Sampling species abundance distributions: Resolving the veil-line debate

Ryan A. Chisholm[a,b,*]

[a]*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08540, USA*
[b]*School of Botany, The University of Melbourne, Parkville, Vic. 3010, Australia*

Abstract

Preston's classic work on the theory of species abundance distributions (SADs) in ecology has been challenged by Dewdney. Dewdney contends that Preston's veil-line concept, relating to the shape of sample SADs, is flawed. Here, I show that Preston's and Dewdney's theories can be reconciled by considering the differing mathematical properties of the sampling process on logarithmic (Preston) versus linear (Dewdney) abundance scales. I also derive several related results and show, importantly, that one cannot reject the log-normal distribution as a plausible SAD based only on sampling arguments, as Dewdney and others have done.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Species abundance distribution; F.W. Preston; A.K. Dewdney; Log-normal; Veil line

1. Introduction

The veil-line concept was introduced by Preston (1948) to describe the relationship between the species abundance distribution (SAD) of an ecological community and the SAD of a sample drawn from that community. The veil-line concept states that, on a logarithmic scale of abundance, the sample SAD resembles a truncated version of the community SAD. The veil-line concept has become central to ecological theory: Hubbell (2001) discusses it in the development of his neutral theory; it is consistent with empirical data (e.g., Longino et al., 2002); and it is described in ecology textbooks (e.g., Molles Jr., 2005). The concept's usefulness, at least as a heuristic, has been accepted by many other authors (Grundy, 1951; Lobo and Favila, 1999; Ludwig and Reynolds, 1988; Magurran, 1988; Williams, 1964).

Dewdney (1997, 1998, 2003) has attacked the veil-line concept, claiming that Preston misunderstood the sampling process. Dewdney conjectures that, instead, sample SADs should closely resemble the form of the community SAD,

without truncation. Aspects of Dewdney's sampling theory have begun to gain traction (Alonso and McKane, 2004; Etienne and Alonso, 2005; McGill, 2003a; Williamson and Gaston, 2005), indicating that this issue requires resolution.

The veil-line concept is linked to the debate over which theoretical SAD best characterizes observed empirical SADs and associated underlying ecological processes. In the 1940s, Fisher and Preston proposed the first theoretical SADs: the log-series distribution (Fisher et al., 1943) and the log-normal distribution (Preston, 1948, 1962a, b, 1980). Subsequent theoretical developments have produced a profusion of different SADs, including the negative binomial (Pielou, 1975), the Zipf-Mandelbrot distribution (Mandelbrot, 1977; Zipf, 1965), the zero-sum multinomial (Hubbell, 2001), the logistic-J (Dewdney, 2003), the logit-normal (Williamson and Gaston, 2005) and the two- and three-group log-normal models (Gray et al., 2005; Ugland and Gray, 1982). Most proposed SADs are shaped similarly to the log-normal or log-series, but are arguably superior in that they are generated by underlying mechanistic models and/or provide better fits to observed data. Of the many criticisms that have been levelled at the log-normal over the years (Hughes, 1986; Ugland and Gray, 1982; Volkov et al., 2003; Williamson and Gaston, 2005), Dewdney's (1997, 1998, 2003) is unique in that he effectively rejects anything even resembling a log-normal

*Corresponding author. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08540, USA.
Tel.: +1 609 258 3987.
E-mail address: chisholm@princeton.edu.

distribution as a possible SAD, because of the supposed invalidity of the veil-line concept and the supposed failure of untruncated log-normals to appear in sample SADs.

My intention here is to evaluate Dewdney's case against the veil-line concept, to reconcile the theories of Dewdney and Preston, and to discuss the resulting implications for the ongoing log-normal debate.

2. Preston's theory

Preston (1948) considers SADs on a logarithmic scale of abundance. On this scale, the lowest abundance class represents species with 1–2 individuals, the next represents species with 2–4 individuals, the next represents species with 4–8 individuals, and so on. Species with exactly 2^i individuals (where i is a non-negative integer) are split between adjacent abundant classes. Justifications for the use of a logarithmic abundance scale are (i) it reduces the noise in SADs, thus illuminating patterns in the data, and (ii) because "(c)ommonness is a *relative* matter" (Preston, 1948), it is natural to use a scale where the abundance of one category of species is approximately double that of the preceding class. Preston's logarithmic abundance scale is widely used (e.g., Hubbell, 2001; Longino et al., 2002), although it should be noted that while Preston does not plot the "<1" abundance class (i.e., the abundance class that contains half the species with one individual), other authors do (e.g., McGill, 2003b; Volkov et al., 2003), and that some authors prefer different abundance class boundaries (for reviews of binning methods see Gray et al., 2006; Williamson and Gaston, 2005).

Preston (1948) notes that the expected sample abundance of a species is proportional to its abundance in the community. To quantify this, define r as the sample proportion n/N, where n is the number of individuals in the sample, and N the number of individuals in the community. A species with j individuals in the community has expected sample abundance rj. Preston (1948) shows that, despite sampling error, most species, especially those in higher abundance classes, achieve sufficiently close to proportional representation such that the sampling process approximately preserves the shape of the SAD on a logarithmic scale—the exception to this is that species in the lowest abundance categories tend not to appear in the sample, so the sample curve appears truncated.

This leads to the concept of the veil line. For a given sample size, the veil line is placed at abundance $1/r$ on the horizontal axis of the community SAD, so that $\log_2(1/r)$ abundance categories lie to the left of the veil line. Species in these abundance categories tend not to appear in the sample (their expected sample abundances are less than one), so the veil line represents the point at which the distribution is apparently truncated by the sampling process. On a $\log_2$ scale, doubling the sample size moves the veil line one abundance category to the left, effectively revealing more of the underlying community SAD.

Preston recognizes that some species to the left of the veil line do, by chance, appear in the sample, just as some species to the right of the veil line sometimes do not appear in the sample. However, these effects largely cancel each other out, and do not undermine the veil line's utility as an approximate tool for studying SADs.

The veil-line concept provides an explanation as to why many sample SADs exhibit a truncated log-normal shape when the underlying community SAD is supposedly log-normal (Preston, 1948).

3. Dewdney's theory

Dewdney (1998) claims that the veil-line concept is based on a misunderstanding of the sampling process. The probability that a species has abundance k in the sample, given that it has abundance j in the community, is given for the case of sampling without replacement (Dewdney, 1998) by the sample transformation function

$$t(k,j) = \frac{\binom{j}{k}\binom{N-j}{n-k}}{\binom{N}{n}} \qquad (1)$$

or for the case of sampling with replacement by

$$t(k,j) = \binom{n}{k}\left(\frac{j}{N}\right)^k\left(1-\frac{j}{N}\right)^{n-k}$$
$$= t^*(k,p) = \binom{n}{k}\left(\frac{p}{n}\right)^k\left(1-\frac{p}{n}\right)^{n-k}. \qquad (2)$$

In this expression, $p = jn/N$ represents the expected abundance of the species in the sample, thereby eliminating any reference to the community size, N. Indeed, as N (and correspondingly, j) becomes large, the expression for $t(k,j)$ in Eq. (1) approaches that of Eq. (2). Intuitively, this means that sampling without replacement is asymptotically equivalent to sampling with replacement as the ratio of community size to sample size increases.

Furthermore,

$$\lim_{n\to\infty} t^*(k,p) = \frac{e^{-p}p^k}{k!}. \qquad (3)$$

This is the Poisson distribution, the expression used by Preston (1948) to represent the relationship between community and sample abundance. This limiting form of the sample transformation function is also used by Dewdney (1998).

The number of species with abundance j in the community that can be expected to have abundance k in the sample is given by

$$f(k,j) = t(k,j)g(j), \qquad (4)$$

where $g(j)$ is the total number of species with abundance j in the community, and $t(k,j)$ is the appropriate form of the sample transformation function (Dewdney, 1998).

The expression $f(0, j)$ then represents the number of species with abundance j in the community that can be expected not to show up in the sample at all. Furthermore,

$$h(j) = g(j) - f(0,j) \tag{5}$$

represents the number of species with abundance j in the community that can be expected to show up in the sample. This leads to the concept of a "veil curve" (Dewdney, 1998). Species under the veil curve can be expected to be present in a sample, those above can be expected to be absent.

Because the veil curve does not coincide with the veil line, Dewdney (1998) claims that he has disproved Preston's veil-line concept, which was allegedly "based on a misunderstanding of the sampling process itself". However, as noted earlier, Preston acknowledges that the veil-line concept is only a useful approximation.

Having supposedly disproved the veil-line concept, Dewdney (1998) rejects the log-normal as a possible SAD because "nothing like an untruncated log-normal distribution ever shows up in species-abundance histograms" (a point that is itself contentious, e.g., Hubbell, 2001; Longino et al., 2002).

The true shape of the sample curve can be investigated mathematically as follows. Using Eq. (4) the expected total number of species with abundance k in the sample is

$$f(k) = \sum_{j=\max(1,k)}^{N-n+k} f(k,j). \tag{6}$$

While neither Preston nor Dewdney expresses this mathematically, Preston gives a full worked example demonstrating his understanding of it, and Dewdney performs the same operation but on a continuous scale of abundance (following Fisher et al., 1943; Pielou, 1969):

$$f(k) = \int_0^\infty \frac{e^{-rx}(rx)^k}{k!} g(x) \, dx. \tag{7}$$

Here, the discrete variable j in (6) has been replaced by the continuous variable x, and $r = n/N$, as before. Throughout this analysis, it is natural to restrict variables representing abundances (x, k, j) to non-negative values. It should be borne in mind that Eq. (7) was derived from Eq. (3), which described sampling with replacement; these equations only describe sampling without replacement in the limit as $r \to 0$.

Dewdney's (1998) analysis of the sample transformation (7) suggests that the sampling process preserves "distribution formulas to within constants and parameters." This prompts him to reject truncation and instead suggest that "species abundance curves that result from sampling echo the shape of the distribution of the particular community being sampled" (Dewdney, 2003).

4. Reconciling the two theories

Dewdney's literal interpretation of the veil-line concept is an immediately obvious source of disagreement between him and Preston. Preston (1948, p. 267) makes it plain that the veil-line concept "is an approximation only, and the postulate that the sample curve is essentially a decapitated version of the universe curve is likewise only an approximation; but it seems that it is sufficiently close for our present purposes." Bearing in mind that the veil-line concept was only ever intended as a useful heuristic, I now present the mathematical arguments that reconcile the approaches of Preston and Dewdney. I then discuss a few finer points of Dewdney's theory.

4.1. The sample transformation on a linear scale

Dewdney (1998) provides the seeds of an analysis of the effect of sampling on a linear abundance scale by observing that the sample transformation function (7) maps polynomials to polynomials and conjecturing that the transformation therefore preserves the form of any analytic function (i.e., he conjectures that the sample SAD differs from the community SAD only in the value of its constants). Here, I show that Dewdney's conjecture as stated is false, but that a weaker version of it holds.

4.1.1. Dewdney's conjecture

Dewdney's (1998) conjecture was that the sample transformation "preserves the form of any analytical function," and that "the formula of the sample distribution will differ from the formula of the parent distribution only in the value of its constants." This conjecture is readily disproved by counterexample, using the precise community SAD that Dewdney considered (this is a Gamma distribution with shape parameter 2, multiplied by the total number of species in the community):

$$g(x) = c_1 x e^{-c_2 x}. \tag{8}$$

To obtain the expected sample SAD we apply the sample transformation (7)

$$\begin{aligned}
f(k) &= \int_0^\infty \frac{e^{-rx}(rx)^k}{k!} g(x) \, dx \\
&= \int_0^\infty \frac{e^{-rx}(rx)^k}{k!} c_1 x e^{-c_2 x} \, dx \\
&= c_1 \int_0^\infty \frac{e^{-(r+c_2)x} r^k x^{k+1}}{k!} \, dx.
\end{aligned} \tag{9}$$

Integrating by parts $k+1$ times yields

$$f(k) = c_1 \frac{(k+1)}{(r+c_2)^2} \left(\frac{r}{r+c_2} \right)^k. \tag{10}$$

If we let $c_1' = c_1/(r + c_2)^2$ and $c_2' = \ln[(r + c_2)/r]$, we can express this as

$$f(k) = c_1'(k + 1)e^{-c_2'k}, \tag{11}$$

where $k \geqslant 0$. (Dewdney fails to produce Eq. (11); instead he has $f(k) = c_1'ke^{-c_2'k}$. Note that Dewdney's solution cannot be converted to the correct form simply by making the change of variables $l = k-1$, because this would violate the condition that variables representing abundances be restricted to non-negative values.)

That the form of Eq. (11) is different from that of Eq. (8) immediately disproves Dewdney's conjecture about the shape of sample SADs. It should be noted that the left-shift induced by the $k + 1$ factor in (11) is not related to the veil-line effect; rather it produces a smoothing effect that Dewdney (1998) recognizes and refers to as the "sigmoid effect".

4.1.2. An asymptotic approximation to the sample transformation

Here I prove that, as k becomes large, the sample SAD $f(k)$ has the same shape as the community SAD $g(j)$, but is compressed by a factor $1/r$ along the horizontal axis and expanded by a factor $1/r$ along the vertical axis.

First, consider the following function:

$$\phi(j) = aj^m,$$

where m is a non-negative integer and a is a real number. Using the basic result

$$\int_0^\infty e^{-rx}\frac{(rx)^k}{k!}\,\mathrm{d}x = \frac{1}{r},$$

the sample transformation can be applied to $\phi(j)$ to yield

$$\begin{aligned}
\theta(k) &= \int_0^\infty e^{-rx}\frac{(rx)^k}{k!}\phi(x)\,\mathrm{d}x \\
&= \int_0^\infty e^{-rx}\frac{(rx)^k}{k!}ax^m\,\mathrm{d}x \\
&= a\frac{(k+m)!}{k!r^m}\int_0^\infty e^{-rx}\frac{(rx)^{k+m}}{(k+m)!}\,\mathrm{d}x \\
&= a\frac{(k+m)!}{k!r^m}\frac{1}{r} \\
&= a\frac{(k+m)(k+m-1)\cdots(k+1)}{r^{m+1}}.
\end{aligned}$$

Taking the limit of this expression as k becomes large (relative to m) gives

$$\begin{aligned}
\lim_{k\to\infty}\theta(k) &= \lim_{k\to\infty}a\frac{(k+m)(k+m-1)\cdots(k+1)}{r^{m+1}} \\
&= a\left(\frac{k}{r}\right)^m\frac{1}{r} = \frac{1}{r}\phi(k/r).
\end{aligned}$$

So, assuming that k is sufficiently large and neglecting higher-order terms, we can write

$$\int_0^\infty \frac{e^{-rx}(rx)^k}{k!}ax^m\,\mathrm{d}x \approx a\left(\frac{k}{r}\right)^m\frac{1}{r}. \tag{12}$$

Now, let $g(j)$ be a general community SAD representable by a Taylor series

$$g(j) = \sum_{m=0}^\infty a_m j^m.$$

Take the sample transformation

$$\begin{aligned}
f(k) &= \int_0^\infty e^{-rx}\frac{(rx)^k}{k!}g(x)\,\mathrm{d}x \\
&= \int_0^\infty e^{-rx}\frac{(rx)^k}{k!}\sum_{m=0}^\infty (a_m x^m)\,\mathrm{d}x \\
&= \sum_{m=0}^\infty\left(\int_0^\infty e^{-rx}\frac{(rx)^k}{k!}a_m x^m\,\mathrm{d}x\right)
\end{aligned}$$

and using (12), obtain

$$f(k) \approx \sum_{m=0}^\infty\left[a_m\left(\frac{k}{r}\right)^m\frac{1}{r}\right] = \frac{1}{r}g(k/r). \tag{13}$$

This gives the desired result. Because approximation (12) was derived by assuming that k is much larger than m, approximation (13) can be expected to hold well if the first few terms of the Taylor series dominate. In practice, the approximation does appear to hold reasonably well (Fig. 1). Also, as expected, the approximation converges to the true SAD as $k\to\infty$, and, because fewer species appear in the lower abundance categories with larger samples, the approximation tends to improve with increasing sample size.

4.2. The sample transformation on a logarithmic scale

Here, I show that the transformation on a linear scale discussed above is equivalent to a veil-line effect on a logarithmic scale. Let $g(j)$ be the community SAD. Then, as given above, the approximation to the sample SAD is

$$f(k) \approx \frac{1}{r}g(k/r). \tag{14}$$

Now transform to a logarithmic abundance scale by letting $u = \log_2 k$:

$$f(2^u) \approx \frac{1}{r}g(2^u/r) = \frac{1}{r}g(2^{u-\log_2 r}) \tag{15}$$

and then let $F(u) = f(2^u)$ and $G(u) = g(2^u)$:

$$F(u) \approx \frac{1}{r}G(u - \log_2 r) = \frac{1}{r}G(u + \log_2(1/r)). \tag{16}$$

Because $r < 1$, $\log_2(1/r) > 0$, so $F(u)$, the sample SAD on a logarithmic abundance scale, is shifted left relative to $G(u)$, the community SAD on a logarithmic abundance scale. Thus, a horizontal compression by a factor $1/r$ on a linear scale of abundance (Eq. (14) and Fig. 2) is equivalent to a left shift of $\log_2(1/r)$ abundance categories on a logarithmic scale of abundance (Eq. (16) and Fig. 3) The left shift produces the veil-line effect, and the extent of the shift is in exact agreement with Preston's theory. This, in
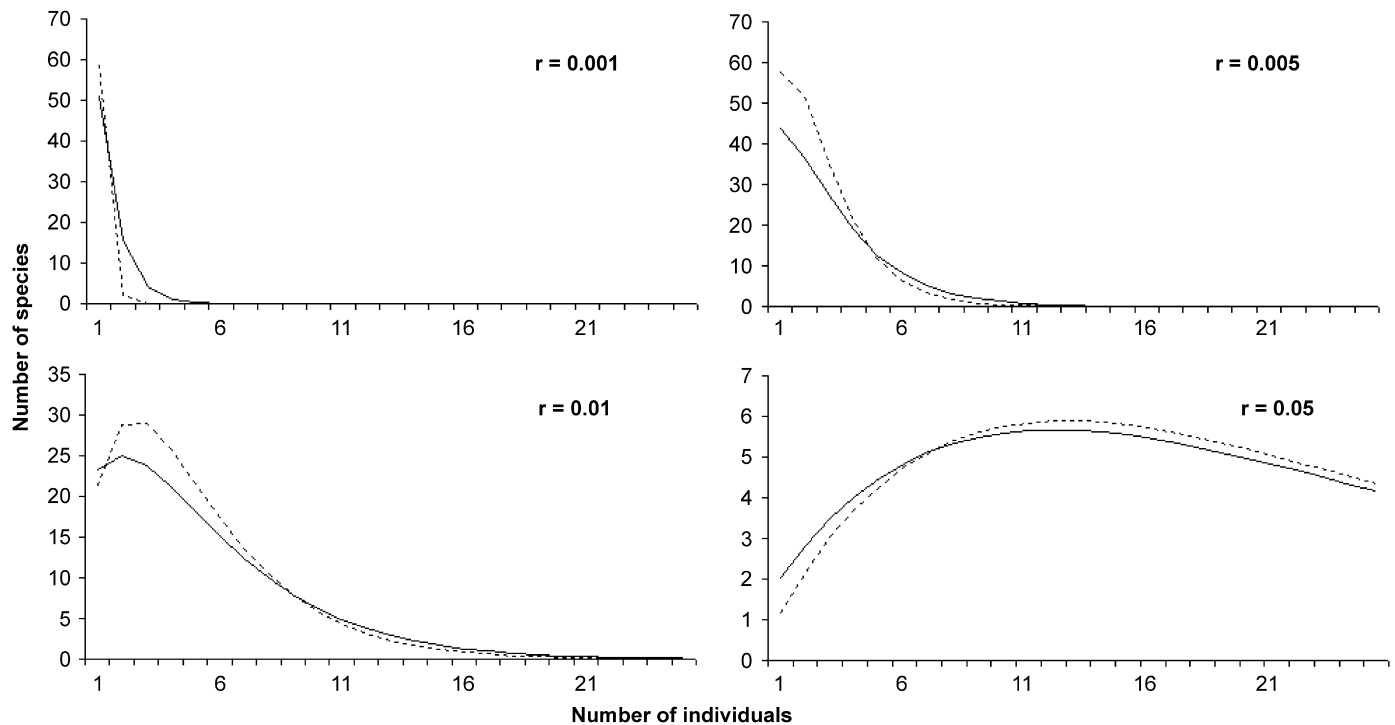
Fig. 1. SADs of samples (solid lines) taken from a community distribution with SAD given by the gamma distribution of equation (8) with $c_1 = 0.0032$ and $c_2 = 0.004$ (this corresponds to a community with 100,000 individuals and 200 species), compared to approximate SADs (dashed lines) as given by Eq. (13). The sample proportion r is shown on each graph.

conjunction with the result from the previous section, reconciles the sampling theory of Preston (a veil-line effect on a logarithmic scale), with the sampling theory of Dewdney (preservation of distribution shape on a linear scale).

4.3. The proportion of species to the left of the veil line

I have proven the general validity of the veil-line concept as an approximation of the sampling process on a logarithmic abundance scale. However, it remains unclear to what extent the species to the left of the veil line actually correspond to those that do not appear in the sample. Dewdney (1998) applies the sample transformation with $k = 0$ to obtain a "veil curve", which shows exactly those species that are not expected to appear in a sample of a given size. Dewdney's veil curve does not coincide with veil line, and this gives him further grounds for rejecting the veil-line concept.

In actual fact, the correspondence between Dewdney's veil curve and the veil line depends upon the sample size (Fig. 4). For small sample sizes, there are a great many species that are absent from the sample, and most of these are indeed to the left of the veil line. For larger sample sizes, the correspondence between species absent from the sample and those to the left of the veil line is much poorer, but there are fewer species absent overall, so the effect on the shape of the sample distribution is less marked.

It should be noted that predicting which species will be absent from a sample and predicting the shape of the sample curve are related but not identical problems. At large sample sizes, the approximations presented here appear to predict the shape of the sample curve well (Fig. 1) but the identity of the absent species poorly (Fig. 4); at small sample sizes the approximations appear to predict the shape of the sample curve less well (Fig. 1) but the identity of the absent species better (Fig. 4). The extent to which this applies to different SADs is a topic for future research.

4.4. The log-normal distribution

Dewdney's rejection of the log-normal distribution as a possible SAD stems from his claims that (i) the veil-line concept is invalid; and (ii) log-normal distributions do not appear in real data. We have seen already that Dewdney's rejection of the veil-line concept was unjustified. To assess the reasoning behind his claim that log-normal distributions do not appear in real data, we must study log-normal distributions on linear abundance scales as used by Dewdney. I show here that, even if the underlying community SAD is log-normal, there are many circumstances under which this is not apparent from the sample SAD because the interior mode of the distribution is not visible.
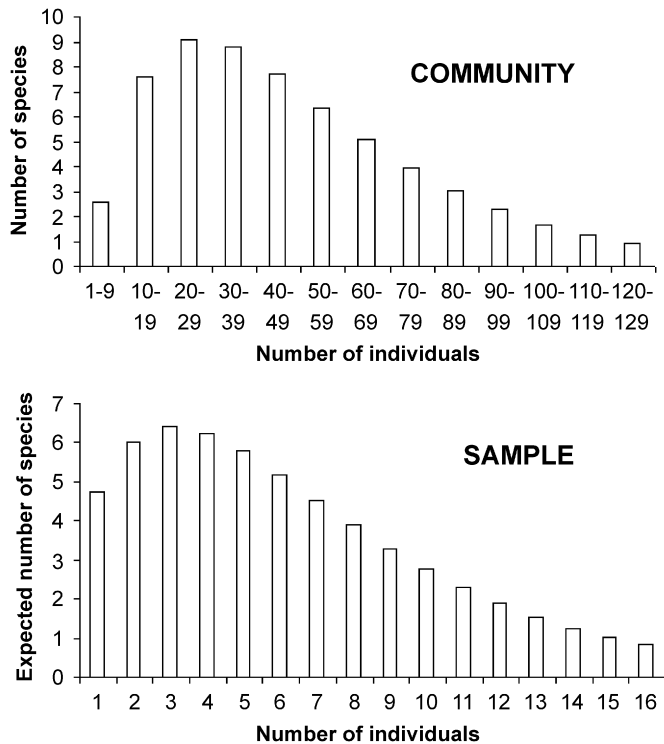
Fig. 2. Top panel: a community SAD as given by Eq. (8), with $c_1 = 0.1$, $c_2 = 0.04$. Bottom panel: the expected SAD of a sample of 1/8 of the community. The sample SAD was computed using Eq. (7) (see also Eq. (11) and accompanying text). The distributions are presented on linear abundance scales. The shape of the sample SAD is similar to that of the community SAD but the distribution is compressed along the abundance axis. Note that the number of species in each category is non-integral because the graphs are based on theoretical distributions.
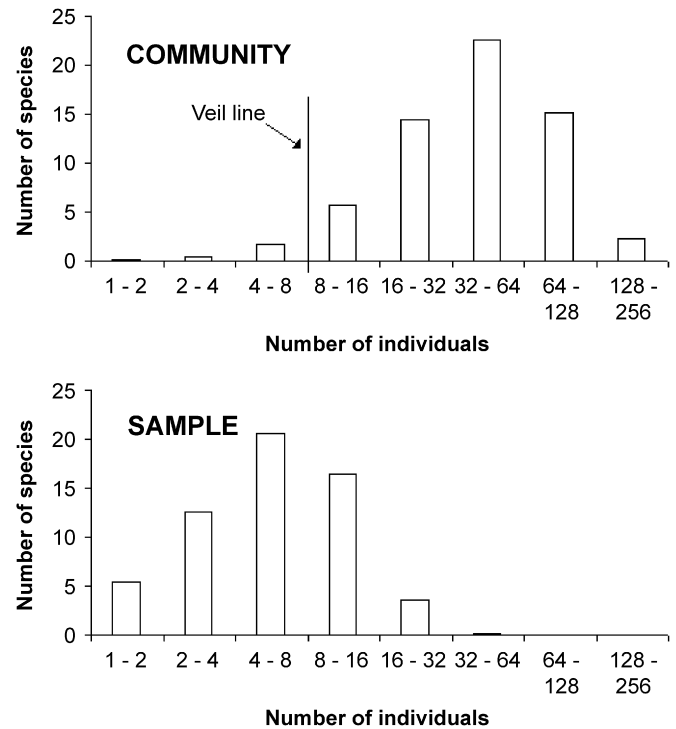

Fig. 3. The distributions from Fig. 2 presented on a logarithmic scale, following Preston (1948). The sample SAD resembles a truncated version of the community SAD—this is Preston's veil-line effect.
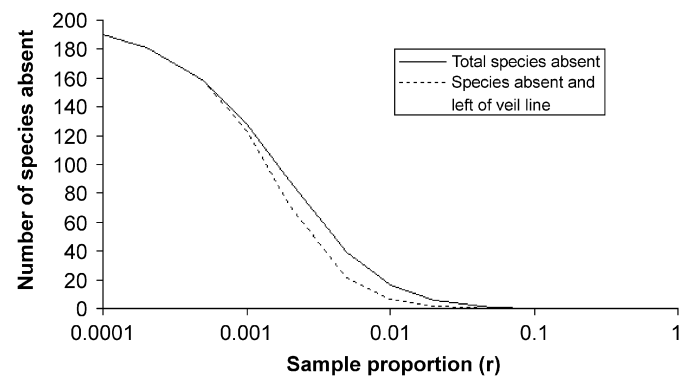

Fig. 4. Number of species to the left of the veil line expected to absent from a sample of a given size (dashed line) compared to the total number of species expected to be absent (solid line). The community SAD from which the samples are drawn is the same as the one from Fig. 1.

Let us consider a log-normal distribution in the form used by Volkov et al. (2003):

$$\langle \Phi_j \rangle = \frac{M}{j} \exp\left[-\frac{1}{2}\left(\frac{\log_2(j/j_0)}{\sigma} \right)^2 \right]. \tag{17}$$

Here, $\langle \Phi_j \rangle$ is the expected frequency of species containing j individuals, and M, j_0, and σ are fitted parameters. The log-normal distribution has an interior mode at

$$j = j_M = j_0 2^{-\sigma^2/\log_2 e}. \tag{18}$$

Even though, mathematically, a community with a log-normal SAD must have an interior mode, in practice the mode may not be visible on a discrete linear scale: this occurs when $j_M < 1$. The condition $j_M < 1$ is satisfied for a broad range of the parameters M, j_0 and σ. For example, consider the log-normal distribution applied to the Barro Colorado Island tree data by Volkov et al. (2003) with parameters $M = 46.29$, $j_0 = 20.82$ and $\sigma = 2.98$. From (18), we see that the mode occurs at $j_M = 0.292$—too low to be observable on a discrete linear scale (Fig. 5). It is worth noting that, on a logarithmic scale, the apparent mode at abundance category 32 (see the figure in Volkov et al., 2003) is an artefact of the binning process and does

not equate to the mode of the continuous mathematical function as computed above.

Even if $j_M > 1$, and the interior mode of the community SAD is visible, there are two factors that make it still unlikely that we observe an interior mode in the sample SAD: (i) the resolution of the sample SAD is lower by a factor r, so we do not observe the interior mode in the sample SAD unless $j_M > 1/r > 1$; and (ii) the sigmoid effect (first noted by Dewdney) causes the interior mode to be slightly displaced to the left in the sample compared to the community.
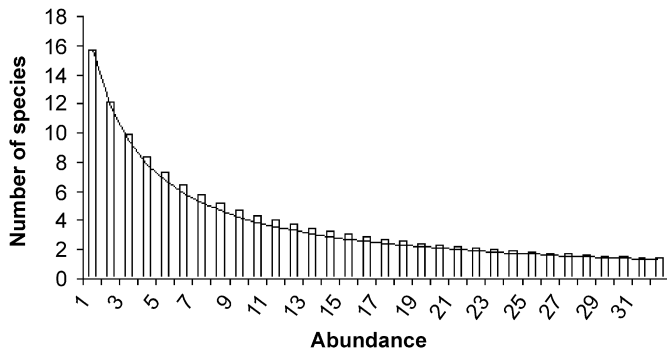
Fig. 5. The bars show the best-fit log-normal to the Barro Colorado Island tree data (Volkov et al., 2003) presented on a linear scale. Note that the interior mode of the distribution is not visible because it lies to the left of the first bar. The curve shows a general logistic-J distribution (Dewdney, 2003) with $\varepsilon = 2.0$ and $\Delta = 860$. The similarity of the two distributions is apparent ($r^2 > 0.99$).

It is clear then, that the absence of interior modes in sample SADs does not constitute evidence against the log-normal or other log-normal-shaped distributions as candidate SADs.

4.5. The logistic-J distribution

The logistic-J distribution is a truncated hyperbola that was proposed by Dewdney (2003) as a model for community SADs following his rejection of the log-normal. On a discrete scale, the logistic-J actually provides a close fit to many log-normal distributions (e.g., Fig. 5; $r^2 > 0.99$). Thus, the logistic-J and log-normal are not mutually exclusive.

5. Conclusions

Here, I have made several contributions to the development of SAD sampling theory. Firstly, I have shown that, on a linear scale, the expected sample SAD is asymptotically equivalent to a compressed version of the community SAD. Secondly, I have shown that, on a logarithmic scale, this is equivalent to a veil-line effect. Thirdly, I have shown that, although the veil-line concept is only a useful heuristic tool, it does accurately predict which species will be absent from small samples. These three developments reconcile the well-established sampling theory of Preston (1948, 1962a b, 1980) with the emerging sampling theory of Dewdney (1997, 1998, 2003) and with the work of others that builds upon the two (e.g., Pielou, 1969, 1975; Williamson and Gaston, 2005).

Lastly, I have shown that sampling arguments alone cannot be used to reject the log-normal distribution as a possible SAD. I have not engaged other aspects of the log-normal debate, such as whether the log-normal distribution is a suitable null model for SADs or whether it has a biological interpretation (Hughes, 1986; Ugland and Gray, 1982; Volkov et al., 2003; Williamson and Gaston, 2005). On these latter points, let me say only that, ideally, theoretical SADs should have underlying mechanistic models that are both biologically realistic and parsimonious, and that when comparing competing SAD models, the Bayesian approach of Etienne and Olff (2004, 2005), which yields a multivariate probability of observing a particular SAD, is preferable to simple goodness-of-fit tests (such as the χ^2) that are difficult to interpret biologically (e.g., McGill, 2003b; Volkov et al., 2003).

Acknowledgments

References

Alonso, D., McKane, A.J., 2004. Sampling Hubbell's neutral theory of biodiversity. Ecol. Lett. 7, 901–910.

Dewdney, A.K., 1997. A dynamical model of abundances in natural communities. COENOSES 12, 67–76.

Dewdney, A.K., 1998. A general theory of the sampling process with applications to the "Veil Line". Theor. Popul. Biol. 54, 294–302.

Dewdney, A.K., 2003. The stochastic community and the logistic-J distribution. Acta Oecol. 24, 221–229.

Etienne, R.S., Alonso, D., 2005. A dispersal-limited sampling theory for species and alleles. Ecol. Lett. 8, 1147–1156.

Etienne, R.S., Olff, H., 2004. A novel genealogical approach to neutral biodiversity theory. Ecol. Lett. 7, 170–175.

Etienne, R.S., Olff, H., 2005. Confronting different models of community structure to species-abundance data: a Bayesian model comparison. Ecol. Lett. 8, 493–504.

Fisher, R.A., Corbet, A.S., Williams, C.B., 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. J. Anim. Ecol. 12, 42–58.

Gray, J.S., Bjørgesæter, A., Ugland, K.I., 2005. The impact of rare species on natural assemblages. J. Anim. Ecol. 74, 1131–1139.

Gray, J.S., Bjørgesæter, A., Ugland, K.I., 2006. On plotting species abundance distributions. J. Anim. Ecol. 75, 752–756.

Grundy, P.M., 1951. The expected frequencies in a sample of an animal population in which the abundances of species are log-normally distributed. 1. Biometrika 38, 427–434.

Hubbell, S.P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton, NJ.

Hughes, R.G., 1986. Theories and models of species abundance. Am. Nat. 128, 879–899.

Lobo, J.M., Favila, M.E., 1999. Different ways of constructing octaves and their consequences on the prevalence of the bimodal species abundance distribution. Oikos 87, 321–326.

Longino, J.T., Coddington, J., Colwell, R.K., 2002. The ant fauna of a tropical rain forest: estimating species richness three different ways. Ecology 83, 689–702.

Ludwig, J.A., Reynolds, J.F., 1988. Statistical Ecology: A Primer on Methods and Computing. Wiley, New York, USA.

Magurran, A.E., 1988. Ecological Diversity and Its Measurement. Princeton University Press, Princeton, NJ, USA.

Mandelbrot, B.B., 1977. Fractals, Fun, Chance and Dimension. Freeman, New York.

McGill, B.J., 2003a. Does mother nature really prefer rare species or are log-left-skewed SADs a sampling artefact. Ecol. Lett. 6, 766–773.

McGill, B.J., 2003b. A test of the unified neutral theory of biodiversity. Nature 422, 881–885.

Molles Jr., M.C., 2005. Ecology. McGraw-Hill, New York, NY.

Pielou, E.C., 1969. An Introduction to Mathematical Ecology. Wiley, New York.

Pielou, E.C., 1975. Ecological Diversity. Wiley, New York.

Preston, F.W., 1948. The commonness and rarity of species. Ecology 29, 254–283.

Preston, F.W., 1962a. The canonical distribution of commonness and rarity: part II. Ecology 43, 410–432.

Preston, F.W., 1962b. The canonical distribution of commonness and rarity: part I. Ecology 43, 185–215.

Preston, F.W., 1980. Noncanonical distributions of commonness and rarity. Ecology 61, 88–97.

Ugland, K.I., Gray, J.S., 1982. Log-normal distributions and the concept of community equilibrium. Oikos 39, 171–178.

Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2003. Neutral theory and relative species abundance in ecology. Nature 424, 1035–1037.

Williams, C.B., 1964. Patterns in the Balance of Nature and Related Problems in Quantitative Ecology. Academic Press, London, UK.

Williamson, M., Gaston, K.J., 2005. The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution. J. Anim. Ecol. 74, 409–422.

Zipf, G.K., 1965. Human Behavior and the Principle of Least Effort. Hafner, New York.