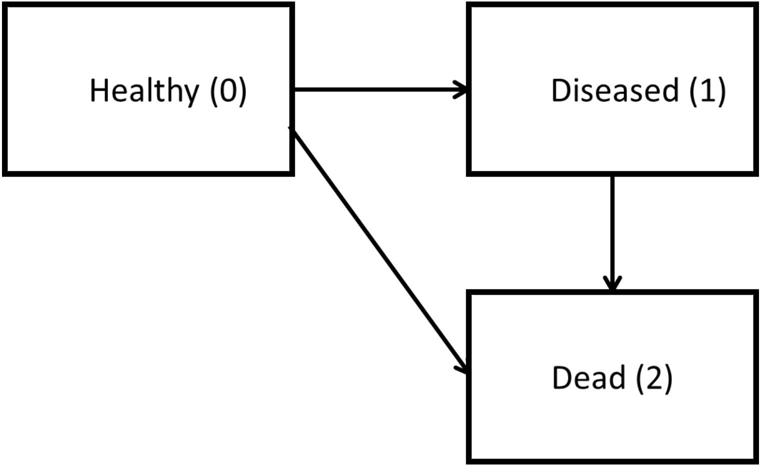
SURVIVAL ANALYSIS

Event-history models
PART VI, HY 2019

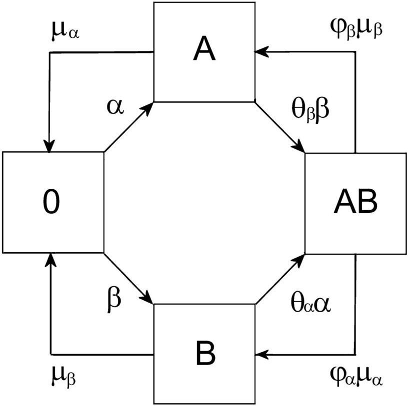
Outline

- Examples on multi-state models
- Model specification
- Likelihood function
- Health-illness model
- Reference: Andersen et al. Multi-state models for event history analysis, Stat Methods Med Res, 2012

Health-illness model



Infection model for two competing strains



Auranen K et al. Am. J. Epidemiol. 2009;171:169-176

American Journal of EPIDEMIOLOGY

Event-history models

- Event-history models or multi-state models are generalisations of simple survival or competing risks models
- There are a number of different states between which the individual can make transitions
- The model can contain an absorbing state(s) and/or it can accommodate repeated transitions between the model states
- The interest is in the estimation of transition-specific hazard functions and quantities that can be derived from those
- Numerous uses in epidemiology, medicine, biosciences, social sciences, engineering, ...

Model specification

- $X_i(t) \in \{1, \ldots, k\}$ is the state of individual i at time t, $t > 0$
- Time τ_i is the termination of observation for individual i
  - This can be the transition time to an absorbing state or censoring time
- Observations on $X_i(t)$ are equivalent to observations from counting processes for each possible transition in the model

$$
\begin{aligned}
N^i_{hj} &= \#(\text{direct transitions from } h \text{ to } j \text{ in } [0, t] \text{ for } i), \quad \text{with data} \\
&\qquad 0 < T^{i1}_{hj} < \cdots < T^{i, N^i_{hj}(\tau_i)}_{hj} \leq \tau_i
\end{aligned}
$$

- The hazard for individual i to make transition $h \to j$ at time t, conditionally on being in state h at $t-$, is denoted by $\alpha^i_{hj}(t)$, for all $h, j \in \{1, \ldots, k\}$
  - Some of the hazards can be 0 (i.e. for those transitions that do not exist in the model)

Likelihood function

- The likelihood is constructed from individual and transition specific contributions
- For each individual i and transition $h \to j$, the likelihood expression consists of contributions from all transition events $h \to j$ of the individual, evaluated at the times of those transitions, as well as from the 'at-risk' time spent in state h before those transitions:

$$\prod_{h \neq j} \prod_{k=1}^{N_{hj}^i(\tau_i)} \alpha_{hj}^i(T_{hj}^{ik}) \exp\left(-\int_0^{\tau_i} \alpha_{hj}^i(t) Y_h^i(t) dt\right)$$

- For simplicity of notation, it is assumed above that the terms $\alpha_{hj}^i(T_{hj}^{ik})$ appear only if such transitions occur
- Note how the the risk indicator takes care of the appropriate at-risk time in each of the states

Remarks

- In the above presentation, only one time variable is used; however, the hazard on making a specific transition may well depend on e.g. the *duration of stay* in that state, or on more than one time variable
- The likelihood function guides in writing the data matrix for statistical analysis
  - Each individual's each transition type $(h \to j)$ produces as many data matrix rows as there are arrows from h to any other states of the model
  - In fact, this is exactly the same situation than we have already seen in competing risks models: the states *to which* there is an arrow *from* state h are competing over which of these events actually occur

Example: health-illness model

- In the model of page 3, there are three possible transitions, with hazard functions $\alpha_{01}(t)$, $\alpha_{02}(t)$ and $\alpha_{12}(t)$
- The likelihood contribution from individual i is now

$$
\begin{aligned}
L_i(\text{parameters}) = \quad & \left(\alpha_{01}^i(T_{01})\right)^{d_{01}^i} \exp\left(-\int_0^{\tau_i} Y_{i0}(t)\alpha_{01}(t)dt\right) \times \\
& \left(\alpha_{02}^i(T_{02})\right)^{d_{02}^i} \exp\left(-\int_0^{\tau_i} Y_{i0}(t)\alpha_{02}(t)dt\right) \times \\
& \left(\alpha_{12}^i(T_{12})\right)^{d_{12}^i} \exp\left(-\int_0^{\tau_i} Y_{i1}(t)\alpha_{12}(t)dt\right)
\end{aligned}
$$

- The likelihood function based on data from n independent individuals is now

$$
L(\text{parameters}) = \prod_{i=1}^{n} L_i(\text{parameters})
$$

Example continues

- A toy example modified from Andersen et al.
- Health-illness model prameterised as follows:

$$\alpha^i_{01}(t) = \alpha_{01}(t)e^{\beta_1 Z_{1i} + \beta_2 Z_{2i}}$$

$$\alpha^i_{02}(t) = \alpha_{02}(t)e^{\beta_3 Z_{1i} + \beta_4 Z_{2i}}$$

$$\alpha^i_{12}(t) = \alpha_{02}(t)e^{\gamma + \beta_1 Z_{1i}}$$

- The death rates are proportional, with the hazard ratio $\exp(\gamma)$
- The covariates Z_{1i} and Z_{2i} have different proportional effects on the baseline rates $\alpha_{01}(t)$ and $\alpha_{02}(t)$
- Covariate Z_{1i} has the same effect on $\alpha_{01}(t)$ and $\alpha_{12}(t)$

Example continues

- For any one individual i, the data matrix contains three rows, where three is now the number of possible transitions in the model
- For fitting the Cox prorportional hazards model, the three data matrix rows from individual i are

Time	Status	Stratum	$\log(\gamma)$	$\log(\beta_1)$	$\log(\beta_2)$	$\log(\beta_3)$	$\log(\beta_4)$
T_{01}^i	d_{01}^i	0	0	Z_{1i}	Z_{2i}	0	0
T_{02}^i	d_{02}^i	1	0	0	0	Z_{1i}	Z_{2i}
T_{12}^i	d_{12}^i	1	1	Z_{1i}	0	0	0

`coxph(Surv(Time,Status)`$\sim$
$logg + logb1 + logb2 + logb3 + logb4 + strata(Stratum), data = A)$