

Inference under a second order Markov model

Guadalupe Gómez Melis¹, Mireia Besalú Mayol²

¹lupe.gomez@upc.edu, Dept. d'Estadística i Investigació Operativa, Universitat Politècnica de Catalunya-BarcelonaTech (UPC)

²mbesalu@ub.edu, Dept. Genètica, Microbiologia i Estadística, Universitat de Barcelona

Multistate modelization (MSM) is a comprehensive approach to study the evolution of individuals along time and through different states. As complementary to standard survival methods where the analysis is centered in the time to one particular event, in MSM a dynamical path across several states is considered. Therefore, disease burden is captured more fully in comparison to conventional survival analysis. MSM models are well developed for continuous and discrete times, under a first order Markov assumption. If the first order Markov condition does not hold, some solutions have been proposed for specific MSM such as the illness-death model with only 3 states.

Motivated by a cohort of more than 3000 COVID-19 patients hospitalized during the first wave of the pandemic (March-April 2020) a complex multistate model for their dynamic evaluation after hospitalization was designed thanks to the fruitful collaboration between biostatisticians and clinicians. Specifically, this MSM is based on 14 possible transitions among the seven following states of a patient: Non Severe Pneumonia (NSP), Severe Pneumonia (SP), Non-Invasive Mechanical Ventilation (NIMV), Invasive Mechanical Ventilation (IMV), Recovery (R), Hospital Discharge (HD) and Death (D). Since a preliminary analysis showed that the first order Markov condition was not met for some transitions, we propose a more general second order Markov model which would lessen some of these restrictions while providing a more realistic description of the reality.

We have developed a second order Markov model where the future evolution not only depends on the current but also on the preceding state. Under a discrete time analysis (days), assuming that past information is restricted to 2 consecutive times and under homogeneity, we have expanded the transition probability matrix to M different matrices of order M (M is the number of states) and have proposed an extension of the Chapman-Kolmogorov equations.

Estimation and inference for the transition probabilities is presented for complete uncensored data. This is used, among others, to compute the transition probability from SP to IMV, after a given number of hospitalized days, and differentiating between patients that arrive to the hospital with SP from those who develop the disease at the hospital. A counting process framework for the multistate model is discussed as a first attempt to define estimators for the transition probabilities under right-censoring.

Keywords: Multistate Models, Non-Markov, COVID-19