Fifth Workshop Dynamical Systems Applied to Biology and Natural Sciences DSABNS 2014 Lisbon, Portugal, February 10-12, 2014

Power Laws, Zipf's Law, and Scaling Laws in Human Language and Music

Álvaro Corral[1]

[1] *Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, Barcelona*
acorral@crm.es

ABSTRACT

Power-law distributions have been a paradigm of complex systems for several decades. But it has been only recently that attention has been devoted to the proper fitting and goodness-of-fit testing of these distributions; in particular, Clauset et al.'s method is widely cited in this regard [1]. In the first part of the talk we present evidence that this method fails in some concrete examples of continuous distributions, rejecting the power-law hypothesis even for simulated data with a power-law tail [2]. We propose an alternative more reliable procedure that, additionally, can be generalized to upper truncated power-law distributions [3].

The second part is devoted to one of the classic examples of power-law distributions: the so-called Zipf's law, in the context of word frequencies in texts. This is considered one of the key statistical regularities of human language. We extend the fitting method to cover this discrete case, showing that, in general, Zipf's law does not hold for the whole domain of word frequencies, but only for the upper tail. In addition, we show how the distribution of word frequencies scales with the size of the text and the size of the vocabulary, providing a recipe for the proper comparison of texts of different size [4]. The distinction between power law and scaling law is fundamental here.

The third and last part is devoted to the extension of Zipf's law to music, drawing parallelisms and differences with texts. The construction of music code-words from the chords defining the pitch in modern popular music reveals the validity of Zipf's law in this case. This law has kept stability for the last 50 years, although other characteristics of music have shown an evolution that seems to indicate a decrease of the complexity of music with time [5].

 ISBN: 978-989-98750-0-5

References

[1] A. Clauset, C. R. Shalizi, and M. E. J. Newman (2009) *Power-law distributions in empirical data*. SIAM Rev., 51, pp. 661–703.

[2] A. Corral, F. Font, and J. Camacho (2011) *Non-characteristic half-lives in radioactive decay*. Phys. Rev. E, 83, 066103.

[3] A. Deluca and A. Corral (2013) *Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions*. Acta Geophys., 61, pp. 1351–1394.

[4] F. Font-Clos, G. Boleda, and A. Corral (2013) *A scaling law beyond Zipf's law and its relation to Heaps' law*. New J. Phys., 15, 093033.

[5] J. Serrà, A. Corral, M. Boguñá, M. Haro, and J. Ll. Arcos (2012) *Measuring the evolution of contemporary western popular music*. Sci. Rep., 2, 521.

 ISBN: 978-989-98750-0-5