Estimation of the Transition Probabilities condition on repeated measures in Multi-state models

Gustavo Soutinho[1], Luís Meira-Machado[2], Pedro Oliveira[1]

[1] EPIUnit, ICBADS, University of Porto, Portugal.
[2] Department of Mathematics and Centre of Molecular and Environmental Biology (CBMA), University of Minho, Portugal.

E-mail for correspondence: lmachado@math.uminho.pt

Abstract: The topic of joint modeling of longitudinal and survival data has received remarkable attention in recent years. In cancer studies for example, these models can be used to assess the impact that a longitudinal marker has on the time to death or relapse. Analyzes of such studies, in which individuals may experience several events, can be successfully performed by multi-state models. The goal of this work is to introduce feasible estimation methods for the transition probabilities conditionally on covariates observed with repeated measures through the use of the landmark methodology and the adaptation of existing methods for joint modeling of longitudinal and survival data. Results of the simulation studies confirm the superiority of the proposed estimator when compared to methods that do not take in consideration the effect of the covariate on the estimated transition probabilities or do not assume all the existence of repeated measures (Breslow estimator).

Keywords: Joint modeling, Markov assumption, Multi-state models, Transition probabilities.

1 Introduction

Multi-state model is a model for a time continuous stochastic process which can be used to describe complex event history data with several events (Meira-Machado and Sestelo, 2019). In medical science studies beyond the times-to-event a main goal is to identify the impact of a set of repeated measures as a time-dependent covariate on the transition among states. In order to produce valid inferences in these cases a joint modeling analysis of longitudinal and multiple survival outcomes are required (Rizopoulos, 2012). The final model is built using two sub-models; a longitudinal sub-model (such as a linear mixed effects model) and a time-to-event sub-model (such as a proportional hazards model) for each transition intensity which are linked through an association structure quantifying the relationship between the outcomes of interest. The background concepts related to the extension of the joint modeling to multi-state models can be found in Ferrer et al (2016). The aim of this paper is propose a feasible estimation method for the transition probabilities conditionally on covariates observed with repeated measures. To this end we will use the subsampling approach, also termed as landmarking, proposed by de Uña-Álvarez and Meira-Machado (2015), combined with methods proposed by Rizopoulos (2012). The landmark methods considers subsamples of individuals of the data that belong in a given state at a pre-specified time point and gives rise to consistent estimators regardless the Markov assumption.

1.1 Joint multi-state model specification

The joint modeling approach for multi-state models can be described by a linear mixed effect model and a survival sub-model for each transition. The longitudinal sub-model follows the gaussian assumptions and the observed measure Y_{ij} at time t_{ij} is given by $Y_{ij} = X_i\left(t_{ij}\right)^T \beta + Z_i\left(t_{ij}\right)^T b_i + \varepsilon_{ij}$, where $X_i\left(t_{ij}\right)$ and $Z_i\left(t_{ij}\right)$ represent the vectors of time-dependent covariates of the individual and b_i is the vector of random effects with $b_i \sim N(0, \Sigma)$. The β parameter is a fixed vector and $\varepsilon_i \sim N\left(0, \sigma^2 I_{n_i}\right)$ where n_i is the number of longitudinal measures by individual (Ferrer (2016)).
The time-to-event outcome at time t from state h to state k, with $h, k \in S$ the finite state space, is modeled by a proportional hazards sub-model which takes the following form $\lambda^i_{hk}(t|b_i)=\lambda_{hk,0}(t)\exp\left\{X^{ST}_{hk,i}\gamma_{hk}+W_{hk,i}(b_i,t)\eta_{hk}\right\}$, where $\lambda_{hk,0}\left(.\right)$ is a parametric baseline intensity (with weibull, exponential or piecewise constant distributions, for instance). The baseline covariates are denoted by w_i with coefficients γ_{hk}. The multivariate function $W_{hk,i}\left(b_i, t\right)$ defines the dependence structure between the longitudinal and multi-state process and represents the true and unobserved value of the longitudinal outcome for patient i at time t. The association between the longitudinal and the times-to-event for each transition is given by η_{hk}.

1.2 Estimation and Dynamic predictions of the transition probabilities

In this study the maximum likelihood estimation for joint models will be used to estimate the parameters of the joint multi-state model under the landmark approach described in Uña-Álvarez and Meira-Machado (2015). The maximization of the log-likelihood function will be done using an EM algorithm coupled with a quasi-Newton algorithm in case of slow convergence. As referred above the aim of this paper is to estimate the transition probabilities $p_{hj}(s,t) = P(X(t) = j|X(s) = h)$ conditional to a vector of covariates that include a covariate with longitudinal measures (as tumor markers measured at different moments) $\widetilde{y}_i(v) = \{y_i(v), 0 \leq u \leq v\}$. For each individual a transition probability is estimated and is assumed that the patient has survived up to the last time point s (Rizopoulos, 2012)

2 Simulation study

The longitudinal and multi-state data were generated through a joint modeling with 1000 replicates with 400 individuals given by $Y_{ij} = \beta_0 + \beta_1 \times t_{ij} + b_{i0} + b_{i1} \times t_{ij} + \varepsilon_{ij}$ and $\lambda^i_{hk}(t|b_i) = \lambda_{hk,0}(t)\exp\{\gamma_{hk} + W_{hk,i}(b_i,t)\,\eta_{hk}\}$, where $h \in \{0,1\}$, $k \in \{0,1,2\}$ and $b_i \sim N\left((0,0)^T, \begin{pmatrix} 20 & 0.2 \\ 0.2 & 0.02 \end{pmatrix}\right)$. The longitudinal times, initially were the same for each individual, given by $t_{ij} = 0.33, 0.66, \cdots, 16.50$ and the $\varepsilon_i \sim N(0,18)$. The parametric baseline intensities were obtained from exponential distributions with rate parameters 3, 1.5 and 0.5. We took the value 2 for the γ_{hk} and for η_{hk} we took the values -0.7, -0.7 and -0.6 for the transitions $0 \to 1$, $0 \to 2$ and $1 \to 2$, respectively. The vector of true transition times, $T^*_i = \left(T^*_{i,01}, T^*_{i,02}, T^*_{i,12}\right)$, were generated following the procedures described in Beyersmann et al. (2011). By comparing T^*_i and C_i, the vector of times $T_i = \min(T^*_i, C_i)$, where C_i denotes the censoring times, which characterizes the multi-state process, was deduced. The longitudinal measurements, generated from the linear mixed sub-model, were truncated at T_{i1} the first observed time of the multi-state process.

2.1 Results

The transition probabilities for the Landmark approach (`LM`), Breslow's method (`BRES`) and Joint Modeling-Landmark estimator (`JMLM`) were obtained through Monte Carlo simulation with 1000 replicas with 400 individuals. For each replica, eight individuals were retained with the purpose to identify the influence of the longitudinal marker on the estimation of the transition probabilities (decreasing, constant, increasing and random values of the marker). Although the variability seems to be quite similar, the results confirm the superiority of the `JMLM` estimator which produced unbiased estimates for all transitions probabilities. This conclusion are in accordance with the ratio between the mean square errors (MSEs) for the transition probabilities $\hat{p}_{00}(4,t)$ (Figure 1) and $\hat{p}_{11}(4,t)$ with $t = \{6, 8, 10\}$.

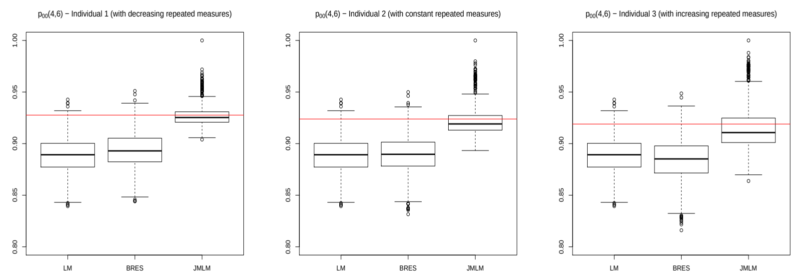

Figure 1: Boxplots of the M = 1000 estimates of the transition probabilities.

It was also possible to observe the ability of the `JMLM` to reflect the evolution of the longitudinal measures of the marker. In fact, for instance, considering $\hat{p}_{00}\,(4,6)$, for individual 3 with an increasing trend of the repeated measures of the marker as the Breslow estimator takes into account the higher value the transition probabilities decrease comparing to the `LM` estimator. However the effect of the previous repeated measures have as consequence the increase of the `JMLM` estimation, following the true values.

3 Conclusions

Results obtained from simulation studies and in the real data application confirmed the good performance of the `JMLM` estimator, providing accurate estimated transition probabilities. The proposed method also demonstrated to have more sensibility to reflect the evolution of the longitudinal measures when comparing to the Breslow's based method which only makes use of a single value of the covariate.

References

Beyersmann, J., Allignol, A. and Schumacher, M (2011). *Competing risks and multistate models with R*. Springer.

de Uña-Álvarez, J., Meira-Machado, L. (2015). Nonparametric estimation of transition probabilities in the non-Markov illness-death model: A comparative study. *Biometrics*, **71**, 141 − 150.

Ferrer, L., Rondeau, V., Dignam, J.J., Pickles, T., Jacqmin-Gadda, H, and Proust-Lima, C. (2016). Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer. *Statistics in Medicine*, **35(22)**: 3933 − 3948.

Meira-Machado, L. and Sestelo, M. (2019). Estimation in the progressive illness-death model: A nonexhaustive review. *Biometrical Journal*, **61(2)**, 245 − 263.

Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R*. CRC Press.