probabilities in multi-state survival data

Luís Machado, DMCT
University of Minho,
Portugal,
lmachado@mct.uminho.pt

aim

Present different approaches for the estimation of transition probabilities in multi-state survival data. The central question here is: what is the probability that a randomly selected person is in stage *j* at time *t*, conditionally on being in state *h* at time *s*:

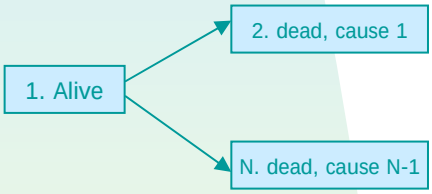
$$p_{hj}(s,t) = P\left(X(t) = j \mid X(s) = h\right)$$

Multi-state models (MSM)

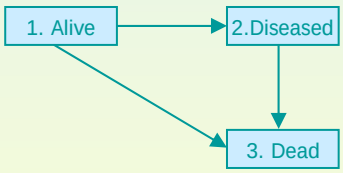
A MSM is a model for a stochastic process in continuous time allowing individuals to move between a finite number of states. An individual moves from one state to another through time. The next state to which the individual moves, and the time of change, are specified through transition intensities that provide the instantaneous hazard for movement out of one state into another.

Some MSM

The mortality model



The Competing risks model



The Illness-death model

Nonparametric approaches

Illness-death model

Aalen–Johansen estimators of transition probabilities: This approach can be thought of as the generalization of the Kaplan-Meier estimate of the simple mortality model (assumes the process to be Markovian):

$$\widehat{p_{11}}(s,t) = \prod_{s<t_{(i)}\le t}\left(1-\frac{d_{12i}+d_{13i}}{n_{1i}}\right) \qquad \widehat{p_{22}}(s,t) = \prod_{s<t_{(i)}\le t}\left(1-\frac{d_{23i}}{n_{2i}}\right)$$

$$\widehat{p_{12}}(s,t) = \sum_{s<t_{(i)}\le t}\widehat{p_{11}}\left(s,t_{(i-1)}\right)\frac{d_{12i}}{n_{1i}}\widehat{p_{22}}\left(t_{(i)},t\right)$$

For a sample of n subjects, $t_{(1)} < t_{(2)} < \ldots < t_{(k)}$ denote the k event times; d_{hji} the nº of transitions from state h to state j at time $t_{(i)}$; n_{hi} the number of individuals at risk at time $t_{(i)}$.

Markov-free estimator: alternative estimators which do not rely on the Markov assumption are expressed as:

$$\widehat{p_{11}}(s,t) = \frac{1-\widehat{H}(t)}{1-\widehat{H}(s)}$$

$$\widehat{p_{12}}(s,t) = \frac{1}{1-\widehat{H}(s)}\sum_{i=1}^{n} W_i \phi_{s,t}\left(Z_{[i]},Y_{(i)}\right)$$

$$\widehat{p_{22}}(s,t) = \sum_{i=1}^{n} W_i \tilde{\phi}_{s,t}\left(Z_{[i]},Y_{(i)}\right) \Big/ \sum_{i=1}^{n} W_i \tilde{\phi}_{s,s}\left(Z_{[i]},Y_{(i)}\right)$$

$Y = \min(T,C)$; T – survival time; C - right-censoring time
Z is the sojourn time in state 1, with distrib. function H, and $\widehat{H}$ its Kaplan-Meier estimator.

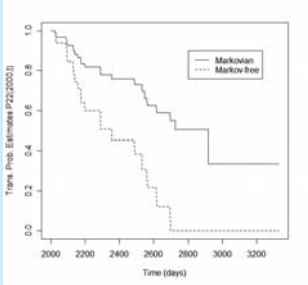
$\phi_{s,t}(u,v) = \mathbb{I}(s<u\le t, v>t)$ and $\tilde{\phi}_{s,t}(u,v) = \mathbb{I}(u\le s, v>t)$

W_i are the Kaplan-Meier weights attached to $Y_{(i)}$;

$Y_{(1)} \le \ldots \le Y_{(n)}$ denote the ordered sample of the Y_i's, and $Z_{[i]}$ for the pair attached (concomitant) to the $Y_{(i)}$ value.

Colon cancer data

For the Colon cancer data (Moertel et al. 2000) we may consider the recurrence as an associated state of risk, and use the illness-death model with states "alive and disease-free", "alive with recurrence" and "dead".



Conclusions

➔The "Markov-free" estimator, have fewer jump points but with bigger steps. The number of jump points and the size of the steps are related to censoring and to the sample size.

➔ With regard to the survival prognosis, we observe serious departures between both survival curves for individuals who have had recurrence.

References

Datta S, Satten GA. Validity of the Aalen–Johansen estimators of stage occupation probabilities and Nelson–Aalen estimators of integrated transition hazards for non-Markov models. Statistics & Probability Letters 55 (2001) 403 – 411

De Uña-Álvarez, Meira-Machado L., Cadarso-Suárez C. Nonparametric estimation of the transition probabilities in a illness-death model. To appear in Lifetime Data Analysis.

Moertel CG et al. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. New Eng. J. Med. (2000) 352-358