis. Making 33:767-779
`msm` package documentation

Contents

Contents

"Panel data" or intermittent observation



- For each patient $i = 1, \ldots, N$, only observe state at a finite series of times. Common in chronic disease modelling.

Likelihood for panel data

(Kalbfleisch and Lawless, JASA 1985)

- Data $\mathbf{x}$: states $(x_{i1}, \ldots, x_{in_i})$ at times $(t_{i1}, \ldots, t_{in_i})$ for person i
  Conditional on state at $t_{i0} = 0$ (e.g. $x_{i0} = 1$)
- Parameters: $\boldsymbol{\theta} = \{q_{rs}\}$: transition intensities of Markov model

Likelihood contribution for person i is product of transition probabilities

$$\begin{aligned} L_i(\boldsymbol{\theta}|\mathbf{x}_i) &= p(x_{i1}|x_{i0})p(x_{i2}|x_{i1})\ldots p(x_{in_i}|x_{i,n_i-1}) \\ &= \prod_{j=1}^{n_i} p_{x_{i,j-1},x_{ij}}(t_{i,j-1}, t_{ij}|\boldsymbol{\theta}) \end{aligned}$$

- Markov assumption $\rightarrow x_{i,j+1}|x_{i,j}$ indep. of $x_{i,1}\ldots x_{i,j-1}$
- Intensities constant within each interval $(t_{i,j-1}, t_{i,j})$: each transition probability is a closed form function of $\boldsymbol{\theta}$.

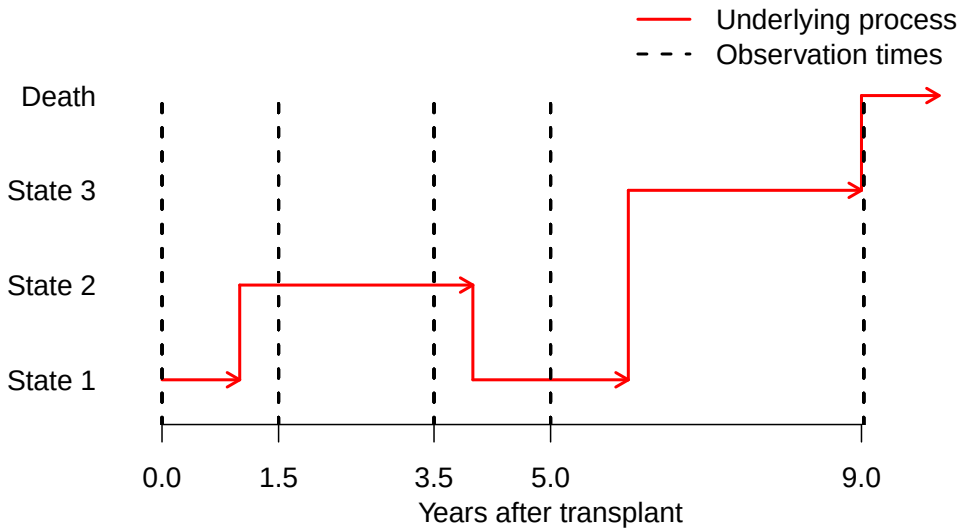
Maximum likelihood estimation

- Maximise the likelihood as a function of transition intensities (and covariate effects on these, see later)
- Then we can compute quantities of interest as functions of intensities:
  - transition probabilities
  - expected total time spent in a state
  - etc.... for different covariate values
- Confidence intervals can be computed by simulation, given the Hessian at the maximum likelihood
- All models from this section implemented in the `msm` package for R (see practical course in September)

Several variants of this basic data and model...

Contents

Panel data with exact death time observed



- For patients who die, day of death known, so assume death time known exactly, but state s at previous instant is unknown
- Get likelihood contribution for final time interval (t_{i,n_i-1}, t_{i,n_i}) by summing over "alive" states s:

$$\sum_s p_{r,s}(t_{in_i} - t_{i,n_i-1})q_{sD}, \qquad r = x_{i,n_i-1}$$

Data with all transition times known?

We could accumulate the likelihood from terms that look like

- P(move out of state r at a time t_{rs} after entering state r)
- P(this move is to state s)

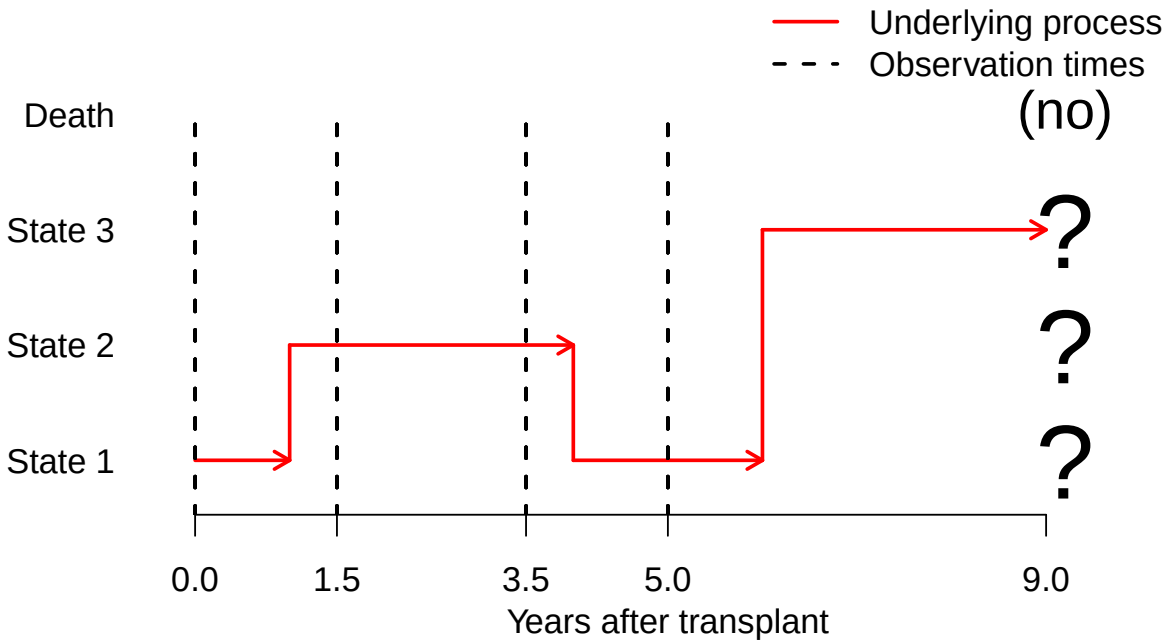
or equivalent formulations.

But these can be calculated easily from any parametric survival distribution — no need to assume piecewise-constant intensities

See framework from other talks based on survival modelling / competing risks ideas.

For fully-parametric modelling of this kind of data, see e.g. `flexsurv` R package, or Stata `multistate` package.

Likelihood with partially-known ("censored") state



- Suppose patient i known to be alive at final time t_{in_i}, but in an unknown disease state x_{in_i}
- Then likelihood contribution for the final transition is $\sum_{s \neq D} p_{r,s}(t_{in_i} - t_{i,n_i-1}) = 1 - p_{r,D}(t_{in_i} - t_{i,n_i-1}), \quad r = x_{i,n_i-1}$
- Generalises easily to partially-known intermediate states

Covariates

Intensities typically given log-linear model

$$q_{rs}(\mathbf{z}_i) = q_{rs}^{(0)} \exp\left(\sum_{m=1}^{M} \beta_m z_{ijm}\right)$$

- $\exp(\beta_m)$ is hazard ratio for mth covariate for ith individual's jth observation
- Covariates could be constant for all times j for individual i
- Any time-varying covariates assumed piecewise-constant:
  - Easiest if constant within each observation interval $(t_{i,j-1}, t_{i,j})$
  - Or if state unknown at a time when covariate changes: integrate over this state using "censored state" likelihood (previous slide)

Time-dependent transition intensities

Covariates could include time itself, giving a time-inhomogeneous model, e.g.

- as a categorical variable, with different intensity estimated for a series of time periods
- through a piecewise-constant approximation to a standard parametric hazard function, e.g. in

$$q_{rs}(\mathbf{z}_i) = q_{rs}^{(0)} \exp\left(\sum_{m=1}^{M} \beta_m z_{ijm}\right)$$

setting $\left.\begin{array}{lcl} z_{ijm} & = & log(t_{ij}) \\ z_{ijm} & = & t_{ij} \end{array}\right\}$ approximates a $\left\{\begin{array}{l} \text{Weibull} \\ \text{Gompertz} \end{array}\right.$ model for the time of transition from state r to s[1]

More general, non piecewise-constant models, possible in principle, though not in `msm` (would need numerical ODE solvers).

[1] van den Hout (2017) *Multi-State Survival Models for Interval-Censored Data*.

Time-dependent transition intensities

Covariates could include time itself, giving a time-inhomogeneous model, e.g.

- as a categorical variable, with different intensity estimated for a series of time periods
- through a piecewise-constant approximation to a standard parametric hazard function, e.g. in

$$q_{rs}(\mathbf{z}_i) = q_{rs}^{(0)} \exp\left(\sum_{m=1}^{M} \beta_m z_{ijm}\right)$$

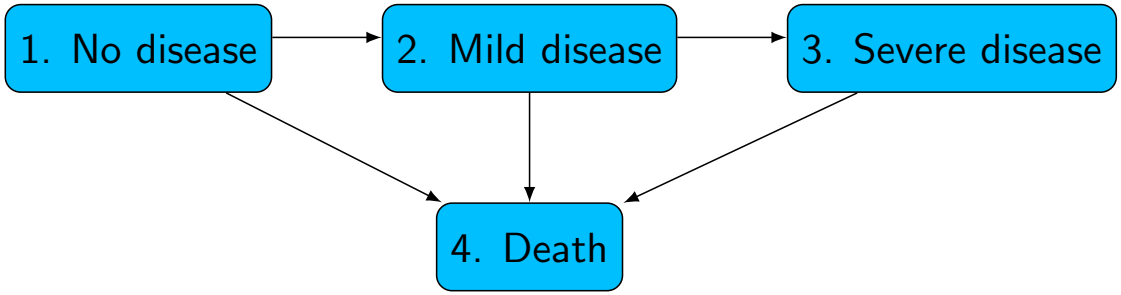
setting $\left.\begin{array}{lcl} z_{ijm} & = & log(t_{ij}) \\ z_{ijm} & = & t_{ij} \end{array}\right\}$ approximates a $\left\{\begin{array}{l} \text{Weibull} \\ \text{Gompertz} \end{array}\right.$ model for the time of transition from state r to s[1]

More general, non piecewise-constant models, possible in principle, though not in `msm` (would need numerical ODE solvers).

[1] van den Hout (2017) *Multi-State Survival Models for Interval-Censored Data*.

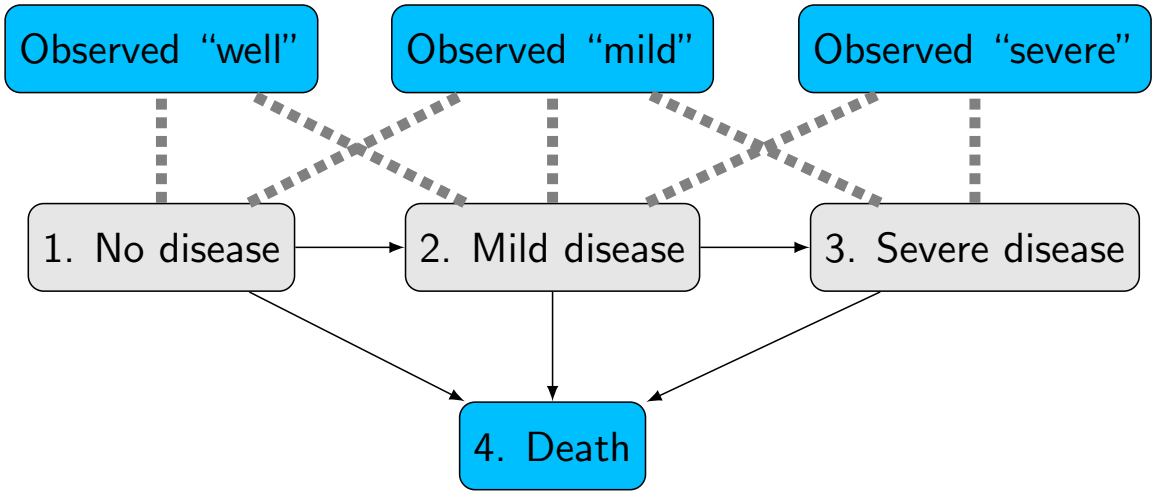
Contents

Hidden Markov models: misclassification of states



- What if disease is known to be an irreversible process...
- ...but transitions from worse to better states observed?
- Screening test may be subject to misclassification

Hidden Markov models: misclassification of states



e.g, could assume
$e_{13} = e_{31} = 0$

- What if disease is known to be an irreversible process...
- ...but transitions from worse to better states observed?
- Screening test may be subject to misclassification

Assume observed state $O(t)$ generated with error given true states $X(t)$

$$e_{rs} = P(O(t) = s | X(t) = r), \qquad e_{rr} = 1 - \sum_{s!=r} e_{rs}$$

True states follow a hidden multi-state model (e.g. hidden Markov)

Misclassification likelihood

Individual i, observed states O_{ij} at time t_{ij}, likelihood contribution is summed over all possible pathways through true states X

$$L_i = P(O_{i0}, \ldots, O_{in_i}) = \sum_{\{X\}} P(O_{i0}, \ldots, O_{in_i} | X_{i0}, \ldots, X_{in_i}) P(X_{i0}, \ldots, X_{in_i})$$

If the $O_{ij}|X_{ij}$ conditionally independent and true process is Markov

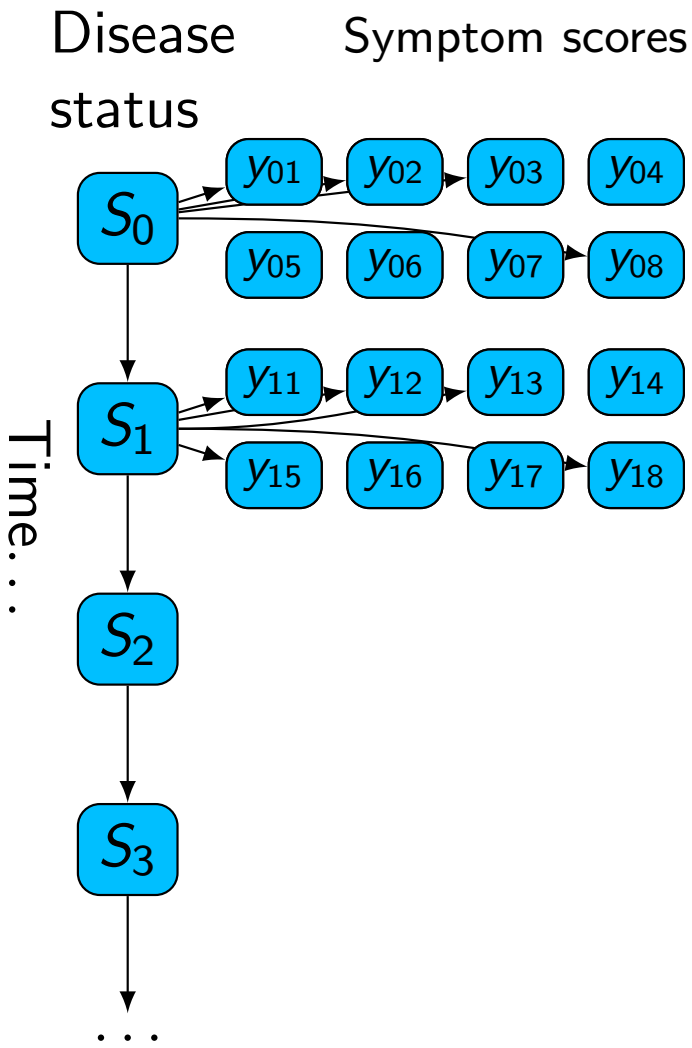
$$L_i = \sum_{X_{i0}} P(O_{i0}|X_{i0}) P(X_{i0}) \sum_{X_{i1}} P(O_{i1}|X_{i1}) P(X_{i1}|X_{i0}) \ldots \sum_{X_{in_i}} P(O_{in_i}|X_{in_i}) P(X_{in_i}|X_{i,n_i-1})$$

Maximise likelihood as a function of transition rates Q, misclassification probabilities $\mathbf{e}$, any covariate effects

- Initial true state probabilities $P(X_{i0})$ can also be estimated, or fixed. For example, could know $X_{i0} = 1$: everyone disease-free at start

Jackson et al (Statistician 2003), after Satten and Longini (Appl. Stat 1997),

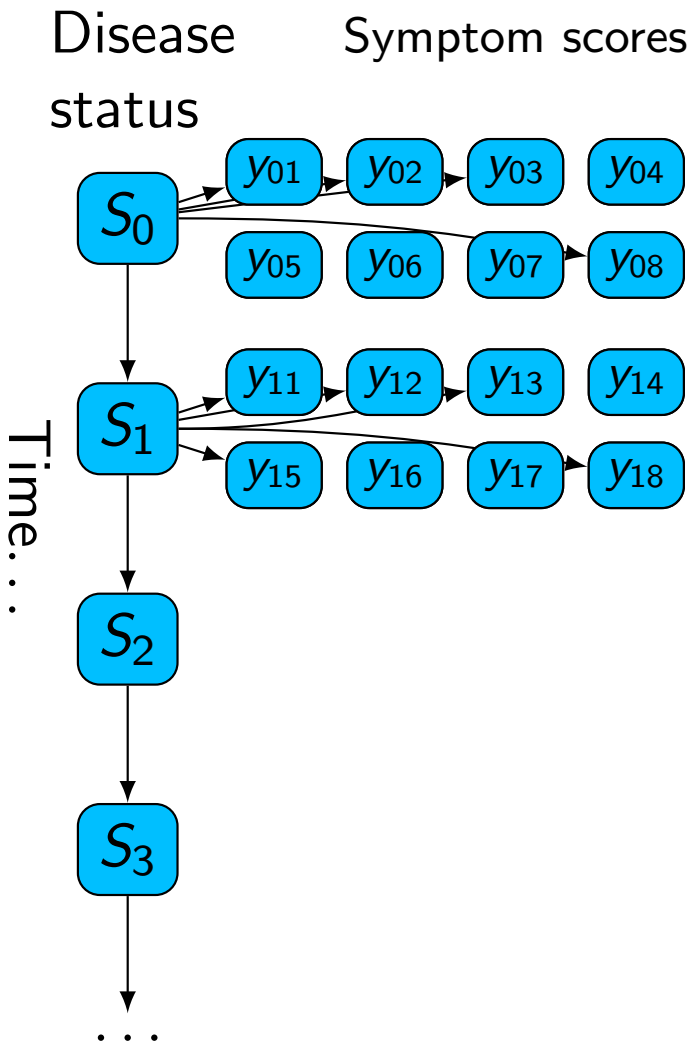
Hidden Markov model with any/multiple outcomes



- Clinical interest in a discrete, latent disease process
  - following a multi-state model
- Vector $\mathbf{y}_{ij}$ of observations, from each person i, at times j
  - e.g. multiple symptoms, biomarkers
- Specify joint distribution of $(\mathbf{y}_{ij}|S)$ for each potential hidden state S
- Likelihood similar to misclassification model
- Usually need constraints for identifiability
  - state known at some times
  - constraints on state-specific parameters

Satten and Longini (Appl. Stat 1997), Jackson and Sharples (Stat Med 2002) for HMM with general outcome
Jackson, Su, Gladman & Farewell (Arthritis Care and Research 2017) for multiple outcomes

Hidden Markov model with any/multiple outcomes



- Clinical interest in a discrete, latent disease process
  - following a multi-state model
- Vector $\mathbf{y}_{ij}$ of observations, from each person i, at times j
  - e.g. multiple symptoms, biomarkers
- Specify joint distribution of $(\mathbf{y}_{ij}|S)$ for each potential hidden state S
- Likelihood similar to misclassification model
- Usually need constraints for identifiability
  - state known at some times
  - constraints on state-specific parameters

Satten and Longini (Appl. Stat 1997), Jackson and Sharples (Stat Med 2002) for HMM with general outcome
Jackson, Su, Gladman & Farewell (Arthritis Care and Research 2017) for multiple outcomes

Contents

Semi-Markov models for intermittently-observed data

Markov models: transition rates independent of

- previous states visited
- time in current state

Semi-Markov models for intermittently-observed data

Semi-Markov models: transition rates

- depend on time in current state, but not previous states

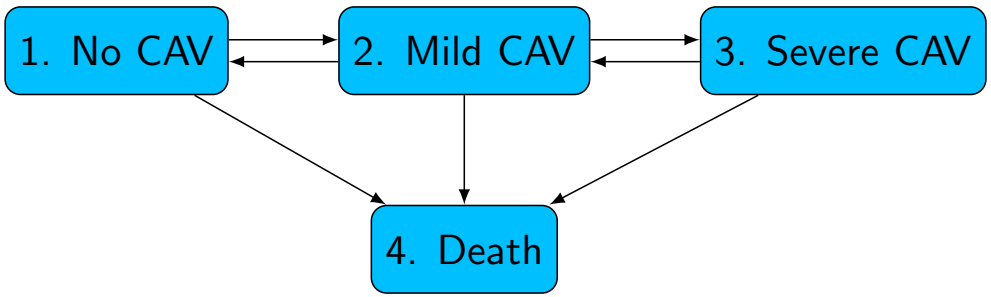
Semi-Markov models for intermittently-observed data

Semi-Markov models: transition rates

- depend on time in current state, but not previous states

Phase-type model

- E.g.: transition rate from state 2 depends on time spent there
- Replace state 2 by a series of hidden "phases".
- Expanded state structure follows a Markov process.



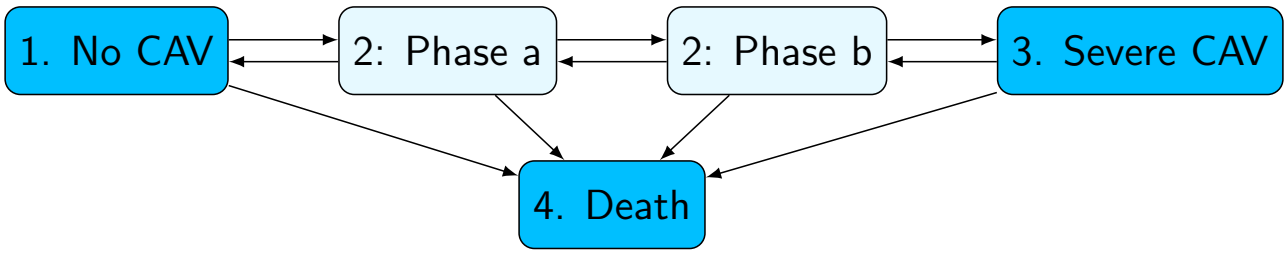
Semi-Markov models for intermittently-observed data

Semi-Markov models: transition rates

- depend on time in current state, but not previous states

Phase-type model

- E.g.: transition rate from state 2 depends on time spent there
- Replace state 2 by a series of hidden "phases".
- Expanded state structure follows a Markov process.



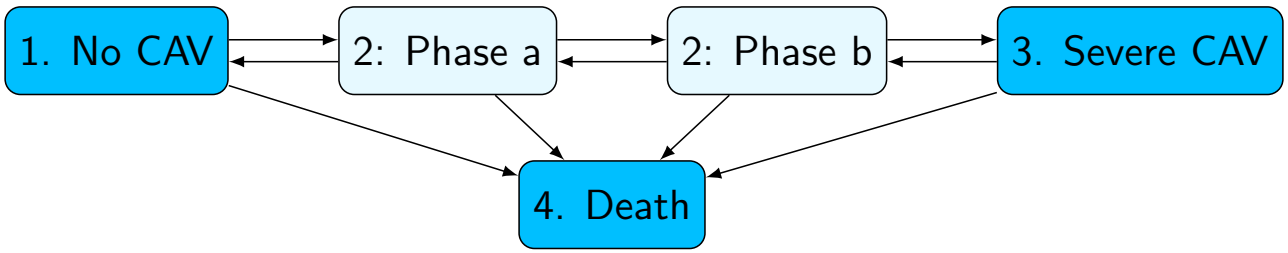
Semi-Markov models for intermittently-observed data

Semi-Markov models: transition rates

- depend on time in current state, but not previous states

Phase-type model

- E.g.: transition rate from state 2 depends on time spent there
- Replace state 2 by a series of hidden "phases".
- Expanded state structure follows a Markov process.



Likelihood the same as that for a hidden Markov model on the expanded state structure $\rightarrow$ can use `msm`.

- See Titman and Sharples (2010) Biometrics 66
- `phase.states` option to `msm()` function

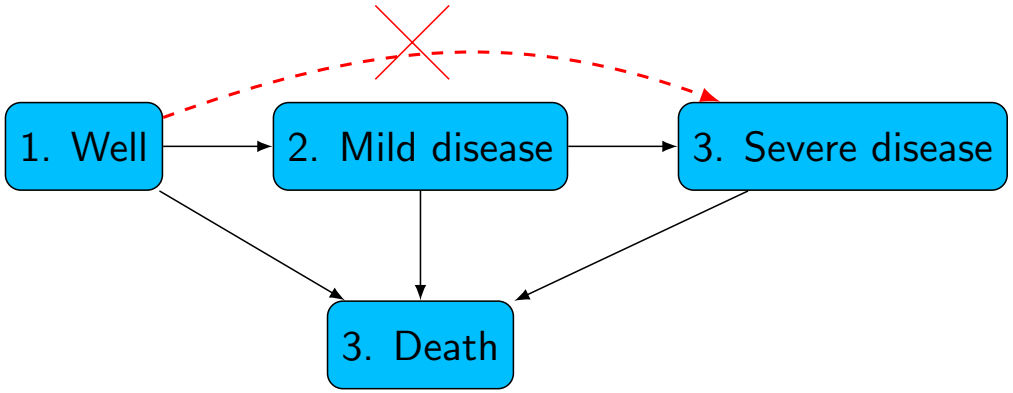
Practical concerns with modelling intermittent observations

It can be tempting to make the model too big

- $\rightarrow$ non-identifiable models, or flat/awkward likelihood function that can't be maximised.

Careful of, e.g.

- choice of state structure. If states are ordered (e.g. disease severity) should only allow transitions between adjacent states. Jumps not plausible in continuous time.



- if several covariates, affecting different transitions in different ways – may need to restrict/constrain effects.
- misclassification models typically need structural constraints. e.g. data can't distinguish "backwards" transitions (e.g. severe to mild disease) from misclassification (e.g. of severe

Practical concerns with modelling intermittent observations

It can be tempting to make the model too big

- $\rightarrow$ non-identifiable models, or flat/awkward likelihood function that can't be maximised.

Careful of, e.g.

- choice of state structure. If states are ordered (e.g. disease severity) should only allow transitions between adjacent states. Jumps not plausible in continuous time.
- if several covariates, affecting different transitions in different ways – may need to restrict/constrain effects.
- misclassification models typically need structural constraints. e.g. data can't distinguish "backwards" transitions (e.g. severe to mild disease) from misclassification (e.g. of severe as mild)

Optimisation routines can sometimes be tweaked to help convergence, but can't rescue an inappropriate model

Practical concerns with modelling intermittent observations

It can be tempting to make the model too big

- $\rightarrow$ non-identifiable models, or flat/awkward likelihood function that can't be maximised.

Careful of, e.g.

- choice of state structure. If states are ordered (e.g. disease severity) should only allow transitions between adjacent states. Jumps not plausible in continuous time.
- if several covariates, affecting different transitions in different ways – may need to restrict/constrain effects.
- misclassification models typically need structural constraints. e.g. data can't distinguish "backwards" transitions (e.g. severe to mild disease) from misclassification (e.g. of severe as mild)

Optimisation routines can sometimes be tweaked to help convergence, but can't rescue an inappropriate model

Practical concerns with modelling intermittent observations

It can be tempting to make the model too big

- $\rightarrow$ non-identifiable models, or flat/awkward likelihood function that can't be maximised.

Careful of, e.g.

- choice of state structure. If states are ordered (e.g. disease severity) should only allow transitions between adjacent states. Jumps not plausible in continuous time.
- if several covariates, affecting different transitions in different ways – may need to restrict/constrain effects.
- misclassification models typically need structural constraints. e.g. data can't distinguish "backwards" transitions (e.g. severe to mild disease) from misclassification (e.g. of severe as mild)

Optimisation routines can sometimes be tweaked to help convergence, but can't rescue an inappropriate model

Contents

Contents

Random effects or hierarchical multi-state models

- "Frailties" shared between groups of individuals, or between different transition types within an individual
- Difficult to distinguish individual frailties from dependence on history (Cook & Lawless 2012)

Implementation and examples:

- Maximum marginal likelihood (e.g. O'Keeffe, Tom & Farewell, JRSSC 2011; Yiu, Farewell & Tom, JRSSC 2018; Cook et al, Biometrics 2004)
- Bayesian methods
  - See e.g. van den Hout, *Multi-state Survival Models for Interval Censored Data* CRC 2017
  - MCMC estimation
  - JAGS and Stan have matrix exponential, allowing general panel data likelihood

Other advanced models for intermittent data (b)

Informative observation times

- When data are observed intermittently, observation time may depend on status at that time
- Jointly model observation process and outcome process, e.g.
  - Lange et al., Biometrics 2015
  - Sweeting, Farewell & De Angelis, Stat. Med. 2010

Joint models for multistate and other longitudinal data

e.g.

- Dantan et al. Biostatistics 2011 (dementia and cognitive function)
- Ferrer et al. Stat. Med. 2016 (cancer recurrence and biomarker series)

Other advanced models for intermittent data (b)

Informative observation times

- When data are observed intermittently, observation time may depend on status at that time
- Jointly model observation process and outcome process, e.g.
  - Lange et al., Biometrics 2015
  - Sweeting, Farewell & De Angelis, Stat. Med. 2010

Joint models for multistate and other longitudinal data

e.g.

- Dantan et al. Biostatistics 2011 (dementia and cognitive function)
- Ferrer et al. Stat. Med. 2016 (cancer recurrence and biomarker series)

Contents

Model checking and comparison

(Titman & Sharples, "Model diagnostics for multi-state models" SMMR, 2010)

- Fit elaborated model with questionable assumption relaxed, then compare using e.g. likelihood ratio, AIC.
- Intermittent data makes it hard to check absolute fit
  - e.g compare fitted transition probabilities $p(t)$ to observed prevalence of a state at time t
  - what if not everyone measured at time t?
  - Current approaches `msm` use imputations/approximations of observed state
- If death times observed, could compare parametric estimates of survival against Kaplan-Meier estimates.
- Pearson-type goodness-of-fit tests have been proposed
  - low power, sensitive to how data grouped to construct the test
  - hard to interpret: what part of the model is most wrong?

Model checking and comparison

(Titman & Sharples, "Model diagnostics for multi-state models" SMMR, 2010)

- Fit elaborated model with questionable assumption relaxed, then compare using e.g. likelihood ratio, AIC.
- Intermittent data makes it hard to check absolute fit
  - e.g compare fitted transition probabilities $p(t)$ to observed prevalence of a state at time t
  - what if not everyone measured at time t?
  - Current approaches `msm` use imputations/approximations of observed state
- If death times observed, could compare parametric estimates of survival against Kaplan-Meier estimates.
- Pearson-type goodness-of-fit tests have been proposed
  - low power, sensitive to how data grouped to construct the test
  - hard to interpret: what part of the model is most wrong?

Model checking and comparison

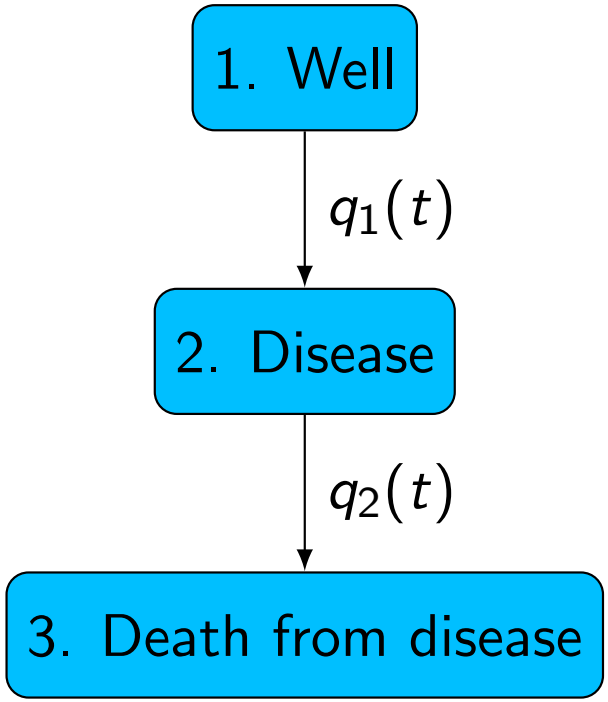
(Titman & Sharples, "Model diagnostics for multi-state models" SMMR, 2010)

- Fit elaborated model with questionable assumption relaxed, then compare using e.g. likelihood ratio, AIC.
- Intermittent data makes it hard to check absolute fit
  - e.g compare fitted transition probabilities $p(t)$ to observed prevalence of a state at time t
  - what if not everyone measured at time t?
  - Current approaches `msm` use imputations/approximations of observed state
- If death times observed, could compare parametric estimates of survival against Kaplan-Meier estimates.
- Pearson-type goodness-of-fit tests have been proposed
  - low power, sensitive to how data grouped to construct the test
  - hard to interpret: what part of the model is most wrong?

Contents

Multi-state models for aggregate data

Age	Disease prevalence	Mortality from disease
0	0	
...		
40	r_1/n_1	s_1/m_1
45	r_2/n_2	s_2/m_2
50	r_3/n_3	s_3/m_3
...		



- Survey of n_i people by age i gives disease prevalence r_i/n_i
- Mortality data: s_i/m_i people die from disease at each age i
- Infer disease incidence $q_1(t)$ and case fatality $q_2(t)$

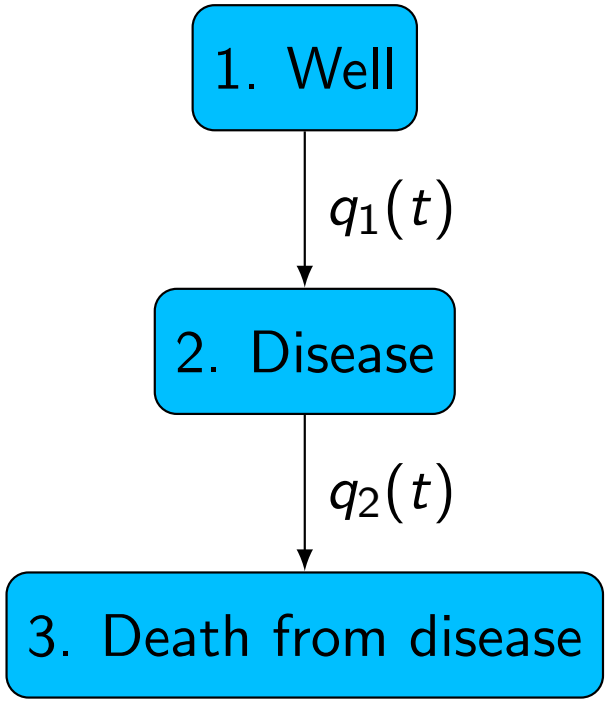
$$r_i \sim Bin(n_i, p_{12}(0, t_i)), \qquad s_i \sim Bin(m_i, p_{23}(t_{i-1}, t_i))$$

- $p_{rs}(.,.)$: transition probability from the multi-state model
  - function of $q_1(t), q_2(t)$ through Kolmogorov equation

(Ongoing work: `github.com/chjackson/disbayes`, Bayesian, in Stan)

Multi-state models for aggregate data

Age	Disease prevalence	Mortality from disease
0	0	
...		
40	r_1/n_1	s_1/m_1
45	r_2/n_2	s_2/m_2
50	r_3/n_3	s_3/m_3
...		



- Survey of n_i people by age i gives disease prevalence r_i/n_i
- Mortality data: s_i/m_i people die from disease at each age i
- Infer disease incidence $q_1(t)$ and case fatality $q_2(t)$

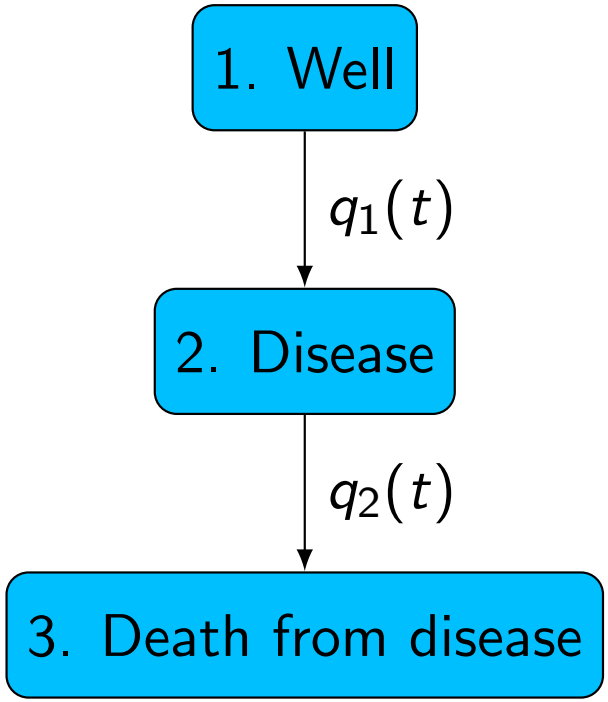
$$r_i \sim Bin(n_i, p_{12}(0, t_i)), \qquad s_i \sim Bin(m_i, p_{23}(t_{i-1}, t_i))$$

- $p_{rs}(.,.)$: transition probability from the multi-state model
  - function of $q_1(t), q_2(t)$ through Kolmogorov equation

(Ongoing work: `github.com/chjackson/disbayes`, Bayesian, in Stan)

Multi-state models for aggregate data

Age	Disease prevalence	Mortality from disease
0	0	
...		
40	r_1/n_1	s_1/m_1
45	r_2/n_2	s_2/m_2
50	r_3/n_3	s_3/m_3
...		



- Survey of n_i people by age i gives disease prevalence r_i/n_i
- Mortality data: s_i/m_i people die from disease at each age i
- Infer disease incidence $q_1(t)$ and case fatality $q_2(t)$

$$r_i \sim Bin(n_i, p_{12}(0, t_i)), \qquad s_i \sim Bin(m_i, p_{23}(t_{i-1}, t_i))$$

- $p_{rs}(.,.)$: transition probability from the multi-state model
  - function of $q_1(t), q_2(t)$ through Kolmogorov equation

(Ongoing work: `github.com/chjackson/disbayes`, Bayesian, in Stan)

References

- Van Den Hout, A. (2016). Multi-state survival models for interval-censored data. CRC Press.
  Methods for multi-state modelling of intermittently-observed data: time/age-dependent models, chronic conditions, advanced topics e.g. random effects, Bayesian, spline models. . . . Examples in R
- Cook, R. J., & Lawless, J. F. (2018). Multistate models for the analysis of life history data. CRC Press.
  Multi-state models for both continuously-observed and intermittently-observed data, mathematical detail, range of examples in medicine. Advanced topics, e.g. nonparametric methods, informative observation, random effects, mixture, hidden Markov models.
- Jackson, C. H. (2011). Multi-state models for panel data: the msm package for R. Journal of Statistical Software, 38(8), 1-29.
  Review of methods for intermittent observation and worked examples of implementing in R

See you in September, hopefully, for more about this package!

Thanks for listening!

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

chris.jackson@mrc-bsu.cam.ac.uk

www.mrc-bsu.cam.ac.uk