Zipf and Related Scaling Laws.
3. Literature Overview of Multidisciplinary Applications
(from Informational Aspects to Energetic Aspects)

Giedrė Būdienė, Alytis Gruodis [a]
Vilnius Business College, Kalvarijų 129, Vilnius, Lithuania

Received 22 October 2016, accepted 30 December 2016

Abstract. Review of the most significant application of scaling laws in communications and natural sciences is presented. Nine themes are observed: human as well as non-human communications, bibliometric impacting problems, website distributions, social activity (including sport and music), physics, geology, geography, chemistry, biological systems.

Citations: Giedrė Būdienė, Alytis Gruodis. Zipf and related scaling laws. 3. Literature overview of multidisciplinary applications (from informational aspects to energetic aspects) – *Innovative Infotechnologies for Science, Business and Education,* ISSN 2029-1035 – **2(21)** 2016 – Pp. 12-19.

Keywords: Zipf law; power law; H-index.
Short title: Zipf law - multidisciplinary - 3.

Introduction

Zipf law belongs to the well-known class of power law which are significant in modelling the communicating as well as economic human activity. Based on the ranked distribution of certain items in the finite set, power law allows estimating the artificial or natural character of the item sequence (informational assumption) or energetically quantized distribution (physical assumption). In our previous review publications [1-2], scaling laws in economics (including urbanistic) and quantitative linguistics (natural and artificial languages) were observed. This review discloses the analysis of the most significant application of scaling laws in communications and natural sciences.

1. Informational vs physical approaches of item sets

In 1949 George Zipf found out word frequency dependence on word rank [3], which, in its simplest case, represents hyperbolic function. Zipfian distribution relates frequency $f(r)$ of item occurrence in finite corpus to item rank $r(w)$ according to Eq. (1).

$$f(r) = \frac{\alpha}{r^{\gamma}} \quad (1)$$

Benoit Mandelbrot [4] proposed the generalized expression of Zipf distribution as a discrete probability distribution:

$$f(r) = \frac{\alpha}{(1+\beta r)^{\gamma}} \quad (2)$$

Adjustable parameters α, β, γ are different for different languages. For finite English corpus, $\alpha \approx 0.1$, $\gamma \approx 1$.

[a] Corresponding author, email: alytis.gruodis@gmail.com

Table 1 represents Zipf and related scaling laws in different applications. There are many fields in informatics as well as in natural sciences where power law distributions are significant in modelling the complex system. By observing the statistical distribution in social studies, Zipf law originally describes the relationship between size and rank of discrete item.

Zipf law represents a variant of Pareto law (1927) known as the 80/20 rule (roughly 80% of the effects come from 20% of the causes, see Ref. [1] for details). Many ranked distributions could be fitted using power law with exception at the tail (long tail problem). Naumis et al [5] investigated this problem to create a general mechanism. Using the concept of macro states, the discrete probability distributions could be treated as complex distributions containing stretched exponential like frequency-rank functions. Tomita et al [6] investigated the random multiplicative process (RMP) which can generate a power law distribution. They have been established that RMP distribution contains two profiles: at head a log-normal distribution and the tail a power law distribution.

Traditionally, power law functions are applicable for growing systems. Ree [7] described the model containing N elements. Main rule of interactions could be formulated as follows: conservative two-body collisions are only allowed (sum of quantities is conserved). This model exhibits the scaling behaviour for some parameter ranges, when the growth of system is not expected. As Ree claims, presented power-law distribution model is useful for generation of scale-free networks when only rewiring is allowed.

Table 1. Zipf and related scaling laws in different applications.

Nr.	Name	Exponent	Ranked distribution	Source
1.	Zipf law	≈1	frequency of words (items)	[1,2]
2.	Zipf-Mandelbrott law	≈1	frequency of words (items)	[1,2]
3.	Heaps law	0.4÷0.6	vocabulary	[2]
4.	Herdan law	<1	signal systems	[2]
5.	Lotka law	2	frequency of publications by authors	[2]
6.	Benford's law	≈1	reference search in scientific databases	this
7.	Bradford law	≈1	search in science journals	this
8.	Piotrowski law	≈1	processes of language changes	[2]
9.	Gibrat law	≈1	grow of cities and firms	[1]
10.	Pareto law	≈1	relationship between effect and cause	[1]
11.	Robert law	≈1	executive compensation	[ ]
12.	Lognormal law	-	probability and statistics	[2]
13.	Weibul distribution	-	probability and statistics	[2]
14.	Manserath-Altmann law	-	influence of constituents	[2]
15.	Aren law	-	process modelling	[2]
16.	Yule law	≈1	probability and statistics	[2]
17.	Stefan-Boltzmann law	4	thermodynamic distribution	this
18.	Gutenberg-Richter law	≈1	magnitude and frequency of earthquakes	this
19.	Steven power law	≈1	distribution of magnitude of physical stimuli	-
20.	Barabasi-Albert law	≈1	network theory	-
21.	Kleiber law	0.75	metabolic parameter of animals on their mass	this
22.	Gompertz Makeham law	≈1	mortality	-

Bashkirov [8] described the possibility to use the equilibrium distributions of probabilities, which was derived from the maximum entropy principle (MEP) for the Renyi and Tsallis entropies. Renyi entropy $H_\alpha(X)$ of order α generalizes the Shannon entropy with estimating of randomness of a system [9]:

$$H_\alpha(X) = \frac{1}{1-\alpha}\Big[\sum_{i=1}^{n} p_i^\alpha\Big] \tag{3}$$

Limitations of order α could be described as follow: $\alpha \geq 0$; $\alpha \neq 1$.

Tsalis entropy [10] allows estimating the degree of disorders in the multifractals. Let us assume the existing of a discrete set of probabilities p_i with the condition

$$\sum_i p_i = 1 \tag{4}$$

and q (any real number) represents an entropic index. Tsallis entropy S_q is defined as

$$S_q(p_i) = \frac{k}{q-1}\Big[1 - \sum_{i=1}^{n} p_i^q\Big] \tag{5}$$

For $q \rightarrow 1$, Eq.(5) transforms to Bolzmann-Gibbs entropy expressed by Eq.(6):

$$S_{BG} = S(p) = -k\sum_{i=1}^{n}\Big[p_i \log p_i\Big] \tag{6}$$

Bashkirov [8] found that the maximum is realized for q within the range [0.25÷0.5] and the exponent is present in interval [1.3÷2].

Question about similarity of advanced aspects of linguistics and natural sciences has been quite old and unanswered until now. Firstly, mathematical approach allows to simulate the distribution by means of hyperbolic, exponential or power law functions to create dictionaries according to frequency principle. Secondly, statistical approach of items (philology, informatics) or events (physics, geology) allows to establish the sequence of cause-consequence with following law making. Due to such circumstances, item ranking distribution related to the certain natural language must be analysed from the statistical point of view.

2. Human communications

By choosing the model, Piantadosi [11] starts from realising the empirical phenomena in several fields including human via machine communication: random typing, organization of any semantic systems, optimization of certain communications, operating systems for computers. In many cases, several levels of organization must be formulated. Aitchison et al [12] provided with intuitive explanation of Zipf's law (explanation applies to a particular dataset).

For human communication, language must be treated as a macrosystem which consists of microsystems. Yoon Mi Oh [13] presented several linguistic microsystems as phonology, morphology, syntax, and semantics) and the mesosystemic interaction between these microsystems. Multiscale approach requires recognising the phenomenon of self-organization (through existence of scaling laws) by means of macro-, meso-, and microsystemic levels [14]. Cognitive aspects of the language learning can be defined in two following statements.

Firstly, Erman et al [15] provided corpus analysis to determine the communication type. They found out the occurrence of prefabricated sequences even though "open choice" among all available words from known dictionary were possible. This fact must be treated like some sort of optimisation of conversation through clusterization.

Secondly, in the time-domain scale, different fused item derivatives could be fixed. Historical changes in vocabulary could be estimated by analysing patterns of co-occurrence. Sentence of Joan Bybee, that "items that are used together fuse together", cited in Ref. [16], expresses the behaviour of a modern language to accumulate all forms originated in the fusion process.

3. Non-human communications

Ferrer-i-Cancho et al [17] investigated the dolphin whistle types to be used in specific behavioural contexts. To establish hypothesis about dolphin whistle as meaning (communicating approach), the presence of Zipf law in dolphin whistle types was found. Suzuki et al [18] investigated the whistle type communication of animals such as bottlenose dolphin - *Tursiops truncatus*. Authors postulate that Zipf-based technique is methodologically inappropriate for investigation the features of human language in the nonhuman communication of animals.

McCowan et al [19, 20] analysed the complexity of nonhuman animal communication systems in comparison with human language. Whistle vocalizations of bottlenose dolphin - *Tursiops truncatus* were classified through the first-order entropic relation in a Zipf-type diagram. Estimation of internal informational structure of animal vocal repertoires using slope of Shannon entropies allows comparing and predicting the organizational complexity of "speech" system across the diversity of species. In terms of behavioural ecology, McCowan discussed the main statements about predictions on the structure and organization of animal communication systems. Presence or changes in n-gram structure in a signalling data set [21] play the main role in noncoded natural communication systems. In that case, by fitting the item dependence using power low, Zipf slope (with exponent of $\approx$1) must be treated as the main condition for recognising the human languages.

4. Bibliometric classification (including impacting)

The impact factor represents a parameter as frequency with which the certain article has been cited in a certain year. Impact factor depends on scientific weight of the journal and allows estimating the importance or rank of a journal in time-dependent scale. In generally, the impact factor dependencies are related to the power-law.

Table 2. Distribution of leading digits $P(n)$

n	$P(n)$
1	0.301
2	0.176
3	0.125
…	…
8	0.051
9	0.046

Bradford's law (or Bradford's law of scattering) describes the distribution of objects (items, text concepts, words, etc) in the finite opus [22]. In 1934 Bradford described the grouping of journals into three groups when the number of journals in each group (C_1, C_2, C_3) is proportional to powered numbers:

$$C_1 : C_2 : C_3 = 1 : n : n^2 \quad (7)$$

$$C_1 : C_2 : C_3 = n^0 : n^1 : n^2 \quad (8)$$

Numbers 1:n:n^2 are called Bradford numbers. Due to clasterization of such type, Bradford expressed the idea of "core" group. Nowadays, different researchers in different fields have different numbers of core journals, and different Bradford multipliers.

Benford's law [23] (or first digit law) expresses the frequency distribution $P(n)$ of leading digits on n. For in set of numerical data $n\in${1,2,3,4,5,6,7,8,9}, the distribution is presented in Table 2.

$$P(n) = \log_{10}\left[\frac{n+1}{n}\right] = \log_{10}\left[1+\frac{1}{n}\right] \quad (9)$$

$$P(n) = \log_{10}[n+1] - \log_{10}[n] \quad (10)$$

Benford law satisfies many functions - especially functions constructed as the result of mathematical production of numbers: e.g. quantity multiplied by price, transaction data of sales, etc.

In 2005, Jorge E. Hirsch [24] presented a tool for estimating of productivity and relative quality in are of theoretical physics. So called Hirsch-index (or H-index) allows measuring the citation impact of the publications of a scientist. As an example, let us take a scientist after 20 years of academic carrier. After having published Z papers each of which has been cited in other papers at least C times, carrier of mentioned academic will be estimated with H-index:

$$H_1 = z_1 = c_1 \quad (11)$$

in case if z_1=c_1. Fig. 1 represents the situation when H_1=12. While comparing the scientific activities of scientists working in the same field, this routine works properly.

Egghe [25] presented critical analysis of experimental data of Mansilla [26]. Mansila found out that rank distribution of the logarithm of these impact factors has a typical S-shape: first a convex decrease, followed by a concave decrease. Egghe tied to explain the S-shape using additional perturbations. These distributions are valid for any type of im-

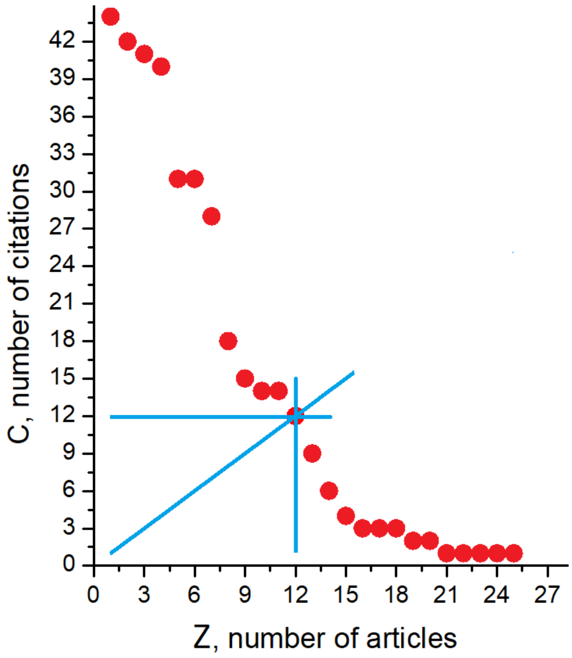

Fig.1. Distribution of article citation frequency C on ranked number of published articles Z: data of certain scientist. H_1=12 (in case if $z_1=c_1$=12.

pact factor (any publication period and any citation period). This means that some psychological factors are significant in selecting the citation behaviour. Egghe [27] also investigated the power law transformations when Lotkaian system is transformed into another Lotkaian system, described by a new Lotka exponent.

De Solla Price [28] critically reinvestigated the empirical results of citation frequency analysis using appropriate underlying probabilistic theory for the Bradford Law, the Lotka Law, the Pareto and Zipf Distributions. The advantages of Beta Function were presented and discussed.

Beirlant et al [29] criticized the usual routines for statistical impacting. They presented the rescaling of journal impact factors on a macro level which requires to categorize the objects into a standardized distribution (in quite different areas of research). Also, Beirant presented an alternative way to estimate the impacting in opposite to the Hirsch index [24]. In many cases, citation distributions could be fitted by Lotkaian-Zipf-Pareto behaviour. Extreme value index can be interpreted as the slope in a Pareto-Zipf quantile plot. Beirlant index, in contrast to the Hirsch index, is not influenced by the number of publications but stresses the decline of the statistical tail of citation counts. It appears to be much less sensitive to the science field than the Hirsch index.

Wallace et al [30] investigated the distribution on not cited papers or of highly cited papers, with respect to the bulk of publications. Over the 1900÷2006 period, data base contains $25 \cdot 10^6$ papers and $600 \cdot 10^6$ references from the Web of Science.

Wallace established several statements that the proportion of cited papers is a function of: i) the number of articles available (the competing papers); ii) the number of citing papers; iii) the number of references they contain. Wallace demonstrates the citation distributions over the 20th century in form of the stretched-exponential function and a form of the Tsallis q-exponential function. Both approaches are empirically differently from power-law fits.

Silagadze [31] described a curious observation with the rank statistics. Ranked scientific citation numbers follows Zipf-Mandelbrot's law in the same manner as some simple random citation models. This means that peculiar characters of the complex process are absent - due to the stochastic nature.

5. Website distribution

Adamic et al [32] analysed the web caching strategies, which are formulated my means of Zipf distribution in the number of requests for web pages. Krashakov et al [33] investigated the logs collected several web caches in Russian academic networks. Several statistical parameters such as duration of data collection, geographical location of the cache server collecting data, and the year of data collection - were analysed. Two-parameter function of the Zipf law type was used for popularity estimation of web site. Krashakov concluded that popularity must be treated as the universal property of Internet. Cattuto et al [34] analysed the social annotation system using the concepts borrowed from statistical physics, such as random walks (RWs), and complex networks theory. Lambiotte et al [35] analysed the word statistics as word occurrences and of the waiting times between such occurrences in Blogs. Consideration of two limiting cases such as the dilute limit the dense limit for frequent word) allows allow to estimate the distribution of waiting times as stretched exponential function.

Levene et al [36] tried to explain the empirically discovered power-law distributions (with exponent >2) for Web evolution. They proposed the extended evolutionary model of the Web to analyse the stochastic processes: a) the distribution of incoming links; b) the distribution of outgoing links; c) the distribution of pages in a Web site; and d) the distribution of visitors to a Web site. This work represents the formal proof of the convergence of the standard stochastic model.

Ebel et al [37] studied the topology of e-mail networks with e-mail addresses. A scale-free link distribution represents the resulting network like small-world in other social networks. Due to established circumstances of random architectures, it is possible to estimate the spreading of e-mail viruses.

6. Social activity

Cox et al [38] examine the commercial success in the popular recorded-music industry, as measured by gold-record output. For the number of artists with one gold record, Lotka's law distribution is not valid due to overestimation. However, for all measures of "successful" records, the fit is excellent. Beltran del Rio et al [39] analysed the power law fit for more than 1800 musical compositions including classical, jazz and rock music. Distribution of notes for each octave and its relationship with the ranking of the notes was estimated as an object for future investigations.

Blasius et al [40] analysed extensive chess databases. They concluded that the frequencies of opening moves are distributed in accordance to the power law. Exponent depends on the game depth. Pooled distribution of all opening weights follows Zipf's law with universal exponent. Blasius found out that in the case of hierarchical fragmentation, the scaling is truly universal and independent from the generating mechanism.

Malacarne et al [41] analyse the goal distributions by goal-players and by games (in football championships). Goal distributions are well adjusted by the Zipf-Mandelbrot law.

7. Physics

In physics, there are many distributions of particles on energy parameters. One of best known **Stefan - Boltzmann law** describes the peculiarities of black-body radiation. Black-body radiant emittance j (as radiated energy E per unit time t per unit surface area A) is directly proportional to the fourth power of absolute temperature T:

$$j = \frac{E}{At} = \sigma{\cdot}T^4 \tag{12}$$

σ represents Stefan-Boltzmann constant. Stefan - Boltzmann law is derived theoretically from Planck distribution as the first assumption for the ideal black objects only. For real objects (grey bodies), grey-factor must be included. Stefan-Bolzmann law is valid approximately as the second assumption.

Kristiansen et al [42] studied a dynamical few-body system made up of micrometer-sized plastic spheres dispersed in a ferroliquid driven by external magnetic fields. Dynamics of a of microparticles was described by mathematical braid theory. Rank-ordering statistics shows a large power-law region in consistency with the Zipf-Mandelbrot relation.

Hernando et al [43] investigated model system through systems' thermal properties. Considering the non-interacting scenario, scale-free ideal gases (SFIGs) model is successful for modelling of distribution such disparate systems as electoral results, city populations and total citations in Physics journals, that seem to indicate that SFIGs do exist. Zipf's law can be understood in a thermodynamic context as the surface of a finite system. Kaniadakis [44] investigated the relationships between the Boltzmann-Shannon entropy and the Maxwell-Bolzmann distribution. He pointed out that Maxwell-Boltzmann distribution (expressed in terms of the exponential function) is obtained by maximizing the Boltzmann-Shannon entropy under proper constraints.

8. Geology

Gutenberg - Richter law, GR law expresses most important relationship in seismology. Let us assume the existence of any seismic active region. Total number of earthquakes N at finite time interval is related to the magnitude of earthquake M:

$$N = 10^{\alpha-\beta * M} \tag{13}$$

$$\log_{10} N = \alpha - \beta * M \tag{14}$$

Values of α and β may be different for different regions and can vary over time due to physical activity of Earth core. In seismic active region, $\beta{\approx}1$. α value represents the total seismicity rate N_{tot} typical for selected region. Derived from basic Eq.(13), number of events N characterized by given magnitude M is expressed by Eq.(15), where total number of events N_{tot} and probability of those events p takes place.

$$N = 10^{\alpha} 10^{\beta * M} = N_{tot} * p \tag{15}$$

$$N_{tot} = 10^{a} \tag{16}$$

$$p = 10^{\beta * M} \tag{17}$$

Crampin et al [45] have reviewed the recent status of seismology related to the applications of Gutenberg - Richter law. Data of earthquakes in Earth and Moon were used to recognize the type of events. Full implications of the GR linearity are not generally recognised. Authors claim that the underlying physics is non-linear and not purely elastic. Any new suggestions are necessary, for example, fluid-rock deformations arise the fluid-saturated micro-cracks. The observation of linear GR dependence in moonquakes suggests that residual fluids exist in depth of the Moon.

Bhattacharyya et al [46] derived common model for two different objects. Firstly, ideal gas-like market model was used for the distribution of money among the agents with random-saving propensities (Pareto law). Secondly, fractal-overlap model for earthquakes was used for distribution of overlaps (GR law). Authors conclude that the power law appears as the asymptotic forms of ever-widening log-normal distributions in both cases.

Analysis of distribution of the calm times (time intervals between successive earthquakes) on arbitrary values of magnitude was done by Abe et al [47] using statistics of seismic time series data in California and Japan. Nature of the earthquake phenomenon was studied in the framework of the Zipf-Mandelbrot power law. Long tail of the distribution does not allow the statistical estimation of calm times.

Nagumo et al [48] investigated the lunar surface area being damaged by collisions of meteorites. The crater size-frequency distribution could be approximated by power law with one exception: at abscise interval about 4 km in diameter, the bending point is present. The above mentioned irregularity could be explained by effects of secondary craters (collision after the primary collision) and/or size-frequency distribution of the impactors.

Merriam et al [49] described the problem of success in resource assessment of mining and petroleum. Distributions of oil- and gas-field size in Kansas, the occurrence of historic earthquakes that affected the state, were estimated by means of Zipf law. Some limits of Zipf law are discussed too.

9. Geography

Primo et al [50] investigated the application of Zipf law in climatology. Usage of the exponential functions derived from the resulting scaling laws allows characterizing the rainfall temporal aggregation patterns. Authors proposed an original model related to the coding of precipitation as a discrete variable with four states. According to the item conception in linguistics, each weekly symbolic sequence of observed precipitation was encoded as "word" or set of items, and each local station defines a "own language" characterized by the observed "words" in the representative period. For "language" characterization, exponents were derived from the Zipf law. Different scaling behaviours for different subclimates (for humid tropical climates and polar climates) were received.

Holmes et al [51] investigated the application of the generalized Pareto distribution (GPD) for the statistical analysis of extreme wind speeds. The GPD is closely related to the generalized extreme value distribution (GEVD) and can be used to determine the appropriate value of shape factor.

Mitchell et al [52] analysed the seascape topology. High-resolution vertical profiles were estimated through Zipf analysis. Zipf exponents ranged within the interval [0.043÷0.83].

10. Chemistry

Benz et al [53] described the power-law distributions in chemoinformatics. The classification objects are rigid segments and ring systems, the distributions of molecular paths and circular substructures, and the sizes of molecular similarity clusters. The characteristic exponents of the power-laws lie in the interval [1.5÷3]. Several unique features also follow Heaps laws. Prediction of the growth of available data in large chemical databases could be estimated using Heaps law. Optimal allocation of experimental or computational resources follows the 80/20 - Pareto rule (see Ref. [1]).

Furusawa et al [54] studied the evolutionary origin of general statistics in a biochemical reaction network. They used cell models with catalytic reaction networks. By analysing the power-law distribution of reaction links and the power-law distribution of chemical abundance, is was concluded that inhomogeneity in chemical abundance is related to the higher growth rates of cells.

11. Biological systems

Mora et al [55] claim that all biological systems are working at a special thermodynamical state, the critical point. Due to that, explanation for Zipf's law is related to the distribution over states of an equilibrium system (Bolzmann statistics).

Kleiber's law [56] claims that metabolic rate of animals q is proportional to the γ power of animal mass M. For animals, γ=0.75, for plants, $\gamma\approx$1.

$$q \approx M^{\gamma} \quad (18)$$

Seuront et al [57] used Zipf and Pareto methods as data analysis methods to classify space-time patterns in marine ecology for identifying and classifying the certain structure in their data sets. Analysis of characteristic shapes according to Zipf's law allows recognising the specific components in mixing processes involving non-interacting and interacting species, for example, phytoplankton growth processes.

Alvarez-Ramirez et al [58] investigated the dependencies related to the world track records (WTR). They concluded that Zipf-Mandelbrot scaling law is useful for accurate fitting. During the race, the human energy release dynamics is limited by metabolic energy balances through several hierarchical relaxation processes. This is the main reason for existing of power-law behaviour of WTR.

Fontanelli et al [59] presented novel probability distribution function containing two fitting parameters a and b and normalization factor C. Authors suggested that parameter a is associated with the behaviour which leads to the power law, whereas b is associated with the fluctuation in noise. Distribution parameter X represents the ranked set, when rank r and N - maximum rank.

$$X = C * \frac{(N+1-r)^b}{r^a} \quad (19)$$

Fontanelli claims that Eq.(19) (so called LaValette distribution) is a good alternative for non-zipfian distributions - when usage of Zipf's law is inaccurate.

Conclusions

1. Power function dependence in Zipf law realization allows us to conclude that popular regularities in natural science (physics, chemistry, geology, biology) can have the common stochastic origin.
2. In communications, power law represents influence of human behaviour where language as a communication tool can be used. Zipfian and lotkaian dependencies are useful for information seach and clastering purposes.

References

1. Artūras Einikis, Giedrė Būdienė, Alytis Gruodis. Zipf and Related Scaling Laws. 1. Literature Overview of Applications in Economics. – *Innovative Infotechnologies for Science, Business and Education* ISSN 2029-1035 - 2(11) 2011 - Pp. 27-35.

2. Giedrė Būdienė, Alytis Gruodis. Zipf and Related Scaling Laws. 2. Literature Overview of Applications in Linguistics. – *Innovative Infotechnologies for Science, Business and Education* ISSN 2029-1035 - 1(12) 2012 - Pp. 17-26.
3. G. K. Zipf. Human Behavior and the Principle of Least Effort. – Addison-Wesley, 1965.
4. Mandelbrot Benoit. Information Theory and Psycholinguistics. – In: R. C. Oldfield and J.C. Marchall. Language. – Penguin Books, 1968.
5. G G Naumis, G Cocho. The tails of rank-size distributions due to multiplicative processes: from power laws to stretched exponentials and beta-like functions. – *New Journal of Physics* 9 (2007) 286.
6. Shinji Tomita, Yukio Hayashi. A controllable model of a random multiplicative process for the entire distribution of population. – *Physica A* 387 (2008) 1345-1351.
7. Suhan Ree. Power-law distributions from additive preferential redistributions. – *Physical Review E* 73 (2006) 026115 .
8. A. G. Bashkirov. On maximum entropy principle, superstatistics, power-law distribution and Renyi parameter. – *Physica A* 340 (2004) 153-162.
9. Alfred Renyi. On measures of information and entropy. – Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, 1960. – 547-561.
10. Constantino Tsallis. Possible generalization of Boltzmann-Gibbs statistics. – *Journal of Statistical Physics* 52(1-2) (1988) 479-487.
11. S. T. Piantadosi. Zipf's word frequency law in natural language: A critical review and future directions. – *Psychonomic Bulletin & Review* 21 (2014) 1112-1130.
12. Laurence Aitchison, Nicola Corradi, Peter E. Latham. Zipf's Law Arises Naturally When There Are Underlying, Unobserved Variables. – *PLOS Computational Biology* (2016) 1-32.
13. Yoon Mi OH. Linguistic Complexity and Information: Quantitative Approaches. – PhD thesis. University of Lyon, 2015.
14. Clay Beckner, Nick C. Ellis, Richard Blythe, John Holland, Joan Bybee, Jinyun Ke, Morten H. Christiansen, Diane Larsen-Freeman, William Croft, Tom Schoenemann. Language Is a Complex Adaptive System: Position Paper. – *Language Learning* 59(1) (2009) 1-26.
15. Erman B., Warren B. The idiom principle and the open choice principle. – *Text* 20(1) (2000) 29-62.
16. Kretzschmar W. A. Complex systems in aggregated variation analyses. – In: Szmrecsanyi, B. and Wälchli, B. (eds.) Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech. ISBN 9783110317558. – Series: Language and litterae (28). De Gruyter: Berlin, 2014. – 150-173.
17. Ramon Ferrer-i-Cancho, Brenda McCowan. A Law of Word Meaning in Dolphin Whistle Types. – *Entropy* 11 (2009) 688-701.
18. Ryuji Suzuki, John R. Buck, Peter L. Tyack. The use of Zipf's law in animal communication analysis. – *Animal Behaviour* 69 (2005) F9-F17.
19. Brenda McCowan, Sean F. Hanser, Laurance R. Doyle. Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires. – *Animal Behaviour*57 (1999) 409-419.
20. Brenda McCowan, Laurance R. Doyle, Sean F. Hanser. Using Information Theory to Assess the Diversity, Complexity, and Development of Communicative Repertoires. – *Journal of Comparative Psychology* 116(2) (2002) 166-172.
21. B. McCowan, L. R. Doyle, J. M. Jenkins, S. F. Hanser. The appropriate use of Zipf's law in animal communication studies. – *Animal Behaviour* 69 (2005) F1-F7.
22. Ruben Urbizagastegui Alvarado. El crecimiento de la literatura sobre la ley de Bradford. – *Investigacion Bibliotecologica: Archivonomia, Bibliotecologia e Informacion* ISSN:0187-358X, 30(68) (2016) 51-72.
23. Mark Nigrini, Joseph T. Wells. Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. – New Jersey: John Wiley & Sons Inc., 2012.
24. J. E. Hirsch. An index to quantify an individual's scientific research output. – *Proc. Nat. Acad. Sci.* 102(46) (2005) 16569-16272.
25. L. Egghe. Mathematical derivation of the impact factor distribution. – *Journal of Informetrics* 3 (2009) 290-295.
26. Mansilla R., Köppen E., Cocho G., Miramontes P. On the behavior of journal impact factor rank-order distribution. – *Journal of Informetrics* 1(2) (2007) 155-160.
27. L. Egghe. General evolutionary theory of information production processes and applications to the evolution of networks. – *Journal of Informetrics* 1 (2007) 115-122.
28. Derek de Solla Price. A General Theory of Bibliometric and Other Cumulative Advantage Processes. – *Journal of the American Society for Information Science* 27 (5-6) (1976) 292-306.
29. Jan Beirlant, Wolfgang Glanzel, An Carbonez, Herlinde Leemans. Scoring research output using statistical quantile plotting. – *Journal of Informetrics* 1 (2007) 185-192.
30. Matthew L.Wallace, Vincent Larivière, Yves Gingras. Modeling a century of citation distributions. – *Journal of Informetrics* 3 (2009) 296-303.
31. Z. K. Silagadze. Citations and the Zipf-Mandelbrot's law. – <arXiv:physics/9901035v2 [physics.soc-ph]>, accessed 2014 02 01.
32. Lada A. Adamic, Bernardo A. Huberman. Zipf's law and the Internet. – *Glottometrics* 3 (2002) 143-150.
33. Serge A. Krashakov, Anton B. Teslyuk, Lev N. Shchur. On the universality of rank distributions of website popularity. – *Computer Networks* 50 (2006) 1769-1780.
34. Ciro Cattuto, Alain Barrat, Andrea Baldassarri, Gregory Schehr, and Vittorio Loreto. Collective dynamics of social annotation. – *Proc. Nat. Acad. Sc.* 106(26) (2009) 10511-10515.
35. R. Lambiotte, M. Ausloos, M. Thelwall. Word statistics in Blogs and RSS feeds: Towards empirical universal evidence. – *Journal of Informetrics* 1 (2007) 277-286.

36. Mark Levene, Trevor Fenner, George Loizou, Richard Wheeldon. A stochastic model for the evolution of the Web. – *Computer Networks* 39 (2002) 277-287.
37. Holger Ebel, Lutz-Ingo Mielsch, and Stefan Bornholdt. Scale-free topology of e-mail networks. – *Physical Review E* 66(2002) 035103.
38. Raymond A. K. Cox, James M. Felton, Kee H. Chung. The Concentration of Commercial Success in Popular Music: An Analysis of the Distribution of Gold Records. – *Journal of Cultural Economics* 19(1995) 333-340.
39. M. Beltran del Rio, G. Cocho, G.G. Naumis. Universality in the tail of musical note rank distribution. – *Physica A* 387 (2008) 5552-5560.
40. Bernd Blasius and Ralf Tonjes. Zipf's Law in the Popularity Distribution of Chess Openings. – *Physical Review Letters* 103 (2009) 218701.
41. L.C. Malacarne, R.S. Mendes. Regularities in football goal distributions. – *Physica A* 286 (2000) 391-395.
42. Kai de Lange Kristiansen, Geir Helgesen, Arne T. Skjeltorp. Experimental observation of Zipf-Mandelbrot relation. – *Physica A* 335 (2004) 413-420.
43. A. Hernando, C. Vesperinas, A. Plastino. Fisher information and the thermodynamics of scale-invariant systems. – *Physica A* 389 (2010) 490-498.
44. G. Kaniadakis. Maximum entropy principle and power-law tailed distributions. – *Eur. Phys. J. B* 70 (2009) 3-13.
45. Stuart Crampin, Yuan Gao. The Physics Underlying Gutenberg-Richter in the Earth and in the Moon. – *Journal of Earth Science* 26(1) (2015) 134-139.
46. Pratip Bhattacharyya, Arnab Chatterjee, Bikas K. Chakrabarti. A common mode of origin of power laws in models of market and earthquake. – *Physica A* 381 (2007) 377-382.
47. Sumiyoshi Abe, Norikazu Suzuki. Scale-free statistics of time interval between successive earthquakes. – *Physica A* 350 (2005) 588-596.
48. Koichiro Nagumo and Akiko M. Nakamura. Reconsideration of crater size-frequency distributionon the moon: effect of projectile population and secondary craters. – *Adv. Space Res.* 28(8) (2001) 1181-1186.
49. Daniel F. Merriam, Lawrence J. Drew, and John H. Schuenemeyer. Zipf's Law: A Viable Geological Paradigm?. – *Natural Resources Research* 13(4) (2004) 265-271.
50. C. Primo, A. Galvan, C. Sordo, J.M. Gutierrez. Statistical linguistic characterization of variability in observed and synthetic daily precipitation series. – *Physica A* 374 (2007) 389-402.
51. J. D. Holmes, W. W. Moriarty. Application of the generalized Pareto distribution to extreme value analysis in wind engineering. – *Journal of Wind Engineering and Industrial Aerodynamics* 83 (1999) 1-10.
52. James G. Mitchell, Laurent Seuront. Towards a seascape topology II: Zipf analysis of one-dimensional patterns. – *Journal of Marine Systems* 69 (2008) 328-338.
53. Ryan W. Benz, S. Joshua Swamidass, and Pierre Baldi. Discovery of Power-Laws in Chemical Space. – *J. Chem. Inf. Model.* 48 (2008) 1138-1151.
54. Chikara Furusawa, Kunihiko Kaneko. Evolutionary origin of power-laws in a biochemical reaction network: Embedding the distribution of abundance into topology. – *Physical Review E* 73 (2006) 011912.
55. Thierry Mora, William Bialek. Are Biological Systems Poised at Criticality? – *J. Stat. Phys.* 144 (2011) 268-302.
56. Max Kleiber. Body size and metabolic rate. – *Physiol. Rev.* 27(4) (1947) 511-41.
57. Laurent Seuront, James G. Mitchell. Towards a seascape typology. I. Zipf versus Pareto laws. – *Journal of Marine Systems* 69 (2008) 310-327.
58. Jose Alvarez-Ramirez, Monica Meraz, Gustavo Gallegos. Zipf-Mandelbrot scaling law for world track records. – *Physica A* 328 (2003) 545-560.
59. Oscar Fontanelli, Pedro Miramontes, Yaning Yang, Germinal Cocho, Wentian Li. Beyond Zipf's Law: The Lavalette Rank Function and Its Properties. – *PLoS ONE* 11(9) (2016) 1-14.