SCALING OF EMPIRICAL COMPARTMENTAL DISTRIBUTIONS

Justas Kvedaravičius[1], dr. Aleksejus Kononovičius[2]

[1]Faculty of Physics, Vilnius university, Lithuania
[2]Institute of Theoretical Physics and Astronomy, Faculty of Physics, Vilnius university, Lithuania
justas.kvedaravicius@ff.stud.vu.lt

Opinion dynamics is by far the most studied topic in sociophysics. Yet some authors note that there's a lack of comparison between voter model's theoretical results and empirical (experimental) data [1]. To overcome this issue some authors have suggested to include space (geography) in modelling opinion dynamics [1, 2]. To understand the implications of space to opinion dynamics, it's worth studying how empirical data can entail structures of opinion or such indicators.

This research analyzes changes in empirical compartmental distributions by changing the scale of empirical data. For this purpose empirical data from United Kingdom 2011 census have been chosen. Different demographic indicators of this data exhibit varying distributions for changes in scale. Yet it is unclear whether distributions stem from existing socio-demographic structures in United Kingdom society or are of statistical nature.

A randomized spatial model is proposed to find the nature of observed distributions. It takes compartments from the most detailed (first) scale of empirical data and joins them to replicate the number of compartments of a less detailed scale. Thus the compartments of the new scale are not consistent with geography but rather randomized. The said procedure is repeated for all non-first data scales. Then two diversity quantities are chosen for which relative changes in their value are calculated. These quantities are, namely, standard deviation and Theil index. Changes of these quantities for randomized spatial model are compared to the ones produced by empirical data. Overall, the procedure used in this model resemble what is well-known in statistics as null-hypothesis testing.

Results of randomized spatial model with the ones produced by empirical data are shown in Fig. 1. It is shown that relative changes of diversity quantities in random model decline more rapidly than the ones of empirical data. Furthermore, we show that changes of proposed model follow the $N^{-1/2}$ power-law for standard deviation whilst for Theil index these correspond to N^{-1}. These results also demonstrate that the changes in selected demographic indicators are not of statistical nature as they follow different power-laws.

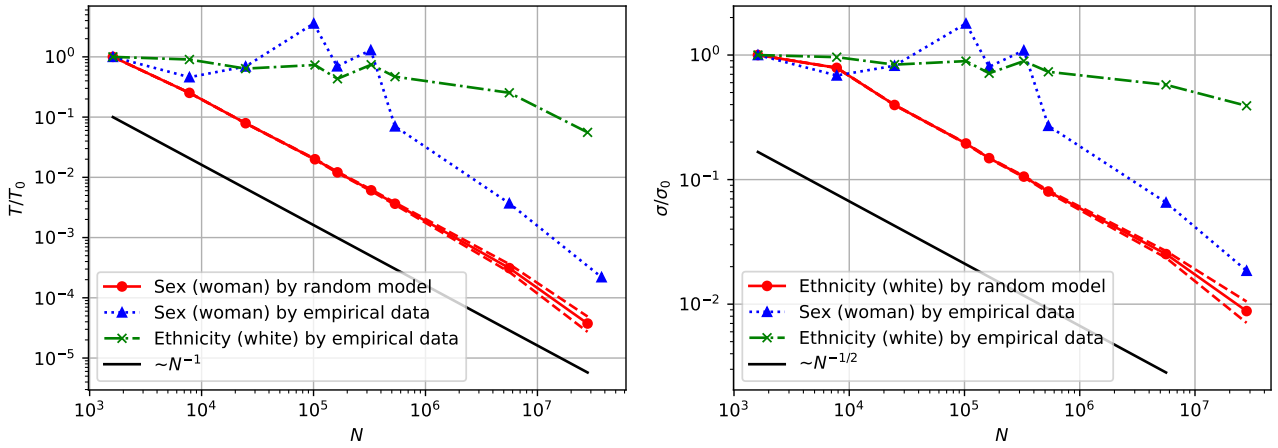

Fig. 1. Changes in relative Theil index (left) and relative standard deviation (right) values. X axis represents the mean population in compartment at different scales of empirical data.

Among other results, it has also been noted that chosen diversity quantities describe changes in different quality. Other two quantities (range and Gini coefficient) have also been tested for results with empirical data. Their relative changes are somewhat similar to the previously mentioned ones: range to Theil index and Gini coefficient to standard deviation.

In conclusion, the proposed randomized spatial model provides researchers with easy yet delicate measure to assess whether or not empirical data (or data generated by spatial models) is of statistical nature. Different diversity and inequality quantities may be chosen for data evaluation as they are bound only to calculability via compartments. Further adaptations to the model may be able to change the nature of data reproduced by it.

[1] J. Fernandez-Gracia, K. Suchecki, J. J. Ramasco, M. Miguel, V. Egu ıluz, *Is the voter model a model for voters?*, Physical Review Letters 112, 158701 (2014)
[2] A. Kononovicius, *Compartmental voter model*, Journal of Statistical Mechanics: Theory and Experiment 2019, 103402 (2019).