Survival Models

Lecture VII. Estimation of transition probabilities and intensities in a multistate model. Cox-Markov model

Aalen-Johanson estimators

Let us recall the Kaplan-Meier ideas and apply them to current setup.

It is easy to see that Kaplan-Meier estimator $\hat{S}(t)$ is in MSM context an estimate for $P_{11}(0,t)$:

$$\widehat{P_{11}}(0,t) = \hat{S}(t) = \prod_{i:y_{(i)}\leq t} \frac{n_i - d_i}{n_i}$$

Now, the Aalen-Johanson estimator for survival probability in interval (s,t) is given by

$$\widehat{P_{11}}(s,t) = \prod_{i:s<y_{(i)}\leq t} \frac{n_i - d_i}{n_i},$$

A matrix $(\widehat{P_{ij}(s,t)})$ is also called empirical transition matrix

Illness-death model (1)

Let us recall the illness-death model (with terminal illness). We have 3 states:

- 1 – alive (healthy)
- 2 – (terminally) ill
- 3 – dead

Then

$$\widehat{P_{11}}(s,t) = \prod_{i: s < y_{(i)} \le t} \frac{n_{1i} - d_{12i} - d_{13i}}{n_{1i}},$$

where

- $y_{(i)}$ is the i-th observed death/failure time
- n_{1i} is the number of subjects at risk right before $y_{(i)}$
- d_{1ki} is the number of subjects that moved from state 1 to state k ($k \in \{1, 2\}$) at $y_{(i)}$

Illness-death model (2)

Similarly, for $P_{22}(s,t)$

$$\widehat{P_{22}}(s,t) = \prod_{i:s<y_{(i)}\le t} \frac{n_{2i} - d_{23i}}{n_{2i}}$$

We can also construct an estimate for transition hazard $h_{12}(u)$:

$$\widehat{h_{12}}(u) = \frac{d_{12j}}{n_{1j}},$$

where $j = \max\{k : y_{(k)} \le u\}$

Illness-death model (2)

To describe the model fully, we also need an estimate for $P_{12}(s,t)$.
Let us substitute the obtained estimates for $\widehat{P_{11}}(s,t)$, $\widehat{h_{12}}(u)$ and $\widehat{P_{22}}(u,t)$ into

$$P_{12}(s,t) = \int_s^t P_{11}(s,u) h_{12}(u) P_{22}(u,t) du,$$

then an estimate for $P_{12}(s,t)$ can be found from

$$\widehat{P_{12}}(s,t) = \sum_{i: s < y_{(i)} \leq t} \widehat{P_{11}}(s, y_{(i)}) \widehat{h_{12}}(y_{(i)}) \widehat{P_{22}}(y_{(i)}, t)$$

Competing risks model (1)

Recall the competing risks model with two risks and let the states be defined as:

- 1 – alive (healthy)
- 2 – dead (cause of risk 1)
- 3 – dead (cause of risk 2)

Similarly to previous, we have simple non-parametric estimates:

$$\widehat{P_{11}}(s,u) = \prod_{i:s<y_{(i)}\leq u} \frac{n_{1i} - d_{12i} - d_{13i}}{n_{1i}}$$

and

$$\widehat{h_{12}}(u) = \frac{d_{12j}}{n_{1j}},$$

where $j = \max\{k : y_{(k)} \leq u\}$

Competing risks model (2)

Let us recall that we also established

$$P_{12}(s,t) = \int_s^t P_{11}(s,u) h_{12}(u) du$$

Using obtained estimates for $\widehat{P_{11}}(s,u)$ and $\widehat{h_{12}}(u)$, we can write

$$\widehat{P_{12}}(s,t) = \sum_{i: s < y_{(i)} \leq t} \widehat{P_{11}}(s, y_{(i)}) \widehat{h_{12}}(y_{(i)})$$

The effect of covariates. Cox-Markov model

The Cox PH model for MSM setup can be written as

$$h_{ij}(t|\underline{x}) = h_{ij0}(t)e^{\underline{x}'\underline{\beta}_{ij}},$$

where

- $h_{ij}(t|\underline{x})$ is transition intensity (hazard) from i to j given $\underline{x}$
- h_{ij0} is the "baseline intensity" from i to j
- $\underline{x}$ is the vector of covariates
- $\underline{\beta}_{ij}$ is the vector of coefficients corresponding to transition from i to j

Stratified Cox model as MSM (1)

We know that if the proportional hazards assumption is not fulfilled for a covariate, one can allow the baseline hazard to vary depending on the value of given covariate.

For example, consider a covariate z with two possible values 0 and 1. Then we can apply the following stratified model:

$$h(t|\underline{x}, z = 0) = h_{0_1}(t)e^{\underline{x}'\underline{\beta}}$$

and

$$h(t|\underline{x}, z = 1) = h_{0_2}(t)e^{\underline{x}'\underline{\beta}},$$

where

- $h_{0_1}(t)$ is the baseline hazard given $z = 0$
- $h_{0_2}(t)$ is the baseline hazard given $z = 1$

Stratified Cox model as MSM (2)

Now we can describe the situation through a MSM as follows:

- state 1 – alive, $z = 0$
- state 2 – alive, $z = 1$
- state 3 – dead

and the corresponding transition intensities are

$$h_{13}(t|\underline{x}) = h(t|\underline{x}, z = 0) = h_{0_1}(t)e^{\underline{x}'\underline{\beta}}$$

and

$$h_{23}(t|\underline{x}) = h(t|\underline{x}, z = 1) = h_{0_2}(t)e^{\underline{x}'\underline{\beta}},$$

Cox model with time-changing covariates as MSM

One can also allow the covariates to change in time.

For example, assuming there is only one time-dependent covariate z, we can write

$$h(t|\underline{x}, z) = h_0(t)e^{\underline{x}'\underline{\beta}+z(t)\gamma}$$

Notice that if we also assume z to have again only two possible values 0 and 1, this situation can be again described as a MSM (similarly to previous):

- state 1 – alive, $z = 0$
- state 2 – alive, $z = 1$
- state 3 – dead

Then

$$h_{13}(t) = h_0(t)e^{\underline{x}'\underline{\beta}}$$

and

$$h_{23}(t) = h_0(t)e^{\underline{x}'\underline{\beta}+\gamma}$$

Additional remarks

Let us consider a competing risks model with k risks

- It is tempting to think that the survival time in such case is minimum of rv-s T_i $(i = 2, \dots, k+1)$, where T_i is the time to death caused by risk i. Unfortunately, since T_i-s are dependent, such approach can not be used
- Moreover, in general we are not able to estimate the marginal distributions T_i – there is no way to obtain data where some risks are eliminated
- Thus, the transition intensities corresponding to T_i are not net hazards (i.e. dependling only on T_i), but cause specific hazards (they depend on the joint behaviour of T_i-s)