Bootstrapping approaches for validation of a spatial microsimulation model in Australia

Azizur Rahman
School of Computing and Mathematics, Charles Sturt University, Wagga Wagga NSW 2078, Australia; Email: azrahman@csu.edu.au

Abstract

Validation and creation of the statistical reliability of small area estimates by microsimulation modelling is challenging due to the fact that the estimated data are typically unavailable from another source for the small areas. The literature suggest few methods including aggregating the small area estimates up to larger levels where reliable data are available to compare the results with the attempt of adopting alternative methods to determine the accuracy of the model (e.g. Tanton et al., 2014; Timmins and Edwards, 2016; Rahman and Harding, 2016; Whitworth et al., 2016). Bootstrapping approaches (Efron and Tibshirani, 1993; Steyerberg et al., 2003) can be another choice for validation of the estimates for small areas with smaller populations. Although bootstrapping is based on re-sampling technique, and widely used as methods of validating statistical prediction models (Gerds, et al., 2008; Kopec et al., 2010), bootstrapping approaches have not, to our knowledge, been applied to validate spatial microsimulation models. This research aims to introduce bootstrapping methodologies in calculating statistical reliability measures for small area estimates from a spatial microsimulation model in Australia. An empirical application demonstrates that the bootstrapping approaches seem most meaningful systems for validation of the spatial microsimulation model for those areas with smaller populations. Some key advantages of such methods are also explored.

Keywords: Bootstrapping; confidence intervals; housing stress; small area estimation; spatial microsimulation; Australia.

References:

Efron, B and Tibshirani, R (1993). *An Introduction to the Bootstrap*. Chapman and Hall/CRC, Boca Raton, Florida.

Gerds, TA, Cai, T and Schumacher, M (2008). The performance of risk prediction models. *Biomedical Journal*, 5(4): 457-479.

Kopec, JA, Finès, P, Manuel, DG, Buckeridge, DL, Flanagan, WM, Oderkirk, J, Abrahamowicz, M, Harper, S, Sharif, B, Okhmatovskaia, A, Sayre, EC, Rahman, MM and Wolfson, MC (2010). Validation of population-based disease simulation models: a review of concepts and methods. *BMC Public Health*, 10(710): 1-13.

Rahman A and Harding A (2016). *Small area estimation and microsimulation modelling.* Chapman and Hall/CRC, Boca Raton, Florida.

Steyerberg, EW, Bleeker, SE, Moll, HA, Grobbee, DE and Moons, KG (2003). Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *Journal of Clinical Epidemiology*, 56(5):441-447.

Tanton R, Williamson P and Harding A (2014). Comparing two methods of reweighting a survey file to small area data. *International Journal of Microsimulation*, 7(1): 76-99.

Timmins, KA and Edwards KL (2016). Validation of spatial microsimulation models: a proposal to adopt the Bland-Altman method. *International Journal of Microsimulation*, 9(2): 106-122.

Whitworth A, Carter E, Ballas D and Moon G (2016). Estimating uncertainty in spatial microsimulation approaches to small area estimation: A new approach to solving an old problem. *Computers, Environment and Urban Systems* (forthcoming), available at http://dx.doi.org/10.1016/j.compenvurbsys.2016.06.004 (accessed on 16/11/2016).