The Journal of Open Source Software

Estimating statistics from multi-state models using simulation with multistateutils

Stuart Lacy[1]

1 Department of Health Sciences, University of York, UK

DOI: 10.21105/joss.00852

Software

- Review
- Repository
- Archive

Submitted: 25 July 2018
Published: 26 July 2018

Summary

`multistateutils` provides functionality for analysing multi-state models using Discrete Event Simulation (DES) in R.

DES is a method of exploring the behaviour of dynamic systems by discretising trajectories into a series of state transitions that occur at specific times. There exists general-purpose DES software packages in several programming languages, such as `simmer` (R) (Ucar, Smeets, and Azcorra 2018) or `simPy` (Python) (Matloff 2008) that can be applied to many systems regardless of the application area. Multi-state modelling is an area of biostatistics that is concerned with the extension of survival analysis to the case where there are multiple time-to-event transitions of interest, such as a patient experiencing various treatments on their disease pathway(Meira-Machado et al. 2009). There are several measures of interest that can be generated from a multi-state model, such as the probability of being in a given state at a certain time, or the estimated time spent in a given state, which can be obtained analytically provided certain assumptions - most notably the Markov property - are met; however, simulation provides a more flexible alternative and can be used to extract any desired summary statistic.

`multistateutils` is designed for the specific use case of multi-state modelling in biostatistics and provides a means for estimating statistics of interest from a multi-state model by simulation without exposing users to DES specific terminology. The functions in this package operate on parametric statistical models of transition times from the widely-used `flexsurv` (Jackson 2016) package, allowing the user to fit transition models in a familiar environment before inputting them directly into the simulation, rather than having to learn the parameterisation of a general purpose DES toolbox. The primary functionality provided by `multistateutils` is an interface to running discrete event simulation over a cohort of patients, thereby providing the simulated outcomes from which users can estimate any required statistics. In addition, for two common statistics - transition probabilities and length of stay estimates - there are functions that will set up and run the simulation and return the estimates directly.

The simulation engine is written in C++ to provide an efficient implementation of a DES. The source code has been archived in Zenodo with the DOI (Lacy 2018).

References

Jackson, Christopher. 2016. "flexsurv: A Platform for Parametric Survival Modeling in R." *Journal of Statistical Software* 70 (8):1–33. https://doi.org/10.18637/jss.v070.i08.

Lacy, Stuart. 2018. "Stulacy/Multistateutils: 1.2.0." https://doi.org/10.5281/zenodo.1320356.

Matloff, Norm. 2008. "Introduction to Discrete-Event Simulation and the Simpy Language."

Meira-Machado, Luís, Jacobo de Uña-Álvarez, Carmen Cadarso-Suárez, and Per K Andersen. 2009. "Multi-State Models for the Analysis of Time-to-Event Data." *Statistical Methods in Medical Research* 18 (2). SAGE Publications Sage UK: London, England:195–222.

Ucar, Iñaki, Bart Smeets, and Arturo Azcorra. 2018. "simmer: Discrete-Event Simulation for R." *Journal of Statistical Software* accepted for publication.