Introduction to sequence analysis for social sciences

Nicola Barban

GENPOP has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 865356).

Assignement

Use the data provided to illustrate life course trajectories of a sample of individuals from the German Family Panel (pairfam) data. Data come from Raab and Struffolino (2022). Additional details and code can be found in the book and in the companion github page

Prepare a 5 page max report in which you:

1. Import the data in R and describe the distribution of life states using relevant sequence analysis tools, including tables, indexes and figures.
2. Calculate life course dissimilarities in the data.
3. Use a clustering methods to derive typical trajectories and describe groups

4. (Optional) Describe group differences in the the distribution of life course clusters

Data Source

The data come from the German Family Panel (pairfam), release 10.0 Brüderl et al. (2019). A description of the study can be found in Huinink et al. (2011).

We gratefully acknowledge the permission of the pairfam team to share a reduced version of their data to illustrate all techniques presented in the book with real-world survey data.

If you are interested in using the complete data sets please turn to: https://www.pairfam.de/en/

Online references

- **Sequence Analysis Association (SAA)** https://sequenceanalysis.org
- **TraMineR R package** http://traminer.unige.ch
- **Sequence Analysis (2022) Marcel Raab and Emanuela Struffolino,** Sage https://sa-book.github.io

Suggested Readings

1. A Aassve, Francesco C Billari, and R Piccarreta. Strings of adulthood: A sequence analysis of young british women's work-family trajectories. Eur J Population, Jan 2007.
2. A Abbott and A Hrycak. Measuring resemblance in sequence data: An optimal matching analysis of musicians' careers. American journal of sociology, 96(1):144–185, 1990.
3. Andrew Abbott. Sequence analysis: new methods for old ideas. Annual Review of Sociology, 21(1):93– 113, 1995.
4. Andrew Abbott and John Forrest. Optimal matching methods for historical sequences. Journal of Interdisciplinary History, 16(3):471–494, Jan 1986.
5. Andrew Abbott and Alexandra Hrycak. Measuring resemblance in sequence data: An optimal matching analysis of musicians' careers.
6. Andrew Abbott and A Tsay. Sequence analysis and optimal matching methods in sociology: Review and prospect. Sociological Methods & Research , 29(1):3, 2000.
7. S Aisenbrey and A Fasang. New life for old ideas: The"second wave"of sequence analysis bringing the"course"back into the life course. Sociological Methods & Research, Jan 2010.
8. N. Barban. Family trajectories and health: A life course perspective. European Journal of Population (2013)

9. N. Barban, De Luna X., Svensson I. and Billari, F.C. “Causal effects of the timing of life course events: age at retirement and subsequent health.” (2018) ”Sociological Methods and Research
10. Barban, N. and M. Sironi (2018) “Sequence Analysis as a Tool for Family Demography,’ ’ in Analytical Family Demography, The Springer Series on Demographic Methods and Population Analysis, Ed. R. Schoen , Germany: Springer.
11. Barban N and F. C. Billari, (2012) “Classifying life course trajectories. A comparison between la- tent class and sequence analysis,” Journal of the Royal Statistical Society Series C (Applied Statistics), 61(5):765–784.
12. Francesco C Billari and R Piccarreta. Analyzing demographic life courses through sequence analysis. Mathematical Population Studies, Jan 2005.
13. Mary Blair-Loy. Career patterns of executive women in finance: An optimal matching analysis. American Journal of Sociology, 104(5):1346–1397, Mar 1999.
14. Christian Brzinsky-Fay, Ulrich Kohler, and M Luniak. Sequence analysis with stata. STATA JOURNAL, Jan 2006.
15. Cees Elzinga. Sequence analysis: Metric representations of categorical time series. Sociological Methods & Research, 2006.
16. Cees Elzinga and Aart C Liefbroer. De-standardization of family-life trajectories of young adults: A cross-national comparison using sequence analysis. Eur J Population, 23(3):225–250, 2007.
17. Brendan Halpin. Optimal matching analysis and life-course data: The importance of duration. Socio- logical Methods & . . . , 38(3):365, Feb 2010.
18. ShinKap Han and Phyllis Moen. Clocking out: Temporal patterning of retirement. American Journal of Sociology, 105(1):191–236, Jul 1999.
19. D.P Hogan. The variable order of events in the life course. American Sociological Review, 43(4):573–586, 1978.
20. Duncan McVicar and Michael Anyadike-Danes. Predicting successful and unsuccessful transitions from school to work by using sequence methods. J Roy Stat Soc A Sta, 165(2):317–334, Jan 2002.
21. T Mouw. Sequences of early adult transitions. On the frontier of adulthood: Theory, Jan 2005.
22. Luca Maria Pesando, Maria Sironi, Nicola Barban, Frank F Furstenberg (2021) A Sequence-Analysis Approach to the Study of the Transition to Adulthood in Low-and Middle-Income Countries Population Development Review
23. R Piccarreta and F.C Billari. Clustering work and family trajectories by using a divisive algorithm. Journal of the Royal Statistical Society: Series A (Statistics in Society), 170(4):1061–1078, 2007.
24. Gary Pollock. Holistic trajectories: a study of combined employment, housing and family careers by using multiple-sequence analysis. J Roy Stat Soc A Sta, 170:167–183, Jan 2007.
25. Sironi M., N. Barban and R. Impicciatore (2015), “The Role of Parental Social Class in the Transition to Adulthood,” Advances in Life Course Research 26, 89-104
26. N Shoval and M Isaacson. Sequence alignment as a method for human activity analysis

in space and time. Annals of the Association of American Geographers, 97(2):282–297, 2007.

27. LL Wu. Some comments on" sequence analysis and optimal matching methods in sociology: Review and prospect". Sociological Methods & Research 29(1):41–64, 2000.