Chaotic Modeling and Simulation (CMSIM) 4: 695–700, 2012

Double power law behavior in everyday phenomena

Carla M.A. Pinto[1], António M. Lopes[2], and J.A. Tenreiro Machado[3]

[1] School of Engineering, Polytechnic of Porto
and Center of Mathematics, University of Porto
Rua Dr António Bernardino de Almeida, 431
4200-072 Porto Portugal (Email: cap@isep.ipp.pt)
[2] UISPA, IDMEC - Polo FEUP
Faculty of Engineering
University of Porto
Rua Dr. Roberto Frias,
4200-465 Porto, Portugal
(Email: aml@fe.up.pt)
[3] Department of Electrical Engineering
School of Engineering, Polytechnic of Porto
Rua Dr António Bernardino de Almeida, 431
4200-072 Porto Portugal (Email: jtm@isep.ipp.pt)

Abstract. We study similar statistical properties observed in distinct real world data. In particular, we focus on the power law (PL) distribution. We find that some data is well fitted by a single PL distribution whereas other phenomena force the use of two distinct PLs. This behavior is similar in, a priori, unrelated phenomena, such as catastrophes (terrorism, earthquakes) and variables associated to man-made systems, such as distribution of the number of words in texts or of the number of hits received by websites.
Keywords: power law, double power law, real world phenomena.

1 Introduction

Pareto [13] and Zipf [19] laws are examples of Power law (PL) distributions. These distributions are characterized by heavy tails and were first studied in 1896 by Pareto [13]. Pareto observed that the relative number of individuals with an annual income larger than a certain value x was proportional to a power of x. The later can be expressed mathematically by the expression (1).

$$F(x) = P(X \geq x) = \frac{C}{\alpha - 1} x^{-(\alpha-1)} \tag{1}$$

where $\alpha > 0$, $C > 0$, and $F(x)$ is the complementary cumulative distribution function of the income x. In the text, we will consider $\tilde{\alpha} = \alpha - 1$ and $\tilde{C} = \frac{C}{\tilde{\alpha}}$. Zipf law is a special case of the Pareto law with exponent $\tilde{\alpha} = 1$.

Received: 9 May 2012 / Accepted: 22 October 2012

ISSN 2241-0503

Variable x in equation (1) has been used to describe quantities in a wide variety of real data. Namely, x may represent the number of: (i) individuals in a population of a city [2,19,6], (ii) articles' citations [11], (iii) hits in webpages [1], (iv) victims in wars, terrorist attacks, or earthquakes [7,16,4], (v) words in texts [5,19], and several other phenomena [12,10]. In the literature interesting reviews on on PL behavior and applications can be found [9,18,14].

Application of PL behavior in natural or human-made phenomena usually comes with a log-log plot, where the axes represent the size of an event and its frequency. The log-log plot is asymptotically a straight line with negative slope.

The paper is organized as follows. In Section 2, we review literature concerning distinct phenomena where PL behavior has been fitted. In Section 3, we present a numerical analysis of real data where PL and double PL behavior is observed. Finally, in Section 4, we state the main conclusions and discuss future research directions.

2 Real events

PL behavior has been used to model the number of casualties in natural and human-made phenomena, such as earthquakes, tornados, terrorist attacks and wars. Understanding patterns of the number of casualties in these events may help to organize rescue operations. [7,16,4]. Other applications of PL behavior, with less impact in terms of human lives, are city and forest fires, words' frequency in texts, or the number of hits in webpages. In what concerns the study of city and forest-fire distributions, results may help to take measures beforehand in view of possible hazards, thus saving natural resources and animal and human lives.

We observe a common underneath behavior considering the number of casualties and the frequency of natural and human-made disasters. Large casualties are less frequent and are associated with low frequency phenomena. Two world wars are two examples of this type. Other wars, not so harmful in terms of preserving human lives [15], are more frequent. Analogously for earthquakes, the frequency of occurrence of terrific earthquakes, that cause a large number of victims, is much lower than that of smaller earthquakes with few casualties [7].

Johnson *et al* [8] studied war and global terrorism patterns, and developed a theory for explaining their similar dynamical evolution. The later was invariant to underlying ideologies, motivations and the terrain in which they operated. They considered each insurgent force as a generic, self-organizing system, which evolved dynamically through the continual coalescence and fragmentation of its constituent groups. Researchers have used wars in Iraq and Afghanistan, and long-term guerrilla war in Colombia, as examples. On global terrorism, attacks to London, Madrid, and New York (September 11) were main choices. Results obtained showed a PL behavior for Iraq, Colombia and Afghanistan, with coefficient value (close to) $\tilde{\alpha} = 2.5$. This value of the coefficient equalized the coefficient value characterizing non-G7 terrorism. In 2007, Clauset *et al* [3] plotted a log-log chart for the frequency versus the severity of terrorist attacks, since 1968, and found a straight line, denoting PL behavior.

In 2003, Song *et al* [17] studied fire distribution in Chinese and Swiss cities. The authors computed the frequency loss and the rank-size plots and verified validity of a PL in both cases. The frequency loss was the frequency of fires with loss L, that is, fire loss L converted into Chinese Yuan. The rank was computed by sorting city fires from large to small, and considering the largest with rank 1. The PL distribution was invariant for scale and time, meaning that fire distribution is common for different places and times.

3 Application to real data

PLs are present in many natural and man-made systems and, for certain cases, a single PL distribution holds over the entire data range. As an example, Figure 1 represents the rank/frequency log-log plot of the largest private American companies, with respect to their annual revenue, in the year 1997, according to Forbes (`http://www.forbes.com/`). The data was collected, sorted and ranked, and then the normalization of the values was carried out. That is, the data (x-axis) was divided by the highest annual revenue, and the rank (y-axis) was divided by the rank of the smallest company. A PL was adjusted to the data using a least squares algorithm. As can be seen in Figure 1, a PL behavior distribution with parameters $(\tilde{C}, \tilde{\alpha}) = (0.0031, 1.3004)$ holds over the entire range of the companies' annual revenue.

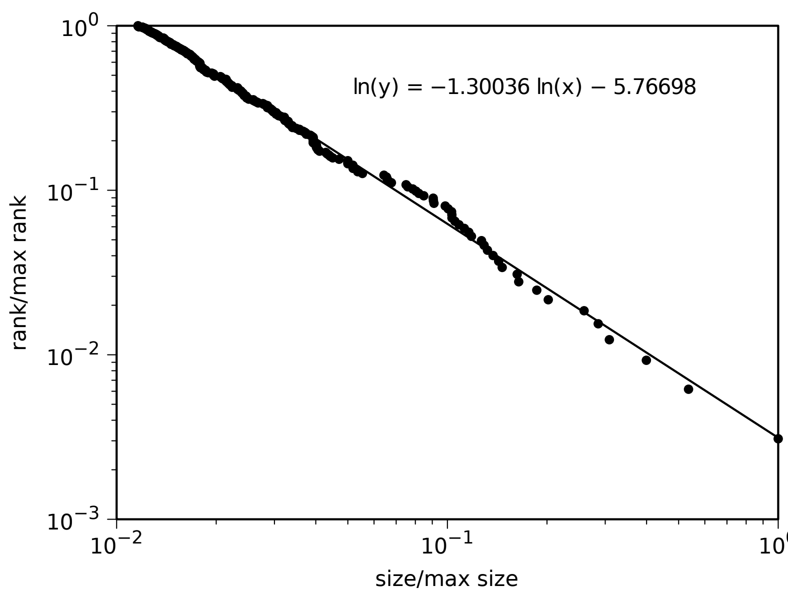

Fig. 1. Rank/frequency log-log plot of the size of the largest American companies in 1997.

In other real applications, different PLs, characterized by distinct parameters, may also be observed. In the sequel, several cases of such behavior are illustrated.

Figure 2 shows the cumulative distribution function of the size of forest fires in Portugal, over the year 2001. The adopted measure for size is the total burned area. Only fires greater than 100 ha in total burned area are considered. The data is available on the Portuguese National Forest Authority (AFN) website (`http://www.afn.min-agricultura.pt/`). For this case, two distinct PLs with parameters $(\tilde{C}_1, \tilde{\alpha}_1) = (0.0383, 0.9232)$ and $(\tilde{C}_2, \tilde{\alpha}_2) = (0.0065, 2.4665)$ fit the data. The change in the behavior occurs at the relative value of 0.35, approximately.

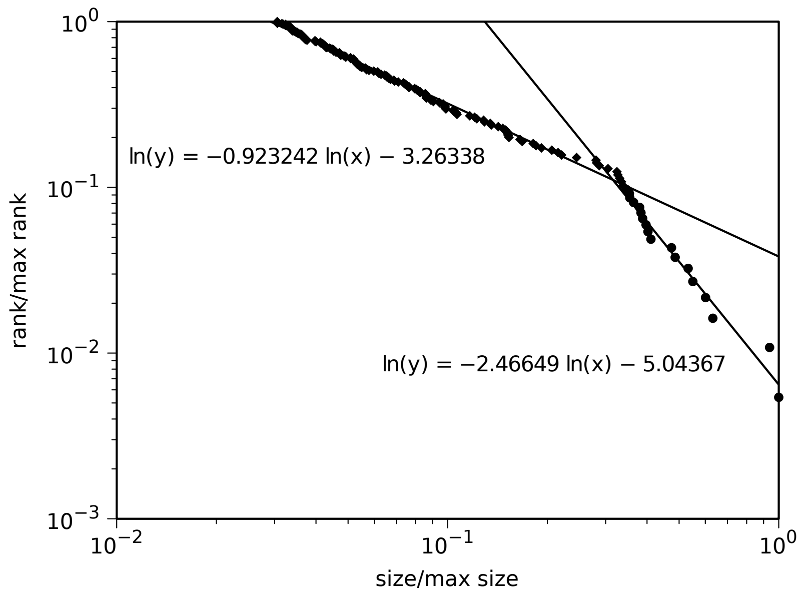

Fig. 2. Rank/frequency log-log plot of a system presenting dual PL behaviour: size of forest fires in Portugal, year 2001.

Figure 3 represents the severity of tornadoes in the USA, during 2003. The total number of human victims (killed and injured) directly related to a given occurrence is used to quantify its severity. The data is available at the U.S. National Oceanic and Atmospheric Administration (`http://www.noaa.gov/`), National Weather Service, Storm Prediction Center website (`http://www.spc.noaa.gov/`). The chart reveals a dual PL behavior with parameters $(\tilde{C}_1, \tilde{\alpha}_1) = (0.0413, 0.5804)$, $(\tilde{C}_2, \tilde{\alpha}_2) = (0.0100, 1.2374)$.

For the finale, we remark that several examples of real world phenomena, where a double PL behavior is observed, were presented. Future work will focus on possible explanation for this peculiarity seen in distinct phenomena, that are described by PLs.

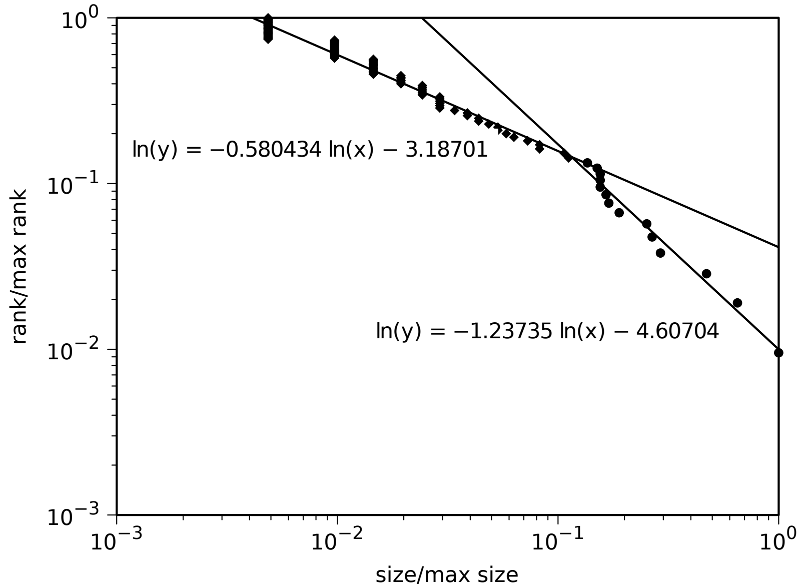

Fig. 3. Rank/frequency log-log plot of a system presenting dual PL behaviour: severity of tornadoes in the USA in 2003.

4 Conclusion

In this paper, we focused on PL distributions as models of sets of real data. We presented examples of data that was well fitted by a single straight line and examples that were best described by two distinct PL distributions. The reason behind this type of behavior in distinct and not related phenomena is still to be found.

References

1.L. A. Adamic and B. A. Huberman. The nature of markets in the World Wide Web. *Quarterly Journal of Electronic Commerce*, 1:512, 2000.
2.F. Auerbach. Das Gesetz der Belvolkerungskoncentration. *Petermanns Geographische Mitteilungen*, 59:74–76, 1913.
3.A. Clauset, M. Young, K.S. Gleditsch. On the Frequency of Severe Terrorist Events. *Journal of Conflict Resolution*, 51: 58–87, 2007.
4.D.R. Davis, D.E. Weinstein. Bones, bombs, and break points: The geography of economic activity. *The American Economic Review*, 92:1269–1289, 2002.
5.J.B. Estoup. *Gammes Stenographiques*. Institut de France, 1916.
6.X. Gabaix. Zipf's Law and the Growth of Cities. *American Economic Review Papers and Proceedings*, LXXXIX:129–132, 1999.
7.B. Gutenberg, R. F. Richter. Frequency of earthquakes in california. *Bulletin of the Seismological Society of America*, 34:185–188, 1944.
8.N.F. Johnson, M. Spagat, J.A. Restrepo, O. Becerra, J.C. Bohorquez, N. Suarez, E. M. Restrepo, R. Zarama. Universal patterns underlying ongoing wars and terrorism. *arXiv:physics/0605035v1*, 2006.

9.W. Li. References on Zipf's law. `http://www.nslij-genetics.org/wli/zipf/`
10.W. Lin, Y. Yang. Zipf's Law in Importance of Genes for Cancer Classification Using Microarray Data. *J. Theor. Biol.* **219** (2002) 539-551.
11.A.J. Lotka. The frequency distribution of scientific productivity. *J Washington Acad Sci*, 16: 317–324, 1926.
12.R.N. Mantegna, H.E. Stanley. The scaling behavior of an economic index. *Nature* **376** (1995) 46–49.
13.V. Pareto. *Cours d'Economie Politique*, Geneva, Switzerland, 1896. Droz.
14.C.M.A. Pinto, A.M. Lopes, and J.A. Tenreiro Machado. A review of power laws in real life phenomena. *Communications in Nonlinear Science and Numerical Simulations*, accepted for publication, January 2012.
15.L.F. Richardson. *Statistics of Deadly Quarrels*. Chicago, 1960. Quadrangle Books.
16.D. C. Roberts, D. L. Turcotte. Fractality and selforganized criticality of wars. *Fractals*, 6: 351–357, 1998.
17.W.G. Song, H.P. Zhang, T. Chen, W.C. Fan. Power-law distribution of city fires. *Fire Safety Journal*, 38: 453–465, 2003.
18.D. Sornette. *Critical Phenomena in Natural Sciences*, chapter 14. Heidelberg, 2nd edition, 2003. Springer.
19.G. Zipf. *Human Behavior and the Priciple of Least Effort*. Cambridge, MA, 1949. Addison-Wesley.