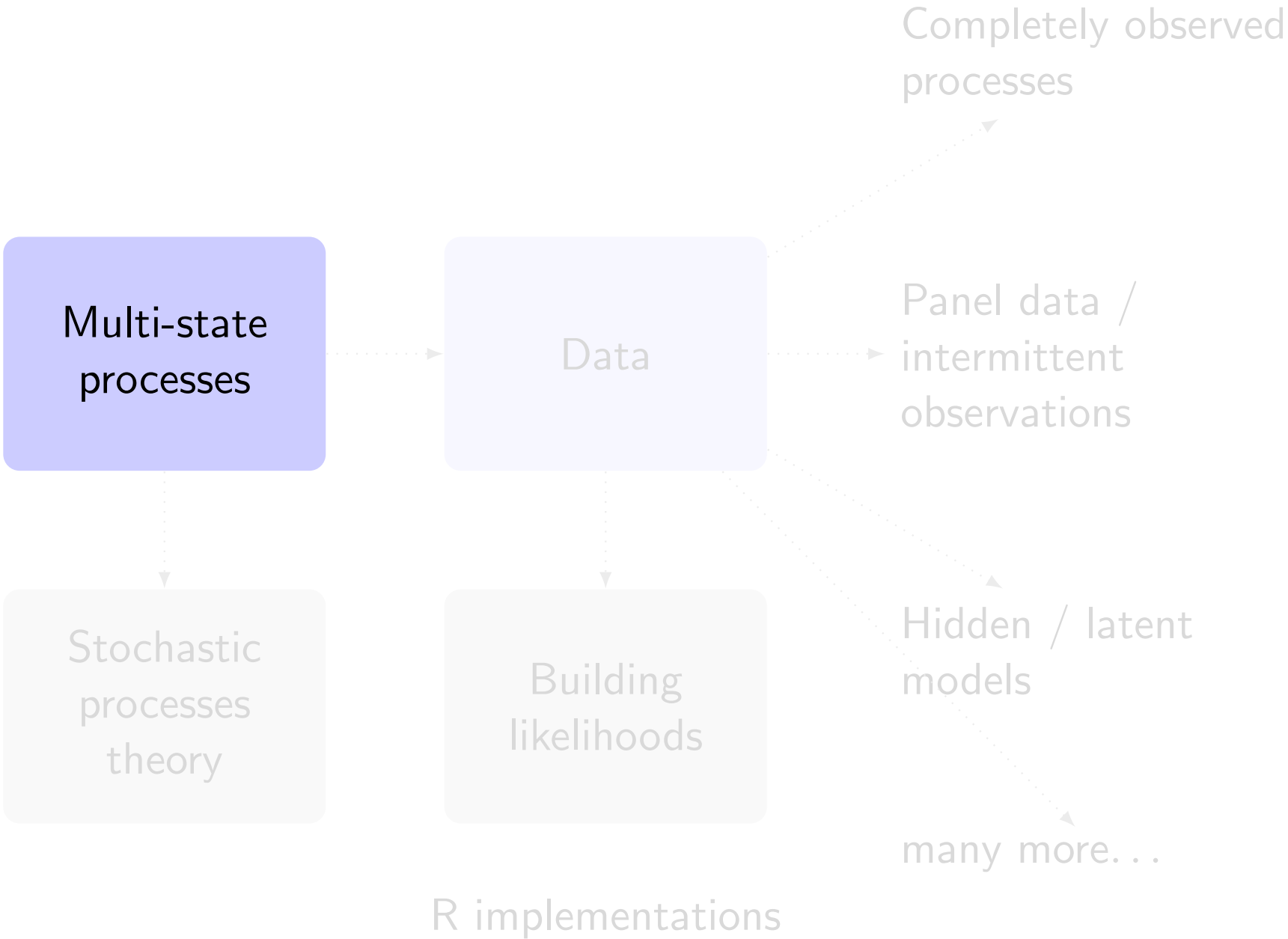
Multi-state Models: Methods and Software

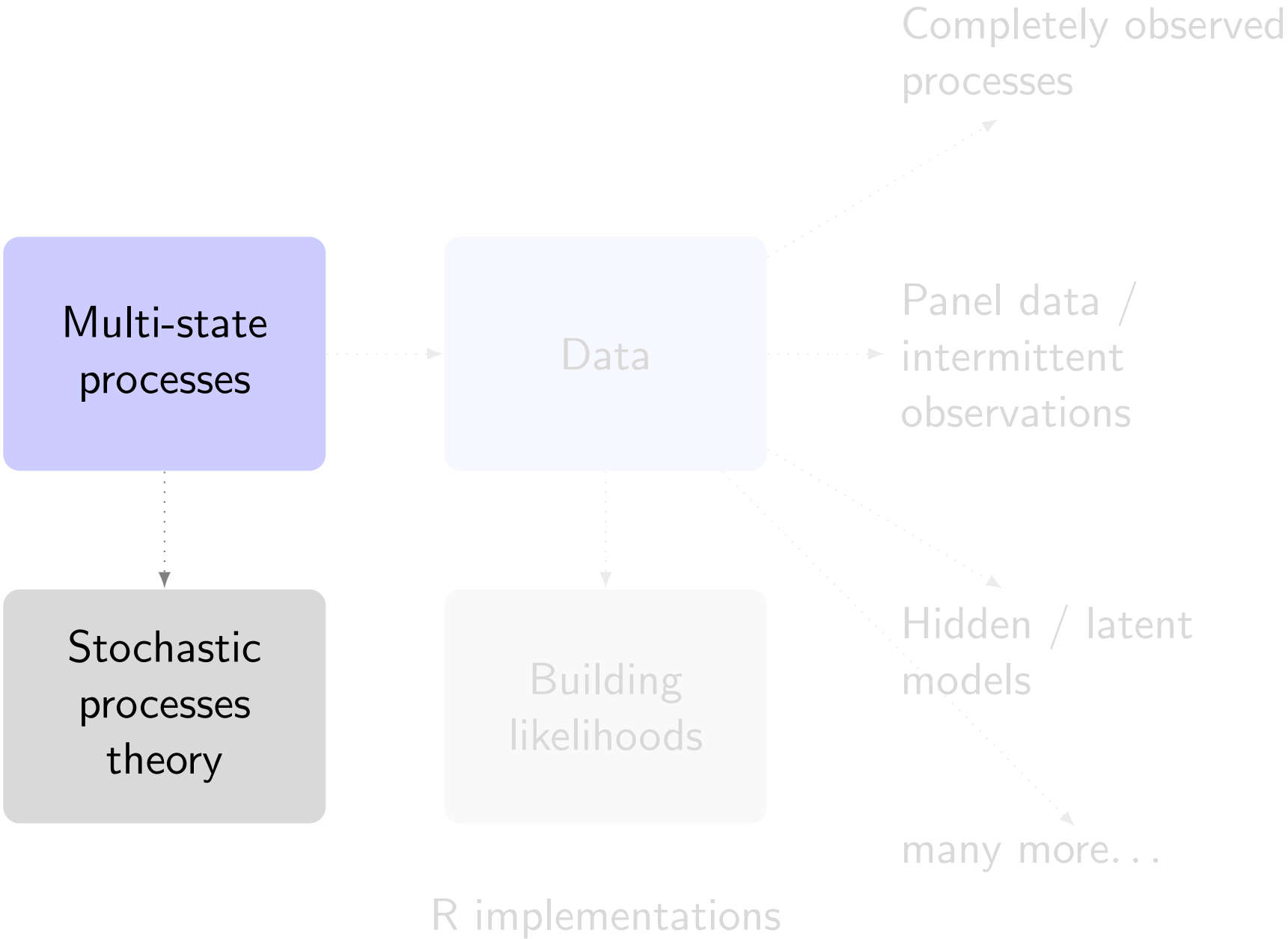
Chris Jackson
MRC Biostatistics Unit, University of Cambridge, UK
chris.jackson@mrc-bsu.cam.ac.uk

ENAR Webinar Series
International Biometric Society
October 20, 2017

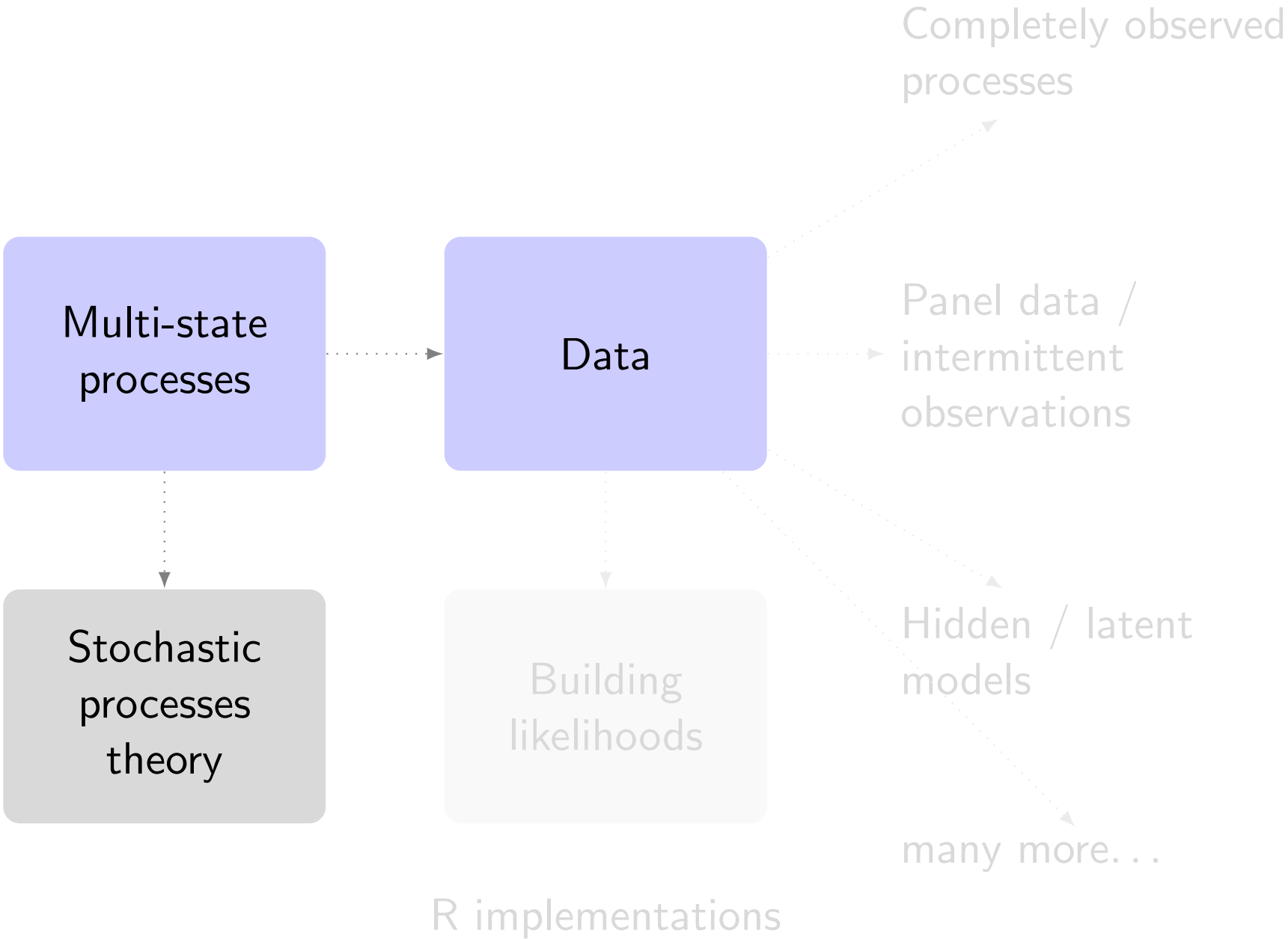
Concepts covered



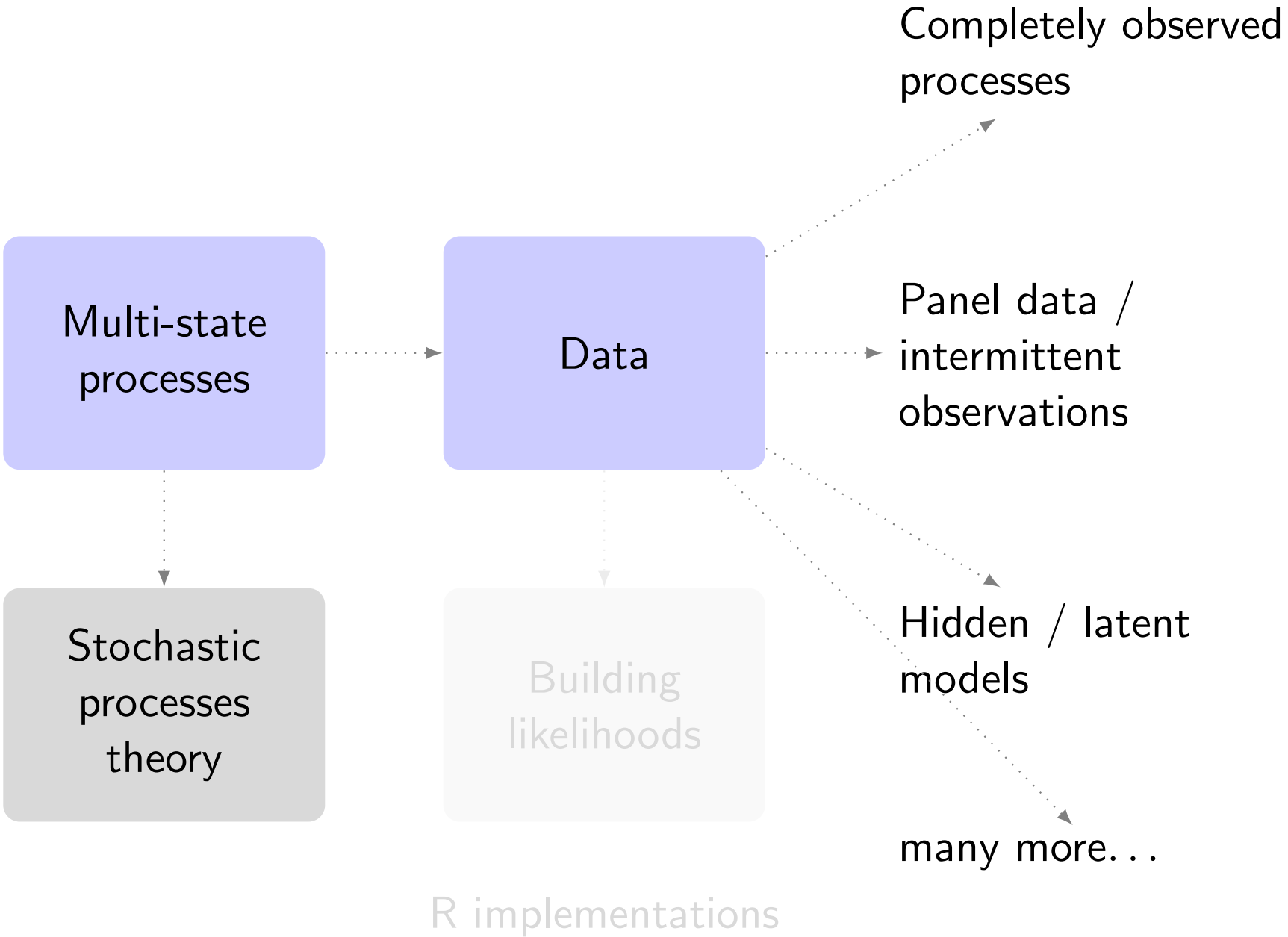
Concepts covered



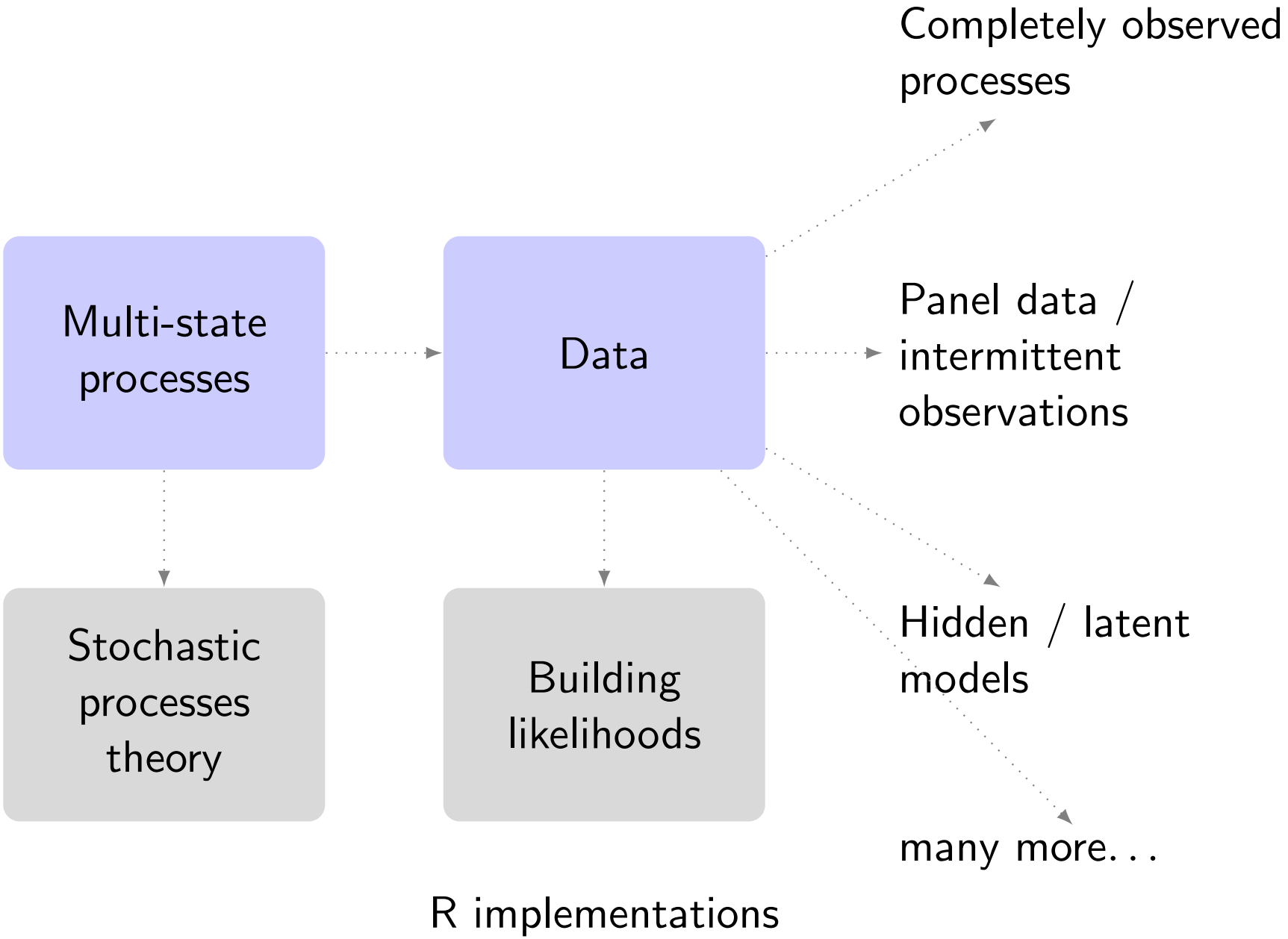
Concepts covered



Concepts covered



Concepts covered



Contents

Introduction 2

Mathematics of multi-state processes 8

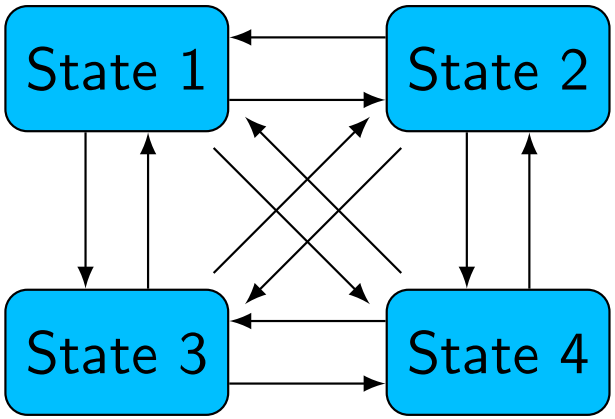
Panel data 19
- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53
- Theory: generalisation of survival models 54
- Implementation and examples 61

Multi-state process

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

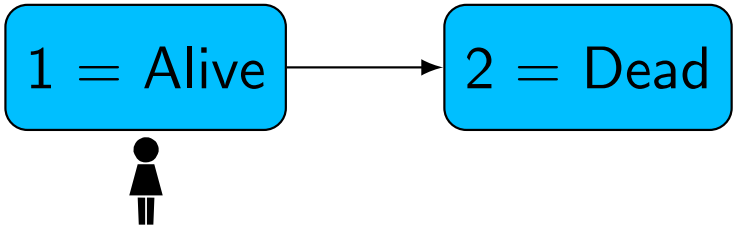
- Person begins in one of the R states
- Spends a random, continuously-distributed time in that state
- Then moves to a random new state
- etc...



(with $R = 4$)

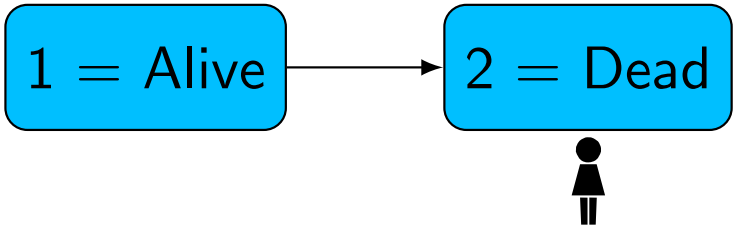
Examples of multi-state processes (a)

Survival model



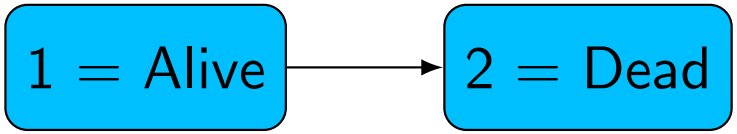
Examples of multi-state processes (a)

Survival model

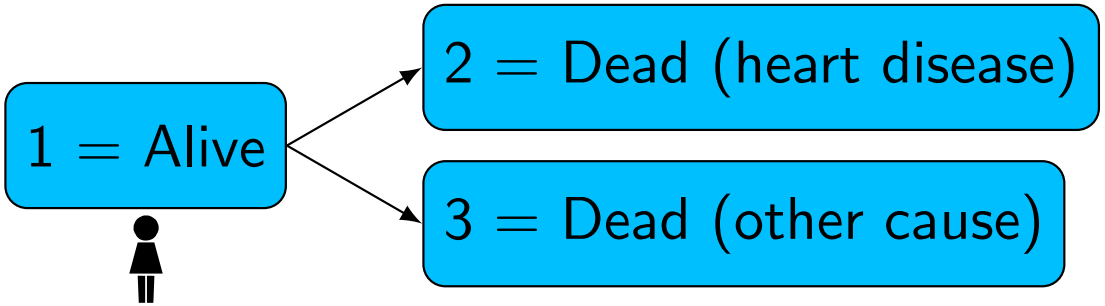


Examples of multi-state processes (a)

Survival model

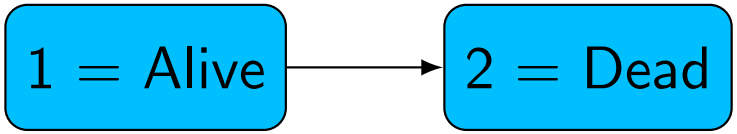


Competing risks model

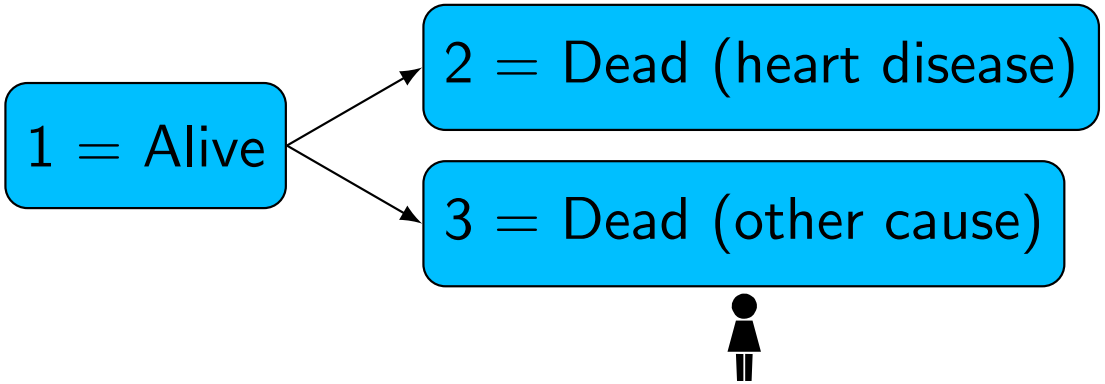


Examples of multi-state processes (a)

Survival model

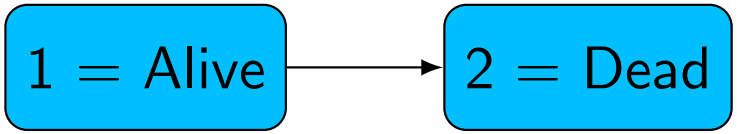


Competing risks model

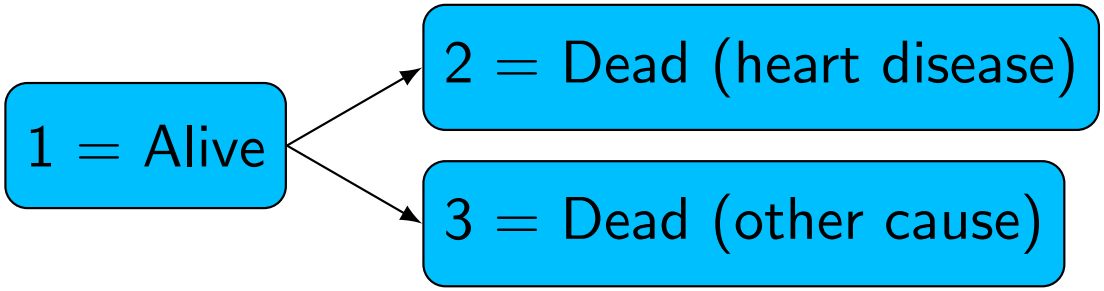


Examples of multi-state processes (a)

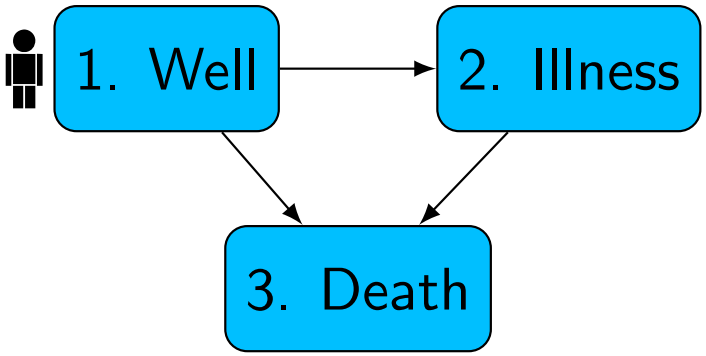
Survival model



Competing risks model

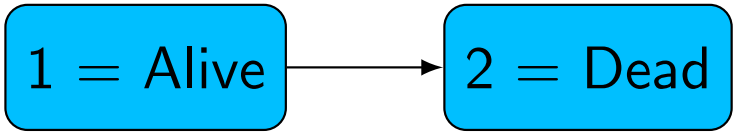


Illness-death model with death from any cause

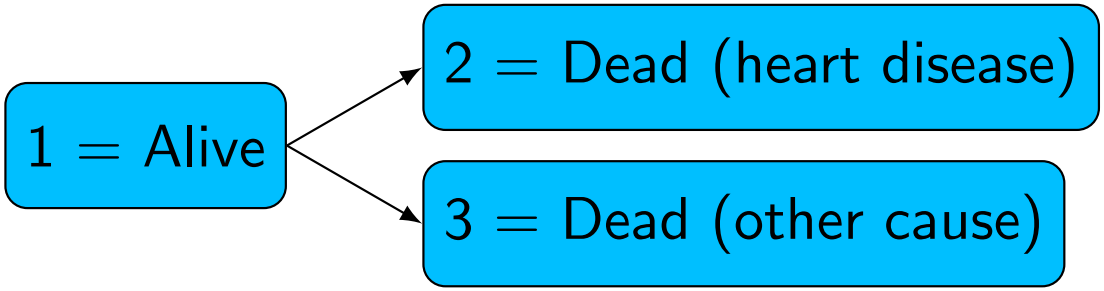


Examples of multi-state processes (a)

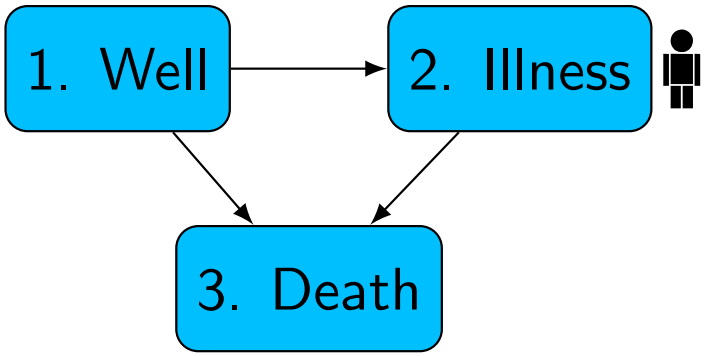
Survival model



Competing risks model

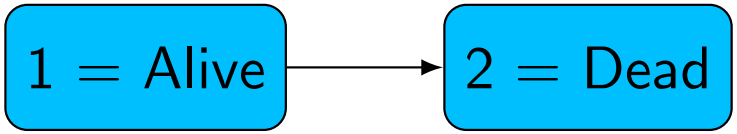


Illness-death model with death from any cause

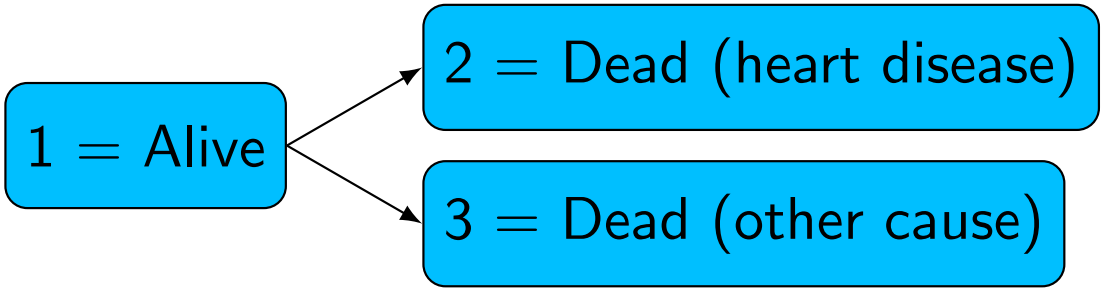


Examples of multi-state processes (a)

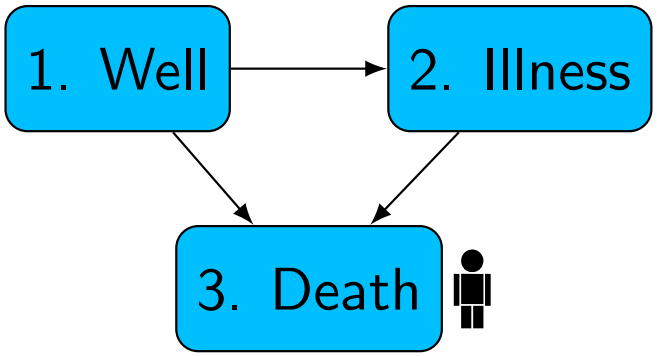
Survival model



Competing risks model

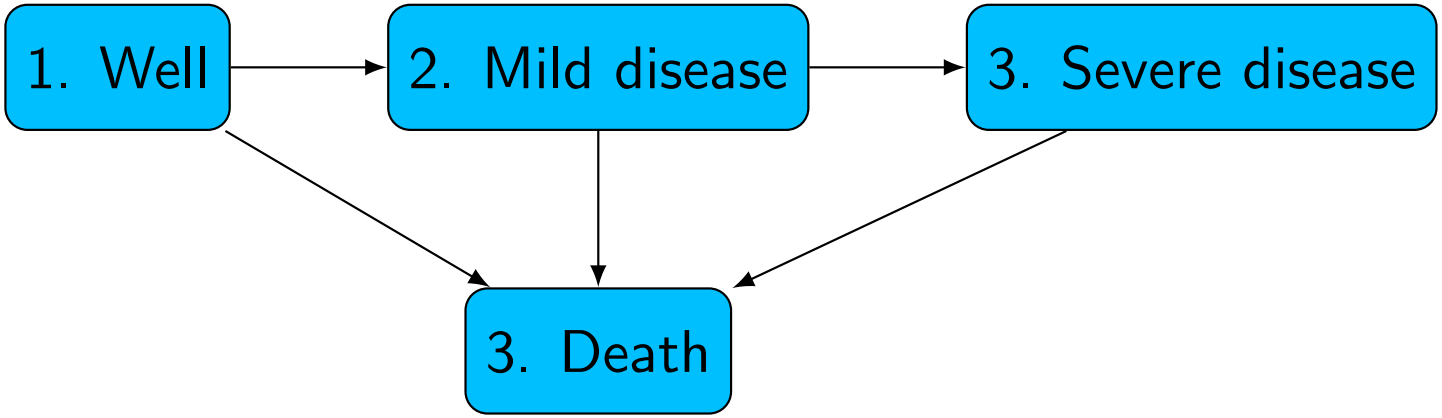


Illness-death model with death from any cause



Examples of multi-state processes (b)

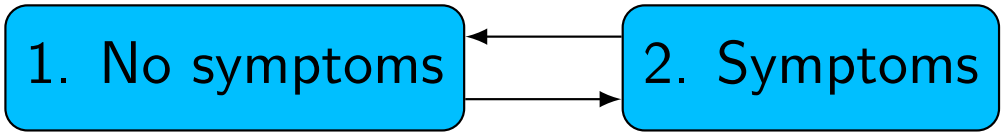
Staged disease progression model



- States typically chosen from some clinical convention
- Often representing underlying continuous "severity"
- Discrete/continuous conflict:
  - discrete model useful for interpretability
  - true process typically continuous
  - data could be discrete, continuous, messy combination...
- See later for models with continuous longitudinal data, and a discrete underlying process...

Examples of multi-state processes (c)

Relapsing and remitting non-fatal condition



- Model doesn't have to include death, or any "absorbing" state
- Process modelled conditionally on surviving

Transition intensities

Hazard in survival analysis for a random death time T:
Rate of death in near future, given survived up to t.

$$h(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t | T > t)}{\delta t}$$

Transition intensity in a multi-state process $X(t)$
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

Hazard in survival analysis for a random death time T:
Rate of death in near future, given survived up to t.

$$h(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t | T > t)}{\delta t}$$

Transition intensity in a multi-state process $X(t)$
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

Hazard in survival analysis for a random death time T:
Rate of death in near future, given survived up to t.

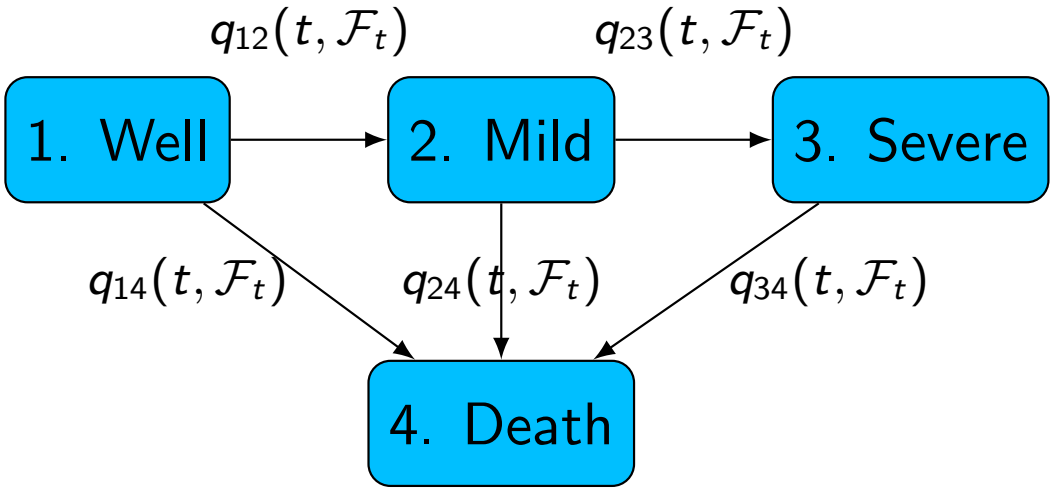
$$h(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t | T > t)}{\delta t}$$

Transition intensity in a multi-state process $X(t)$
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 $q_{rs}(t, \mathcal{F}_t) : r \neq s, r, s = 1, ..., R$.

Consider only continuous-time processes here: discrete-time models are parameterised by transition probabilities (governing state at next step). Beware of confusion!

Markov processes

If $q_{rs}(t; \mathcal{F}_t) = q_{rs}(t)$, 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 $X(t)$.
  - Most general model considered in this talk.
- Time-homogeneous Markov model $q_{rs}(t; \mathcal{F}_t) = q_{rs}$
  - constant transition rate.
- Dependence on covariates $\mathbf{z}(t)$: $q_{rs}(\mathbf{z}(t); \mathcal{F}_t)$
  - covariates could be constant or time-varying
  - proportional hazards common
    $q_{rs}(\mathbf{z}(t); \mathcal{F}_t) = q_{rs}^{(0)}(t, \mathcal{F}_t) \exp(\beta' \mathbf{z}(t))$

Time scale t chosen so that become at risk of first transition at time origin $t = 0$ (e.g. birth, disease onset...)

Transition intensity matrix in a time-homogeneous Markov model

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

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

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)? $-q_{rs}/q_{rr} = q_{rs}/\sum_{j \neq r} q_{rj}$
- Could define richer class of semi-Markov models by replacing the exponential by other distributions

Transition intensity matrix in a time-homogeneous Markov model

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

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

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)? $-q_{rs}/q_{rr} = q_{rs}/\sum_{j \neq r} q_{rj}$
- Could define richer class of semi-Markov models by replacing the exponential by other distributions

Transition intensity matrix in a time-homogeneous Markov model

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

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

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)? $-q_{rs}/q_{rr} = q_{rs}/\sum_{j!=r} q_{rj}$
- Could define richer class of semi-Markov models by replacing the exponential by other distributions

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)? $-q_{rs}/q_{rr} = q_{rs}/\sum_{j!=r} q_{rj}$
- Could define richer class of semi-Markov models by replacing the exponential by other distributions

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 $Q(t)$

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

In a time-homogeneous model

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

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

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 $Q(t)$

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

In a time-homogeneous model

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

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

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 $Q(t)$

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

In a time-homogeneous model

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

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

Computing the transition probabilities

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

For certain smaller/simpler models (2,3,4 states)

- can calculate $p_{rs}(t)$ for each r, s as an explicit function of the elements of Q.
- See e.g. `msm` R package, or van den Hout *Multi-State Models for Interval Censored Data* CRC 2017.

Numerical computation required in general, see, e.g.

- `expm` R package (Goulet et al.)
- Moler and van Loan *Nineteen Dubious Ways to Calculate the Exponential of a Matrix* (SIAM Review 1978, revisited 2003)

Time-varying transition intensities

Transition intensities Q commonly not constant, but 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

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

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

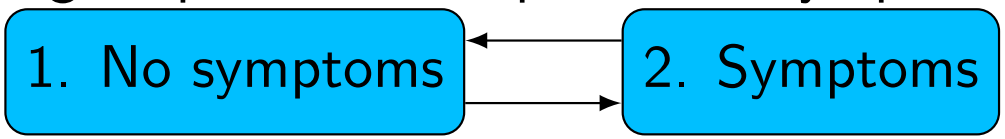
- e.g. Titman, Biometrics 67:780–7, 2011
- `deSolve` R package

Total time expected to spend in a state

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 time spent with symptoms in



Analytic solution for vector $\mathbf{x}$ with sth element $x_s =$

- total expected time spent in all stays in s before time t
- given probabilities π_0 of being in each state at time 0

$$\mathbf{x} = \begin{bmatrix} 1 & \mathbf{0}_K \end{bmatrix} Exp(tQ') \begin{bmatrix} \mathbf{0}_K \\ I_K \end{bmatrix} \qquad Q' = \begin{bmatrix} 0 & \pi_0 \\ \mathbf{0}_K & Q - rI_K \end{bmatrix}$$

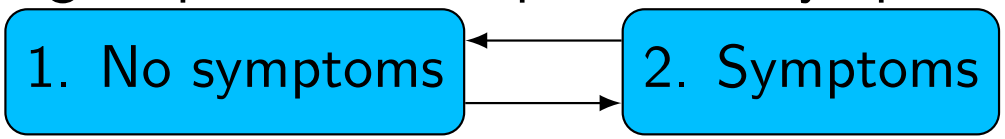
van Loan (1978) IEEE Trans Automatic Control 23(3)395–404
van Rosmalen et al. (2013) Med. Decis. Making 33:767-779

Total time expected to spend in a state

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 time spent with symptoms in



Analytic solution for vector $\mathbf{x}$ with sth element $x_s =$

- total expected time spent in all stays in s before time t
- given probabilities π_0 of being in each state at time 0

$$\mathbf{x} = \begin{bmatrix} 1 & \mathbf{0}_K \end{bmatrix} Exp(tQ') \begin{bmatrix} \mathbf{0}_K \\ I_K \end{bmatrix} \qquad Q' = \begin{bmatrix} 0 & \pi_0 \\ \mathbf{0}_K & Q - rI_K \end{bmatrix}$$

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

Other quantities of interest

All functions of the transition intensities. For example

- expected first passage time (first visit time to a state)
  - also probability of ever visiting this state
  - derived by setting the outward transition intensities from that state to zero,
  - then calculating the time before "absorption" into this state, or the transition probability to this absorbing state
- expected number of visits to a state
  - corollary of total time expected to spend in a state
  - see `msm` documentation, `envisits.msm` function

Types of multi-state data

Data from sample of n individuals, each one following a multi-state process

Estimate the

- transition intensities, and / or
- their dependence on time and / or other covariates
- ... hence estimate transition probabilities, total expected time in state, other aspects of process...

What do the data look like?

Multi-state processes are often partially-observed in various ways...

Contents

Introduction 2

Mathematics of multi-state processes 8

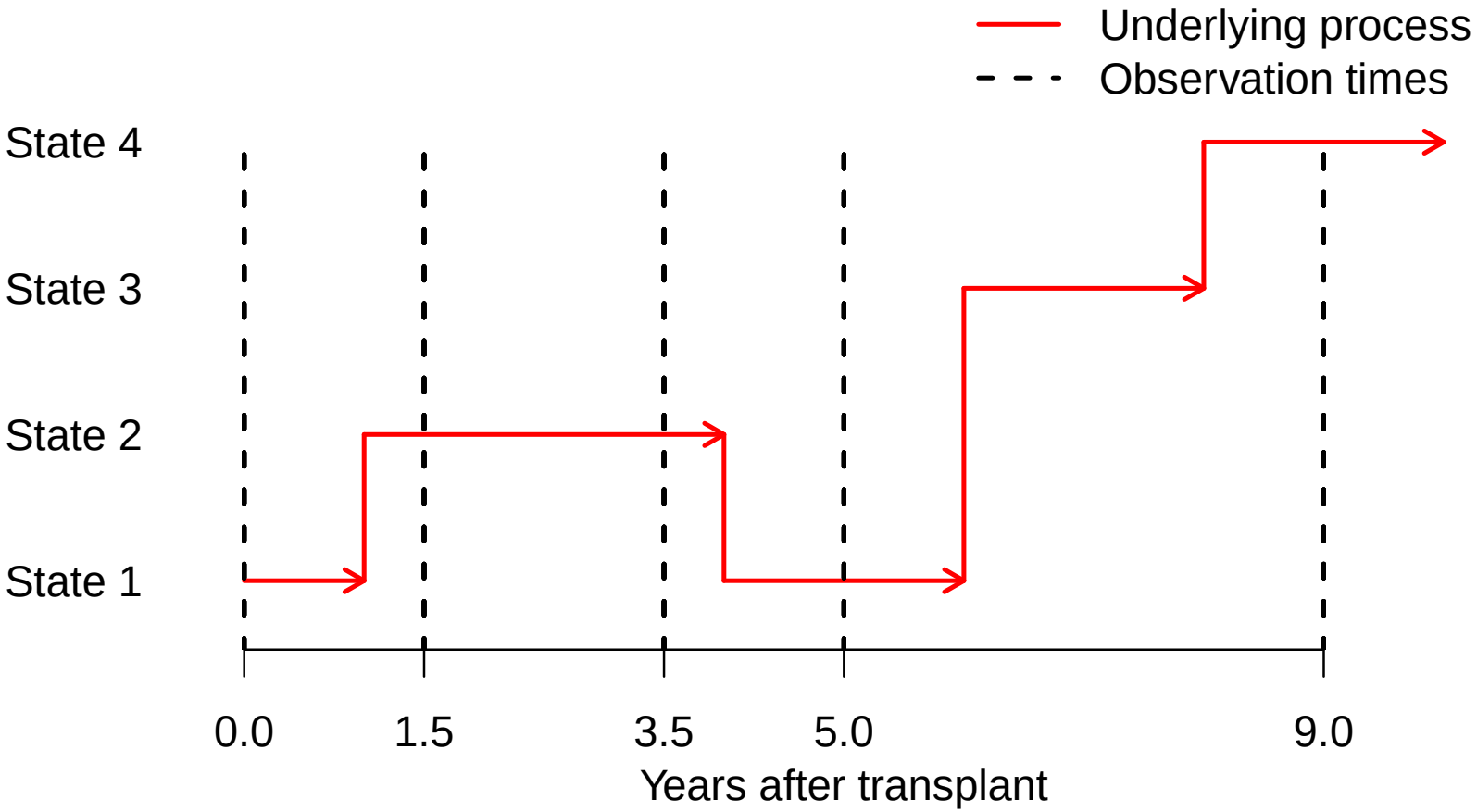
Panel data 19

- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53

- Theory: generalisation of survival models 54
- Implementation and examples 61

"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

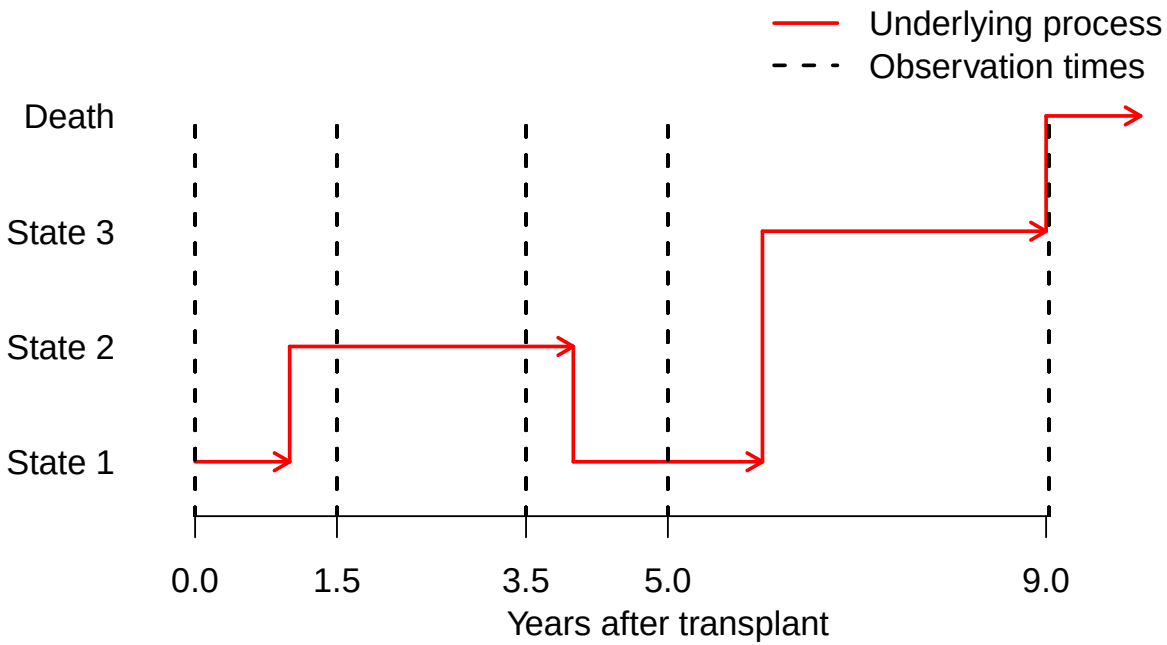
- Data $\mathbf{x}$: states $(x_{i1}, \ldots, x_{in_i})$ at times $(t_{i1}, \ldots, t_{in_i})$ for person i
  Conditional on state at $t_{i0} = 0$ (e.g. $x_{i0} = 1$)
- Parameters: $\boldsymbol{\theta} = \{q_{rs}\}$: transition intensities of Markov model

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 $x_{i,1}\ldots x_{i,j-1}$
- Time homogeneity over each interval $(t_{i,j-1}, t_{i,j})$: each transition probability is a closed form function of $\boldsymbol{\theta}$.

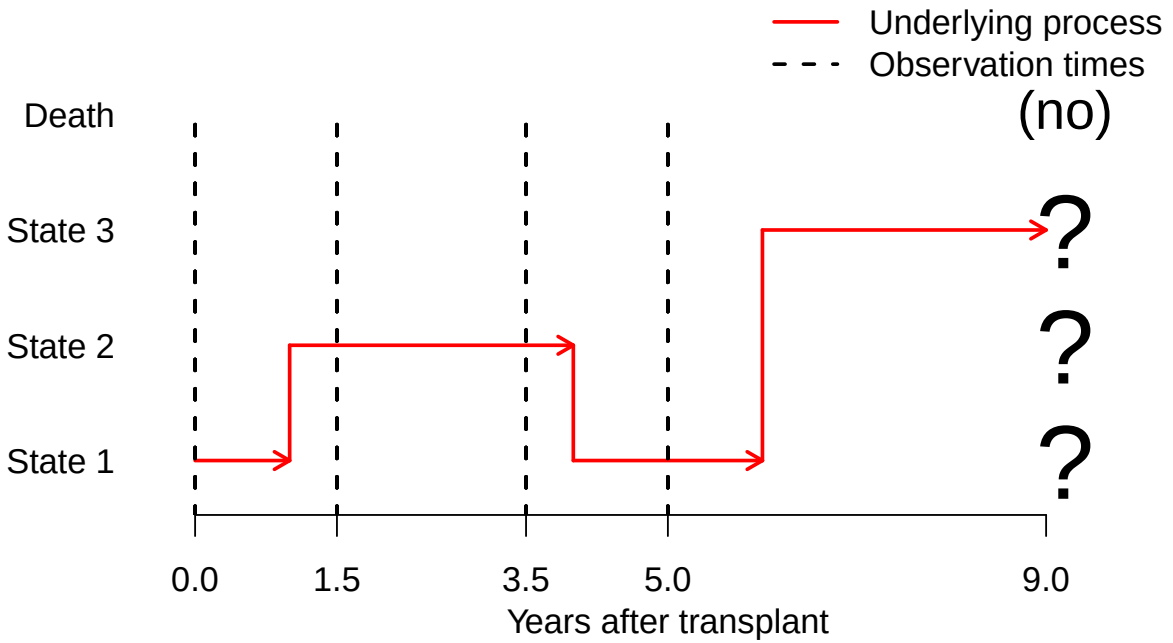
Panel data with exact death time observed



- 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:

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

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

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

- $\exp(\beta_m)$ is rate ratio or hazard ratio for mth covariate for ith individual's jth observation
- interpreted as the relative *instantaneous* risk of transition (recall definition of q_{rs})
- Covariates could be constant for all times j for individual i
- Any time-varying covariates assumed piecewise-constant, that is, constant within each observation interval $(t_{i,j-1}, t_{i,j})$.

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]

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

Contents

Introduction 2

Mathematics of multi-state processes 8

Panel data 19

- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53

- Theory: generalisation of survival models 54
- Implementation and examples 61

The msm R package for multi-state modelling

http://CRAN.R-project.org/package=msm

- Designed for panel data and its variations
- Any transition structure permitted
- General loglinear model for covariates on intensities
- Standard numerical maximum likelihood estimation

Documentation and worked examples:

- Jackson C (2011) Multi-state models for panel data: the msm package for R. *Journal of Statistical Software*.
- Detailed vignette in package, with similar material to paper

Example data in msm

Coronary artery vasculopathy (CAV) after heart transplants

- state: 1 (no CAV), 2 (moderate CAV), 3 (severe CAV), 4 (death)
- years: years after transplant
- PTNUM: patient number

```
library(msm)
head(cav[,c("state","PTNUM","years","age","sex")])

##   state  PTNUM years age sex
## 1     1 100002   0.0  52   0
## 2     1 100002   1.0  53   0
## 3     2 100002   2.0  54   0
## 4     2 100002   3.1  56   0
## 5     2 100002   4.0  56   0
## 6     3 100002   5.0  57   0
```

Summary of multi-state data

Summary of transitions over an interval

Number of times person observed in

- state r at one observation time
- followed by state s at next observation time

```
statetable.msm(state, PTNUM, cav)
```

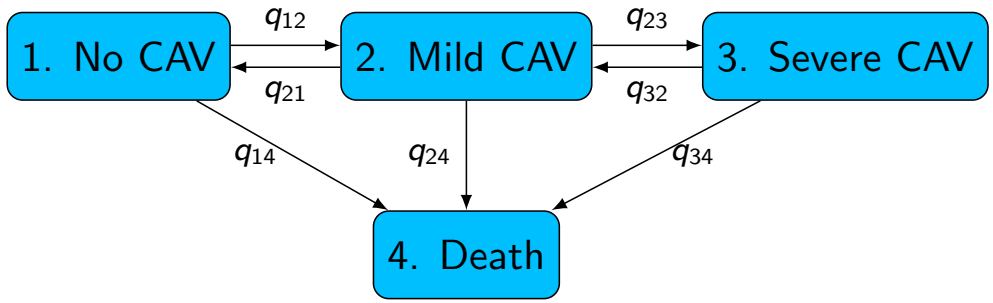
```
##      to
## from    1    2    3    4
##    1 1367  204   44  148
##    2   46  134   54   48
##    3    4   13  107   55
```

Defining the multi-state model structure

Define allowed transitions as a matrix of binary indicators

- rows: from state, columns: to state (diagonal ignored)

```
Q <- rbind(
c(0, 1, 0, 1),
c(1, 0, 1, 1),
c(0, 1, 0, 1),
c(0, 0, 0, 0)
)
```



Defines allowable instantaneous transitions in continuous time, NOT interval transitions in discrete time

- 1–3 transitions were observed for 44 intervals in data
  - but instantaneous 1-3 transition not allowed
- These people assumed to pass unobserved through state 2 (mild) during interval

Estimating the transition intensity matrix

Either then call

- `msm(..., gen.inits=TRUE, ...)` to auto-generate initial values for maximum likelihood estimation using a heuristic
- Usually works, unless model badly misspecified

Or...

- Supply initial values for optimisation alongside the transition structure (0: transition disallowed)

```
Q  <-  rbind (
c(0, 0.25, 0, 0.25),
c(0.166, 0, 0.166, 0.166),
c(0, 0.25, 0, 0.25),
c(0, 0, 0, 0)
)
```

e.g. first row: guess mean of $1/(0.25+0.25) = 2$ years in state 1, and equal chance that next state is 2 or 4

msm function for fitting models: examples (1)

(a) Specify state, time, and subject IDs, and transition structure.

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q)  # assumes all panel data
```

(b) Entry times into State 4 known, but state at previous instant unknown

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q,
                deathexact = 4)
```

(c) Covariates: same covariates on all transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
                covariates = ~ sex + age)
```

(d) Different covariates affect different transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
              covariates = list("1-2" = ~ sex + dage, "2-3"= ~dage))
```

msm function for fitting models: examples (1)

(a) Specify state, time, and subject IDs, and transition structure.

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q)  # assumes all panel data
```

(b) Entry times into State 4 known, but state at previous instant unknown

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q,
                deathexact = 4)
```

(c) Covariates: same covariates on all transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
                covariates = ~ sex + age)
```

(d) Different covariates affect different transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
              covariates = list("1-2" = ~ sex + dage, "2-3"= ~dage))
```

msm function for fitting models: examples (1)

(a) Specify state, time, and subject IDs, and transition structure.

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q)  # assumes all panel data
```

(b) Entry times into State 4 known, but state at previous instant unknown

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q,
                deathexact = 4)
```

(c) Covariates: same covariates on all transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
                covariates = ~ sex + age)
```

(d) Different covariates affect different transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
              covariates = list("1-2" = ~ sex + dage, "2-3"= ~dage))
```

msm function for fitting models: examples (1)

(a) Specify state, time, and subject IDs, and transition structure.

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q)  # assumes all panel data
```

(b) Entry times into State 4 known, but state at previous instant unknown

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q,
                deathexact = 4)
```

(c) Covariates: same covariates on all transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
                covariates = ~ sex + age)
```

(d) Different covariates affect different transition rates

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
              covariates = list("1-2" = ~ sex + dage, "2-3"= ~dage))
```

msm function for fitting models: examples (2)

(e) censor: Observations of 999 in state variable could be states 1, 2, 3 (alive) but not 4 (dead)

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
                covariates = list("1-2" = ~ sex),
                censor=999, censor.states=c(1,2,3)
                )
```

Printing the fitted model object

```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
                qmatrix = Q, deathexact = 4,
                covariates = list("1-2" = ~ sex)
)
cav.msm


##
## Call:
## msm(formula = state ~ years, subject = PTNUM, data = cav, qmatrix = Q,     covariates = list(`1-2` = ~s
##
## Maximum likelihood estimates
## Baselines are with covariates set to their means
##
## Transition intensities with hazard ratios for each covariate
##                   Baseline                       sex
## State 1 - State 1 -0.16937 (-0.18933,-0.15151)
## State 1 - State 2  0.12666 ( 0.11012, 0.14569) 0.5266 (0.3271,0.8478)
## State 1 - State 4  0.04271 ( 0.03434, 0.05312) 1.0000
## State 2 - State 1  0.22672 ( 0.16872, 0.30466) 1.0000
## State 2 - State 2 -0.60839 (-0.70952,-0.52167)
## State 2 - State 3  0.34225 ( 0.27295, 0.42913) 1.0000
## State 2 - State 4  0.03942 ( 0.01090, 0.14256) 1.0000
## State 3 - State 2  0.13052 ( 0.07945, 0.21440) 1.0000
## State 3 - State 3 -0.43661 (-0.55232,-0.34514)
## State 3 - State 4  0.30609 ( 0.23796, 0.39374) 1.0000
##
## -2 * log-likelihood:  3960
## [Note, to obtain old print format, use "printold.msm"]
```

Model predictions

Transition probability matrix $P(t|\mathbf{z}) = Exp(tQ(\mathbf{z}))$

- over a time interval t ($t = 10$ here)
- for covariates $\mathbf{z}$ (sex=1 is women here)

```
p <- pmatrix.msm(cav.msm, t=10, covariates=list(sex=1), ci="normal")
print(p, digits=2)
```

```
##         State 1              State 2              State 3
## State 1 0.423 (0.324,0.501) 0.070 (0.048,0.095) 0.060 (0.038,0.084)
## State 2 0.221 (0.158,0.282) 0.050 (0.034,0.070) 0.065 (0.042,0.089)
## State 3 0.072 (0.043,0.113) 0.025 (0.014,0.040) 0.043 (0.023,0.069)
## State 4 0                   0                   0
##         State 4
## State 1 0.447 (0.398,0.520)
## State 2 0.664 (0.606,0.734)
## State 3 0.860 (0.802,0.913)
## State 4 1.000 (1.000,1.000)
```

Confidence interval for $P(t|\mathbf{z})$ produced by

- simulating from the asymptotic normal distribution of the MLEs for $log(Q), \beta$
- calculating $P(t|\mathbf{z})$ for each simulation
- taking 95% quantiles of the simulated sample of $P(t|\mathbf{z})$

Other output functions

- `qmatrix.msm` for intensity matrix
- `sojourn.msm` for mean sojourn time
- `totlos.msm` for total length of stay
- ... and more.

All take `covariates` argument to return the output corresponding to a particular set of covariate values.

Contents

Introduction 2

Mathematics of multi-state processes 8

Panel data 19

- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53

- Theory: generalisation of survival models 54
- Implementation and examples 61

Model building advice

Easy to over-parameterise models when multi-state process is only intermittently-observed

- ... can lead to flat likelihood, parameters not identifiable, convergence failure
- Make sure state structure represents allowed instantaneous (not interval) transitions
- Simplify the model if necessary
  - Exclude covariate effects on certain transitions
  - Constrain covariate effects, e.g. same effect of age on rate of death from any state: `msm(..., constraint=)`
  - Fix parameters at their initial values instead of optimising: `msm(..., fixedpars=)`
- Merge adjacent states e.g. mild/moderate disease

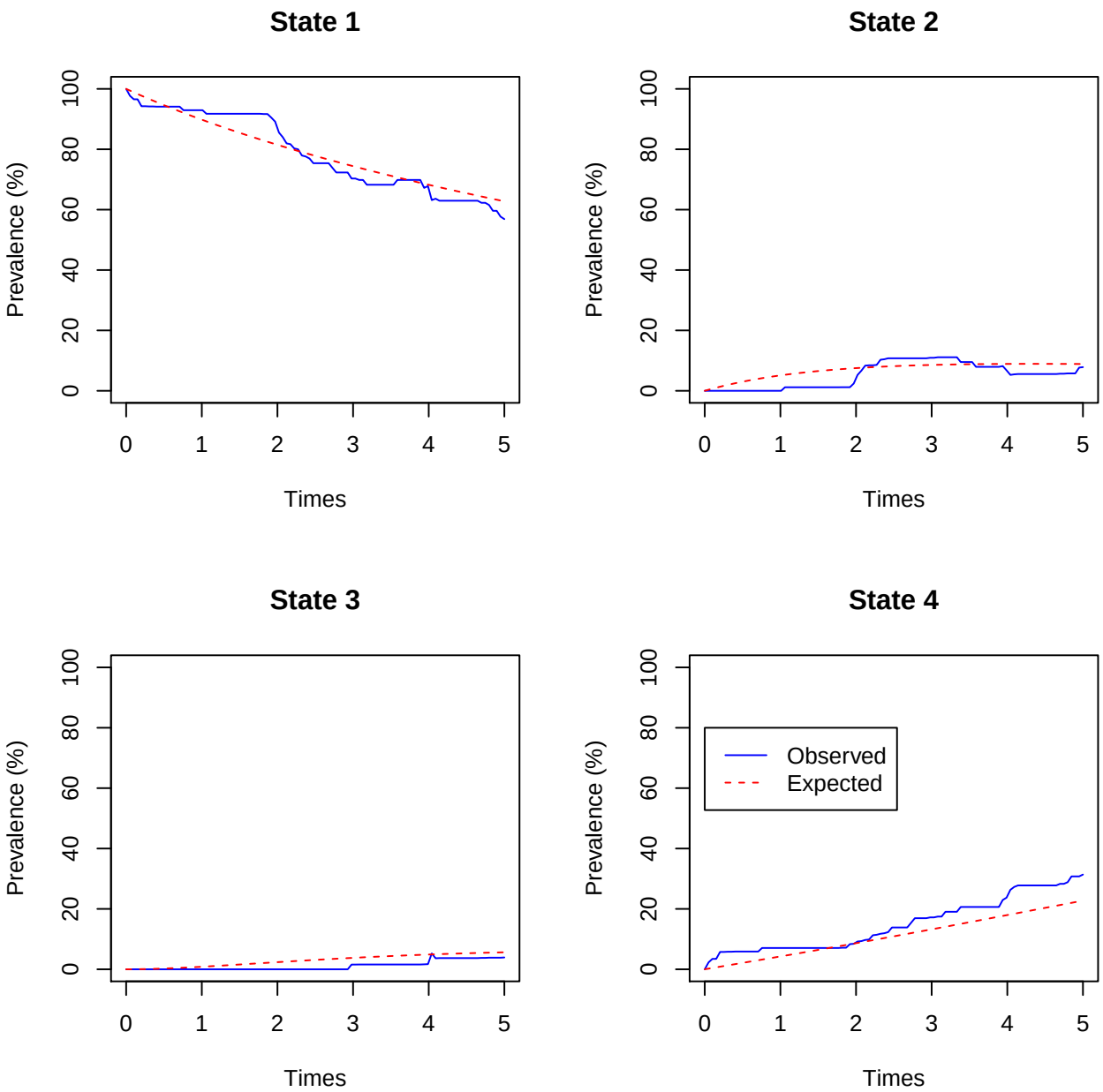
Checking predictions from fitted models

Compare the

- observed proportion of people in each state at each time t
  - relies on interpolation if everybody observed at different times
- expected proportion, using $P(t|\mathbf{z})$ from the fitted model, given, e.g. everyone in state 1 at time 0.
- for specific subgroup of data $\mathbf{z}$ (here, women)

```
women <- unique(cav$PTNUM[cav$sex==1])
plot.prevalence.msm(cav.msm, maxtime=5,
                    covariates=list(sex=1),
                    subset=women)
```

Checking predictions from fitted models



Model checking (other methods)

Pearson-type goodness-of-fit test

- Compares observed and expected transition counts
  - grouped in various ways (covariates, time interval length...)
  - generalized χ^2-type test (Aguirre-Hernandez and Farewell (2002) Stat. Med. 21, Titman and Sharples (2008) Stat. Med. 27.)
- Implemented in `msm` as `pearson.msm`
- Can be tricky to interpret
  - p-value sensitive to choice of grouping
  - if fit is poor, difficult to judge what model may be better.
  - statistically $\neq$ practically significant poor fit

See Titman and Sharples (Stat. Meth. Med. Res. 2010) for discussion of other model diagnostic methods

- e.g. if death times known, compare model predictions of survival with Kaplan-Meier estimates

Model checking (other methods)

Pearson-type goodness-of-fit test

- Compares observed and expected transition counts
  - grouped in various ways (covariates, time interval length...)
  - generalized χ^2-type test (Aguirre-Hernandez and Farewell (2002) Stat. Med. 21, Titman and Sharples (2008) Stat. Med. 27.)
- Implemented in `msm` as `pearson.msm`
- Can be tricky to interpret
  - p-value sensitive to choice of grouping
  - if fit is poor, difficult to judge what model may be better.
  - statistically $\neq$ practically significant poor fit

See Titman and Sharples (Stat. Meth. Med. Res. 2010) for discussion of other model diagnostic methods

- e.g. if death times known, compare model predictions of survival with Kaplan-Meier estimates

Model comparison

Assess model by comparing it against an expanded model with a questionable assumption relaxed

- examine parameter estimates in expanded model
- use likelihood ratio test or AIC
- e.g. different covariates, affecting different transitions
- time-(in)homogeneity of transition rates
- Markov vs semi-Markov models, hidden Markov models (see later...)

Contents

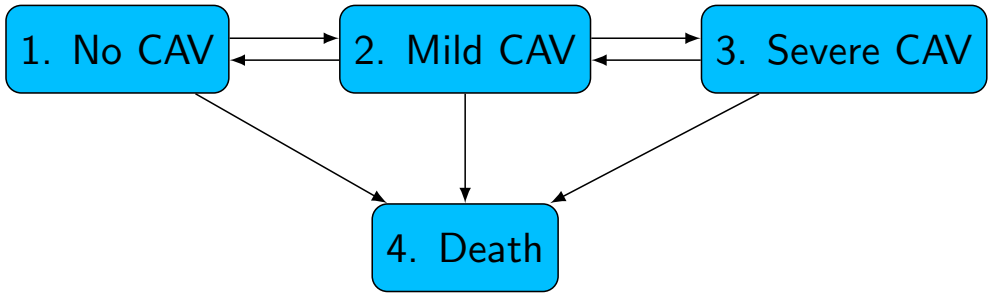
Introduction 2

Mathematics of multi-state processes 8

Panel data 19
- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

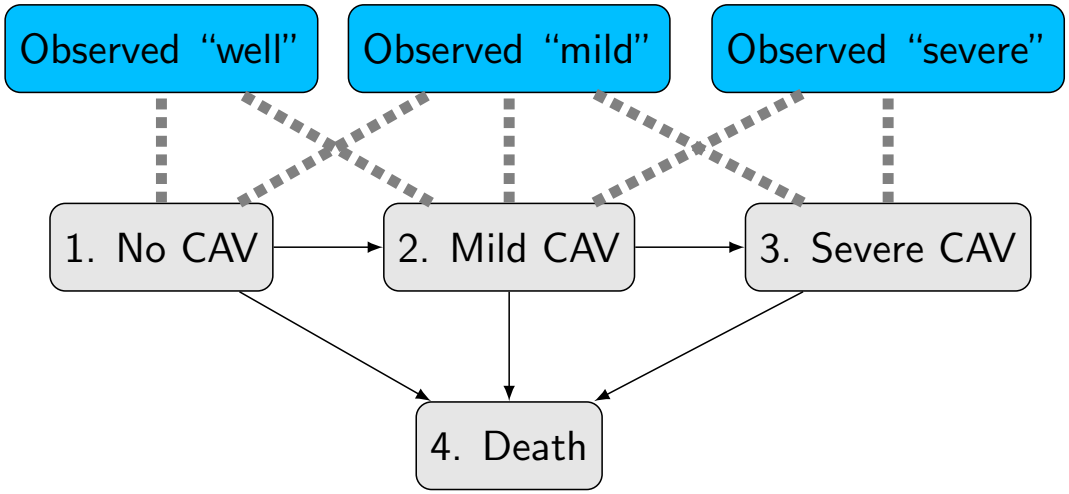
Continuously-observed processes 53
- Theory: generalisation of survival models 54
- Implementation and examples 61

Hidden Markov models: misclassification of states



- Transitions from worse to better states observed in data, but implausible
- CAV is known to be an irreversible process
- Screening test (angiography) is subject to misclassification

Hidden Markov models: misclassification of states



In CAV example, could assume $e_{13} = e_{31} = 0$

- Transitions from worse to better states observed in data, but implausible
- CAV is known to be an irreversible process
- Screening test (angiography) is 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 multi-state model: hidden Markov model

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 Markov

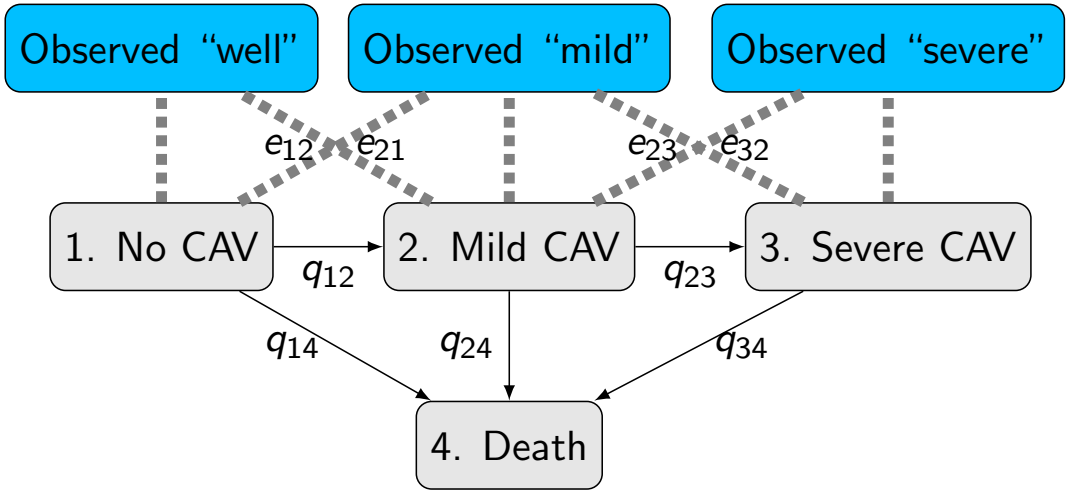
$$L_i = \sum_{X_{i0}} P(O_{i0}|X_{i0}) P(X_{i0}) \sum_{X_{i1}} P(O_{i1}|X_{i1}) P(X_{i1}|X_{i0}) \ldots \sum_{X_{in_i}} P(O_{in_i}|X_{in_i}) P(X_{in_i}|X_{i,n_i-1})$$

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. In CAV example, know true $X_{i0} = 1$: CAV free at transplant
- Covariates can also be included on the misclassification probabilities using multinomial logistic regression

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

Misclassification models in msm



Transition intensities

```
Q  <-  rbind(
c(0, 0.2, 0, 0.2),
c(0, 0, 0.2, 0.2),
c(0, 0, 0, 0.2),
c(0, 0, 0, 0)
)
```

Misclassification probabilities

```
E  <-  rbind(
c(0, 0.1, 0, 0),
c(0.1, 0, 0.1, 0),
c(0, 0.1, 0, 0),
c(0, 0, 0, 0)
)
```

Non-zero values estimated starting from these initial values

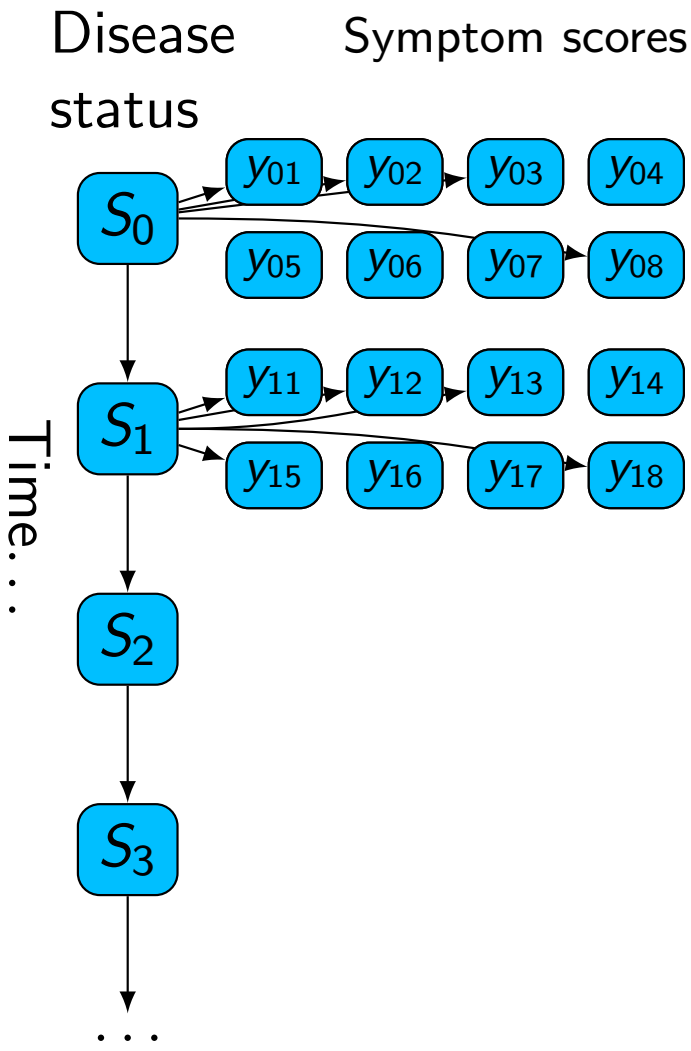
```
cav.msm <- msm( state ~ years, subject=PTNUM, data = cav,
               qmatrix = Q, ematrix=E, deathexact=4)
```

Misclassification models in msm

```
cav.msm


##
## Call:
## msm(formula = state ~ years, subject = PTNUM, data = cav, qmatrix = Q,      ematrix = E, deathexact = 4)
##
## Maximum likelihood estimates
##
## Transition intensities
##                     Baseline
## State 1 - State 1 -0.14202 (-0.159907,-0.12613)
## State 1 - State 2  0.10134 ( 0.086613, 0.11857)
## State 1 - State 4  0.04068 ( 0.032529, 0.05087)
## State 2 - State 2 -0.26067 (-0.316165,-0.21492)
## State 2 - State 3  0.22673 ( 0.169044, 0.30410)
## State 2 - State 4  0.03394 ( 0.008594, 0.13404)
## State 3 - State 3 -0.30847 (-0.390575,-0.24362)
## State 3 - State 4  0.30847 ( 0.243624, 0.39058)
##
## Misclassification probabilities
##                         Baseline
## Obs State 1 | State 1 0.992337 (0.982062,0.99675)
## Obs State 2 | State 1 0.007663 (0.003254,0.01794)
## Obs State 1 | State 2 0.245021 (0.154227,0.36613)
## Obs State 2 | State 2 0.703943 (0.551335,0.82145)
## Obs State 3 | State 2 0.051036 (0.029710,0.08631)
## Obs State 2 | State 3 0.124326 (0.062491,0.23219)
## Obs State 3 | State 3 0.875674 (0.767807,0.93751)
##
## -2 * log-likelihood:  3952
## [Note, to obtain old print format, use "printold.msm"]
```

Hidden Markov model underlying any outcome



- 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
  - could be continuous or discrete, e.g. different symptoms
- Specify joint distribution of $(\mathbf{y}_{ij}|S)$ for each hidden disease state S
  - Many standard distributions in `msm`
- 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), `msm` package.
Jackson, Su, Gladman & Farewell (Arthritis Care and Research 2017)

Semi-Markov models

Markov models: transition rates independent of

- previous states visited
- time in current state

Semi-Markov models

Semi-Markov models: transition rates

- independent of previous states visited
- depend on time in current state

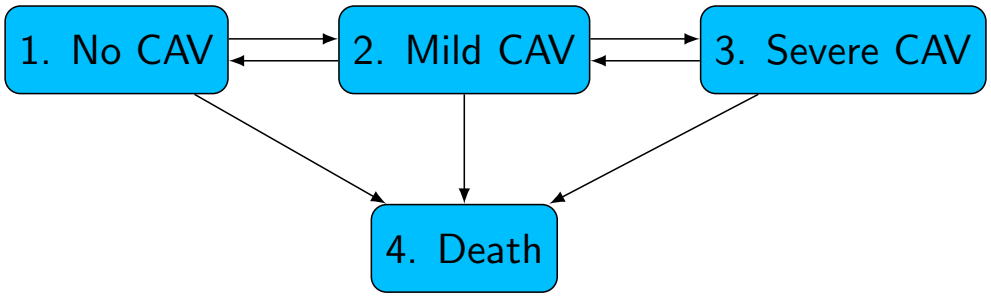
Semi-Markov models

Semi-Markov models: transition rates

- independent of previous states visited
- depend on time in current state

Phase-type model

- Replace the state by a series of hidden "phases".
- Expanded state structure follows a Markov process.
- Sojourn time in state 2 follows "Coxian" distribution instead of exponential



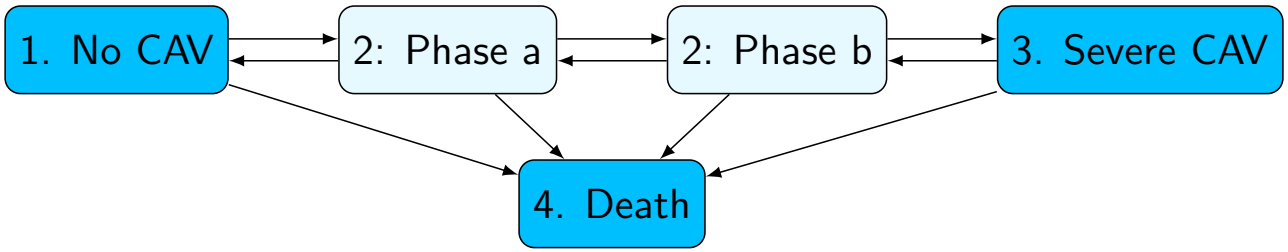
Semi-Markov models

Semi-Markov models: transition rates

- independent of previous states visited
- depend on time in current state

Phase-type model

- Replace the state by a series of hidden "phases".
- Expanded state structure follows a Markov process.
- Sojourn time in state 2 follows "Coxian" distribution instead of exponential



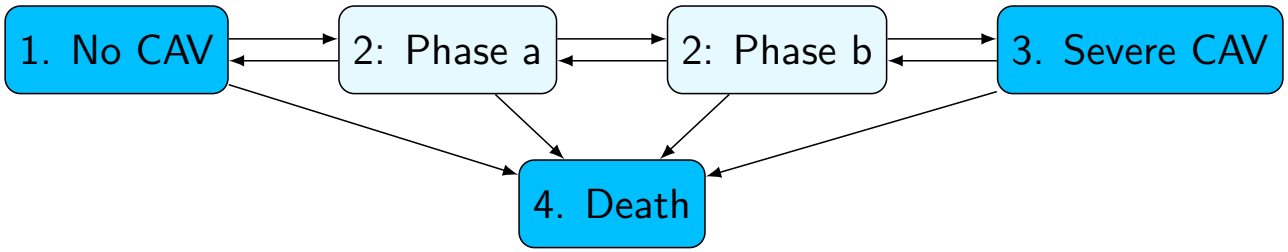
Semi-Markov models

Semi-Markov models: transition rates

- independent of previous states visited
- depend on time in current state

Phase-type model

- Replace the state by a series of hidden "phases".
- Expanded state structure follows a Markov process.
- Sojourn time in state 2 follows "Coxian" distribution instead of exponential



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

Other advanced models (a)

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)
- Maximum marginal likelihood (see e.g. O'Keeffe, Tom & Farewell, JRSSC 2011; Yiu, Tom & Farewell, Stat Med. 2016 in press; Cook et al, Biometrics 2004)
- Bayesian approaches
  - See e.g. van den Hout, *Multi-state Survival Models for Interval Censored Data* CRC 2017
  - JAGS and Stan have matrix exponential, allowing general panel data likelihood
  - MCMC estimation

Other advanced models (b)

Informative observation times

- Observation time may depend on status at that time
- Jointly model observation process and outcome process
  - 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 (b)

Informative observation times

- Observation time may depend on status at that time
- Jointly model observation process and outcome process
  - 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)

Further resources: multi-state models for panel data

- A. van den Hout, *Multi-state Survival Models for Interval Censored Data*, CRC Press, 2017
- `msm` package: `http://CRAN.R-project.org/package=msm`
- Cook & Lawless (2014) Statistical issues in modeling chronic disease in cohort studies *Statistics in Biosciences* 6(1):127–261 (nice review paper)

Contents

Introduction 2

Mathematics of multi-state processes 8

Panel data 19
- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53
- Theory: generalisation of survival models 54
- Implementation and examples 61

Contents

Introduction 2

Mathematics of multi-state processes 8

Panel data 19
- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53
- Theory: generalisation of survival models 54
- Implementation and examples 61

Continuously-observed processes

N individuals, each followed up for a different length of time

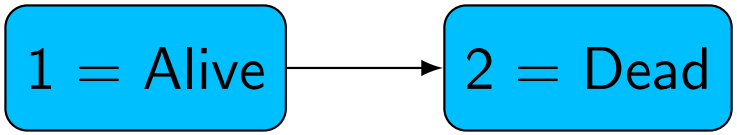
- Each person's state known throughout their follow-up
- Times of all transitions known

Continuously-observed processes

N individuals, each followed up for a different length of time

- Each person's state known throughout their follow-up
- Times of all transitions known

Example: standard survival analysis with right-censoring.



At end of each person's follow-up, they are known to be

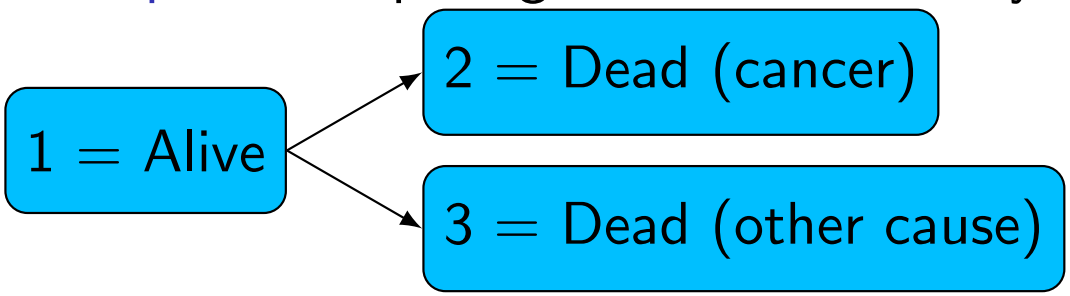
- dead, with known death time
- alive, with right-censored death time.

Continuously-observed processes

N individuals, each followed up for a different length of time

- Each person's state known throughout their follow-up
- Times of all transitions known

Example: competing risks survival analysis.

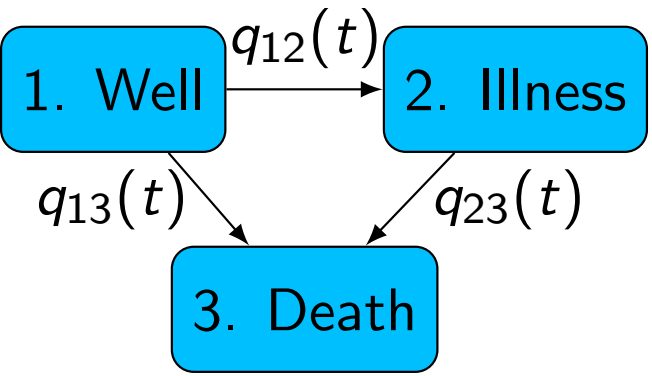


At end of each person's follow-up, they are known to be

- alive, with right-censored death time, cause unknown.
- dead, with cause known, at known time

Time to death from other cause is censored — gives information about the other-cause hazard

Example: illness-death model



Typical data:

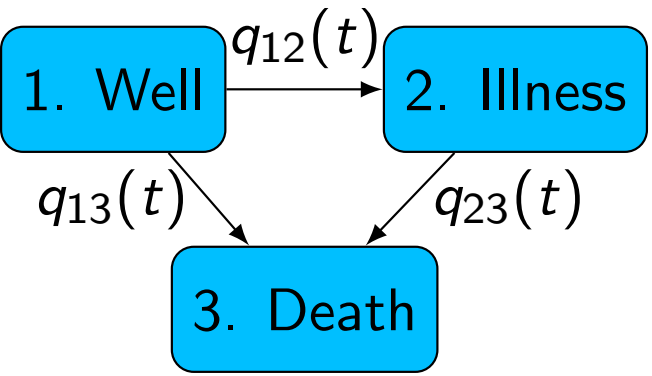
Person	Time	Event	State
1	0	Start of process	1
1	45	Alive and well	1
2	0	Start of process	1
2	65	Illness onset	2
2	85	Death	3
3	0	Start of process	1
3	25	Death without illness	3
...			

Model governed by transition intensities (hazards, rates) for each potential transition: $q_{12}(t), q_{23}(t), q_{13}(t)$.

Alternative time scales

- Semi-Markov model / clock-reset:
  t is time of entry into current state
  - death rate from illness $q_{23}(t)$ depends on duration of illness t
- Inhomogeneous Markov model / clock-forward:
  t is start of the multi-state process
  - e.g. $q_{23}(t)$ depends on age t, but not duration of illness

Example: illness-death model



Typical data:

Person	Time	Event	State
1	0	Start of process	1
1	45	Alive and well	1
2	0	Start of process	1
2	65	Illness onset	2
2	85	Death	3
3	0	Start of process	1
3	25	Death without illness	3
...			

Model governed by transition intensities (hazards, rates) for each potential transition: $q_{12}(t), q_{23}(t), q_{13}(t)$.

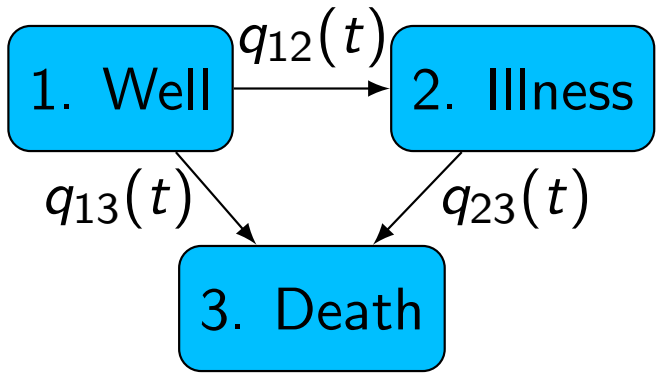
Alternative time scales

- Semi-Markov model / clock-reset:
  t is time of entry into current state
  - death rate from illness $q_{23}(t)$ depends on duration of illness t
- Inhomogeneous Markov model / clock-forward:
  t is start of the multi-state process
  - e.g. $q_{23}(t)$ depends on age t, but not duration of illness

Converting data to time-to-event format

Arrange from longitudinal format (one row per event)...

Person	Time	Event	State
1	0	Start of process	1
1	45	Alive without illness	1
2	0	Start of process	1
2	65	Illness onset	2
2	85	Death	3
3	0	Start of process	1
3	25	Death without illness	3



to time-to-event format (one row per potential transition)...

Person	Start time	Stop time	Transition	Status
1	0	45	1–2	Censored
1	0	45	1–3	Censored
2	0	65	1–2	Observed
2	0	65	1–3	Censored
2	65	85	2–3	Observed
3	0	25	1–2	Censored
3	0	25	1–3	Observed

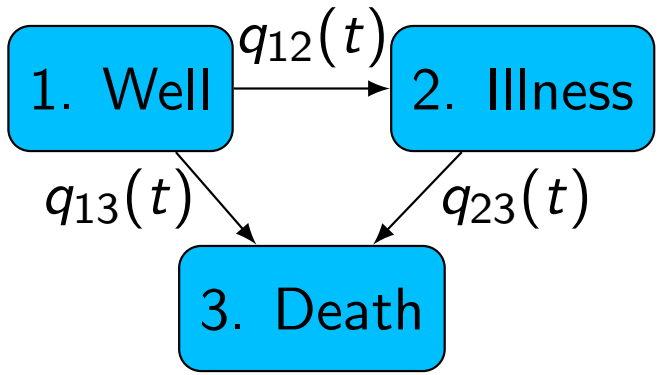
Start time: time when become at risk of the transition event

Each row informs the rate of the corresponding transition...

Converting data to time-to-event format

Arrange from longitudinal format (one row per event)...

Person	Time	Event	State
1	0	Start of process	1
1	45	Alive without illness	1
2	0	Start of process	1
2	65	Illness onset	2
2	85	Death	3
3	0	Start of process	1
3	25	Death without illness	3



to time-to-event format (one row per potential transition)...

Person	Start time	Stop time	Transition	Status
1	0	45	1–2	Censored
1	0	45	1–3	Censored
2	0	65	1–2	Observed
2	0	65	1–3	Censored
2	65	85	2–3	Observed
3	0	25	1–2	Censored
3	0	25	1–3	Observed

Start time: time when become at risk of the transition event

Each row informs the rate of the corresponding transition...

Transition-specific time-to-event models

Each r, s transition is governed by its own time-to-event model

- PDF $f_{rs}(t)$, CDF $F_{rs}(t)$ defined in terms of intensities $q_{rs}(t)$

Cumulative transition-specific hazard

$$H_{rs}(t) = \int_0^t q_{rs}(u)du$$

"Survivor" function, $P(T > t)$ where T is the time to entering state s from state r

$$S_{rs}(t) = \exp(-H_{rs}(t))$$

CDF and PDF

$$F_{rs}(t) = 1 - S_{rs}(t), \quad f_{rs}(t) = \frac{d}{dt}F_{rs}(t)$$

Forming the likelihood for a semi-Markov ("clock-reset") model

Person	t_i^{start}	t_i^{stop}	$dt_i = t_i^{stop} - t_i^{start}$	Transition	Status
2	0	65	65	1–2	Observed
2	0	65	65	1–3	Censored
2	65	85	20	2–3	Observed

Transition rate $q_{rs}(t)$ depends on time t since entering state r

- Outcome is the time interval $dt_i = t_i^{stop} - t_i^{start}$
  - either an observed event time or a censoring time
- Each dt_i contributes one term to the likelihood for the parameters governing the transition rate $q_{rs}(t)$
  - $f_{rs}(dt_i)$ if r, s transition observed
  - $S_{rs}(dt_i) = 1 - F_{rs}(t_i)$ if r, s transition censored

  assuming independent censoring.
- Here likelihood contribution is $f_{12}(65) \times S_{13}(65) \times f_{23}(20) \ldots$
- Collect together similar terms from the entire dataset...

Implementation as a standard survival model

Full likelihood: $\ell_{12}(\boldsymbol{\alpha}_{12}|\mathbf{t}_{12}) \times \ell_{13}(\boldsymbol{\alpha}_{13}|\mathbf{t}_{13}) \times \ell_{23}(\boldsymbol{\alpha}_{23}|\mathbf{t}_{23})$

- $\mathbf{t}_{rs}$ are the data informing the r, s transition model
- $\boldsymbol{\alpha}_{rs}$ are the basic parameters governing the transition intensities $q_{rs}(t)$
  - $q_{rs}(t) = \boldsymbol{\alpha}_{rs}$, exponential time-to-event $\equiv$ constant hazard
  - or $\boldsymbol{\alpha}_{rs}$ could include the shape and scale of a Weibull model
  - $\boldsymbol{\alpha}_{rs}$ could also include covariate effects

If parameters α_{rs} independent for different r, s transitions

Just fit a standard survival model to data for each r, s pair in turn

If some parameters shared between different α_{rs}

- e.g. identical effects of covariates on different transitions
- Fit one survival model to entire dataset
  Transition type can be included as a categorical covariate

Implementation as a standard survival model

Full likelihood: $\ell_{12}(\boldsymbol{\alpha}_{12}|\mathbf{t}_{12}) \times \ell_{13}(\boldsymbol{\alpha}_{13}|\mathbf{t}_{13}) \times \ell_{23}(\boldsymbol{\alpha}_{23}|\mathbf{t}_{23})$

- $\mathbf{t}_{rs}$ are the data informing the r, s transition model
- $\boldsymbol{\alpha}_{rs}$ are the basic parameters governing the transition intensities $q_{rs}(t)$
  - $q_{rs}(t) = \boldsymbol{\alpha}_{rs}$, exponential time-to-event $\equiv$ constant hazard
  - or $\boldsymbol{\alpha}_{rs}$ could include the shape and scale of a Weibull model
  - $\boldsymbol{\alpha}_{rs}$ could also include covariate effects

If parameters $\boldsymbol{\alpha}_{rs}$ independent for different r, s transitions

Just fit a standard survival model to data for each r, s pair in turn

If some parameters shared between different $\boldsymbol{\alpha}_{rs}$

- e.g. identical effects of covariates on different transitions
- Fit one survival model to entire dataset
  Transition type can be included as a categorical covariate

Implementation as a standard survival model

Full likelihood: $\ell_{12}(\boldsymbol{\alpha}_{12}|\mathbf{t}_{12}) \times \ell_{13}(\boldsymbol{\alpha}_{13}|\mathbf{t}_{13}) \times \ell_{23}(\boldsymbol{\alpha}_{23}|\mathbf{t}_{23})$

- $\mathbf{t}_{rs}$ are the data informing the r, s transition model
- $\boldsymbol{\alpha}_{rs}$ are the basic parameters governing the transition intensities $q_{rs}(t)$
  - $q_{rs}(t) = \boldsymbol{\alpha}_{rs}$, exponential time-to-event $\equiv$ constant hazard
  - or $\boldsymbol{\alpha}_{rs}$ could include the shape and scale of a Weibull model
  - $\boldsymbol{\alpha}_{rs}$ could also include covariate effects

If parameters $\boldsymbol{\alpha}_{rs}$ independent for different r, s transitions

Just fit a standard survival model to data for each r, s pair in turn

If some parameters shared between different $\boldsymbol{\alpha}_{rs}$

- e.g. identical effects of covariates on different transitions
- Fit one survival model to entire dataset
  Transition type can be included as a categorical covariate

Contents

Introduction 2

Mathematics of multi-state processes 8

Panel data 19
- Likelihoods for panel data models 19
- Fitting multi-state models to panel data with `msm` 26
- Model checking and comparison 37
- Advanced panel data models 43

Continuously-observed processes 53
- Theory: generalisation of survival models 54
- Implementation and examples 61

Options for modelling and software

Cox regression: $q_{rs}(t) = q_{rs}^{(0)}(t)\exp(\beta'_{rs}\mathbf{x})$

- estimate hazard ratios $\exp(\beta)$ on risk of r, s transition
- baseline hazard $q_{rs}^{(0)}(t)$ unspecified
- `coxph` in standard `survival` package in R
- `mstate` package (de Wreede et al., see later) for multi-state interpretation of fitted Cox model(s)

Parametric survival modelling

- $q_{rs}^{(0)}(t)$ is the hazard function of a parametric survival distribution
- `flexsurv` R package: wide choices of distributions, multi-state modelling features
- `survreg` in `survival` R package (but limited, see later)

See also Stata package `multistate` (Crowther & Lambert)

Options for modelling and software

Cox regression: $q_{rs}(t) = q_{rs}^{(0)}(t) \exp(\beta'_{rs}\mathbf{x})$

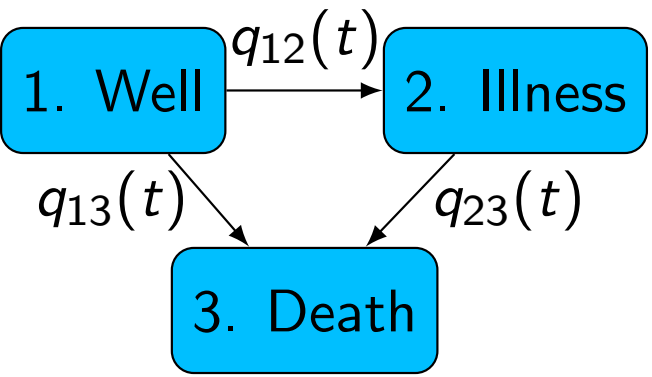
- estimate hazard ratios $\exp(\beta)$ on risk of r, s transition
- baseline hazard $q_{rs}^{(0)}(t)$ unspecified
- `coxph` in standard `survival` package in R
- `mstate` package (de Wreede et al., see later) for multi-state interpretation of fitted Cox model(s)

Parametric survival modelling

- $q_{rs}^{(0)}(t)$ is the hazard function of a parametric survival distribution
- `flexsurv` R package: wide choices of distributions, multi-state modelling features
- `survreg` in `survival` R package (but limited, see later)

See also Stata package `multistate` (Crowther & Lambert)

Semi-Markov models: examples



person	tstart	tstop	years	trans	status
2	0	65	65	1–2	Observed
2	0	65	65	1–3	Censored
2	65	85	20	2–3	Observed
...					

Time from state r entry to state s entry could be

1. Weibull with different shape and scale for each transition type: $q_{rs}(t) = \lambda_{rs}\alpha_{rs}t^{\alpha_{rs}-1}$
2. Weibull with proportional hazards, common shape: $q_{rs}(t) = \lambda_{rs}\alpha t^{\alpha-1}$
3. Cox with different baseline hazards: $q_{rs}(t)$ nonparametric
4. Cox with proportional hazards, common baseline: $q_{rs}(t) = q^{(0)}(t)\exp(\beta_{rs})$

```
bwei <- flexsurvreg(Surv(years, status) ~ trans + shape(trans),
                    data=bosms3, dist="weibullPH") # 1
cwei <- flexsurvreg(Surv(years, status) ~ trans, data=bosms3,
                    dist="weibullPH") # 2
bcox <- coxph(Surv(years, status) ~ strata(trans), data=bosms3) # 3
ccox <- coxph(Surv(years, status) ~ trans, data=bosms3) # 4
```

Covariates could be included — interaction with transition type if necessary

Likelihood for an inhomogeneous Markov ("clock-forward") model

Person	t_i^{start}	t_i^{stop}	$dt_i = t_i^{stop} - t_i^{start}$	Transition	Status
2	0	65	65	1–2	Observed
2	0	65	65	1–3	Censored
2	65	85	20	2–3	Observed

Outcome is t_i^{stop}: time since beginning of the process (rather than time dt_i since last transition)

- 2–3 transition time here is left-truncated at $t_i^{start} = 65$
  - since 2–3 transition impossible until enter state 2
  - condition on this event not happening by $t_i^{start} = 65$
- Contribution of ith data row to likelihood for $q_{rs}(t)$ is
  - $f_{rs}(t_i^{stop})/S_{rs}(t_i^{start})$ if $r - s$ transition observed
  - $S_{rs}(t_i^{stop})/S_{rs}(t_i^{start})$ if $r - s$ transition censored

For transitions from state 1, same as semi-Markov model ($S_{rs}(0) = 1$)

Software for inhomogeneous Markov models

Left-truncated time-to-event data specified in R model formula as, e.g.

```
fcox <- coxph(Surv(Tstart, Tstop, status) ~ strata(trans),
              data=bosms3)
```

- `coxph` supports left-truncation
- `survreg` doesn't support left-truncation.
  `flexsurv` supports left-truncated fully parametric survival models — and allows multi-state interpretation (see later)

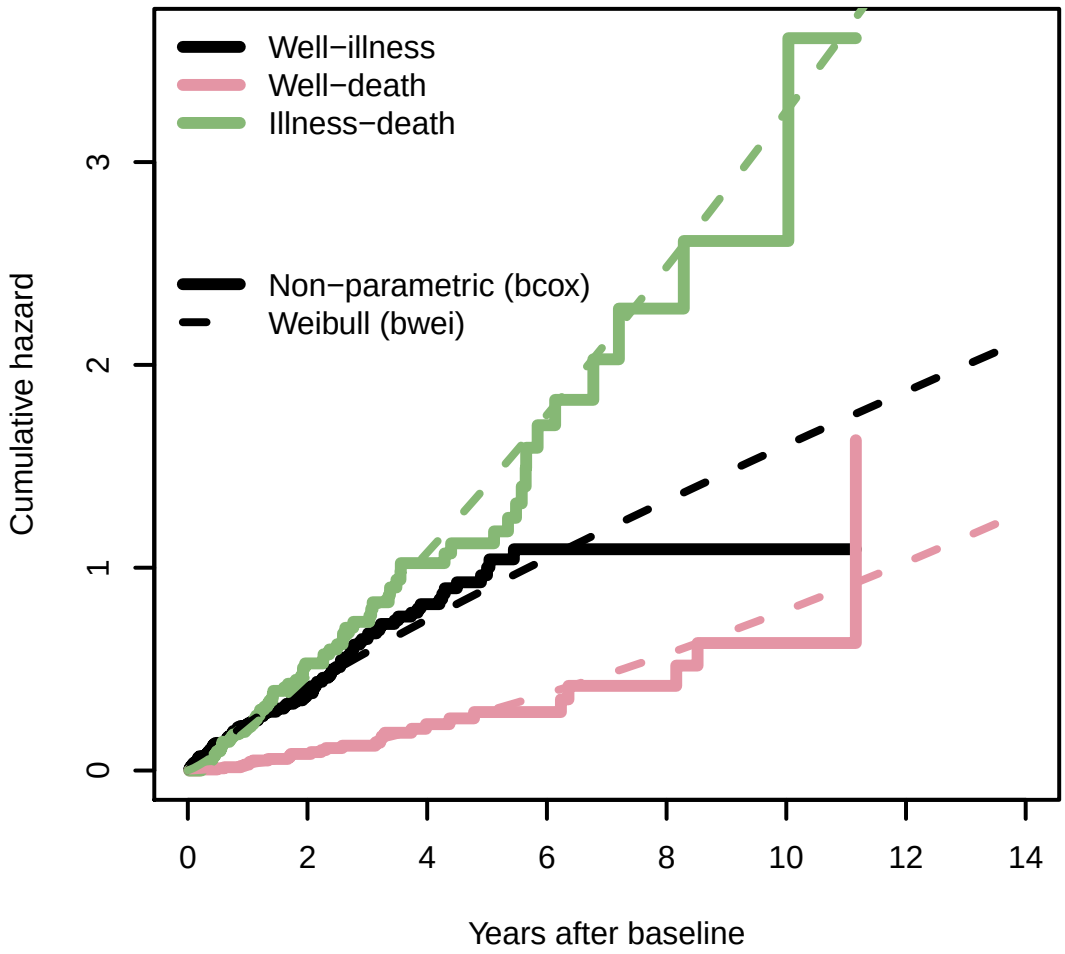
Comparing models: cumulative transition-specific hazards

Cox models

- Cumulative baseline hazard can be estimated nonparametrically: Breslow estimator
- see `mstate` package, de Wreede et al (2010) Comp Meth Prog Biomed 99:261–274, for theory.

Fully-parametric models

- Cumulative hazards explicit in model specification
- Check fit against nonparametric estimates



Example code in the vignette for `flexsurv`, and in `mstate`

Transition probabilities

Probability of being in certain state at given point in the future

$$P(t_0, t) = P(X(t_0 + t) = s | X(t_0) = r)$$

Inhomogeneous Markov models: Solve the Kolmogorov equation

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

- Nonparametric: Aalen-Johansen estimator

$$\hat{P}(t_0, t) = \prod_{i=0}^{m-1} \{I + Q(t_i)dt\}$$

given fitted hazard increments $Q(t_i)dt$ on grid of survival times:

`probtrans` in `mstate`

- Parametric: numerically solve. `flexsurv` calls `deSolve` package (Soetaert et al 2010) for this.

Transition probabilities

Probability of being in certain state at given point in the future

$$P(t_0, t) = P(X(t_0 + t) = s | X(t_0) = r)$$

Semi-Markov models

- $Q(t)$ is no longer a deterministic function of t: depends on the transition history!
- Kolmogorov equation no longer applies
- Simulate transition history for a large number of individuals
  - from the nonparametric estimates of the cumulative hazards `mssample` in `mstate`
  - from the fitted parametric model `pmatrix.simfs` in `flexsurv`: typically faster

  and summarise

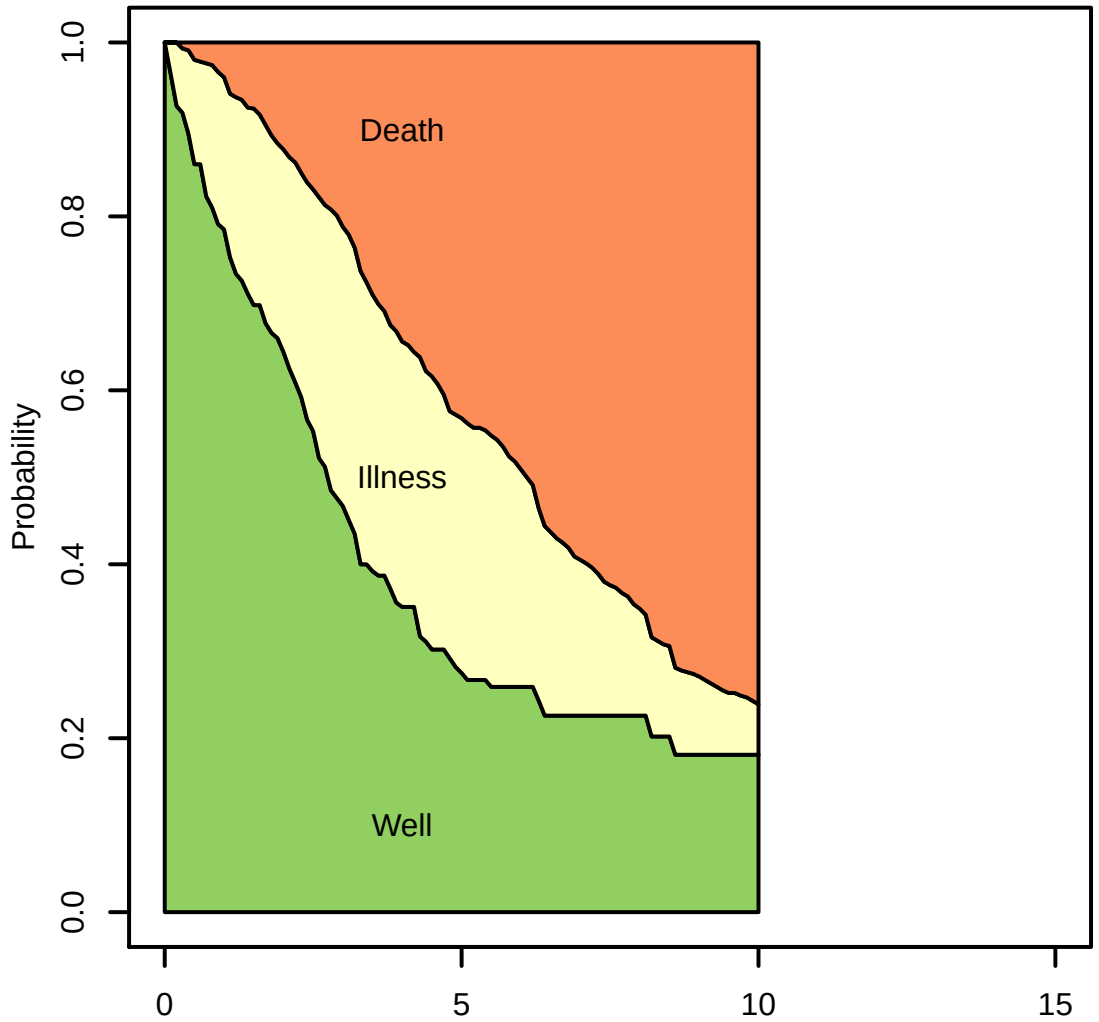
Transition probabilities: example

For someone well at time 0, simulated probabilities of being in each state s, $p_{0s}(t) = P(X(t) = s|X(0) = 1)$, against t

Semi-Markov models:

- Nonparametric estimates (`bcox`)...
- compared with estimates from Weibull parametric model (`bwei`).
- Standard errors require a second level of simulation (available in `flexsurv`)

Could integrate $p_{0s}(t)$ to give expected length of time spent in state s (`totlos.simfs`)



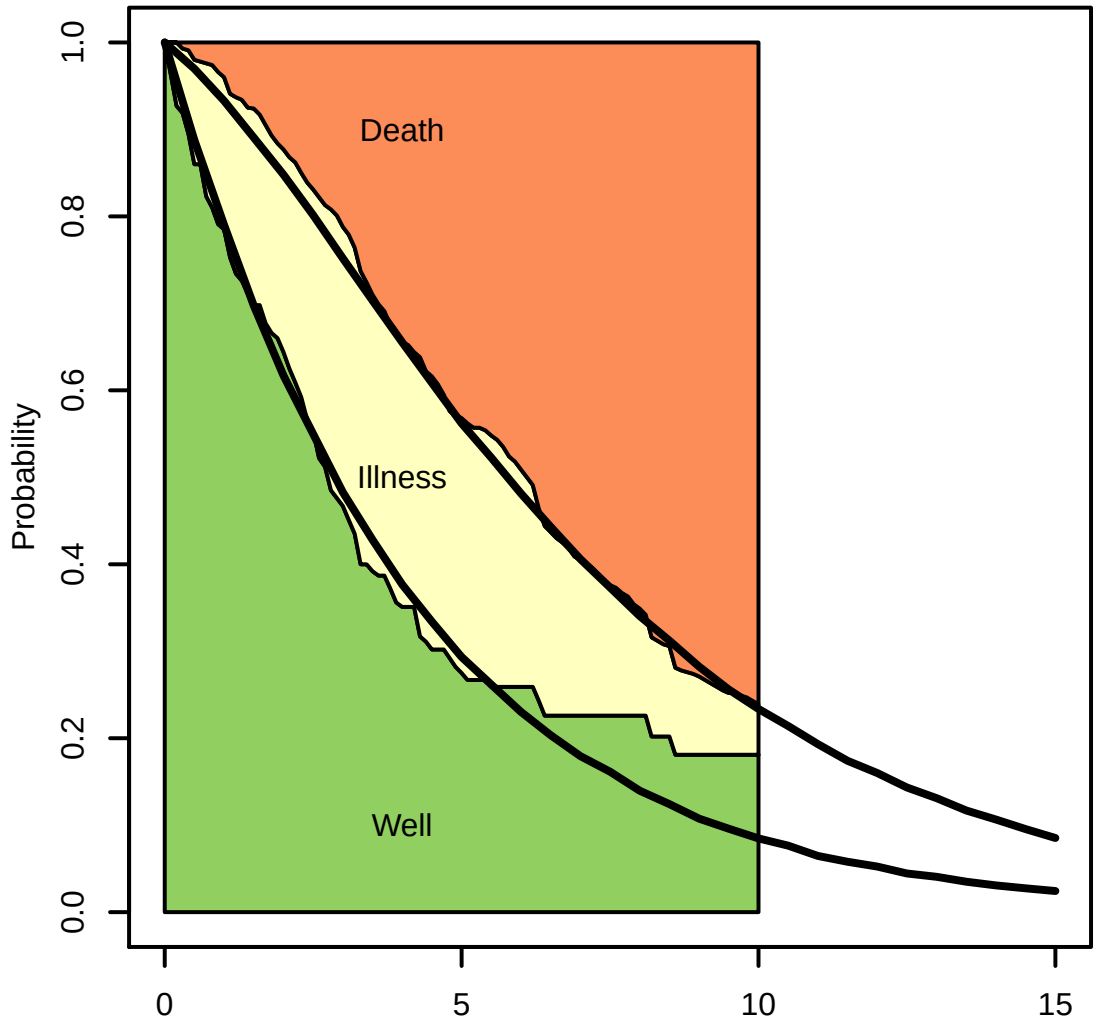
Transition probabilities: example

For someone well at time 0, simulated probabilities of being in each state s, $p_{0s}(t) = P(X(t) = s|X(0) = 1)$, against t

Semi-Markov models:

- Nonparametric estimates (`bcox`)...
- compared with estimates from Weibull parametric model (`bwei`).
- Standard errors require a second level of simulation (available in `flexsurv`)

Could integrate $p_{0s}(t)$ to give expected length of time spent in state s (`totlos.simfs`)



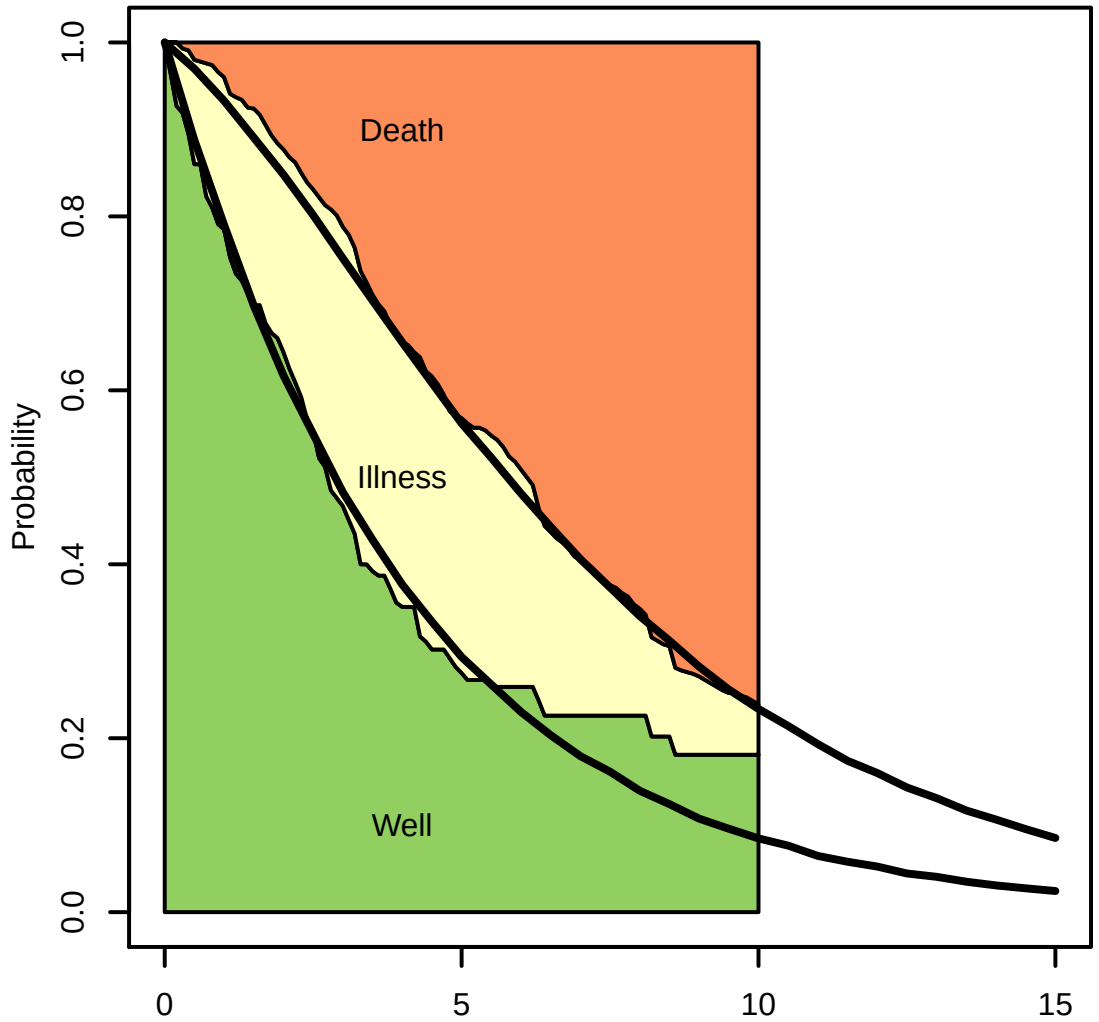
Transition probabilities: example

For someone well at time 0, simulated probabilities of being in each state s, $p_{0s}(t) = P(X(t) = s|X(0) = 1)$, against t

Semi-Markov models:

- Nonparametric estimates (`bcox`)...
- compared with estimates from Weibull parametric model (`bwei`).
- Standard errors require a second level of simulation (available in `flexsurv`)

Could integrate $p_{0s}(t)$ to give expected length of time spent in state s (`totlos.simfs`)



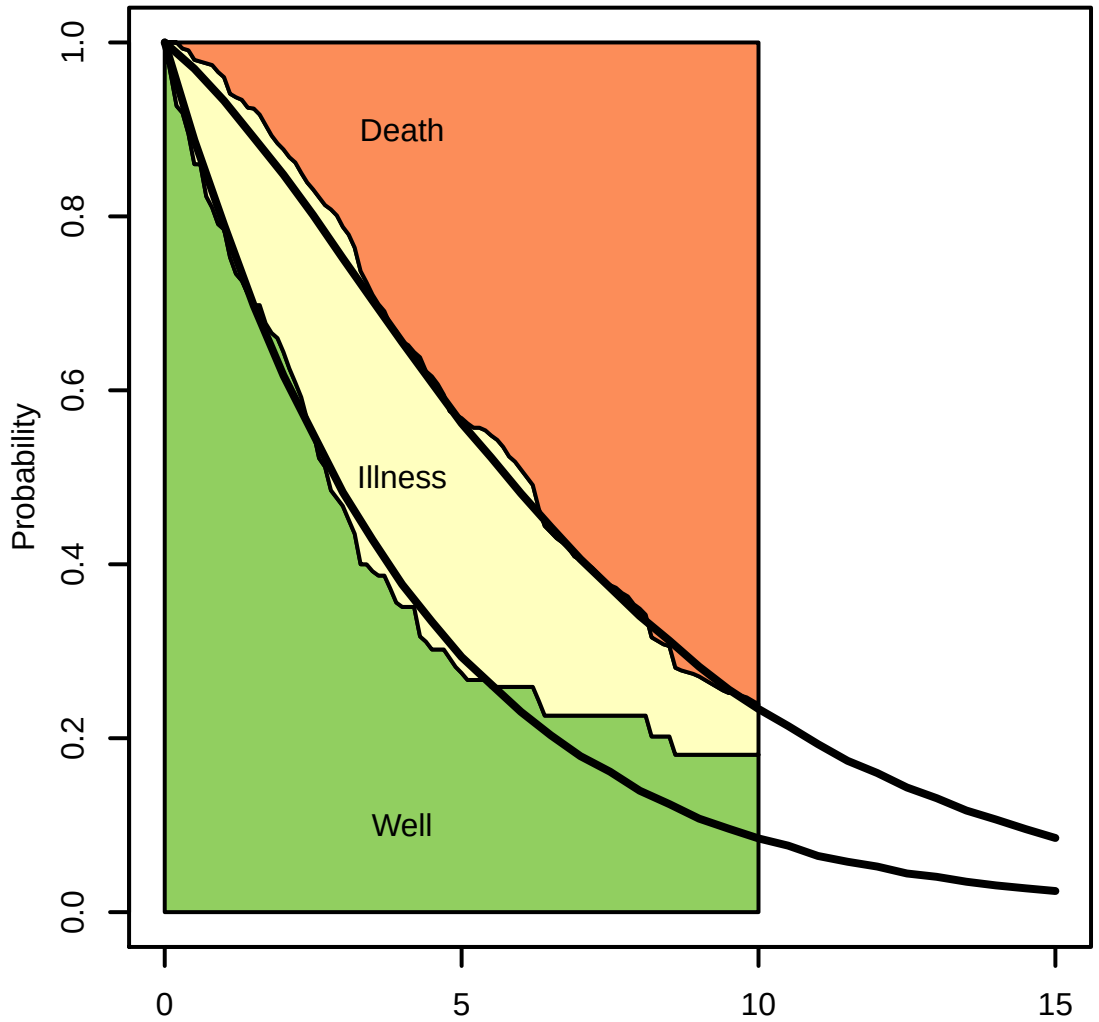
Transition probabilities: example

For someone well at time 0, simulated probabilities of being in each state s, $p_{0s}(t) = P(X(t) = s|X(0) = 1)$, against t

Semi-Markov models:

- Nonparametric estimates (`bcox`)...
- compared with estimates from Weibull parametric model (`bwei`).
- Standard errors require a second level of simulation (available in `flexsurv`)

Could integrate $p_{0s}(t)$ to give expected length of time spent in state s (`totlos.simfs`)



Continuously-observed multi-state models: further resources

Non/semi-parametric:

- Putter et al. (2007) Tutorial in biostatistics: competing risks and multistate models. *Stat Med.*
  - general tutorial: modelling principles, data formatting
- de Wreede et al. (2011) `mstate`: an R package for the analysis of competing risks and multi-state models *J Stat. Soft.*
  - introduction to package and worked examples
- de Wreede et al. (2010) The `mstate` package for estimation and prediction in non-and semi-parametric multi-state and competing risks models *Comp. Meth. Prog. Biomed.*
  - more theory behind the package
- Fiocco et al. (2008) Reduced rank proportional hazards regression and simulation based prediction for multistate models *Stat. Med*
- Carstensen and Plummer (2011) Using Lexis Objects for Multi-State Models in R. *J. Stat. Soft*
  - tools for summarising and manipulating data for modelling

Continuously-observed multi-state models: further resources

Fully parametric:

- Jackson (2016) `flexsurv`: a platform for parametric survival modelling in R. *J. Stat. Soft*
- `flexsurv` vignette, `http://CRAN.R-project.org`
  - same as the paper, but will be kept up to date
- Crowther and Lambert (2017) Parametric multistate survival models: Flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences *Stat. Med*
  - Stata perspective: with `multistate` Stata package.
- Iacobelli and Carstensen (2013) Multiple time scales in multistate models. *Stat. Med*
  - Fully parametric hazard models, implemented using Poisson regression
  - Comparing inhomogeneous (clock-forward) Markov with semi-Markov (clock-reset) models empirically

Thanks for listening!

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

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

www.mrc-bsu.cam.ac.uk