Analytic state episode decomposition (proof of concept)

Tim Riffe*[1]

[1]Max Planck Institute for Demographic Research

June 19, 2018

[This note will be circulated among a few people, and the idea advertised at a lab talk in September, but I do not expect to have time to develop the idea further until REVES abstracts are due for the 2019 meeting in Barcelona.]

Abstract

On the basis of a discrete time Markov matrix population model, we describe how to decompose expected state occupancies into the expected contributions of state episodes by duration. Results can be summarized in terms of short (acute) and long (chronic) episodes, into first, intermediate, and terminal contributions. Additional decomposition steps are able to isolate the effects of

Transition rates or probabilities are usually used to calculate state expectancies, such as healthy or unhealthy life expectancy, working life expectancy, and so forth. The discrete life trajectories that are implied by state transition probabilities are not usually invoked, but they are straightforward to calculate or imagine. The reason why researchers do not calculate the full trajectory space implied by transition probabilities is that this object is often too large for present computational facilities to efficiently handle. For example, a population model with s states (not including dead) and a age classes (closed out after the final age class) implies a total of $s^1 + s^2 + \ldots + s^a$ possible discrete life trajectories. These trajectories are at our fingertips, but somehow out of reach. So much useful and interesting information lies in this trajectory space, however: how does the character of state episodes change over age? Are terminal spells (those ending in death) shorter or longer than other spells? Is a change in expectancy due to having more acute or more chronic disability spells? Such questions can be answered with transition probabilities, and not necessarily Markovian ones. Examples here will however be Markovian.

In a previous treatment I toyed with such trajectories, skirting the large numbers problem by sampling from the trajectory space (Riffe 2018b). With trajcetory sets of manageable size, one is free to play with different notions of lifecourse alignment, and time counting, which enables the generation of a diverse set of macro patterns from the same set of transition probabilities. Analytical approaches to the episode statistics of matrix population models have been worked out to calculate the mean number and duration of episodes (Dudel and Myrskylä 2018). Here I'd like to offer a proof of concept for an analytical approach to work out the full episode duration distribution for states. Some prior knowledge is assumed.

This method is cheap and available with some data objects that we already are in the habit of calculating: the fundamental matrix $\mathbf{N}$ and transition probabilities, $\mathbf{p}_x^{i \to j}$, where x is the lower bound of the discrete age class, i is the origin state, and j is the destination state, where death is also a potential destination, and where the case of $i = j$ denotes remaining in the state. Recall that the elements of $\mathbf{N}$ are the conditional age-state occupancies (conditional on starting at the youngest age and in a given state). Say we have $s = 3$ (say, healthy, mildly, and severly disabled), where 4 is dead, and age classes are single years. Say we want to know the episode duration distribution for episodes of state 2 that start at age 50. Then we'd need to extract from $\mathbf{N}$ the probability of not being in state 2 at age 49 π_{49}^{-2}

$$\pi_{49}^{-2} = \pi_{49}^{1} + \pi_{49}^{3} \qquad . \tag{1}$$

The probability of an episode of state 2 starting at age 50 $\epsilon_{50}^{2,new}$ is:

$$\epsilon_{50}^{2,new} = \frac{\pi_{49}^{1} \cdot p_{49}^{1 \to 2} + \pi_{49}^{3} \cdot p_{49}^{3 \to 2}}{\pi_{49}^{-2}} \tag{2}$$

*riffe@demogr.mpg.de

Episodes starting at age 50 can assume a range of durations, and the distribution of durations can be derived by taking cumulative products of the probability of simply staying in state 2. Let's use demographers' interval notation to denote ${}_{d}e^{2}_{50}$ as the episodes of state 2 starting at age 50 and of duration d. Then

$$ {}_{d}\epsilon^{2}_{50} = \epsilon^{2,new}_{50} \cdot \prod_{n=0}^{d} p^{2\rightarrow 2}_{50+n} \tag{3} $$

And indeed ${}_{d}\epsilon^{2}_{50}$ is a probability, and $e(a)^{s} = \sum_{n=0}^{\omega} \sum_{d=0}^{\omega-a} {}_{d}\epsilon^{s}_{a+d}$. An empirial example may need to wait until I need to prepare the September lab talk. It may involve downstream decompositions, á la Riffe (2018a) and Riffe et al. (2018).

References

C. Dudel and M. Myrskylä. Estimating the expected number and length of episodes using markov chains with rewards. (under review), 2018.

Tim Riffe. Flow decompositions in multistate markov models. manuscript in preparation, 2018a. URL `https://github.com/timriffe/ArrowDecomp`.

Tim Riffe. Alignment, clocking, and macro patterns of episodes in the life course. manuscript in preparation, 2018b. URL `https://github.com/timriffe/Spells`.

Tim Riffe, Neil Mehta, Daniel Schneider, and Mikko Myrskylä. Healthy lives: Delayed onset, improved recovery, or mortality change? manuscript in preparation, 2018. URL `https://github.com/timriffe/hledecomp`.