Competing risks models

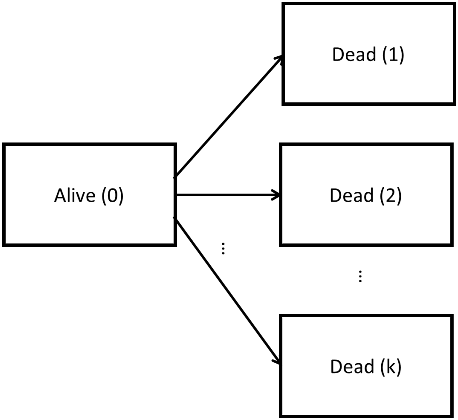
HY 2019

Outline

- Cause-specific hazards
- The probability of net survival
- Cause-specific subdensities and cumulative incidence functions
- Likelihood function and estimation in competing risks
- A couple of references
  - Andersen: Competing risks as a multi-state model (Stat Methods Med Res 2012)
  - Andersen and Keiding: Multistate models for event-history analysis (Stat Methods Med Res 2012)

Competing risks

- We consider models for a total of k competing risks, i.e. k possible causes of failure, only one of which can occur



Notations

- With k competing causes of failure, denote the failure time by T
- The cause of the failure is indicated by a *mark* $M \in \{1, \ldots, k\}$; only one cause can be marked
- For individual i, denote the data as (t_i, d_i, m_i) where t_i and d_i are defined as before as the observation time and the event indicator, and $m_i \in \{1, \ldots, k\}$ indicates the event that actually occurred
  - We return to data and the likelihood construction later; for the time being, let's concentrate on building the necessary concepts (and omit indexing individuals for convenience)
- N.B. The theory of competing risks is sometimes based on modelling k failure times $T_1, \ldots, T_k$, only one of which may occur in reality
  - However, that approach relies on assumptions on counterfactual times whereas building the theory on cause-specific hazards avoids such problems

Cause-specific hazards

- Following the example of a single-event hazard in standard survival analysis (see lecture notes, part I, page 13), define cause-specific hazards as follows:

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{\mathrm{P}(T \in [t, t + \Delta t[\text{ and } M = j | T \geq t)}{\Delta t}, \; j = 1, \ldots, k$$

- The jth cause-specific hazard is the intensity (rate) of failure of cause j occurring at time t, given that *none* of the k possible failure events has not yet occurred

Total hazard

- A natural first question is to ask what the total hazrd is, i.e. the hazard for *any* of the k events occurring, given that none of them has yet occurred; the time to this event is T
- Because only one of the failures can occur, cause-specific hazards are *additive* so that the total hazard is simply a sum of the cause-specific hazards (remember also that hazards are not probabilities but rates):

$$\begin{aligned}
\lambda(t) &= \lim_{\Delta t \to 0} \frac{\mathrm{P}(T \in [t, t+\Delta t[\mid T \geq t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\sum_{j=1}^{k} \mathrm{P}(T \in [t, t+\Delta t[\text{ and } M=j \mid T \geq t)}{\Delta t} \\
&= \sum_{j=1}^{k} \lambda_j(t)
\end{aligned}$$

Net survival

- We can now write the expression for the probability $\mathrm{P}(T > t)$ of the event of surviving all k causes up to time t
- Using the total hazard $\sum_{j=1}^{k} \lambda_j(t)$ we can apply the standard formula of the survival function for time T:

$$S(t) = \mathrm{P}(T > t) = \exp(-\int_0^t \lambda(u)du) = \exp\left(-\int_0^t \sum_{j=1}^{k} \lambda_j(u)du\right)$$

A note on "independence"

- The formula on the previous page is sometimes written in the following equivalent form:

$$S(t) = \prod_{j=1}^{k} \exp\left(-\int_0^t \lambda_j(u)du\right)$$

- The above appears to be constructed as the joint probability for observing independent events "$T_j > t$" with probabilities $\exp(-\int_0^t \lambda_j(u)du)$, $j = 1, \ldots, k$
- However, one would then rely on the independence of times T_j, only one of which can occur in actual fact
- Therefore, it is advised not to interpret the formula as suggesting independence of basically non-observable times
- Moreover, the correct expression for the probability of event j not having occurred by time t is the complement of its cumulative incidence (cf. page 10)

$$1 - \mathrm{P}(T \le t \text{ and } M = j) = 1 - \int_0^t \lambda_j(u)S(u)du.$$

Subdensity functions

- The density function of failing of cause h ($\in \{1, \ldots, k\}$) is

$$\begin{aligned} f_h(t) &= \lim_{\Delta t \to 0} \frac{\mathrm{P}(T \in [t, t+\Delta t[\text{ and } M = h)}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{\mathrm{P}(T \in [t, t+\Delta t[\text{ and } M = h | T > t)}{\Delta t} \mathrm{P}(T > t) \\ &= \lambda_h(t) S(t), \quad h = 1, \ldots, k \end{aligned}$$

- These are called subdensities because each of them integrates (from 0 to ∞) to the probability of failure h ever occurring; this probability is < 1 (see the next page)
- Note that it holds

$$\lambda_h(t) = \frac{f_h(t)}{S(t)}, \; h = 1, \ldots, k$$

Cumulative incidence functions

- The next question is about the *cumulative incidences*, i.e. the probabilities that failure j, $j = 1, \ldots, k$, occurs by time t
- This means avoiding failure by any of the k causes up to failure by cause j

$$\mathrm{P}(T \leq t \text{ and } M = j) = \int_0^t f_j(u)du = \int_0^t \lambda_j(u)S(u)du$$

- The *cumulative incidence* at time t is thus the integral of the corresponding subdensity (from 0 to t)
- When integrated from 0 to ∞ we obtain the life-time risk of failing of cause j: $\int_0^\infty \lambda_j(u)S(u)du$
- It easy to verify that the sum of life-time risks over all k causes is one:

$$\sum_{j=1}^{k} \int_0^\infty \lambda_j(u)S(u) = \int_0^\infty \sum_{j=1}^{k} \lambda_j(u)S(u) = \int_0^\infty \lambda(u)S(u)du = 1$$

A note on Kaplan-Meier plots

- It appears tempting to treat events by all other causes, except the one of interest, as censorings and plot estimates of "cause-specific survival functions" using the Kaplan-Meier method and stating that they are estimates of

$$S_j(t) = \exp(-\int \lambda_j(u)du)$$

- However, expressions like $S_j(t)$ do not have proper interpretation in a competing risks situation (cf. page 8)
- $S_j(t)$ would be interpretable as a probability only if the other causes of failure would be eliminated
- This is different from censoring which only removes the study subject from the risk set (from being followed)
- The same warning applies to parametric estimates of "cause-specific survival functions", based on parameter estimates of cause-specific hazards

Likelihood function

- Denote the data as (t_i, d_i, m_i), where t_i and d_i are as before, and $m_i \in \{1, \ldots, k\}$ gives the cause of failure when $d_i = 1$
- A censored observation at time t_i contributes the probability of net survival up to time t_i:

$$L_i(\theta) = \mathrm{P}(T > t_i) = S(t_i) = \exp(-\int_0^{t_i} \sum_{j=1}^{k} \lambda_j(u)du) = \exp\left(-\int_0^{t_i} \lambda(u)du\right)$$

- Failure of cause j at time t_i contributes the subdensity at t_i:

$$L_i(\theta) = \lambda_j(t_i)S(t_i) = \lambda_j(t_i)\exp(-\int_0^{t_i} \lambda(u)du)$$

- So we obtain the likelihood function:

$$L(\theta) = \prod_{i=1}^{n} L_i(\theta) = \prod_{i=1}^{n} \left([\lambda_{m_i}(t_i)]^{d_i} \times e^{-\int_0^{t_i} \lambda(u)du}\right)$$

Likelihood function (2)

- The parameter vector θ includes the k cause-specific hazards and effects of the covariates
- Making the dependence of the parameters explicit, the likelihood function is

$$L(\theta) = \prod_{i=1}^{n} \left([\lambda_{m_i}(t_i; Z_i, \theta]^{d_i} \times e^{-\int_0^{t_i} \lambda(u; Z_i, \theta) du} \right)$$

Estimation in practice

- Irrespective of the warnings on not *interpreting* failure by competing causes as censoring, they can be technically accomodated as censoring
- This means that when estimating the cause-specific hazard $h \in \{1, \ldots, k\}$, failure by any of the other causes are taken into account as censorings
- In particular, consider the part of the likelihood expression that depends on cause-specific hazard λ_h:

$$L(\lambda_h) = \prod_{i=1}^{n} [\lambda_h(t_i; Z_i, \theta]^{\mathbf{1}(d_i=1 \text{ and } m_i=h)} \times e^{-\int_0^{t_i} \lambda_h(u; Z_i, \theta) du}$$

- This means that, by appropriate coding of censoring, standard methods can be applied in the estimation of cause-specific hazards and cumulative cause-specific hazards

Example

- For convenience, assume only two competing risks, with constant cause-specific baseline hazards, λ_1 and λ_2
- Assume there is a binary covariate Z_i and a proportional hazard $\theta = \exp(\beta Z_i)$
- The model is thus

$$\lambda_1(t; Z_i, \beta) = \lambda_1 \exp(\beta Z_i),$$

$$\lambda_2(t; Z_i, \beta) = \lambda_2 \exp(\beta Z_i)$$

- Introduce new notation for cause-specific event indicators $d_{0j}^{(i)} = \mathbf{1}(d_i = 1 \text{ and } m_i = j)$
- The likelihood function of the parameters λ_1,λ_2 and β is

$$\begin{aligned} L(\lambda_1, \lambda_2, \beta) &= \prod_{i=1}^{n} \left\{ [\lambda_1 e^{\beta Z_i}]^{d_{01}^{(i)}} [\lambda_2 e^{\beta Z_i}]^{d_{02}^{(i)}} \exp\left(-[\lambda_1 e^{\beta Z_i} + \lambda_2 e^{\beta Z_i}] t_i\right) \right\} \\ &= \prod_{i=1}^{n} \left\{ [\lambda_1 e^{\beta Z_i}]^{d_{01}^{(i)}} \exp\left(-[\lambda_1 e^{\beta Z_i}\right) \right\} \\ &\quad \times \prod_{i=1}^{n} \left\{ [\lambda_2 e^{\beta Z_i}]^{d_{02}^{(i)}} \exp\left(-[\lambda_2 e^{\beta Z_i}\right) \right\} \end{aligned}$$

Example cont.

- If there were no common covariates (in the above example, there is one, i.e. β), the likelihood function would completely separate to two parts, one for λ_1 and another for λ_2
- In presence of common covariates (i.e. β in the above example), the estimation of all parameters must be carried out simultaneously
- The data matrix for the above example is given on the next page, with columns for `time`, `status` (i.e. cause-specific event indicator), and the model matrix inputs for the explanatory variables $log(\lambda_1)$, $\log(\lambda_2)$ and $\log(\beta)$
- In particular, note that with k competing risks, there will be k rows in the data matrix per each observation time t_i

Example cont.

Time	Status	$\log(\lambda_1)$	$\log(\lambda_2)$	$\log(\beta)$
t_1	$d_{01}^{(1)}$	1	0	Z_1
t_1	$d_{02}^{(1)}$	0	1	Z_1
t_2	$d_{01}^{(2)}$	1	0	Z_2
t_2	$d_{02}^{(2)}$	0	1	Z_2
t_3	$d_{01}^{(3)}$	1	0	Z_3
t_3	$d_{02}^{(3)}$	0	1	Z_3
...	...	...	...	...