Physica A 496 (2018) 171–188

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Statistical analysis of Brazilian electoral campaigns via Benford's law

Daniel Gamermann *, Felipe Leite Antunes

Universidade Federal do Rio Grande do Sul (UFRGS)-Instituto de Física, Av. Bento Gonçalves 9500, Porto Alegre, Rio Grande do Sul, Brazil

HIGHLIGHTS

- Study of public data on Brazil's electoral campaign.
- The total amount of money a candidate receives predicts if he is elected or not.
- The first digit distribution of the amounts does not match Benford's law.
- The amounts from a model with similar statistic does match Benford's law.

ARTICLE INFO

Article history:
Received 21 July 2017
Received in revised form 25 October 2017
Available online 29 December 2017

Keywords:
Benford's law
Logistic regression
Electoral campaign
Politics

ABSTRACT

The principle of democracy is that the people govern through elected representatives. Therefore, a democracy is healthy as long as the elected politicians do represent the people. We have analyzed data from the Brazilian electoral court (*Tribunal Superior Eleitoral, TSE*) concerning money donations for the electoral campaigns and the election results. Our work points to two conclusions that combined may be in conflict with the democratic principle: money is the determining factor on whether a candidate is elected or not (opposed to representativeness); secondly, the use of Benford's Law to analyze the declared donations received by the parties and electoral campaigns shows either possible manipulations in the declarations or a significant number of donations that might not have been spontaneous from the donors. The better term that describes Brazil's government system is plutocracy (govern by the wealthy).

1. Introduction

Modern society dependence on technologies, in particular the Internet and mobile phones, has as consequence the generation of huge amounts of raw data. Apart from the problematic involved in the processing and storage of this data, the data's volume, structure and variety call for the development of new analysis methodologies in order to extract the important information (knowledge) behind it. Also, as scientific fields that have traditionally adopted qualitative approaches slowly tackle quantitative analyses, a vast new horizon opens to new applications of methodologies long known to the physics community.

This interaction of physics with other sciences has been fruitful in apparently distant fields such as economics [1–4], biology [5], medicine [6] or political sciences [7,8]. In this context, Statistical Physics has much to offer, particularly in understanding, quantifying and modeling the dynamics and properties of a large number of elements. Big Data [9] with

* Corresponding author.
E-mail address: danielg@if.ufrgs.br (D. Gamermann).

https://doi.org/10.1016/j.physa.2017.12.120
0378-4371/

its unprecedented scale and much finer resolution, provides a powerful experimental apparatus to challenge our existing models, explore new tools and frameworks, and lead research to new areas [10].

At the moment, the sector that most benefits from the rising data science field is the private sector. Companies invest heavily in studding customer profiles and needs in order to offer more attractive services and increase their profits or optimizing decision making process minimizing risks. On the other hand, the public sector should benefit from knowledge obtained with these new information technologies. Objective data analysis could guide public policies preventing the spread of epidemics [11,12], minimize traffic jams [13], decreasing unemployment [14], fighting corruption [15–17], crime [18,19] and violence [20].

An interesting result applied in the detection of anomalies or manipulations in financial declarations is the Benford's Law. Noted for the first time by the astronomer and mathematician Simon Newcomb [21], and empirically postulated by Benford when comparing data collected from a variety of sources, ranging from the statistics of the American baseball league to the atomic weights of the elements, the law of probability of occurrence of numbers, as observed by Newcomb, is such that the logarithm of the mantissas[1] of numbers in a large set are equiprobable. This observation can be put as follows[2] [23]:

$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(1 + \frac{1}{d}\right), \tag{1}$$

$$d \in \{1, 2, .., 9\}. \tag{2}$$

Despite its simplicity, the first rigorous proof was only developed by Hill in 1995 [24]. In the original work, Hill proves, based on probability theory, that scale invariance implies base invariance and base invariance, in turn, implies the Benford's Law.

Sets of numbers tend to follow this law given that they are naturally occurring (random) numbers, coming from multiple different distributions and expanding many orders of magnitude. By naturally occurring numbers, it is meant numbers that are not sequential, man made, as would be for example, serial numbers or license car plates, which would not be random, but cover a given range uniformly. Another set of number that is expected to follow Benford's law (and is actually the one which motivated Newcomb's argument [21]) is a set of numbers that are the result of multiplications or quotients of others.

It is interesting to note that Benford's distribution is scale invariant, so it does fell as a natural law (independent of man made measurement systems or concepts): i.e. take the measurement of the heights of all mountains in a country, if they tend to follow Benford's law, they will do so no matter if the measurements are made in meters, feet or inches. The distribution of the first digit will have approximately the same shape no matter the unit system used. Were the distribution uniform in a given measurement system, it would have a complete different shape in another system, the distribution would then be measurement system dependent.

Benford's law may be an important tool in order to search big amounts of data for anomalies. It is interesting to note that Benford's law has already been used in order to detect evidence of manipulation in electoral results [25] and in revenue tax declarations [26,27]. Moreover, Benford's Law has already been applied to analyze campaign financing using FEC (US Federal Election Commission) data[3] to detect irregularities on in-kind contributions [28]. In the US, donations can be done directly to political parties with no limits attached, these are known as "soft money". The US law says that this money can only be used for "party-building activities" such as advocating the passage of a law or voter registration, for example, but not for a particular candidate campaign during the elections. On the other hand, if a donation is made directly to a political candidate the money is called "hard money" and can only come from an individual or a political action committee. This money must follow the strict limits set by the FEC but, in contrast to the "soft money", can be used for a candidate's promotion. In [28], the authors investigate whether candidates may have used the soft and hard money distinction to bend FEC rules and the tool they use to detect possible anomalies is Benford's law. In fact, Benford's law is nowadays an important tool of forensic accounting [29–31].

In this work we analyze publicly available data on Brazilian elections. Brazil's superior electoral court (TSE from *Tribunal Superior Eleitoral*) freely provides the financial declarations made by parties, candidates and electoral committees and all statistics on the election results. This information can be downloaded from the TSE webpage [32] (see also the Appendices A and B).

Ideally, in a democracy, the people elects its leaders based on representativeness. Those politicians that better represent the population or groups within the population and better defend their interests should end up elected. The electoral campaign is the opportunity the candidates to offices have to express their ideas and the voters to get acquainted with the candidates and to chose those that better represent their interests. In practice, Brazil's system faces many problems. On one hand, not all candidates have the same opportunity to appear in front of the population and express their plans; on the other hand, no matter what a politician promises during the campaign, once elected he can follow a completely different line. The first problem, we believe, can be traced to a single factor: money. Electoral campaigns are much closer to plain publicity than to ideological debate. The more money a candidate or a party has, the better the marketing professionals he can hire and the more time he can buy in private media and consequently, the more he is remembered by the voters. The public media time is shared by the candidates and parties, but it is proportional to the number of congressmen each party

[1] The mantissa m of a given number x is such that $x = m \times 10^n$, Where n is an integer and $m \in [1, 10)$.

[2] More accurately, the uniform distribution of the log of the mantissa is equivalent to the generalized Benford's distribution for n-digits [22].

[3] https://www.fec.gov/.