Statistical patterns of Lithuanian municipality elections

Aleksejus Kononovicius
Institute of Theoretical Physics and Astronomy, Vilnius University, Sauletekio al. 3, 10257 Vilnius
email: aleksejus.kononovicius@tfai.vu.lt; www: http://kononovicius.lt, http://rf.mokslasplius.lt

Introduction

Many models in sociophysics literature consider binary opinions (e.g., yes or no, democrat or republican) [1], which likely natural having in mind importance of the Ising model in statistical physics. Most of these models also are convergent, meaning that some fixed point is reached by the system after some time, which allows to use well established tools to study critical exponents and phase transitions [1]. Yet in reality societies are much more complex than binary options are able to describe. Furthermore societies seem to be constantly undergoing change, yet retaining their inherent diversity and showing no signs of convergence. Such context raises a very interesting question, which was already asked by Axelrod [2], "if people tend to become more alike in their beliefs, attitudes, and behavior when they interact, why do not all such differences eventually disappear?"

There are few empirical works, which observe similar dove patterns we will discuss here. Namely, that the vote share distribution is rather broad and is poorly describe by a single fixed point. Some of these works have proposed to use various distributions to fit the empirical data (e.g., normal, log-normal, Weibull or Beta) and provided models, which generate those distributions [3, 4, 5, 6]. We are strongly in favor of Beta distribution, because it has multinomial alternative, which is well defined on simplex, and can be reproduced by an extremely simple model – the classical voter model with noise [7], which is equivalent to Kirman's herding model [8, 9, 10].

In this contribution we will present an analysis of statistical patterns observed during Lithuanian municipality elections. We consider parties' vote share at the polling station level. We perform a comparison across the different municipalities with an aim to understand whether the vote share samples could have come from the same distribution. This approach is based on our earlier works in which we have considered Lithuanian parliamentary elections [9, 10].

Empirical analysis

We analyze the empirical data from two perspectives – using probability density functions (PDF) and rank-size (RS) distributions. These two approaches are equivalent, but using both of them reveals a more detailed picture.

In both cases we analyze the same parameter – parties' vote shares over the different polling stations. We define vote share as a number of votes for a party in a given polling station divided by a total of votes cast for all parties within the same polling station. We consider each election as a separate data set, in which each polling station provides a different data point.

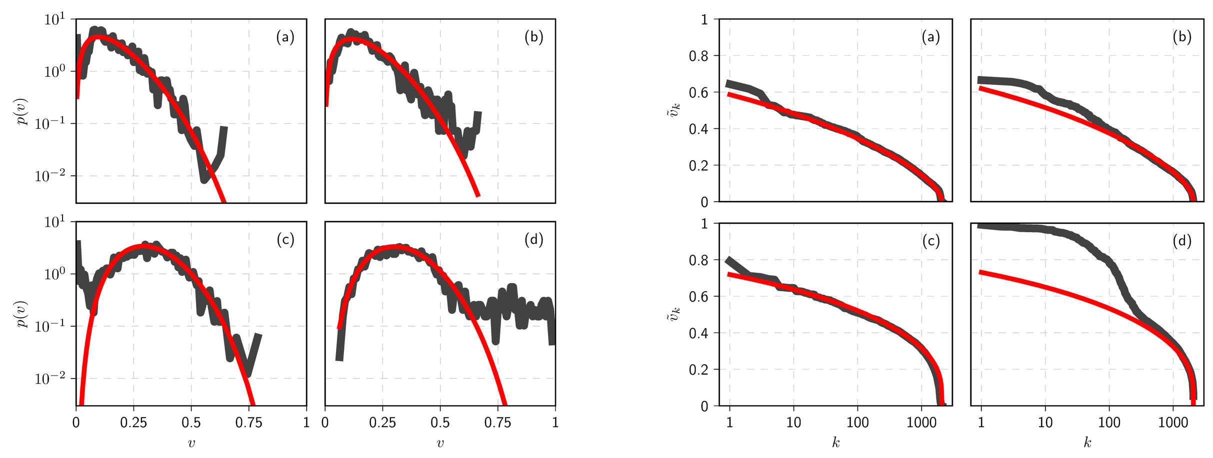

Fig 1: Party vote share PDF (left) and RS (right) from the 1995 municipality elections. Parties considered: (a) - LKDP, (b) - LDDP, (c) - TS, (d) - others combined. Empirical curves are gray, while red curves are fits (assuming Beta distribution with α_1 and β_1 parameters from Table 1).

The fit is not perfect in a similar way as it was with 1992 parliamentary elections [9, 10]. Namely some parties, such as Electoral Action of Poles in Lithuania, have already established their own "strongholds". Note that this most strongly impacts the combined party, because it is composed of smaller weaker parties, which do not try to win whole country, but instead focus on specific regions. Here we have assumed that 90% of the polling stations can be described by $\mathcal{B}e\left(\alpha_1, \beta_1\right)$ and 10% of the polling stations can be described by $\mathcal{B}e\left(\alpha_2, \beta_2\right)$ (parameter values are given in Table 1.

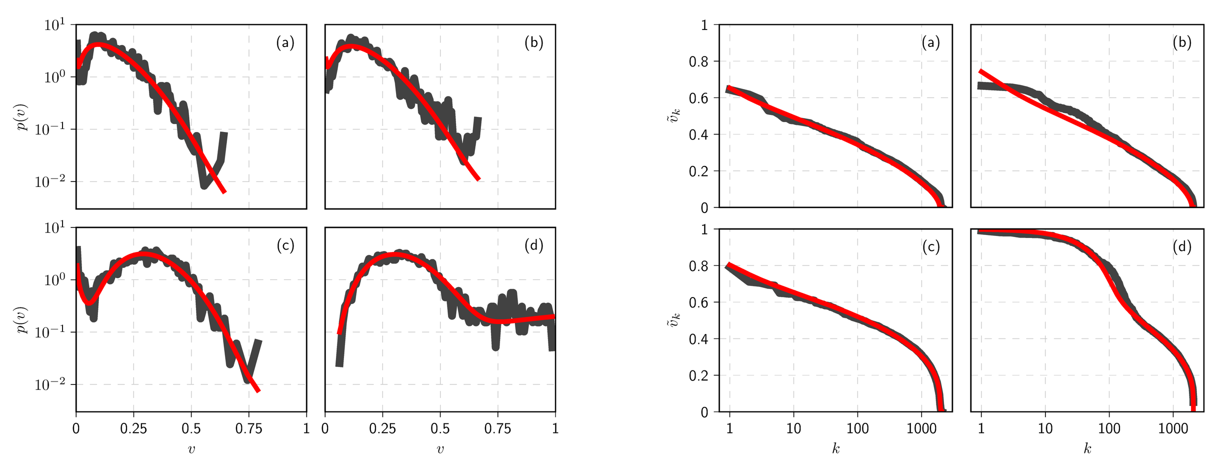

Fig 2: Party vote share PDF (left) and RS (right) from the 1995 municipality elections. Parties considered: (a) - LKDP, (b) - LDDP, (c) - TS, (d) - others combined. Empirical curves are gray, while red curves are fits (assuming mixture of two Beta distributions using all parameters from Table 1).

Table 1: Parameters of the fits. ρ quantifies the weight of the second Beta distribution.

Party	α_1	β_1	ρ	α_2	β_2
LKDP	2.1	11.12	0.1	0.1	2.5
LDDP	2.2	10.3	0.1	0.2	2.4
TS	2.5	5.77	0.1	0.3	2.3
Other	2.8	4.5	0.1	2	1

Agent-based model

Here we generalize a well known two state model [8] to account for the switching between multiple states. In Fig 4 we illustrate the model for the three state case. This model is equivalent to the classical voter model with noise [7] on complete or random graph [6, 9, 10].

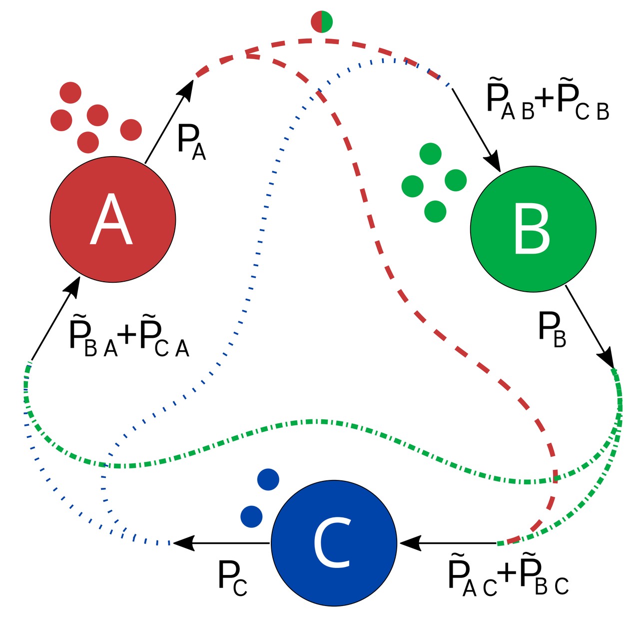

Fig 4: Model schema with the three available states (parties). Notation explained in text.

Probability for a single agent to switch from party i to j during time interval Δt is given by:

$$\tilde{P}_{ij} = X_i \left(\varepsilon_{ij} + X_j\right) h \Delta t.$$

Here ε_{ij} is a realitive impact of idiosyncratic behavior, independent "discovery" of the state j (party, voting option) from state i, while h adjusts the event time scale. This process can be better modeled as a birth–death process using Gillespie approach [11].

If there are only two states, then it is trivial to show that $x_j = X_j/N$ is distributed according to the Beta distribution, $x_j \sim \mathcal{B}e(\varepsilon_{ij}, \varepsilon_{ji})$. If there are more than two state, unless $\varepsilon_{ij} = \varepsilon_j$, x_j distribution will not be the Beta distribution. Thus we simply the model by making this assumption. Though there will be some times this assumption must be violated to get better agreement with empirical data [9, 10].

Under the simplifying assumption we can write the exit, P_i, and the entry, $\tilde{P}_j$, probabilities as follows:

$$P_i = \sum_{j \neq i} \tilde{P}_{ij} = X_i \sum_{j \neq i} \left(\varepsilon_j + X_j\right) h \Delta t = X_i \left(\sum_{j \neq i} \varepsilon_j + [N - X_i] \right) h \Delta t,$$
$$\tilde{P}_j = \sum_{i \neq j} \tilde{P}_{ij} = \sum_{i \neq j} X_i \left(\varepsilon_j + X_j\right) h \Delta t = [N - X_j] \left(\varepsilon_j + X_j\right) h \Delta t.$$

Then the distribution of x_j is also the Beta distribution, $x_j \sim \mathcal{B}e\left(\varepsilon_j, \sum_{i \neq j} \varepsilon_i\right)$. Parameters of which we have estimated earlier as α_1 and β_1 respectively.

Is this diversity really about the opinion dynamics?

In [12] a group of prominent sociophysicists raised a question whether the voter model is truly a model of voter behavior. Here we show that Beta distribution is also observed in census data, therefore all of our data and modeling could be related to socio–economic segregation processes instead. Below we show that you can indeed fit various census data using Beta distribution, though notably not all categories can be fitted as well as shown.

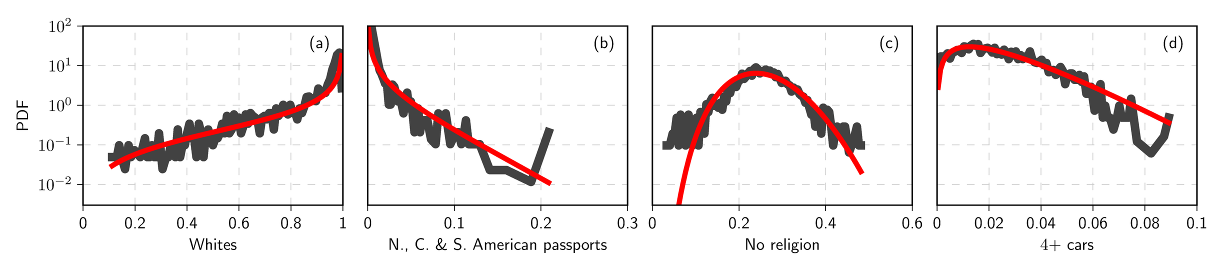

Fig 5: Fitting UK 2011 Census data with Beta distributions.

References

[1] C. Castellano, *et al.*, Rev. Mod. Phys. **81**: 591–646 (2009).
[2] A. Flache, *et al.*, JASSS **20**: 2 (2017).
[3] M. C. Mantovani, *et al.*, EPL **96**: 48001 (2011).
[4] S. Fortunato, C. Castellano, PRL **99**: 138701 (2007).
[5] T. Fenner, *et al.*, Quality & Quantity **52**: 1 (2017).
[6] F. Sano, *et al.*, in APEC-SSS2016, 011016 (2016).
[7] P. Clifford, A. Sudbury, Biometrika **60**, 581–588 (1973).
[8] A. P. Kirman, Q. J. Eco. **108**, 137–156 (1993).
[9] A. Kononovicius, Complexity **2017**, 73354642 (2017).
[10] A. Kononovicius, Acta Physica Polonica A **133** (6): 1450 (2018).
[11] A. Carro, *et al.*, arXiv: 1302.6477 [physics.soc-ph].
[12] J. Fernandez–Garcia, *et al.*, PRL **112**: 158701 (2014).

Acknowledgement: This research was funded by the European Social Fund under the No 09.3.3-LMT-K-712 "Development of Competences of Scientists, other Researchers and Students through Practical Research Activities" measure.

2014-2020 Operational Programme for the European Union Funds Investments in Lithuania