Ecology, 89(4), 2008, pp. 905–912

ON ESTIMATING THE EXPONENT OF POWER-LAW FREQUENCY DISTRIBUTIONS

Ethan P. White,[1,2,4] Brian J. Enquist,[2] and Jessica L. Green[3]

[1]*Department of Biology and the Ecology Center, Utah State University, Logan, Utah 84322 USA*
[2]*Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721 USA*
[3]*Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, Oregon 97403 USA*

Abstract. Power-law frequency distributions characterize a wide array of natural phenomena. In ecology, biology, and many physical and social sciences, the exponents of these power laws are estimated to draw inference about the processes underlying the phenomenon, to test theoretical models, and to scale up from local observations to global patterns. Therefore, it is essential that these exponents be estimated accurately. Unfortunately, the binning-based methods traditionally used in ecology and other disciplines perform quite poorly. Here we discuss more sophisticated methods for fitting these exponents based on cumulative distribution functions and maximum likelihood estimation. We illustrate their superior performance at estimating known exponents and provide details on how and when ecologists should use them. Our results confirm that maximum likelihood estimation outperforms other methods in both accuracy and precision. Because of the use of biased statistical methods for estimating the exponent, the conclusions of several recently published papers should be revisited.

Key words: *binning; distribution; exponent; maximum likelihood estimation; power laws.*

Introduction

Power laws have a long history in ecology and other disciplines (Bak 1996, Brown et al. 2002, Newman 2005). Power-law relationships appear in a wide variety of physical, social, and biological systems and are often cited as evidence for fundamental processes that underlie the dynamics structuring these systems (Bak 1996, Brown et al. 2002, Newman 2005). There are two major classes of power laws commonly reported in the ecological literature. The first are bivariate relationships between two variables. Examples of this type of relationship include the species–area relationship and body-size allometries. Standard approaches to analyzing this type of data are generally reasonable and discussions of statistical issues related to this kind of data are presented elsewhere (e.g., Warton et al. 2006). The second type of power law, and the focus of this paper, is the frequency distribution, where the frequency of some event (e.g., the number of individuals) is related to the size, or magnitude, of that event (e.g., the size of the individual).

Frequency distributions of a wide variety of ecological phenomena tend to be, at least approximately, power-law distributed. These phenomena include distributions of species body sizes (Morse et al. 1985), individual body sizes (Enquist and Niklas 2001), colony sizes (Jovani and Tella 2007), abundance among species (Pueyo 2006), trends in abundance of species through time (Keitt and Stanley 1998), step lengths in animal search patterns (i.e., Levy flights; Reynolds et al. 2007), fire magnitude (Turcotte et al. 2002), island size (White and Brown 2005), lake size (Wetzel 1991), flood magnitude (Malamud and Turcotte 2006), landslide magnitude (Guzzetti et al. 2002), vegetation patch size (Kefi et al. 2007), and fluctuations in metabolic rate (Labra et al. 2007). Frequency distributions are usually displayed as simple histograms of the quantity of interest. If a distribution is well-characterized by a power law then the frequency of an event (e.g., the number of individuals with mass between 10 and 20 g), f, is related to the size of that event, x, by a function of the following form:

$$f(x) = cx^{\lambda} \tag{1}$$

where c and λ are constants, and λ is called the exponent and is typically negative (i.e., $\lambda < 0$). Because $f(x)$ is a probability density function (PDF) the value of c is a simple function of λ and the minimum and maximum values of x (Table 1). The specific form of the PDF depends on whether the data are continuous or discrete, on the presence of minimum and maximum values, and on whether λ is <-1 or >-1. The different forms are often given distinct names for clarity (see Table 1).

There is substantial interest in using the parameters of these power-law distributions to make inferences about the processes underlying the distributions, to test mechanistic models, and to estimate and predict patterns and processes operating beyond the scope of the observed data. For example, power-law species abundance distributions with $\lambda \approx -1$ are considered to represent evidence for the primary role of stochastic

Manuscript received 6 August 2007; revised 9 November 2007; accepted 20 November 2007. Corresponding Editor: A. M. de Roos.

[4] E-mail: epwhite@biology.usu.edu

TABLE 1. Descriptions of different power-law frequency distributions, including the name of the distribution, the range of data and parameter values over which it applies, its probability density function (PDF; or probability mass function) $f(x)$, its cumulative distribution function $F(x)$, and the maximum likelihood estimate (MLE) for λ based on the PDF.

Distribution†	$f(x)$	$F(x)$	MLE for λ
1) Pareto Range $a \leq x < \infty$ Parameters $\lambda < -1$, $a > 0$	$-(\lambda + 1)a^{-(\lambda+1)}x^{\lambda}$	$1 - a^{-(\lambda+1)}x^{\lambda+1}$	$\hat{\lambda} = -1 - \left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{x_i}{a}\right)\right]^{-1}$
2) Truncated Pareto‡ Range $a \leq x \leq b$ Parameters $\lambda \neq -1$, $a \geq 0$, $b \geq 0$	$(\lambda + 1)(b^{\lambda+1} - a^{\lambda+1})^{-1}x^{\lambda}$	$\frac{x^{\lambda+1} - a^{\lambda+1}}{b^{\lambda+1} - a^{\lambda+1}}$	$\overline{\ln x} = \frac{-1}{(\hat{\lambda}+1)} + \frac{b^{\hat{\lambda}+1}\ln b - a^{\hat{\lambda}+1}\ln a}{b^{\hat{\lambda}+1} - a^{\hat{\lambda}+1}}$
3) Discrete Pareto‡,§ Range $x = a, a+1, a+2, \ldots \infty$ Parameters $\lambda < -1$, $a \geq 1$	$\frac{x^{\lambda}}{\zeta(-\lambda, a)}$	$\frac{\sum_{j=a}^{x} j^{\lambda}}{\zeta(-\lambda, a)}$	$\overline{\ln x} = \frac{-\zeta'(-\hat{\lambda}, a)}{\zeta(-\hat{\lambda}, a)}$
4) Power Function|| Range $0 \leq x \leq b$ Parameters $\lambda > -1$, $b > 0$	$(\lambda + 1)b^{-(\lambda+1)}x^{\lambda}$	$(x/b)^{\lambda+1}$	$\hat{\lambda} = \left[\log(b) - \frac{1}{n}\sum_{i=1}^{n}\log(x_i)\right]^{-1} - 1$

Notes: The minimum value of x for which a distribution is valid is given by a, which is defined to be greater than zero. The maximum value of x for which a distribution is valid is given by b, which is defined to be less than infinity.

† Sources are as follows: Pareto (Johnson et al. 1994); Truncated Pareto (Page 1968); Discrete Pareto (Clauset et al. 2007); Power Function (Evans et al. 2000). There is an error in the MLE solution given by Evans et al. (2000) that has been corrected. Note that MLEs are only guaranteed to be minimum variance unbiased estimators in the limit of large n. If n is small, corrections to the MLE are available (Johnson et al. 1994, Clark et al. 1999, Clauset et al. 2007). All solutions assume that a and b are known.

‡ The MLE equations for these distributions cannot be solved analytically for $\hat{\lambda}$, so they must be solved using numerical methods.

§ $\zeta(-\lambda, a) = \Sigma_{k=0}^{\infty}(k + a)^{\lambda}$ is the generalized zeta function and $\zeta'(-\lambda, a)$ is its derivative with respect to $-\lambda$.

|| The Power Function distribution is often ignored in discussions of power-law distributions because it rarely occurs in natural systems (Newman 2005). We include it here for completeness and because it has been suggested that in some groups individual size distributions based on mass may be approximately power-law distributed with $\lambda > -1$ (e.g., Enquist and Niklas 2001).

birth–death processes, combined with species input, in community assembly (Pueyo 2006, Zillio and Condit 2007); quantitative models of tree size distributions make specific predictions (e.g., $\lambda = -2$; Enquist and Niklas 2001) that can be used to test these models (Coomes et al. 2003, Muller-Landau et al. 2006); and power-law frequency distributions of individual size have been used to scale up from individual observations to estimate ecosystem level processes (Enquist et al. 2003, Kerkhoff and Enquist 2006).

One concern when interpreting the exponents of these distributions is that there are a wide variety of different approaches currently being used to estimate the exponents (Sims et al. 2007, White et al. 2007). These include techniques based on: (1) binning (e.g., Enquist and Niklas 2001, Meehan 2006, Kefi et al. 2007); (2) the cumulative distribution function (e.g., Rinaldo et al. 2002); and (3) maximum likelihood estimation (e.g., Muller-Landau et al. 2006, Edwards et al. 2007, Zillio and Condit 2007). There has been little discussion in the ecological literature of how the choice of methodology influences the parameter estimates, and methods other than binning are rarely used. If different methods produce different results this could have consequences for the conclusions drawn about the ecology of the system (Edwards et al. 2007, Sims et al. 2007).

Here, we: (1) describe the different approaches used to quantify the exponents of power-law frequency distributions; (2) show that some of these approaches give biased estimates; (3) illustrate the superior performance of some approaches using Monte Carlo methods; (4) make recommendations for best estimating parameters of power-law distributed data; and (5) show that some of the conclusions of recent studies are effected by the use of biased statistical techniques.

Methods for Estimating the Exponent

Linear binning

Perhaps the most intuitive way to quantify an empirical frequency distribution is to bin the observed data using bins of constant linear width. This generates the familiar histogram. Specifically, linear binning entails choosing a bin i of constant width ($w = x_{i+1} - x_i$), counting the number of observations in each bin (i.e., with values of x between x_i and $x_i + w$), and plotting this count against the value of x at the center of the bin ($x_i/2 + x_{i+1}/2$). If the counts are divided by the sum of all the counts, this plot is an estimate of the probability density function, $f(x)$. The traditional approach to estimating the power-law exponent is to fit a linear regression to log-transformed values of $f(x)$ and x, with the slope of the line giving an estimate of the exponent, λ. Bins with zero observations are excluded, because log(0) is undefined, and sometimes bins with low counts are also excluded (e.g., Enquist and Niklas 2001). While in practice the choice of bin width is normally arbitrary, this choice represents a trade-off between the number of bins analyzed (i.e., the resolution of the frequency distribution) and the accuracy with which each value of $f(x)$ is estimated (fewer observations per bin provide a poorer density estimate; Pickering et al. 1995).

Logarithmic binning

Simple logarithmic binning.—This approach is similar to linear binning, except that instead of the bins having constant linear width, they have constant logarithmic width, $b = \log(x_{i+1}) - \log(x_i)$. The estimate of λ is obtained by log-transforming the values of x and following the procedure described in the previous section. Since the x data are transformed to begin with, it is not necessary to transform the bin centers again prior to fitting the regression. For power-law-like distributions, an advantage of logarithmic binning is the reduction of the number of zero and low-count bins at larger values of x because the linear width of a bin increases linearly with x; i.e., $w_i = x_i(e^b - 1)$. However, this means that the number of observations within each bin is determined not only by x, but also by the linear width of the bin. Therefore, the slope of the regression will give an estimate of $\lambda + 1$, not λ (Appendix A; Han and Straskraba 1998, Bonnet et al. 2001, Sims et al. 2007).

Normalized logarithmic binning.—The problem of increasing linear width of logarithmic bins can be dealt with by normalizing the number of observations in each bin by the linear width of the bin, w. This converts the counts into densities (number of observations per unit of x; Bonnet et al. 2001, Christensen and Moloney 2005). The linear width of a logarithmic bin can be calculated as $x_i(e^b - 1)$ (Appendix A). This normalization approach is typically used in the characterization of aquatic size-spectra and power-law distributions in physics (Kerr and Dickie 2001, Christensen and Moloney 2005). It removes the artifact from traditional logarithmic binning while maintaining the advantage of using larger bins where there are fewer values of x. An alternative approach is to use simple logarithmic binning and subtract one from the estimated exponent (Han and Straskraba 1998, Bonnet et al. 2001).

Fitting the cumulative distribution function

An alternative to binning methods is to work with the cumulative distribution function (CDF). The CDF describes the probability that a random variable, X, drawn from $f(x)$ is $\leq x$. The CDF is straightforward to construct for a set of observed data, and no binning is required. To construct the CDF, first rank the n observed values (x_i) from smallest to largest ($i = 1 \ldots n$). The probability that an observation is less than or equal to x_i (the CDF) is then estimated as i/n (this is the Kaplan-Meier estimate; Evans et al. 2000). Analyzing the CDF avoids the subjective influence of the choice of bin width and the problem of empty bins. Having determined the CDF for a power-law distribution, the exponent, λ, of the probability density function (PDF) can be estimated using regression. The traditional approach is to transform the equation for the CDF such that the slope of a linear equation is a function of λ. The linearized equation differs among distributions (Appendix A). The slope of the regression will be equal to $\lambda + 1$, making it necessary to subtract 1 to obtain λ (Bonnet et al. 2001, Rinaldo et al. 2002).

Maximum likelihood estimation

Maximum likelihood estimation (MLE) is one of the preferred approaches for estimating frequency distribution parameters (e.g., Rice 1994). MLE determines the parameter values that maximize the likelihood of the model (in this case, a power law with an unknown exponent) given the observed data. Specifically, MLE finds the value of λ that maximizes the product of the probabilities of each observed value of x (i.e., the product of $f(x)$ evaluated at each data point; see Rice [1994] for a good introduction to maximum likelihood methods). The specific solution for the maximum likelihood estimate of λ and whether the solution is closed form or requires numerical methods to solve depends on the minimum and maximum values of x and on the value of λ (Table 1). Alternatively, the likelihood can be maximized directly using numerical methods (Clauset et al. 2007, Zillio and Condit 2007). While MLE does not provide an opportunity for visual inspection of the distribution to determine if the assumption of the power-law functional form is reasonable, the validity of this assumption can be assessed using simple goodness-of-fit tests such as the Chi-square on binned data (Clark et al. 1999, Clauset et al. 2007, Edwards et al. 2007), or by visually assessing the linearity of binned data, or the CDF (Benhamou 2007), under the appropriate transformation.

Comparing the Methods

While uncorrected simple logarithmic binning clearly provides incorrect estimates of λ, the alternative approaches discussed above all seem reasonable and intuitive. However, the different approaches do not perform equally well, and some produce biased estimates of the exponent (e.g., Pickering et al. 1995, Clark et al. 1999, Sims et al. 2007). We applied Monte Carlo methods to illustrate the advantages and disadvantages of the various approaches and to explore cases relevant to ecology that have not been previously addressed. Monte Carlo methods generate data that are, by definition, power-law distributed with known exponents, making it possible to compare the performance of the different techniques in estimating the value of λ.

We generated power-law distributed random numbers using the inverse transformation method for the Pareto distribution (Ross 2006), and using the rejection method for the discrete Pareto distribution (Devroye 1986). Each analysis consisted of the following: (1) generating 10 000 Monte Carlo data sets for each point in the analysis (e.g., for each sample size), (2) estimating the exponent for each data set using the methods described previously, and (3) comparing the performance of the methods based on bias (i.e., accuracy) and on the variance in the estimate (i.e., precision). We report on simulated distributions generated using $\lambda = -2$ and $a = 1$.

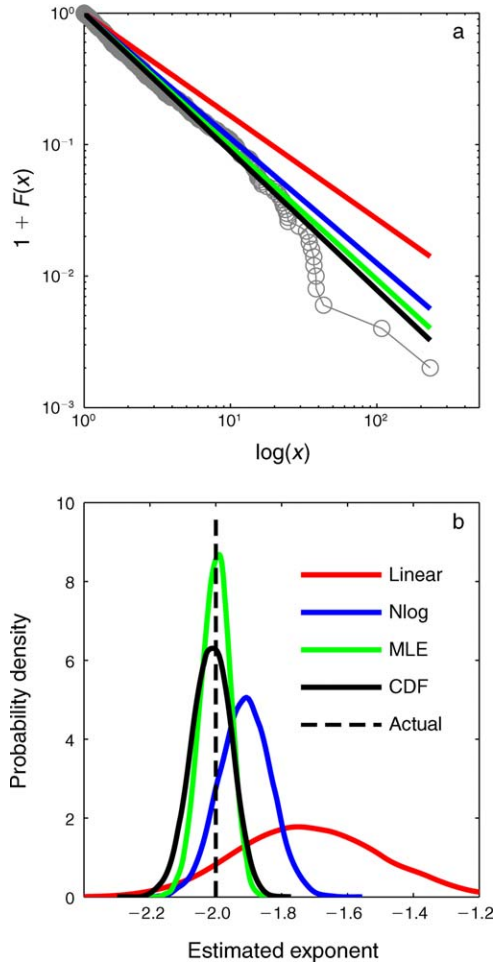

FIG. 1. Example of Monte Carlo results for the different methods of fitting the power-law exponent. (a) A single Monte Carlo sample from a Pareto distribution plotted as 1 minus the cumulative distribution, $F(x)$. Data are plotted as gray open circles along with the fits to the data using the four different methods: linear binning (red; linear), normalized logarithmic binning (blue; Nlog), maximum likelihood estimation (green; MLE), and cumulative distribution function fitting (black; CDF). (b) Kernel density estimates of the distribution of exponents from 10 000 Monte Carlo runs. Line colors are the same as for (a), and the value of λ used to generate the data is indicated by the dashed line. Parameter values: $n = 500$ random points in each simulation; $\lambda = -2$; $1 \leq x < \infty$; linear bin width = 3; logarithmic bin width = 0.3. The binning analyses used the minimum value of x and excluded the last bin and bins containing ≤ 1 individual. Exclusion of the last bin is not necessary, but it improves the performance of binning-based approaches and is thus conservative in the context of our conclusions. The single sample for (a) was chosen to illustrate the general results shown in (b). Binning methods generate biased estimates of the exponent and result in more variable estimates than approaches based on MLE and CDF.

The results for other combinations of parameters are qualitatively similar. We also evaluated the influence of sample size on the various estimation techniques, and for binning-based approaches we evaluated the effect of bin width on the analysis.

General Rules

Uncorrected simple logarithmic binning gives the wrong exponent.—Non-normalized logarithmic binning does not estimate λ; it estimates $\lambda + 1$ (Han and Straskraba 1998, Bonnet et al. 2001, Sims et al. 2007). Therefore if simple logarithmic binning is used, and an estimate of λ is the desired result, then it is necessary to subtract 1 from the slope of the logarithmically binned data. Not doing so will give the wrong value for the exponent.

Binning-based approaches perform poorly.—Linear binning performs poorly by practically any measure. In most cases it produces biased estimates of the exponent and its estimates are highly variable (Figs. 1 and 2). In addition, the estimated exponent is highly dependent on the choice of bin width, and this dependency varies as a function of sample size (Fig. 3). While normalized logarithmic binning performs better than linear binning, its estimates are also dependent on the choice of bin width and are more variable than alternate approaches. Our results are based on recommended practices in binning analyses (following Pickering et al. 1995). Many alternative approaches to constructing bins and performing regressions on binned data are conceivable, and it is possible that some of these may improve the performance of the estimates. However, this highlights the fact that binning-based methods are sensitive to a variety of decisions, and it appears that no amount of tweaking will be able to produce a consistent binning-based method for estimating the exponent. In general, binning results in a loss of information about the distributions of points within a bin and is thus expected to perform poorly (Clauset et al. 2007, Edwards et al. 2007). Therefore, while binning is useful for visualizing the frequency distribution, and normalized logarithmic binning performs well at this task, binning-based approaches should be avoided for parameter estimation (Clauset et al. 2007).

Maximum likelihood estimation performs best.—While fitting the cumulative distribution function (CDF) generally produces good results, estimates of λ using the CDF approach are often biased at small sample sizes and are consistently more variable than those using maximum likelihood estimation (MLE; Fig. 2; Clark et al. 1999, Newman 2005). This probably results because the logarithmic transformation used in fitting the CDF weights a small number of points more heavily, and because the points in the CDF are not independent thus violating regression assumptions (see Clauset et al. [2007] for other issues with regression-based approaches). While alternative approaches to fitting the CDF (e.g., nonlinear regression) could improve the performance of this estimator, MLE has been shown mathematically to be the single best approach for estimating power-law exponents (i.e., it is the minimum variance unbiased estimator; Johnson et al. 1994, Clark et al. 1999, Newman 2005). In addition, MLE produces valid confidence intervals for the estimated exponent

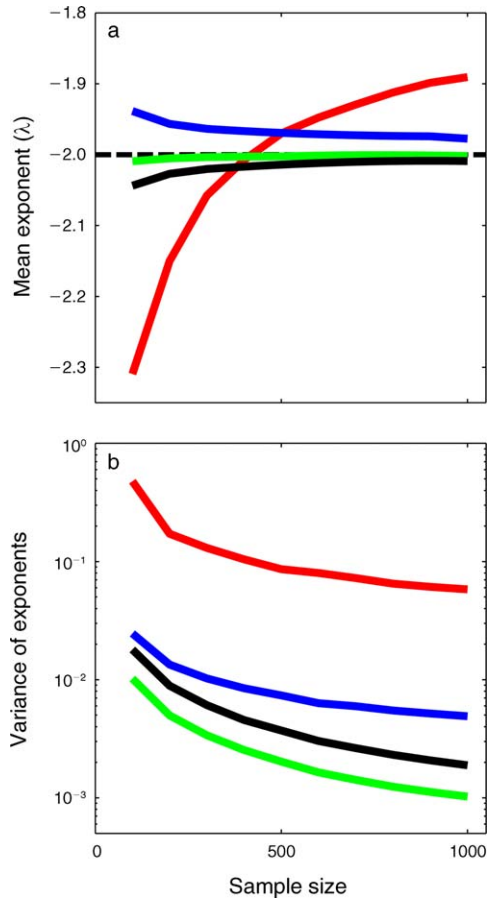

FIG. 2. (a) Effect of sample size on the mean estimated exponent and (b) the variance of that exponent, for the four estimation methods: linear binning (red), normalized logarithmic binning (blue), maximum likelihood estimation (green), and cumulative distribution function fitting (black). Values for each sample size were generated using 10 000 Monte Carlo runs from the Pareto distribution with parameter values of $\lambda = -2$ (dashed line), and $1 \leq x < \infty$; linear bin width = 7.5, and logarithmic bin width = 0.75. Other binning methods are as in Fig. 1. Linear binning fails to converge to the correct estimate. While the other methods all appear to converge at large sample sizes, maximum likelihood estimation always yields the lowest variance in the estimated exponent.

(Appendix A), which the other methods do not (Clark et al. 1999, Newman 2005, Clauset et al. 2007).

COMPLICATIONS

Minimum and maximum values.—Minimum and maximum attainable values of ecological quantities can result either from natural limits on the quantity being measured (e.g., trees cannot grow above some maximum size), or from methodological limits on the values that can be observed (e.g., fires <1 ha are not recorded). In addition, the power-law form of the distribution may not hold over the entire range of x, making it necessary to select a restricted range of x on which to estimate the exponent. While binning-based approaches do not assume particular limits on x (but see Pickering et al. 1995), CDF and MLE approaches assume the minimum and maximum attainable values of x given in Table 1. In some cases these limits may be known, but if not, it may be necessary to estimate them (e.g., Kijko 2004, Clauset et al. 2007). Because maximum likelihood estimation for the truncated Pareto requires numerical methods, it has been suggested that in some cases with both a minimum and maximum value that the error introduced by assuming that there is no maximum is small enough that it is reasonable to estimate the exponent using the maximum likelihood estimate for the Pareto distribution. Clark et al. (1999) suggest this approximation in cases where the maximum value is at least two orders of

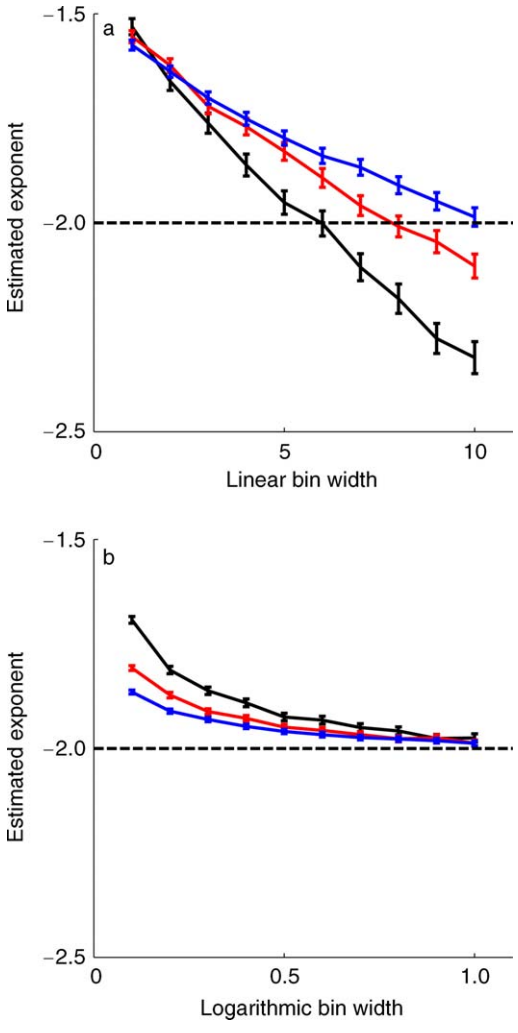

FIG. 3. Effect of bin width on the estimated exponents for (a) linear and (b) normalized logarithmic binning for three different sample sizes: $n = 200$ (solid black line), $n = 500$ (solid red line), and $n = 1000$ (solid blue line); based on 1000 Monte Carlo runs from the Pareto distribution per point. Parameter values were $\lambda = -2$ (dashed black line) and $1 \leq x < \infty$. Error bars are ± 2 SE. Other binning methods are as in Fig. 1. Changing bin width changes the estimated exponent for all sample sizes.

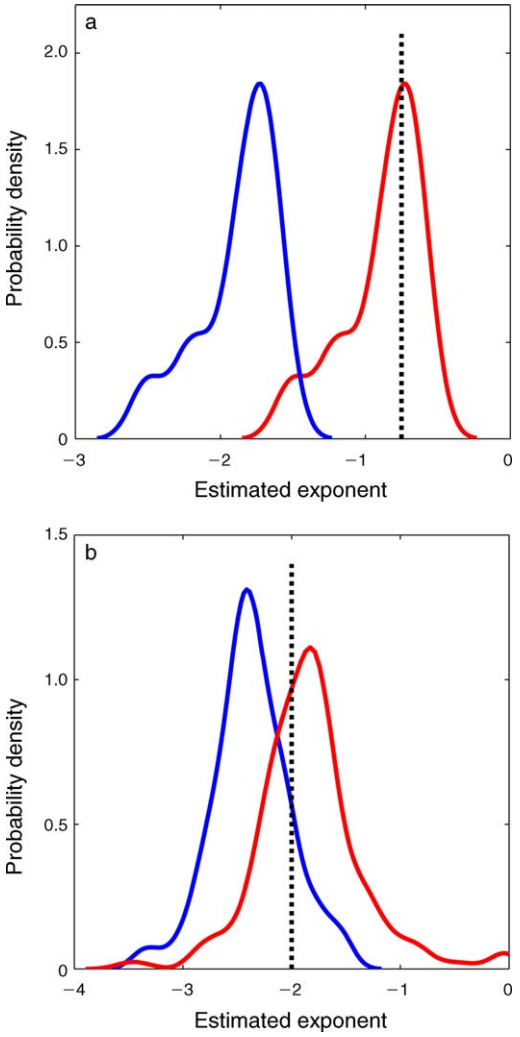

FIG. 4. Reanalysis of individual size distribution data from (a) Meehan (2006) and (b) Enquist and Niklas (2001), using less biased methods. Plots are probability densities of the estimated exponents using the original methodology from these studies (red line; simple logarithmic binning in Meehan, linear binning in Enquist and Niklas), and using less biased methods (blue line; normalized logarithmic binning for Meehan, maximum likelihood estimation for Enquist and Niklas). Both studies purported to support a theoretically derived exponent (dotted line). However, when the exponents are estimated using less biased methods it becomes clear that the observed data deviate significantly from the theoretical prediction.

magnitude greater than the minimum; that is, max $(x) > 100 \times \min(x)$.

Deviations from the power law.—Empirical data are rarely perfectly power-law distributed over the entire range of x (Brown et al. 2002, Newman 2005). MLE and CDF approaches respond to deviations differently because the traditional MLE analysis implicitly weights data on a linear scale while the traditional CDF approach weights it on a logarithmic scale (McGill 2003). The CDF approach will therefore respond more strongly to deviations from the power law at large values of x (such as those observed in individual size distributions; e.g., Coomes et al. 2003) than the MLE approach, whereas MLE will respond more strongly to deviations at small values of x (commonly observed in many power-law distributions; e.g., Newman 2005). It is common to truncate data in the tails that exhibit deviations from the power law before fitting the exponent (e.g., Newman 2005). However, these deviations also should not be ignored, as they may help identify important biological processes (e.g., Coomes et al. 2003). In some cases deviations may suggest that the power law is in fact not the appropriate model for the data. This can be evaluated using goodness-of-fit tests on binned data (Clark et al. 1999, Clauset et al. 2007, Edwards et al. 2007) or by using model selection techniques to compare the power-law to alternative distributions (Muller-Landau et al. 2006, Clauset et al. 2007, Edwards et al. 2007).

Discrete data.—Most of the MLE and CDF methods presented here assume that the data are continuously distributed, as is often the case (e.g., body size). However, some ecological patterns (e.g., species-abundance distributions) are comprised of discrete observations (e.g., it is impossible to census 4.3 individuals). It is therefore necessary to use analogous discrete distributions. In the case of the Pareto distribution a discrete analog exists in the form of the aptly named discrete Pareto distribution (Johnson et al. 2005, Newman 2005; Table 1; also called the Zipf or Riemann-zeta distribution). In some cases continuous distributions can reasonably approximate discrete data; but in the case of the Pareto, using the continuous maximum likelihood estimate instead of that derived from the discrete distribution produces strongly biased results and should be avoided (Appendix C; Clauset et al. 2007).

IMPLICATIONS FOR PUBLISHED RESULTS

One of the most important implications for published results is that studies that have estimated exponents using uncorrected simple logarithmic binning (e.g., Morse et al. 1988, Meehan 2006) have reported the wrong exponent. This is particularly important in cases where the exponent is used to test quantitative predictions. For example, an analysis in Meehan (2006) evaluates whether observed individual size distribution exponents were consistent with those predicted, using simple logarithmic binning. Meehan concluded that the empirical data matched the predictions (Fig. 4a). However, since the reported exponents are equal to $\lambda + 1$, the analysis suggests that the size distribution is substantially steeper than expected, thus refuting, rather than supporting, the hypothesized mechanism (Fig. 4a; Appendix B).

Analyses based on linearly binned data should also be revisited due to the potential for biased estimates and the strong influence of bin width on the estimated exponent. In particular, studies that have used linear binning to test the predictions of theoretical models or

compare exponents from different data sets (e.g., Enquist and Niklas 2001, Coomes et al. 2003, Niklas et al. 2003, Kefi et al. 2007) may have reached incorrect conclusions. We reanalyzed the original data from Enquist and Niklas (2001) and found that while the original linear binning analyses suggested that observed diameter distribution exponents were near the theoretical prediction of −2, MLE suggests that the observed exponents are actually closer, on average, to −2.5 (Fig. 4b; Appendix B). Our reanalysis indicates that the size–frequency distributions in Gentry's plots are not, in general, adequately represented by a power law with an exponent of −2, as originally claimed by Enquist and Niklas (2001; see Appendix B for an important caveat).

While normalized logarithmic binning performs better than linear binning, it can still introduce biases of ~10% depending on the bin width. While many analyses based on normalized logarithmic binning are probably reasonable, the recent suggestion that normalized logarithmic binning is the best approach for fitting exponents (Sims et al. 2007) is unwarranted, and MLE should be used whenever possible (Clark et al. 1999, Clauset et al. 2007).

Compared to binning-based approaches, results from fitting the CDF are probably reasonable. In cases with low sample sizes, where small errors in the estimated exponent could influence the conclusions of the study, or where minimum or maximum attainable values of x have been ignored (see Pickering et al. 1995), it may be worth checking the results using MLE. Regardless, MLE is the single best method for estimating exponents and should be used in future studies.

Conclusions

The vast majority of ecological studies that estimate exponents for power-law-like distributions use approaches based on binning the empirical data (e.g., Morse et al. 1988, Enquist and Niklas 2001, Coomes et al. 2003, Niklas et al. 2003, Meehan 2006, Jovani and Tella 2007, Kefi et al. 2007, Reynolds et al. 2007, Sims et al. 2007). These binning-based methods tend to produce results that are biased, have high variance, and are contingent on the choice of bin width. Instead of binning, maximum likelihood estimation should be used when fitting power-law exponents to empirical data (Clark et al. 1999, Newman 2005, Edwards et al. 2007).

We have focused on power laws because they, at least approximately, characterize a number of distributions of interest to ecologists. The issues raised here, and the conclusions discussed, should apply broadly to frequency distributions in general, and in particular to other distributions with heavy tails. Paying careful attention to fitting methodologies and consultation of statistical references (e.g., Johnson et al. 1994) should help improve the estimation of distributional parameters.

Acknowledgments

We particularly thank Tim Meehan for generously sharing his data with us and for comments on the manuscript. For the Gentry data we thank Alwyn Gentry, the Missouri Botanical Garden, and collectors who assisted Gentry or contributed data for specific sites. We also thank David Coomes, Susan Durham, Jim Haefner, Brian McGill, and Tommaso Zillio for comments on the manuscript. This work was funded by the National Science Foundation through a Postdoctoral Fellowship in Biological Informatics to E. P. White (DBI-0532847). We used the first–last author emphasis approach (sensu Tscharntke et al. 2007) for the sequence of authors.

Literature Cited

Bak, P. 1996. How nature works: the science of self-organized criticality. Springer-Verlag, New York, New York, USA.

Benhamou, S. 2007. How many animals really do the Levy walk? Ecology 88:1962–1969.

Bonnet, E., O. Bour, N. E. Odling, P. Davy, I. Main, P. Cowie, and B. Berkowitz. 2001. Scaling of fracture systems in geological media. Reviews of Geophysics 39:347–383.

Brown, J. H., V. K. Gupta, B. L. Li, B. T. Milne, C. Restrepo, and G. B. West. 2002. The fractal nature of nature: power laws, ecological complexity and biodiversity. Philosophical Transactions of the Royal Society of London, Series B 357: 619–626.

Christensen, K., and N. R. Moloney. 2005. Complexity and criticality. Imperial College Press, London, UK.

Clark, R. M., S. J. D. Cox, and G. M. Laslett. 1999. Generalizations of power-law distributions applicable to sampled fault-trace lengths: model choice, parameter estimation and caveats. Geophysical Journal International 136:357–372.

Clauset, A., C. R. Shalizi, and M. E. J. Newman. 2007. Power-law distributions in empirical data. ⟨http://arxiv.org/abs/0706.1062v1⟩

Coomes, D. A., R. P. Duncan, R. B. Allen, and J. Truscott. 2003. Disturbances prevent stem size–density distributions in natural forests from following scaling relationships. Ecology Letters 6:980–989.

Devroye, L. 1986. Non-uniform random variate generation. Springer-Verlag, New York, New York, USA.

Edwards, A. M., R. A. Phillips, N. W. Watkins, M. P. Freeman, E. J. Murphy, V. Afanasyev, S. V. Buldyrev, M. G. E. da Luz, E. P. Raposo, H. E. Stanley, and G. M. Viswanathan. 2007. Revisiting Levy flight search patterns of wandering albatrosses, bumblebees, and deer. Nature 449: 1044–1048.

Enquist, B., E. Economo, T. Huxman, A. Allen, D. Ignace, and J. Gillooly. 2003. Scaling metabolism from organisms to ecosystems. Nature 423:639–642.

Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. Nature 410: 655–660.

Evans, M., N. Hastings, and B. Peacock. 2000. Statistical distributions. Third edition. John Wiley and Sons, New York, New York, USA.

Guzzetti, F., B. D. Malamud, D. L. Turcotte, and P. Reichenbach. 2002. Power-law correlations of landslide areas in central Italy. Earth and Planetary Science Letters 195:169–183.

Han, B. P., and M. Straskraba. 1998. Size dependence of biomass spectra and population density – I. The effects of size scales and size intervals. Journal of Theoretical Biology 191:259–265.

Johnson, N. L., A. W. Kemp, and S. Kotz. 2005. Univariate discrete distributions. Third edition. Wiley-Interscience, New York, New York, USA.

Johnson, N. L., S. Kotz, and N. Balakrishnan. 1994. Continuous univariate distributions. Volume 1. Second edition. Wiley-Interscience, New York, New York, USA.

Jovani, R., and J. L. Tella. 2007. Fractal bird nest distribution produces scale-free colony sizes. Proceedings of the Royal Society of London B 274:2465–2469.

Kefi, S., M. Rietkerk, C. L. Alados, Y. Pueyo, V. P. Papanastasis, A. ElAich, and P. C. de Ruiter. 2007. Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. Nature 449:213–217.
Keitt, T. H., and H. E. Stanley. 1998. Dynamics of North American breeding bird populations. Nature 393:257–260.
Kerkhoff, A. J., and B. J. Enquist. 2006. Ecosystem allometry: the scaling of nutrient stocks and primary productivity across plant communities. Ecology Letters 9:419–427.
Kerr, S. R., and L. M. Dickie. 2001. Biomass spectrum. Columbia University Press, New York, New York, USA.
Kijko, A. 2004. Estimation of the maximum earthquake magnitudes, mmax. Pure and Applied Geophysics 161: 1655–1681.
Labra, F. A., P. A. Marquet, and F. Bozinovic. 2007. Scaling metabolic rate fluctuations. Proceedings of the National Academy of Sciences (USA) 104:10900–10903.
Malamud, B. D., and D. L. Turcotte. 2006. The applicability of power-law frequency statistics to floods. Journal of Hydrology 322:168–180.
McGill, B. 2003. Strong and weak tests of macroecological theory. Oikos 102:679–685.
Meehan, T. D. 2006. Energy use and animal abundance in litter and soil communities. Ecology 87:1650–1658.
Morse, D. R., J. H. Lawton, M. M. Dodson, and M. H. Williamson. 1985. Fractal dimension of vegetation and the distribution of arthropod body lengths. Nature 314:731–733.
Morse, D. R., N. E. Stork, and J. H. Lawton. 1988. Species number, species abundance and body length relationships of arboreal beetles in bornean lowland rain-forest trees. Ecological Entomology 13:25–37.
Muller-Landau, H. C., et al. 2006. Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. Ecology Letters 9:589–602.
Newman, M. E. J. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46:323–351.
Niklas, K. J., J. J. Midgley, and R. H. Rand. 2003. Tree size frequency distributions, plant density, age and community disturbance. Ecology Letters 6:405–411.
Page, R. 1968. Aftershocks and microaftershocks of the great Alaska earthquake of 1964. Bulletin of the Seismological Society of America 58:1131–1168.
Pickering, G., J. M. Bull, and D. J. Sanderson. 1995. Sampling power-law distributions. Tectonophysics 248:1–20.
Pueyo, S. 2006. Diversity: between neutrality and structure. Oikos 112:392–405.
Reynolds, A. M., A. D. Smith, R. Menzel, U. Greggers, D. R. Reynolds, and J. R. Riley. 2007. Displaced honey bees perform optimal scale-free search flights. Ecology 88:1955–1961.
Rice, J. A. 1994. Mathematical statistics and data analysis. Duxbury Press, Pacific Grove, California, USA.
Rinaldo, A., A. Maritan, K. K. Cavender-Bares, and S. W. Chisholm. 2002. Cross-scale ecological dynamics and microbial size spectra in marine ecosystems. Proceedings of the Royal Society of London Series B 269:2051–2059.
Ross, S. 2006. A first course in probability. Seventh edition. Pearson Prentice Hall, Upper Saddle River, New Jersey, USA.
Sims, D. W., D. Righton, and J. W. Pitchford. 2007. Minimizing errors in identifying Levy flight behavior of organisms. Journal of Animal Ecology 76:222–229.
Tscharntke, T., M. E. Hochberg, T. A. Rand, V. H. Resh, and J. Krauss. 2007. Author sequence and credit for contributions in multiauthored publications. PLoS Biology 5:e18.
Turcotte, D. L., B. D. Malamud, F. Guzzetti, and P. Reichenbach. 2002. Self-organization, the cascade model, and natural hazards. Proceedings of the National Academy of Sciences (USA) 99:2530–2537.
Warton, D. I., I. J. Wright, D. S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. Biological Reviews 81:259–291.
Wetzel, R. G. 1991. Land–water interfaces: metabolic and limnological regulators. Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie 24:6–24.
White, E. P., and J. H. Brown. 2005. The template: patterns and processes of spatial variation. Pages 31–47 *in* G. M. Lovett, C. G. Jones, M. G. Turner, and K. C. Weathers, editors. Ecosystem function in heterogeneous landscapes. Springer, New York, New York, USA.
White, E. P., S. K. M. Ernest, A. J. Kerkhoff, and B. J. Enquist. 2007. Relationships between body size and abundance in ecology. Trends in Ecology and Evolution 22:323–330.
Zillio, T., and R. Condit. 2007. The impact of neutrality, niche differentiation and species input on diversity and abundance distributions. Oikos 116:931–940.

APPENDIX A

Details of logarithmic binning, cumulative distribution function fitting, and maximum likelihood estimation (*Ecological Archives* E089-052-A1).

APPENDIX B

Details of reanalysis of results published in Meehan (2006) and Enquist and Niklas (2001) (*Ecological Archives* E089-052-A2).

APPENDIX C

Example of bias resulting from fitting discrete data with continuous maximum likelihood solutions (*Ecological Archives* E089-052-A3).

SUPPLEMENT

Matlab files for fitting power-law exponents using different methods (*Ecological Archives* E089-052-S1).

Ecological Archives **E089-052-A1**

Ethan P. White, Brian J. Enquist, and Jessica L. Green. 2008. On estimating the exponents of power-law frequency distributions. Ecology 89:905-912.

Appendix A.Details of logarithmic binning, cumulative distribution function fitting, and maximum likelihood estimation.

Why simple logarithmic binning estimates λ+1

By definition a bin of constant logarithmic width means that the logarithm of the upper edge of a bin (x_{i+1}) is equal to the logarithm of the lower edge of that bin (x_i) plus the bin width (b). That is,

$$\log\left(x_{i+1}\right) = \log\left(x_i\right) + b$$
$$\Rightarrow x_{i+1} = e^{\left(\log(x_i)+b\right)} = e^{\log(x_i)}e^b = x_i e^b$$

Since the linear bin width of bin i, w_i, is defined as

$$w_i = x_{i+1} - x_i$$

the linear bin width is directly proportional to x_i because

$$w_i = x_i e^b - x_i = x_i\left(e^b - 1\right).$$

The number of observations in a bin (n) is equal to the density of observations in that bin times the width of that bin. Therefore if the probability density function, $f\left(x\right) \propto x^{\lambda}$ and the width of the bin, $w \propto x$, then

$$n \propto x^{\lambda} x = x^{\lambda+1} = x^{\lambda+1}$$

and regressing log(n) against log(x) yields a slope equal to λ+1, not λ. If n is divided by the linear width of the bin then,

$$\frac{n}{w} \propto \frac{x^{\lambda} x}{x}$$
$$\propto x^{\lambda}$$

and thus a regression of the normalized logarithmic bin counts against the logarithm of x will estimate λ.

Linearizations of the cumulative distribution functions

Distribution	CDF	Linearization of CDF
Pareto[1]	$1-a^{-(\lambda+1)}x^{\lambda+1}$	$\log\left(1-F\left(x\right)\right)=-\left(\lambda+1\right)\log\left(a\right)+\left(\lambda+1\right)\log\left(x\right)$
Truncated Pareto[2]	$\frac{x^{\lambda+1}-a^{\lambda+1}}{b^{\lambda+1}-a^{\lambda+1}}$	------
Discrete Pareto[2]	$\frac{\sum_{j=a}^{x} j^{\lambda}}{\zeta\left(-\lambda,a\right)}$	------
Power Function[1]	$\left(x/b\right)^{\lambda+1}$	$\log\left(F\left(x\right)\right)=-\left(\lambda+1\right)\log\left(b\right)+\left(\lambda+1\right)\log\left(x\right)$

[1] Typically the linearization of the CDF is then fit using simple linear regression. This ignores the fact that the intercept is also a function of lambda, or alternatively that the CDF at the minimum attainable value of *x* must be equal to 0. This could be dealt with using non-linear regression, but this is not done in the ecological literature.
[2]There is no obvious way to isolate λ as a simple component of the slope for these distributions.

Confidence intervals for maximum likelihood estimates

Solutions for the standard error (SE) of the estimated value of λ[1]. Estimates of the SE can also be obtained using the bootstrap or jackknife techniques (Newman 2005).

Distribution	SE of MLE
Pareto	$\frac{\left(-\hat{\lambda}-1\right)}{\sqrt{n}}$
Truncated Pareto	$\frac{1}{\sqrt{n}}\left(\frac{1}{\hat{\lambda}^2}-\frac{\left(b/a\right)^{-\lambda}\left[\ln\left(b/a\right)\right]^2}{\left[1-\left(b/a\right)^{-\hat{\lambda}}\right]^2}\right)^{-\frac{1}{2}}$
Discrete Pareto[2]	$\frac{1}{\sqrt{n\left[\frac{\zeta''\left(-\hat{\lambda},a\right)}{\zeta\left(-\hat{\lambda},a\right)}-\left(\frac{\zeta'\left(-\hat{\lambda},a\right)}{\zeta\left(-\hat{\lambda},a\right)}\right)^2\right]}}$
Power Function	$\frac{\hat{\lambda}+1}{\sqrt{n}}$

[1]Source: Pareto (Clauset et al. 2007); Truncated Pareto (Aban et al. 2006); Discrete Pareto (Clauset et al. 2007); Power Function (JLG). Solutions for SE are in the limit of

large *n*. It is possible to correct the SE for small *n* and solutions for this correction are available for the Pareto distribution (Johnson et al. 1994, Newman 2005, Clauset et al. 2007).

[2]The estimates of the SE for the Pareto distribution can be used as an approximation for the Discrete Pareto for reasonably large *n* and *a* (Clauset et al. 2007).

Aban, I. B., M. M. Meerschaert, and A. K. Panorska. 2006. Parameter estimation for the truncated pareto distribution. Journal Of The American Statistical Association **101**:270-277.

Clauset, A., C. R. Shalizi, and M. E. J. Newman. 2007. Power-law distributions in empirical data. arXiv:0706.1062v1 [physics.data-an].

Johnson, N. L., S. Kotz, and N. Balakrishnan. 1994. Continuous Univariate Distributions: Volume 1, 2nd edition. Wiley-Interscience, New York.

Newman, M. E. J. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics **46**:323-351.

Ethan P. White, Brian J. Enquist, and Jessica L. Green. 2008. On estimating the exponents of power-law frequency distributions. *Ecology* 89:905–912.

Appendix B. Details of reanalysis of results published in Meehan (2006) and Enquist and Niklas (2001).

General Approach

In Fig. 4 of our paper we report reanalyses of data originally reported by Meehan (2006) and Enquist and Niklas (2001). We first analyzed the data following the protocols reported in the original paper. We replicated the original analysis as closely as possible including bin widths, positions of bin edges, and exclusion of data points. Results are reported as kernel density estimates (basically histograms) of the fitted exponents.

Reanalysis of Meehan (2006)

Our reanalysis of Meehan (2006) focuses on the "local scale" size distribution results reported in his Fig. 3. Meehan's original analysis used simple logarithmic binning to fit size distributions in litter and soil community data originally published by Petersen and Luxton (1982). Meehan's original analysis included additional variables in a multivariate framework (trophic level and temperature). We ignore these here for simplicity and our estimates of the size distribution exponents appear to be unaffected by excluding these variables (i.e., his simple logarithmic binning estimates including the over variables are very similar to our estimates excluding them). While Meehan's final model fit all sites and trophic levels using a single model, the exponent derived using all of the data is similar to those fit to each site and trophic level separately (Meehan 2006). Therefore, we fit each distribution separately yielding a total of 12 distributions.

We used the bin edges reported in the original paper, but changed the analysis slightly to make it consistent with the standard description of logarithmic binning given in our paper. Meehan (2006) calculated the average mass of an individual occurring in a body size bin based on the reported body size data and used this as the value of x instead of simply using the bin center. In our reanalysis we used the mass at the center of the bin (i.e., the geometric mean of the linearly scaled bin edges) for both our replication of Meehan's analysis and our reanalysis. Results using Meehan's original definition of average mass were qualitatively similar.

The raw data in Meehan (2006) consists of an average mass and a total abundance for each species at each site. Not having the actual masses of each individual will affect maximum likelihood estimation in unknown ways. Therefore we reanalyzed this data by following all of the protocols reported in Meehan (2006), but normalized the counts by the linear width of the logarithmic bins. This allowed us to illustrate the dangers of simple logarithmic binning despite the limitations of the data.

Data for this reanalysis was kindly provided by Timothy Meehan, who may be contacted regarding access to the data by email at tdmeehan@entomology.wisc.edu.

Reanalysis of Enquist and Niklas (2001)

Our reanalysis of Enquist and Niklas (2001) focuses on size distribution results reported in their Figs. 2 and 3. The original analyses consisted of fitting diameter distributions for Alwyn Gentry's tree communities (Phillips and Miller 2002) using linear binning. We used the same data used in Enquist and Niklas 2001, provided by BJE. When binning the data we follow all of their original protocols including the exclusion of diameter bins with less than 5 individuals.

We reanalyzed this data using maximum likelihood estimation based on the Pareto distribution. The minimum diameter recorded in the Gentry data is 2.5cm, so $a = 2.5$. The results are qualitatively similar using the Truncated Pareto distribution with the maximum size set to the maximum size of an individual tree observed at the site.

Caveat: These results are not intended to comment directly on the form of the tree-size distribution for two reasons. First, tree size distributions may deviate from simple power laws (Coomes et al. 2003, Muller-Landau et al. 2006, Coomes and Allen 2007), and the analyses to establish whether the power-law is the best model for the distribution are beyond the scope of this paper. Second, the Gentry data (Phillips and Miller 2002) on which these results are based are actually are not ideal for evaluating the form of the tree size distribution. Gentry's data include single trees that branch below breast height and therefore have multiple stem measurements recorded for an individual. Unfortunately, the Gentry data does not identify the stems with the individual that they came from making it impossible to back calculate the basal stem diameter for an individual. As a result these data have typically been treated as if every stem is it's own individual (e.g., this is how Enquist and Niklas analyze the data in their original paper, and therefore how we do so here). As a result, there is likely a bias towards an overrepresentation of smaller individuals as small multiple stems from a single tree are counted as separate individuals. Thus, caution should be taken when interpreting these results in terms of actual individual tree diameter distributions. Regardless, these results clearly demonstrate the importance of using unbiased methods for estimating power law exponents.

The Gentry data is available from the Missouri Botanical Garden (http://www.mobot.org/MOBOT/research/gentry/transect.shtml).

LITERATURE CITED

Coomes, D. A., and R. B. Allen. 2007. Mortality and tree-size distributions in natural mixed-age forests. Journal of Ecology 95:27–40.

Coomes, D. A., R. P. Duncan, R. B. Allen, and J. Truscott. 2003. Disturbances prevent stem size-density distributions in natural forests from following scaling relationships. Ecology Letters 6:980–989.

Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. Nature 410:655–660.

Meehan, T. D. 2006. Energy use and animal abundance in litter and soil communities. Ecology 87:1650–1658.

Muller-Landau, H. C., R. S. Condit, K. E. Harms, C. O. Marks, S. C. Thomas, S. Bunyavejchewin, G. Chuyong, L. Co, S. Davies, R. Foster, et al. 2006. Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. Ecology Letters 9:589–602.

Petersen, H., and M. Luxton. 1982. A Comparative-Analysis Of Soil Fauna Populations And Their Role In Decomposition Processes. Oikos 39:287–388.

Phillips, O., and J. S. Miller. 2002. Global Patterns of Plant Diversity: Alwyn H. Gentry's Forest Transect Data Set. Missouri Botanical Garden Press, St. Louis, Missouri.

[Back to E089-052]

Ethan P. White, Brian J. Enquist, and Jessica L. Green. 2008. On estimating the exponents of power-law frequency distributions. *Ecology* 89:905–912.

Appendix C. Example of bias resulting from fitting discrete data with continuous maximum likelihood solutions.

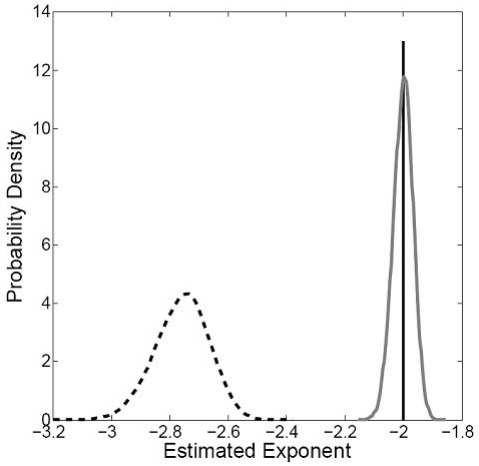

FIG. C1. Comparison of discrete Pareto (gray) and continuous Pareto (dashed black) MLE fits to discrete Pareto data based on 10,000 Monte Carlo runs with $\lambda = -2$, $a = 1$, and $n = 1000$. Using the continuous Pareto significantly overestimates λ. In cases where the minimum value of x is >>1 the two estimates should eventually converge, but in most cases in ecology where data is discrete the minimum values of x will be too small for this approximation to be valid.

[Back to E089-052]