

Self-Similar Processes Follow a Power Law in Discrete Logarithmic Space

Mitchell G. Newberry*

Center for the Study of Complex Systems, University of Michigan, Ann Arbor, Michigan 48109-1042, USA

Van M. Savage

*Department of Biomathematics, University of California at Los Angeles, Los Angeles, California 90095, USA,
Department of Ecology and Evolutionary Biology, University of California at Los Angeles, Los Angeles, California 90095, USA,
and Santa Fe Institute, Santa Fe, New Mexico 87501, USA* (Received 27 September 2018; revised manuscript received 7 March 2019; published 19 April 2019)

Cities, wealth, and earthquakes follow continuous power-law probability distributions such as the Pareto distribution, which are canonically associated with scale-free behavior and self-similarity. However, many self-similar processes manifest as discrete steps that do not produce a continuous scale-free distribution. We construct a discrete power-law distribution that arises naturally from a simple model of hierarchical self-similar processes such as turbulence and vasculature, and we derive the maximum-likelihood estimate (MLE) for its exponent. Our distribution is self-similar, in contrast to previously studied discrete power laws such as the Zipf distribution. We show that the widely used MLE derived from the Pareto distribution leads to inaccurate estimates in systems that lack continuous scale invariance such as branching networks and data subject to logarithmic binning. We apply our MLE to data from bronchial tubes, blood vessels, and earthquakes to produce new estimates of scaling exponents and resolve contradictions among previous studies.

DOI: [10.1103/PhysRevLett.122.158303](https://doi.org/10.1103/PhysRevLett.122.158303)

For power-law distributions [1], the frequency of events or data points that exceed a value x is proportional to $x^{-\alpha}$. The value of the scaling exponent α determines the frequency of large events and is used to discriminate between models of stock price fluctuations [2], the timing and magnitude of earthquakes [3,4], internet topology [5], turbulence [6], tree-limb branching [7], and pulmonary and vascular morphology [8]. For instance, word frequencies are characterized by $\alpha = 1$, corresponding to Zipf's law. So the 10th most common word in English "it" appears 5 times more often than the 50th most common word "up." Likewise, the distribution of blood vessel radii approximately follows a power law with $\alpha = 3$ in a laminar flow regime and $\alpha = 2$ under pulsatile flow [9,10].

The classic power-law distribution—developed to study the distribution of wealth—is called the Pareto distribution [11] or continuous power-law distribution [12]. The probability density of measuring a continuous value x with minimum x_m for a Pareto distribution is $p_c(x) = \alpha x_m^\alpha x^{-(\alpha+1)}$. The probability an observation exceeds x —the tail distribution $\int_x^\infty p_c(x') dx'$ —is thus the power law [13]

$$P_c(x) = \left(\frac{x}{x_m}\right)^{-\alpha}. \quad (1)$$

Power laws and scale invariance arise from diverse phenomena [1] including self-similarity and criticality [14,15]. Many physical systems are scale invariant—rescaling by an arbitrary constant preserves relative proportions. For the Pareto distribution [16], for example, incomes of \$80 000 occur less frequently than \$50 000 in just the same proportion as incomes of \$800 000 and \$500 000 or \$160 000 and \$100 000.

However, many self-similar systems are scale invariant only in discrete steps. A blood vessel tends to branch into two smaller vessels, a fluid vortex into two or three smaller vortices, and the Sierpinski triangle is self-similar only by powers of two [8,17,18]. These systems preserve relative proportions upon rescaling from one step to the next, but not upon arbitrary rescaling. This property is termed discrete-scale invariance [15] or discrete renormalizability [19]. It is a weaker condition than the continuous scale invariance underlying the Pareto distribution. Whereas strict scale invariance implies a power law and vice versa, discrete-scale invariance allows log-periodic modulations in the frequencies of observations [15] that deviate from a pure power law such as Eq. (1). Such modulations are indeed observed in bronchial tube diameter, vortex enstrophy, and financial asset prices [18,20,21].

Discrete self-similar phenomena are often nonetheless fit to power laws. Power-law distributions have been studied

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

extensively [11,12,22], yet there is still difficulty detecting and interpreting them [12,23,24]. Well-motivated methods to estimate α by maximum likelihood have long been available [1,12,25,26], yet many empiricists still use linear regression on log-log plots [27] or visual comparison despite their known problems [12]. Canonically, power-law distributions are categorized as either continuous or discrete [11,12]. Continuous distributions with power-law behavior include the classical [11] Pareto distribution, several generalizations (types II–IV and the generalized Pareto [11,22]), and related distributions such as the log-normal [22]. Discrete power-law distributions [28] are commonly assumed to take integer values and have been called the Zipf, Riemann-zeta, or the discrete Pareto distribution [11,12,28], sometimes interchangeably [28].

However, none of these distributions cover the case of discrete self-similarity: The continuous power-law distributions cannot admit log-periodic fluctuations, whereas the discrete distributions over integers cannot admit rational values such as sizes in the Sierpinski triangle.

We remedy this by describing a discrete distribution that is explicitly constructed to be self-similar. This distribution is qualitatively different from previously described distributions because, rather than to integers, we assign probability to the set of values $x_m \lambda^k$ (for some $x_m, \lambda > 1$) that result from rescaling x_m by any k factors of λ (Fig. 1). Thus our distribution is both discrete and able to accommodate noninteger values. Cases of our distribution appear often, as in ideal models of vasculature [29] or as the degree distribution of Apollonian networks [30].

The qualitative difference means that our distribution corresponds to a different maximum-likelihood estimator (MLE) of α than Pareto or any other power-law distribution. Moreover, when strict scale invariance is violated, the Pareto MLE can produce inaccurate or unstable estimates (Fig. 2). This finding may explain empiricists' reluctance to

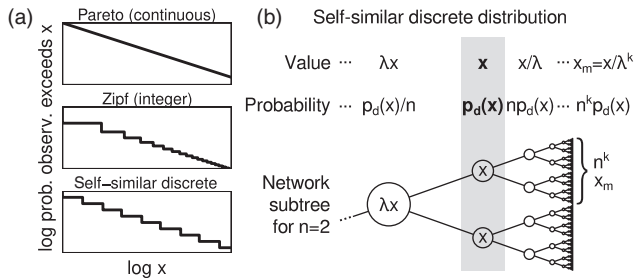


FIG. 1. (a) Tail distributions of power laws [cf. Eq. (1)]. Power laws and self-similarity appear on log-log plots as regular slopes and repeating patterns. The self-similar discrete distribution shows stair steps (allowed values) at regular intervals in logarithmic space, in contrast to Zipf's law, the classical discrete power law. (b) Our distribution draws node values $x > x_m$ uniformly from a self-similar branching network wherein each node with value x has one parent with value λx and n children of value x/λ .

adopt existing maximum-likelihood methods as instability in practice signals deeper problems. The MLE we derive for our distribution is simple and intuitive, and it has a closed-form solution in contrast to power-law distributions over the integers [12,31]. Our MLE limits to the Pareto MLE as $\lambda \rightarrow 1^+$. A general MLE for binned Pareto samples [32] can be specialized to recover our MLE for the case of logarithmic binning because logarithmic binning is self-similar, illustrating that discrete-scale invariance can arise from the physics of a system or as an artifact of measurement. Finally, our distribution provides a measurement and interpretation of α in continuous systems that are not strictly scale-free but have meaningful power laws evident in log-log plots.

These results enable us to resolve previous paradoxes. For instance, the Richter scale uses logarithmic discretization. Through a re-analysis of earthquake data, we identify and correct errors this discretization introduces in estimates of the Gutenberg-Richter b value (Fig. 3) from the Pareto MLE. Our MLE further reconciles data with theoretical predictions in bronchial diameter measurements where discrete self-similarity arises from branching and the Pareto MLE cannot be trusted.

Deriving the distribution.—We derive the self-similar discrete distribution in two ways: first intuitively from a model of self-similar branching, and then more abstractly from discrete-scale invariance plus one assumption that shows the connection to logarithmic binning.

We first consider a branching network in which each node (e.g., vessel) with value x (e.g., radius) has one parent node of value λx and n identical offspring nodes of value $\lambda^{-1}x$ ($\lambda > 1$) [Fig. 1(b)]. The network is self-similar

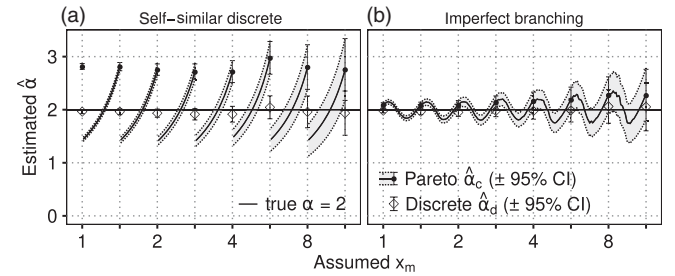


FIG. 2. Comparison of estimated $\hat{\alpha}$ and the true value ($\alpha = 2$) versus assumed x_m in synthetic data. The data are (a) 10 000 samples from our distribution (discrete) with $x_m = 1$, $n = 2$, $\lambda = \sqrt{2}$, and (b) 10 000 samples with replacement from a simulated imperfect branching network (continuous) where each child node introduces a random 2.5% proportional error, $\lambda \sim \sqrt{2} \text{Lognormal}[0, \log(1 + \epsilon)]$, independent, $\epsilon = 0.025$, starting from node value 100. The Pareto MLE $\hat{\alpha}_c$ [Eq. (5), dots and lines] over or under estimates α for many assumed values of x_m . Our MLE $\hat{\alpha}_d$ [Eq. (6), diamonds] CIs include the true α for all feasible assumed x_m . The CIs of the Pareto $\hat{\alpha}_c$ often exclude the true value, suggesting more data would not improve the estimate. These parameter choices represent the Da Vinci model for tree limbs.

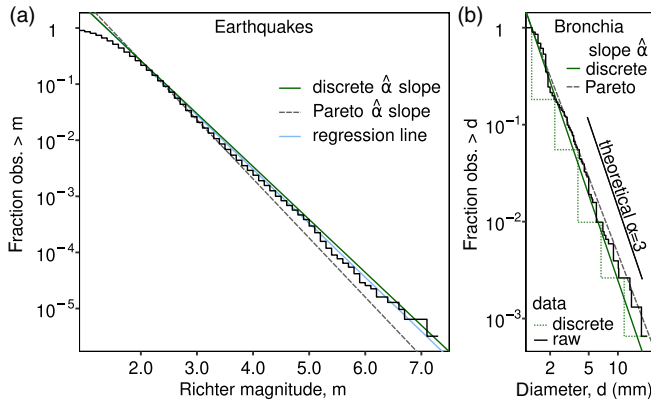


FIG. 3. Tail distributions of (a) earthquake magnitudes and (b) bronchial diameters with slopes derived from different estimates of α (Table I). On earthquake magnitudes (a) the Pareto MLE slope $-\hat{\alpha}_c$ (gray dashed line) is visibly incorrect due to failure to account for discrete data. Bronchial measurements (b) are continuous but deviate from continuous scale invariance. Thus, discretizing by powers of 2 (dotted green line) and using the discrete MLE $\hat{\alpha}_d$ produces a slope (green line) that is visually consistent with the data and consistent with theory (black bar, Table I) in contrast to the continuous $\hat{\alpha}_c$ (dashed gray line).

because an identical network is obtained by multiplying all values x by λ . This branching relationship is a toy model of earthquake aftershocks producing aftershocks [3], one blood vessel branching into two smaller vessels [29], or any case where as little fleas [33,34] and little whirls [35] both feed on larger versions, recursively the larger so proceed in like proportion [36]. If nodes represent tree limbs and x represents their diameter, then $n = 2$ and $\lambda = \sqrt{2}$ give the model—due to Leonardo Da Vinci [37]—that cross-sectional area is preserved when one limb divides into two.

The probability to sample x — $p_d(x)$ —derives from counting nodes in the network [Fig. 1(b)]: $p_d(x) = p_d(x_m)n^{-k}$ whenever $x = \lambda^k x_m$ for some integer $k \geq 0$ and minimum observable node value x_m . We can write k in terms of x as $k = \log_\lambda(x/x_m)$ so that n^{-k} becomes $(x/x_m)^{-1/\log_n \lambda}$. Labeling this exponent $-\alpha$ gives $n = \lambda^\alpha$ and $p_d(x) \propto (x/x_m)^{-\alpha}$ with $p_d(x_m)$ the constant of proportionality. We solve for $p_d(x_m)$ by noting that total probability must sum to 1. Hence,

$$p_d(x) = (1 - \lambda^{-\alpha}) \left(\frac{x}{x_m} \right)^{-\alpha}, \quad (2)$$

if $x = x_m \lambda^k$ for some integer $k \geq 0$ and otherwise $p_d(x) = 0$. The tail distribution [Fig. 1(a)] is a stair-step function that is equal to the Pareto tail distribution [Eq. (1)] at possible values of x .

An alternative derivation shows the simplicity of our distribution among all possible distributions with discrete-scale invariance. The property [15] of any discrete-scale

invariant function $f(x)$ that $f(\lambda x) = \lambda^{-\alpha} f(x)$ allows us to extrapolate a discrete-scale invariant distribution from any normalizable function defined on a short segment $[x_m, \lambda x_m)$. Choosing this segment to be a delta function at x_m —a point-mass that could represent some more complicated density distribution—gives $p_d(\lambda^k x_m) = p_d(x_m) \lambda^{-k\alpha}$ which, identifying $x = \lambda^k x_m$, gives $p_d(x) \propto (x/x_m)^{-\alpha}$ and implies Eq. (2) as above.

The generality of this latter derivation furthermore encompasses logarithmic binning of self-similar data. When aggregating observations into bins of exponentially increasing size, the bin boundaries are $x_m \lambda^k$ for some x_m and λ (e.g., for bins 0.1–1, 1–10, 10–100, ..., $x_m = 0.1$ and $\lambda = 10$). Therefore, Eq. (2) results whenever the underlying distribution is self-similar upon rescaling by λ . Indeed, a general MLE for binned Pareto samples [32] reduces to Eq. (2) whenever the binning scheme is logarithmic.

By extension, logarithmic binning reduces many unknown distributions to our known distribution by flattening within-scale modulations. We can thus interpret and measure α as the slope of repeating patterns on a log-log plot [e.g., the stair steps of Fig. 1(a)], including damped log-periodic oscillations [Fig. 2(b)]. Logarithmic binning also attenuates measurement errors that correspond to noise within a bin by counting only how many data points occur in each. For this reason our distribution provides better estimates from noisy data than Pareto.

Estimating α .—Here we derive a general form of the MLE for any power-law distribution to show that the form depends only on the support of the distribution, i.e., which values are possible. We then reduce this general form to MLEs for Pareto, Zipf, and our self-similar discrete distribution as special cases.

The likelihood \mathcal{L} of a set of N independent observations x_j is the product of their probabilities, $\mathcal{L} = \prod_{j=1}^N p(x_j)$. For any power law, $p(x)$ can be written in the form $p(x) = Cx^{-(\alpha+d)}$ where the normalizing constant C is required for the total probability to sum to 1 and d adjusts the dimension ($d = 1$ for probability density and $d = 0$ for probability mass). The normalizing constant C may depend on α and any other parameters such as x_m that determine the support but does not depend on x .

The log-likelihood

$$\log \mathcal{L} = N \log C - (\alpha + d) \sum_{j=1}^N \log x_j \quad (3)$$

is monotonic with the likelihood so it has the same maxima. We therefore maximize $\log \mathcal{L}$ with respect to α by setting $\partial \log \mathcal{L} / \partial \alpha$ equal to zero. This gives

$$\frac{1}{C} \frac{\partial C}{\partial \alpha} = \langle \log x_j \rangle, \quad (4)$$

where $\langle \log x_j \rangle$ denotes the average $(1/N) \sum_{j=1}^N \log x_j$.

Because $1/C$ is a sum over all values, C depends on the support and each power-law distribution has a unique MLE.

In the Pareto distribution, $C = [\int_{x_m}^{\infty} x^{-(\alpha+1)} dx]^{-1} = \alpha x_m^{\alpha}$. Substituting this C into Eq. (4) yields the classical MLE for Pareto samples [12,25]

$$\hat{\alpha}_c = \langle \log x_j - \log x_m \rangle^{-1}. \quad (5)$$

In Zipf's distribution, $C = [\sum_{k=x_m}^{\infty} k^{-\alpha}]^{-1} = 1/\zeta(\alpha, x_m)$, where $\zeta(\alpha, x_m)$ is the Hurwitz zeta function. Hence we recover the condition for maximum likelihood $\zeta'(\hat{\alpha}_z)/\zeta(\hat{\alpha}_z) = -\langle \log x_j \rangle$ that has been used to solve numerically for $\hat{\alpha}_z$ [12,31].

In our case, $C = [\sum_{k=0}^{\infty} (x_m \lambda^k)^{-\alpha}]^{-1} = x_m^{\alpha} (1 - \lambda^{-\alpha})$ and so Eq. (4) gives

$$\hat{\alpha}_d = \log_{\lambda} [1 + \langle \log_{\lambda} x_i - \log_{\lambda} x_m \rangle^{-1}]. \quad (6)$$

Our MLE [Eq. (6)] resembles the Pareto MLE [Eq. (5)] but with logarithms to the base λ . Indeed, our MLE recovers the Pareto MLE in the limit that λ goes to 1 from above.

Notably, as Fig. 2 shows, the continuous and discrete estimators $\hat{\alpha}_c$ and $\hat{\alpha}_d$ [Eqs. (5) and (6)] may produce drastically different estimates given the same data. Furthermore [Fig. 2(b)], deviations from a pure power law cause the Pareto estimator to become unstable—depending sensitively on the choice of x_m which otherwise should not matter [12]. Yet researchers to date have often used the Pareto MLE on any data containing noninteger values simply because Pareto has been the only recognized power law to allow nonintegers at all.

Our MLE also applies to discretized continuous data. For example, real branching networks produce error at each step and converge to a continuous power law after many branches. We simulate error [Fig. 2(b)] by recursively constructing child nodes, assigning each the value $x\lambda^{-1}\eta$ given its parent value x . The error factor η is drawn independently for each node from a log-normal distribution with geometric mean 1 and geometric standard deviation $1 + \epsilon$ so that η lies between $(1 + \epsilon)^2$ and $(1 + \epsilon)^{-2}$ 95% of the time. The distribution of node values is then neither discrete nor a continuous power law. The Pareto MLE oscillates as a function of x_m as we might expect from

Fig. 2(a). After logarithmic binning, however, our distribution is an accurate enough approximation that our MLE recovers the $\alpha = 2$ corresponding to the geometric mean λ .

Failures of the Pareto MLE.—The Pareto distribution is the continuous analog of the discrete power law and indeed, $\lim_{\lambda \rightarrow 1^+} \hat{\alpha}_d = \hat{\alpha}_c$. Consequently, one might expect Pareto to be a continuum-limit approximation to our distribution. However, $\hat{\alpha}_d \approx \hat{\alpha}_c$ only when λ is very close to 1. Common values of λ —such as Da Vinci's $\sqrt{2}$ or logarithmic binning by powers of 10—do not approach this limit.

An earthquake example shows that even $10^{0.1}$ is not sufficiently close to $\lambda = 1$ to well approximate the Pareto. Earthquake magnitudes have long been thought to follow a power-law distribution. Its exponent—the Gutenberg-Richter b value—was first measured in 1944 as $b = 0.88 \pm 0.03$ by linear regression on scant data [38]. A modern study [3] on an extensive dataset also used regression to conclude that $b = 0.95 \pm 0.01$. In contrast, the Pareto MLE produces $\hat{\alpha}_c = 1.054 \pm 0.007$ using the same dataset [3], in visible disagreement with the slope of the tail distribution [Fig. 3(a)]. Clauset *et al.* [12] also concluded that earthquake magnitudes were not well fit by the continuous power law, paradoxical to long-held belief.

Our findings resolve these contradictory results for earthquake data. Magnitudes are recorded with two digits of precision on the Richter scale. Thus, the data are logarithmically binned from the start with $\lambda = 10^{0.1}$. Using our MLE with $x_m = 2.0$ and $\lambda = 10^{0.1}$, we estimate $b = \hat{\alpha}_d = 0.943 \pm 0.006$, visually consistent with the data and in line with previous measures based on regression [Fig. 3(a), Table I]. Although 0.1 is relatively fine-grained over the range of data (from 2.0 to 7.3), the value $10^{0.1} = 1.26$ is not sufficiently close to $\lambda = 1$ for the Pareto approximation to apply. Thus the continuous estimator yields biased estimates. Furthermore, the $\approx 11.7\%$ error the Pareto MLE introduces does not diminish with more observations or a greater range of observations. This error also far exceeds the statistical random error 0.7% (95% CI) in this high-quality dataset, producing confidence intervals that exclude the true value.

Fitting continuous data.—Even in continuous systems, deviations from continuous scale invariance can cause inaccuracy in the Pareto MLE. In contrast, our MLE in conjunction with logarithmic binning applies to data with

TABLE I. Pareto (continuous $\hat{\alpha}_c$) and our (discrete $\hat{\alpha}_d$) MLEs of α from earthquake magnitudes [3], bronchial tube diameters [39], and blood vessel radii [40], compared with slopes of regression to the tail distribution and theoretical predictions. Likely errors are marked (*). The Pareto MLE $\hat{\alpha}_c$ fails to account for discretization in earthquake data and thus leads to a $> 10\%$ bias [Fig. 3(a)]. For bronchial data the continuous $\hat{\alpha}_c$ and discrete $\hat{\alpha}_d$ are inconsistent due to deviations from continuous scaling. In vascular data—which is approximately scale invariant— $\hat{\alpha}_c$ and $\hat{\alpha}_d$ are consistent while regression likely fails [12].

Data	x_m	λ	Continuous $\hat{\alpha}_c$	Discrete $\hat{\alpha}_d$	Regression slope	Theoretical α
Earthquakes	2.0	$10^{0.1}$	* 1.05 ± 0.01	0.94 ± 0.01	0.96 ± 0.01	NA
Bronchia	1.6 mm	2	* 3.4 ± 0.2	2.9 ± 0.2	2.78 ± 0.07	$\alpha = 3$ [29,41]
Vessels	0.7 mm	2	2.3 ± 0.1	2.3 ± 0.2	* 2.56 ± 0.02	$2 \leq \alpha \leq 3$ [29]

discrete or continuous scale invariance, and it performs well in the face of within-scale measurement noise. Therefore, we can detect deviations from continuous scale invariance by looking for discrepancies between the estimators. If the unknown distribution that generates the data is truly scale invariant, both MLEs will produce mutually consistent results regardless of λ . A discrepancy between the estimators indicates a violation of scale invariance, where there is no guarantee the Pareto MLE is correct. In that case, our MLE procedure can offer more reliable estimates.

We illustrate these alternatives with bronchial and vascular data (Table I). Bronchia tend to branch symmetrically into two similarly sized bronchia with statistical self-similarity around powers of $\lambda = 2^{1/3}$ as a consequence of Murray's law [29,41]. A database of 1097 bronchial diameter measurements from a single human lung (Raabe *et al.* [39], file D6) binned with $\lambda = 2$ gives $\hat{\alpha}_d = 2.9 \pm 0.2$ —in agreement with theory [Fig. 3(b)]. The Pareto MLE $\hat{\alpha}_c = 3.4 \pm 0.2$ differs significantly from the data and predicted values. We conclude that bronchial measurements deviate from strict scale invariance due to discrete self-similarity arising from branching or measurement error, and that our MLE therefore is more accurate.

Although blood vessels also branch dichotomously, random asymmetry in branching disperses vessel diameters from any specific ratios. On a database of 1569 blood vessel radii measured from a single mouse lung [40], $\hat{\alpha}_c$ and $\hat{\alpha}_d$ produced statistically indistinguishable estimates (Table I), independent of the chosen λ , and are therefore both likely accurate. The mutual consistency between the estimators suggests that the distribution of blood vessel measurements is effectively scale invariant despite the underlying branching.

Conclusion.—These theoretical and empirical arguments show that modeling discrete self-similar processes with a continuous power-law distribution can lead to errors, and a discrete distribution is essential to properly describe many biological and physical systems. We derive a novel estimation procedure that is intuitive, computationally simple, and more reliable than previous methods [12,24,27] because it applies to data with any kind of self-similarity—continuous or discrete. When faced with the unknown, we recommend the empiricist take a version of Pascal's wager: discretize to avoid unknown errors at the cost of a known and modest decrease in statistical power. Our prescription is then to bin logarithmically and use the self-similar discrete estimator.

Reproducibility: All results and figures herein can be reproduced from public data and software [42].

We thank Tom Chou and Jakub Otwinowski for comments on the manuscript. We are grateful for funding from the U.S. National Science Foundation Division of Biological Informatics CAREER Grant (ID: 1254159).

*mgnew@umich.edu

- [1] M. E. Newman, *Contemp. Phys.* **46**, 323 (2005).
- [2] X. Gabaix, *Annu. Rev. Econ.* **1**, 255 (2009).
- [3] P. Bak, K. Christensen, L. Danon, and T. Scanlon, *Phys. Rev. Lett.* **88**, 178501 (2002).
- [4] Y. Y. Kagan, *Tectonophysics* **490**, 103 (2010).
- [5] T. Bu and D. Towsley, *Proc. IEEE INFOCOM* **2**, 638 (2002).
- [6] Z.-S. She and E. Leveque, *Phys. Rev. Lett.* **72**, 336 (1994).
- [7] L. P. Bentley, J. C. Stegen, V. M. Savage, D. D. Smith, E. I. Allmen, J. S. Sperry, P. B. Reich, and B. J. Enquist, *Ecol. Lett.* **16**, 1069 (2013).
- [8] B. J. West, V. Bhargava, and A. L. Goldberger, *J. Appl. Physiol.* **60**, 1089 (1986).
- [9] M. Zamir, P. Sinclair, and T. Wonnacott, *J. Biomech.* **25**, 1303 (1992).
- [10] M. G. Newberry, D. B. Ennis, and V. M. Savage, *PLoS Comput. Biol.* **11**, e1004455 (2015).
- [11] B. C. Arnold, *Pareto Distributions*, 2nd ed. (Chapman and Hall/CRC, Boca Raton, FL, 2015).
- [12] A. Clauset, C. R. Shalizi, and M. E. Newman, *SIAM Rev.* **51**, 661 (2009).
- [13] We define α following Ref. [11] as the log-log slope of the tail distribution whereas Ref. [12] uses α to denote the slope of the probability density function. Thus our $\alpha = \alpha_{[11]} = \alpha_{[12]} - 1$.
- [14] B. B. Mandelbrot, *Fractals: Form, Chance, and Dimension* (W.H. Freeman, San Francisco, 1977).
- [15] D. Sornette, *Phys. Rep.* **297**, 239 (1998).
- [16] V. Pareto, *Giornale degli economisti* **10**, 59 (1895).
- [17] H. Saleur, C. G. Sammis, and D. Sornette, *J. Geophys. Res.* **101**, 17661 (1996).
- [18] W.-X. Zhou and D. Sornette, *Physica (Amsterdam)* **165D**, 94 (2002).
- [19] H. Saleur and D. Sornette, *J. Phys. I* **6**, 327 (1996).
- [20] B. J. West, *Ann. Biomed. Eng.* **18**, 135 (1990).
- [21] J. A. Feigenbaum and P. G. Freund, *Int. J. Mod. Phys. B* **10**, 3737 (1996).
- [22] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. (John Wiley & Sons, Inc., New York, NY, 1994), Vol. 1.
- [23] M. L. Goldstein, S. A. Morris, and G. G. Yen, *Eur. Phys. J. B* **41**, 255 (2004).
- [24] E. P. White, B. J. Enquist, and J. L. Green, *Ecology* **89**, 905 (2008).
- [25] A. N. M. Muniruzzaman, *Calcutta Stat. Assoc. Bull.* **7**, 115 (1957).
- [26] J. L. Doob, *Trans. Am. Math. Soc.* **36**, 759 (1934).
- [27] D. I. Warton, I. J. Wright, D. S. Falster, and M. Westoby, *Biol. Rev. Camb. Philos. Soc.* **81**, 259 (2006).
- [28] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions* (John Wiley & Sons, Hoboken, NJ, 2005).
- [29] G. West, J. Brown, and B. Enquist, *Science* **276**, 122 (1997).
- [30] J. S. Andrade Jr., H. J. Herrmann, R. F. S. Andrade, and L. R. da Silva, *Phys. Rev. Lett.* **94**, 018702 (2005).
- [31] H. L. Seal, *J. Inst. Actuaries* **78**, 115 (1952).
- [32] Y. Virkar, A. Clauset *et al.*, *Ann. Appl. Stat.* **8**, 89 (2014).
- [33] J. Swift, *On Poetry: A Rhapsody* (1733).

- [34] A. De Morgan, *A Budget of Paradoxes* (Logmans, Green, and Co., London, 1872).
- [35] L. Richardson, *Weather Prediction by Numerical Methods* (Cambridge University Press, Cambridge, England, 1922).
- [36] Lupe Fiasco, Dots & Lines, Tetsuo & Youth, Atlantic Records, 2015.
- [37] H. S. Horn, in *Scaling in Biology*, edited by J. H. Brown and G. B. West (Oxford University Press, New York, NY, 2000), pp. 199–220.
- [38] B. Gutenberg and C. F. Richter, *Bull. Seismol. Soc. Am.* **34**, 185 (1944).
- [39] O. Raabe, H. C. Yeh, G. M. Schum, and R. F. Phalen, Tracheobronchial geometry: Human, dog, rat, hamster, U.S. Government Printing Office, Washington, DC, 1976.
- [40] E. Tekin, D. Hunt, M. G. Newberry, and V. M. Savage, *PLoS Comput. Biol.* **12**, e1005223 (2016).
- [41] C. Murray, *Proc. Natl. Acad. Sci. U.S.A.* **12**, 207 (1926).
- [42] <https://github.com/mnewberry/dplfit>.