



Critical Truths About Power Laws

Michael P. H. Stumpf and Mason A. Porter

Science **335**, 665 (2012);

DOI: 10.1126/science.1216142

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of June 26, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/335/6069/665.full.html>

This article **cites 16 articles**, 7 of which can be accessed free:

<http://www.sciencemag.org/content/335/6069/665.full.html#ref-list-1>

This article has been **cited by 7 articles** hosted by HighWire Press; see:

<http://www.sciencemag.org/content/335/6069/665.full.html#related-urls>

This article appears in the following **subject collections**:

Computers, Mathematics

http://www.sciencemag.org/cgi/collection/comp_math

during memory formation. This revealed an acute induction of KAEDA expression in DAL neurons but not in MBs during memory consolidation. Thus, protein synthesis occurs and is required in DAL neurons but is neither required nor detected in MBs. It is possible that low levels of protein synthesis in MBs escape the RICIN^{CS} effect and are below detection with KAEDA. However, Chen *et al.* further investigated this question by genetic manipulations of CREB function. They demonstrated that CREB-mediated gene transcription is required for LTM in DAL neurons but not MBs as previously thought (11, 12).

The role for CREB-dependent gene expression in DAL neurons but not MBs is at first hard to understand for two reasons. Expression of the adenylyl cyclase *rutabaga* in MBs is sufficient to support both STM (13, 14) and LTM (5, 15). Thus, CREB function in DAL neurons is not downstream in a signaling sense from *rutabaga* action in MB. There is much evidence that olfactory stimuli

are represented as a pattern of activity within MBs and that the associative memory forms there. In principle, olfactory memory could be initiated in one place and then transferred to another during consolidation. But it seems implausible that a few DAL neurons can represent the olfactory percept and also “store” LTM. Instead, a view emerges in which the consolidated memory is distributed within a neural circuit that includes MBs, DAL neurons, and ellipsoid neurons (6). Indeed, Chen *et al.* report that DAL axons likely are presynaptic to a subset of MB neurons called pioneer α/β neurons. Although this suggests that MBs are postsynaptic to DAL neurons, the role of gene transcription and protein synthesis within DAL neurons that is driven by coincidence detection in MB neurons also suggests that MBs are upstream of DAL neurons (presynaptic). Taken together, the simplest interpretation is a MB-DAL-MB feedback loop, perhaps including ellipsoid body neurons (6). The next task will be to integrate

the established biochemical and emerging systems views of consolidation.

References

1. A. C. Keene, S. Waddell, *Nat. Rev. Neurosci.* **8**, 341 (2007).
2. S. E. McGuire, M. Deshazer, R. L. Davis, *Prog. Neurobiol.* **76**, 328 (2005).
3. S. H. Wang, R. G. Morris, *Annu. Rev. Psychol.* **61**, 49, (C1) (2010).
4. P. K. Dash, A. E. Hebert, J. D. Runyan, *Brain Res. Brain Res. Rev.* **45**, 30 (2004).
5. A. L. Blum, W. Li, M. Cressy, J. Dubnau, *Curr. Biol.* **19**, 1341 (2009).
6. C. L. Wu *et al.*, *Nat. Neurosci.* **10**, 1578 (2007).
7. M. J. Krashes, A. C. Keene, B. Leung, J. D. Armstrong, S. Waddell, *Neuron* **53**, 103 (2007).
8. C.-C. Chen *et al.*, *Science* **335**, 678 (2012).
9. J. C. Yin, T. Tully, *Curr. Opin. Neurobiol.* **6**, 264 (1996).
10. A. J. Silva, J. H. Kogan, P. W. Frankland, S. Kida, *Annu. Rev. Neurosci.* **21**, 127 (1998).
11. D. Yu, D. B. Akalal, R. L. Davis, *Neuron* **52**, 845 (2006).
12. D. B. Akalal *et al.*, *J. Neurosci.* **30**, 16699 (2010).
13. T. Zars, M. Fischer, R. Schulz, M. Heisenberg, *Science* **288**, 672 (2000).
14. S. E. McGuire, P. T. Le, A. J. Osborn, K. Matsumoto, R. L. Davis, *Science* **302**, 1765 (2003).
15. S. Trannoy, C. Redt-Clouet, J. M. Dura, T. Preat, *Curr. Biol.* **21**, 1647 (2011).

10.1126/science.1218171

MATHEMATICS

Critical Truths About Power Laws

Michael P. H. Stumpf¹ and Mason A. Porter²

The ability to summarize observations using explanatory and predictive theories is the greatest strength of modern science. A theoretical framework is perceived as particularly successful if it can explain very disparate facts. The observation that some apparently complex phenomena can exhibit startling similarities to dynamics generated with simple mathematical models (1) has led to empirical searches for fundamental laws by inspecting data for qualitative agreement with the behavior of such models. A striking feature that has attracted considerable attention is the apparent ubiquity of power-law relationships in empirical data. However, although power laws have been reported in areas ranging from finance and molecular biology to geophysics and the Internet, the data are typically insufficient and the mechanistic insights are almost always too limited for the identification of power-law behavior to be scientifically useful (see the figure). Indeed, even most statistically “successful”

calculations of power laws offer little more than anecdotal value.

By power-law behavior, one typically means that some physical quantity or probability distribution $y(x)$ satisfies (2, 3)

$$y(x) \propto x^{-\lambda} \text{ for } x > x_0,$$

where λ is called the “exponent” of the power law. In the equation, the power-law behavior occurs in the tail of the distribution (i.e., for $x > x_0$). A power-law distribution has a so-called “heavy tail,” so extreme events are far more likely than they would be in, for example, a Gaussian distribution. Examples of such relationships have been reported in a wide range of situations, including the Gutenberg-Richter law in seismology (4), allometric scaling in animals (5), the distribution of hyperlinks on the World Wide Web (6), the sometimes vehemently refuted (7) “scale-free” nature of the Internet (8), a purported unified theory of urban living (9), patterns of insurgent and terrorist activity (10), and (ironically) the paper publication rates of statistical physicists (11). A subtlety to note is that this list includes two different types of reported power laws: bivariate power laws like allometric scaling and

power-law probability distributions like the paper publication rates.

Power laws in statistical physics emerge naturally from microscopic theories and can be related to observable macroscopic phenomena. A good example is magnetization (3). The derivation of a power law suggests that—in a certain (“critical”) regime—phenomena do not possess a preferred scale in space, time, or something else: They are, in a sense, “scale free.” However, as Philip Anderson pointed out in 1972 (12), one must be cautious when claiming power-law behavior in finite systems, and it is not clear whether power laws are relevant or useful in so-called “complex systems” (13, 14). It is important to take a nuanced approach and consider not only whether or not one has or can derive a detailed mechanistic model of a system’s driving dynamics, but also the extent of statistical support for a reported power law. One additionally needs to consider empirical support, as theories for power-law behavior arise from infinite systems, and real systems are finite.

The power law reported for allometric scaling stands out as genuinely good (see the figure) (5): Not only is there a sound theory underlying why there should be a power-law

¹Centre for Integrative Systems Biology and Bioinformatics, Imperial College London, London SW7 2AZ, UK. ²Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute and CABDyN Complexity Centre, University of Oxford, Oxford OX1 3LB, UK. E-mail: m.stumpf@imperial.ac.uk; porterm@maths.ox.ac.uk

relationship between body size x and metabolic performance y , but this relationship has been supported empirically over many orders of magnitude (from bacteria to whales). The clear dependence of various biological characteristics on body size is, of course, insufficient by itself to infer a causal relationship, but few people would dispute the reality of such a relationship.

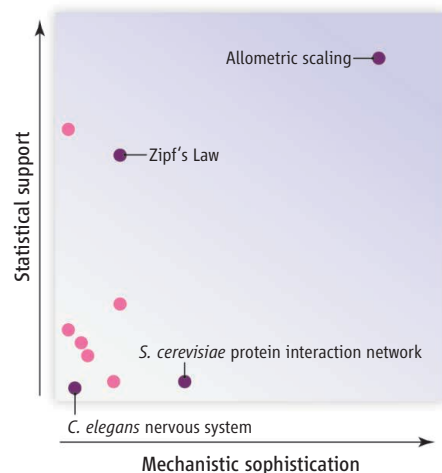
Purported power laws fall loosely into two categories: those with statistical support—by itself a nontrivial task (15)—and those without it. Numerous scholars have neglected to apply careful statistical tests to data that were reported to exhibit power-law relationships; so-called “scale-free” networks are perhaps the best known and most widely discussed examples (2, 6, 13). However, when formal statistical tools have been applied to network data, evidence favoring power-law relationships has almost always been negligible (7, 15, 16).

As a rule of thumb, a candidate power law should exhibit an approximately linear relationship on a log-log plot over at least two orders of magnitude in both the x and y axes. This criterion rules out many data sets, including just about all biological networks. Examination (15) of the statistical support for numerous reported power laws has revealed that the overwhelming majority of them failed statistical testing (sometimes rather epically). For example, a recent study found (17) that the number of interacting partners (i.e., the degree) of proteins in yeast is power-law distributed, but careful statistical analysis refutes this claim (18). Noise or incomplete data can further distort the picture (19). Trying to discern a power-law relationship by eyeballing straight lines (or even trying to find them using, for example, least-squares fitting) on log-log plots of data can be appealing, but the human ability to detect patterns from even the flimsiest of evidence might lead researchers to conclude the existence of a bona fide power law based on purely qualitative criteria.

Even if a reported power law surmounts the statistical hurdle, it often lacks a generative mechanism. Indeed, the same power law (that is, with the same value of λ) can arise from many different mechanisms (3). In the absence of a mechanism, purely empirical fitting does have the potential to be interesting, but one should simply report such results in a neutral fashion rather than provide unsubstantiated suggestions of universality. The fact that heavy-tailed distributions occur in complex systems is certainly important (because it implies that extreme events occur more frequently than would otherwise be the case), and statistically sound empiri-

cal fits of event data, when used with caution, can help in data interpretation (as it is certainly useful to estimate how often extreme events occur in a given system). However, a statistically sound power law is no evidence of universality without a concrete underlying theory to support it. Moreover, knowledge of whether or not a distribution is heavy-tailed is far more important than whether it can be fit using a power law.

Suppose that one generates a large number of independent random variables x_i drawn from heavy-tailed distributions, which need not be power laws. Then, by a version of the central limit theorem (CLT), the sum of these random variables is generically power-law distributed (20). Few people today would express amazement at finding that the CLT holds in a given context (when one adds up random variables drawn from distributions with finite moments), and the CLT is a vital tool in statistics, providing the basis for many rigorous scientific analyses. It also holds ubiquitously, including in situations in which random variables are drawn from heavy-tailed distributions; in such cases, however, power laws replace the Gaussian distribution as the limiting situation. One thus expects power laws to emerge naturally for rather unspecific reasons, simply as a by-product of mixing multiple (potentially rather disparate) heavy-tailed distributions. For example, it is possible to decompose a supposedly “power-law” degree distribution of a metabolic network into separate distributions of metabo-



How good is your power law? The chart reflects the level of statistical support—as measured in (16, 21)—and our opinion about the mechanistic sophistication underlying hypothetical generative models for various reported power laws. Some relationships are identified by name; the others reflect the general characteristics of a wide range of reported power laws. Allometric scaling stands out from the other power laws reported for complex systems.

lites of different types (16). The degree distribution for each of these metabolite classes is different, reflecting the different roles that they play in the organism.

Finally, and perhaps most importantly, even if the statistics of a purported power law have been done correctly, there is a theory that underlies its generative process, and there is ample and uncontroversial empirical support for it, a critical question remains: What genuinely new insights have been gained by having found a robust, mechanistically supported, and in-all-other-ways superb power law? We believe that such insights are very rare.

Power laws do have an interesting and possibly even important role to play, but one needs to be very cautious when interpreting them. The most productive use of power laws in the real world will therefore, we believe, come from recognizing their ubiquity (and perhaps exploiting them to simplify or even motivate subsequent analysis) rather than from imbuing them with a vague and mistakenly mystical sense of universality.

References and Notes

1. R. M. May, *Nature* **261**, 459 (1976).
2. M. E. J. Newman, *Am. J. Phys.* **79**, 800 (2011).
3. J. P. Sethna, *Entropy, Order Parameters, and Complexity* (Oxford Univ. Press, Oxford, 2010).
4. B. Gutenberg, C. F. Richter, *Seismicity of the Earth and Associated Phenomena* (Princeton Univ. Press, Princeton, NJ, ed. 2, 1954).
5. G. B. West, J. H. Brown, B. J. Enquist, *Science* **276**, 122 (1997).
6. A.-L. Barabási, R. Albert, *Science* **286**, 509 (1999).
7. W. Willinger, D. Alderson, J. C. Doyle, *Nat. Am. Math. Soc.* **56**, 586 (2009).
8. R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
9. L. Bettencourt, G. West, *Nature* **467**, 912 (2010).
10. N. Johnson *et al.*, *Science* **333**, 81 (2011).
11. N. Li *et al.*, <http://arxiv.org/abs/1108.0833> (2011).
12. P. W. Anderson, *Science* **177**, 393 (1972).
13. E. Fox, Keller, *Bioessays* **27**, 1060 (2005).
14. D. Avnir, O. Biham, D. Lidar, O. Malcai, *Science* **279**, 39 (1998).
15. A. Clauset, C. R. Shalizi, M. E. J. Newman, *SIAM Rev.* **51**, 661 (2009).
16. R. Tanaka, M. Csete, J. Doyle, *Syst. Biol.* **152**, 179 (2005).
17. H. Yu *et al.*, *Science* **322**, 104 (2008).
18. A. Clauset, Power laws in the mist (15 October 2008); Structure & Strangeness, www.cs.unm.edu/~aaron/blog/archives/2008/10/power_laws_in_t_1.htm.
19. M. P. Stumpf, C. Wiuf, R. M. May, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4221 (2005).
20. W. Willinger, D. Alderson, J. C. Doyle, L. Li, in *Proceedings of the 2004 Winter Simulation Conference*, R. G. Ingalls, M. D. Rossetti, J. S. Smith, B. A. Peters, Eds. (Institute for Operations Research and the Management Sciences, Hanover, MD, 2004), pp. 130–141; paper available at www.informs-sim.org/wsc04papers/016.pdf.
21. M. P. H. Stumpf, P. J. Ingram, *Europhys. Lett.* **71**, 152 (2005).
22. We thank J. Carlson, A. Clauset, and A. Lewis for useful discussions and Ch. Barnes, A. MacLean, and C. Wiuf for helpful comments on the manuscript.

10.1126/science.1216142