COENOSES 12(2-3):67-76, 1997

A DYNAMICAL MODEL OF ABUNDANCES IN NATURAL COMMUNITIES

A. K. Dewdney

Environmental Science Program & Department of Computer Science, The University of Western Ontario,
London, Ontario, Canada N6A 5B7

Keywords: Logistic distribution, Model, Species abundance.

Abstract: A simple dynamical community model mimics species interactions and produces, when it reaches equilibrium, a distribution of species abundances that is indistinguishable from the "J-curves" regularly seen by field biologists in samples and surveys. Analysis of the model reveals an equilibrium solution that is inverse linear in form. The theoretical distribution that results, here called the "logistic" distribution, is defined formally as the inverse linear function under a translation of axes which forms its two parameters.

The initial hypothesis, that the resemblance between computer and natural abundance distributions is more than coincidental, has led to a near-cottage industry of new hypotheses, the three main ones being:

1. The logistic distribution describes abundance in natural communities.
2. The sampling process transforms a community abundance distribution into a sample abundance distribution by a scale transformation in both axes.
3. In natural communities a major influence on species abundance is the random (i.e., unpredictable) fluctuation in populations.

This report describes the computer model, analysis of the model, and a brief summary of progress toward the establishment of Hypothesis 1. Hypothesis 2 is susceptible of mathematical proof. It abolishes the idea of a "veil-line" as a sampling phenomenon. The third hypothesis is discussed briefly.

Introduction

In a study of stream benthic microbiota, the author (1996) noticed that the 222 species sampled followed an abundance distribution that field biologists call the "J-curve." Although this curve arose from a sample and did not necessarily represent the benthic population of microorganisms as a whole, it led to the question of whether they might have this abundance distribution generally.

Curiosity about the phenomenon led to a computer model in which stream benthic microorganisms would encounter each other at random. In the dog-eat-dog world of the stream benthos, it seemed reasonable to designate one of the microorganisms as predator, the other as prey and to specify that the ingesting organism would enjoy a reproductive boost immediately following the trophic act.

An initial model that incorporated these behaviors was defined and implemented by a computer program which was also capable of displaying the current species/abundance histogram every 100 encounter cycles. Running the model with 200 species, it was something of a surprise to see the stream abundance distribution, or something closely resembling it, appear after a few hundred display cycles. The curve appeared on every occasion.

To probe the dynamical system for its potential as a general population model, the system was complicated by the addition of food webs and multiple habitats (patches). It was an even greater surprise to discover how robust the algorithmic process was. Under a wide variety of complications, all leading in the direction of increased realism, it continued to produce the "J-curve" reliably. The reason for the robustness proved to lie not with the details of this or that implementation, but with a more general phenomenon related to seemingly random fluctuations in component populations.

At the time of writing, some 30 biosurveys (drawn at random from the literature) have been examined. They all show characteristic J-curve distributions that not only resemble the distribution of my stream benthic community, but resemble equally the productions of the dynamical system in any of its incarnations. These results, in spite of their preliminary nature, have made it necessary to propose a new model of species abundance in communities, joining several models already proposed (Magurran 1988).

A multi-species dynamical model

Let a community C consist of N individuals distributed among m species. The computer model described here is dynamical, stochastic and individual-based. The basic dynamical step selects two individuals at random (the stochastic element) from the general population of N individuals.

To guarantee randomness at the individual level, selections are made on a proportional basis by the interval division technique: A procedure invoked by the main program selects a random number between 1 and N, then counts through the species by their abundances until it comes to the species where the count first equals or exceeds the randomly drawn number. In the algorithmic conventions employed here, variable names are set in italics, the assignment operator is indicated by a left-arrow, and pseudo-language elements are set in bold face:

```
procedure Select Species
  select a random number k from [1, N]
  Count ← 0, s ← 1
  repeat
    Count ← Count + abundance of s
    s ← s + 1
  until count ≥ k
  return s
```

The basic algorithm for the dynamical system is extremely simple. Within a user-terminated loop, the procedure just defined makes a random selection of two individuals, returns their species id numbers, and then alters their abundances according to the following scheme.

```
repeat
  Predator ← Select Species
  Prey ← Select Species
  increment abundance of Predator
  decrement abundance of Prey
until key pressed
```

Abundances of the m species are stored in an array called *Abundance*.

Anyone with programming experience should be capable of writing a program based on this algorithmic description. The program was written by the author in Turbo Pascal. It runs on a 486 computer and incorporates the basic algorithm within a display loop of 100 iterations. It also includes a histogram-drawing procedure that uses text graphics to show the number of species at each abundance value.

At the end of a run, the program also compiles the average distribution of all abundances from the onset of equilibrium until the user presses the termination key. It displays average abundances in the same graphic format as the inner display cycle does. The histogram shown in Figure 1 shows a typical final output of averaged species numbers for this version of the model.

The distributions produced within individual display cycles while the model is running are rarely this smooth, not surprisingly, since they reflect random events directly and, if anything, more closely resemble field samples.

Called the multi-species logistical (MSL) model, this very simple system reliably produces the curve shown in Figure 1 after a few hundred iterations of the basic display cycle.

In a typical run, the program is initialized to 200 species with, say, 20 individuals each. At first, the histogram appears as a single large spike at abundance 20. By a few hundred iterations, this spreads out into a binomial distribution that, as it collapses further, produces a tail into the high end of the abundance axis and begins to pile up at the low end. The resulting asymmetry eventually dominates the binomial form and ends by erasing it completely as the system nears

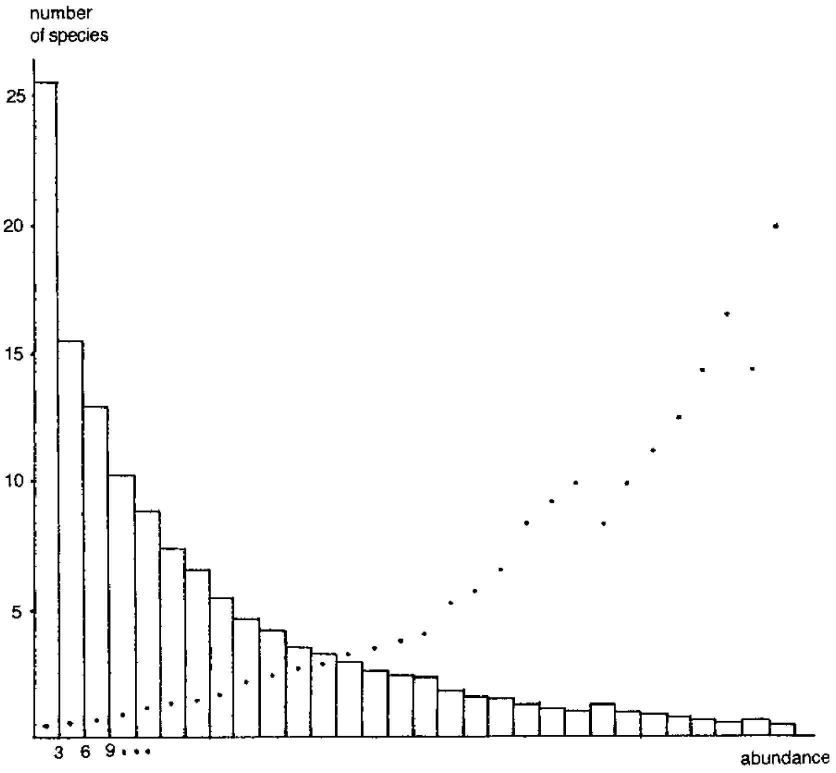


Figure 1. A typical output of the basic dynamical model running with 135 species. Species frequencies for each abundance class are averaged from the onset of equilibrium until the end of a run. Accompanying the histogram is a point-plot of the inverted frequencies.

the equilibrium shape shown in Figure 1. The same shape emerges if the initial populations are set to a uniform random distribution instead of the spike distribution. In other words, the shape in question is a genuine property of the system, and not an artifact of the initial distribution. The model can also be run with much larger abundances that would be more typical of the community inhabiting a particular region or patch. But the equilibrium distribution of this community is far more "strung out" with rarely more than one species enjoying a particular abundance at a time. Such distributions can only become visually accessible under an operation of high-order grouping in which each display category spans 10 or more abundances.

The histogram in Figure 1 uses low-order grouping to make the display readable, but the shape is not an artifact of grouping. It appears if one displays abundance categories 1, 2, 3, etc, instead of categories 1-3, 4-6, and so on, as was done for Figure 1.

Complications of this simple form of the dynamical system produce results that, at a glance, are indistinguishable from the MSL model.

The first of these employs an abstract food web in which species s is allowed to prey only on species s+1, s+2, . . . , s+15, for example (Figure 2). In such a cyclic food web, the system takes somewhat longer to reach equilibrium, but the resulting shape is the same.

In the next version of the model (Figure 2), all species are divided into three large categories, heterotrophs, autotrophs, and saprobes. In this version, heterotrophs are allowed to eat each other as well as the other two kinds of species, autotrophs receive energy automatically (sunlight), but are allowed to "eat" each other (think of competitive exclusion), and saprobes are allowed to eat either of the other two, following simulated death events. This food web version also produces the same species abundance distribution as the basic MSL model.

Additional trials of the system involved fractional trophism in which the number of individuals did not increase or decrease, but the biomass increased (fractionally), instead. Again the ramp-shaped curve appeared.

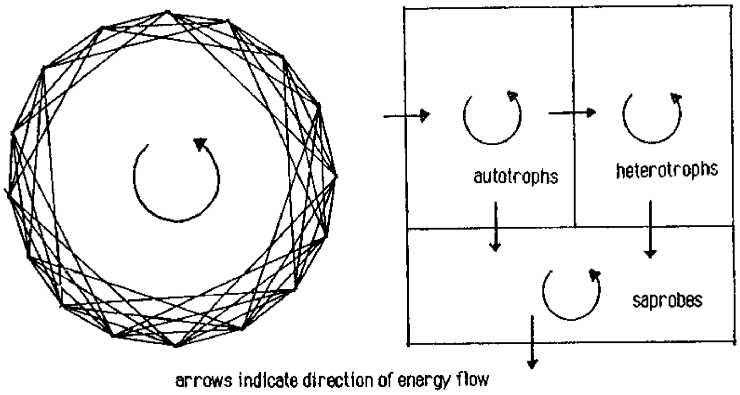
The program contains a switch that enables the user to specify whether extinctions will be allowed or not. With no extirpations allowed, the pile-up of species at the low abundance end soon comes into equilibrium with the rest of the distribution and the number of species in the lowest abundance category stabilizes.

With the extinction switch on, species naturally disappear, leaking out the low end of the distribution, so to speak. The rate of extirpation, even at equilibrium, is surprisingly low. This is mainly the result of such species being chosen more and more rarely for interactions as they decline in numbers.

Yet another version of the basic system employs two or three patches, each a duplicate system of the original. Each patch begins with the same number and kind of species. As the separate communities develop over time, occasional randomly selected individuals are allowed to migrate from one patch to another. Although the migration rates were pre-set to quite modest levels (one per thousand predation events), extirpations slowed by approximately 30% in a two-patch model and by 40% in the three-patch model.

These variations on the basic trophic theme create the suspicion that the basic model, in spite of its extreme simplicity, conceals some subtle features that explain its (apparent) closeness to real multi-species populations. If one replaces classical predation by the more general idea of competition for (energy) resources, the model may be re-interpreted without change. When one organism preys on another, it generally ingests it. Yet one plant may "prey" on another too, in a sense, by stealing it's sunlight. Energy that would have gone to the second plant, goes instead to the first. In the compartmentalized model of trophism, for example, the energy awarded to one autotroph is taken by another, causing an increase in numbers (or biomass) for the first species and causing a loss for the second.

The possibility remains that the seeming success of this model as a predictor of community abundances lies not in its robustness nor in its subtlety (as such), but in the fact that each individual has an equal chance of reproducing or dying.



arrows indicate direction of energy flow

Figure 2. The dynamical system with cyclic (left) and compartmentalized (right) food webs (see explanation in the main text).

A frequency formula for species abundance

As it happens, the frequency formula for the steady state (average or expected) species abundances in the MSL model is relatively easy to derive from first principles. In what follows, k will represent an abundance and f(k) will represent the number of species in the system that currently have abundance k.

The derivation assumes that the model has reached equilibrium. Among the species of abundance k there will be $k \cdot f(k)$ individuals and among the species of abundance k+1 there will be $(k+1) \cdot f(k+1)$ individuals. The probability that the next trophic event will involve a given individual is

$$p = 1/N.$$

In other words, the probability that one of the species in the k-class will increase in abundance is

$$p \cdot k \cdot f(k)$$

By the same reasoning, the probability that one of the species in the (k+1)-class will decrease is

$$p \cdot (k+1) \cdot f(k+1)$$

The probability of being prey is, after all, the same as the probability of being predator. Under the equilibrium assumption, stability of the function f will only be guaranteed by the equality of these two quantities. The number of species migrating from column k to column k+1 must equal the number of species migrating in the opposite direction:

$$p \cdot k \cdot f(k) = p \cdot (k+1) \cdot f(k+1)$$

The general solution of this equality is c/k, where c is some constant. Thus, to within a constant factor, there is only one integer-valued function that satisfies this recurrence relation:

$$f(k) = 1/k$$

One way to check this analysis is to accumulate and average the frequencies in each abundance value at the end of a run of the MSL system, and plot their values. The purpose of inversion is to make the inverse linear component of the curve into a straight line, while magnifying departures from this shape at the high abundance end of the histogram. The results are invariably a straight line over the lower half of the range of abundances, but then the line begins a gentle, but accelerating trend upward, albeit in increasingly ragged fashion. In longer runs of the program, the plot remains linear through a slightly longer portion of its range. The latter observation provides a clue that explains the upward trend in the inverted abundances.

The inverse linear nature of the equilibrium state is a local condition only. Globally, over the entire range of abundances that prevails, a logistic factor operates. Since the MSL preserves the total number N of individuals, as one species becomes increasingly abundant, the others perforce grow less so. In non-extinction mode, the extreme case obtains when all but one species have abundance 1. If N = 4000 and m = 200 for example, and all but one species had abundance 1, the 200th species would have abundance 3,801. This would represent an extreme logistic limit that would have only an infinitesimal probability of actually occurring during any run. A more realistic logistic limit would nevertheless have to be based on probability. For example, it might be specified that the limit be exceeded only 5 percent of the time. Solving the equilibrium equation with the added condition that the total number of individuals must always be N, we obtain that the maximum abundance, max, would have the associated species frequency of 1/20, resulting in the equation,

$$\max = \sqrt{20N}$$

$$= 283$$

Since we will not know N in practice, it will be reasonable to build the limit into the equation as a multiplicative factor, echoing the well-known logistic equation of May (1974).

$$f(k) = k^{-1}(1 - k\delta);\ 1 \leq k \leq \Delta,$$

where Δ is the maximum abundance and $\delta = 1/\Delta$. An equivalent formulation that is somewhat easier to use has the form,

$$f(k) = 1/k - \delta;\ 1 \leq k \leq \Delta \qquad \text{(i)}$$

When this function is inverted and plotted, it looks like a straight line over the lower part of its range, but then begins to curve upward as the logistic factor comes increasingly into play. The curve closely resembles plots of inverted MSL frequencies (see Figure 1).

The MSL system is part of a new trend in ecological modeling towards individual based systems (Judson 1994). It may be that the extensive use of differential equations in ecological modeling has limited progress by insisting on a determinism that is not in fact present in nature to a degree that would warrant the use of such models. Individual based models are ideally suited to the portrayal of stochastic events that undoubtedly occur routinely at all levels in natural communities.

The logistic distribution

Like many distributions that have been proposed as models of species abundance in communities, the *logistic distribution* is defined as a continuous function. It has two parameters, ε and δ.

$$f(x) = c(1/x - \delta);\ \varepsilon \leq x \leq \Delta \qquad \text{(ii)}$$

The quantity Δ is called the *logistic limit* and is defined by $\Delta = 1/\delta$. The parameter ε will be called the *inner limit*. The *logistic coefficient* c may be obtained by integration of the function over its domain $[\varepsilon, \Delta]$.

$$c = \ln(\Delta/\varepsilon) - 1 \qquad \text{(iii)}$$

Figure 3 illustrates the relationship between the logistic function and the inverse linear curve. In a more formal approach, the distribution could be viewed as the result of a translation of the axes

$$x = x' + \varepsilon$$

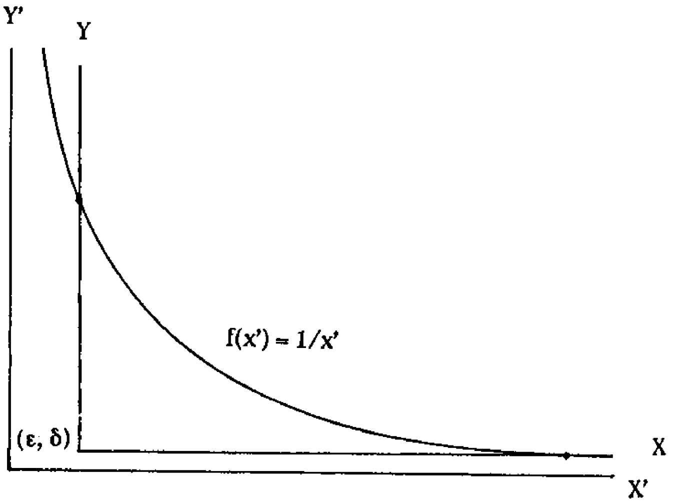


Figure 3. Relationship between the logistic function and the inverse linear curve. The logistic distribution is obtained by a translation of the axes.

$y = y' + \delta$

The resulting formula would then become

$f(x) = c((x + \varepsilon)^{-1} - \delta);\ 0 \leq x \leq \Delta - \varepsilon,$

but expression (ii) is equivalent to this formula and much easier to use.

The formulas for both the logistic coefficient c (iii) and for the mean μ (below) are close approximations obtained by dropping vanishingly small terms during the derivation process.

$\mu = \Delta/2c$

A complete description of the logistic distribution and its use in sampling milieus is beyond the scope of this paper but a complete description of theory and practice is available in Dewdney (1997).

A brief theory of sampling

Crucial to the application of the logistic distribution, as well as all others, is a valid sampling theory that relates the distribution g of abundances in a community to the distribution f of abundances in a sample. In view of the failure of current theory to do this explicitly, it has been necessary to elaborate the elements of such a theory already present in Pielou (1969).

Briefly, let a random, unbiased sample of size n be taken of the community C and let

$r = n/N$

be called the sampling ratio. A species of abundance j in C will have a probability p(k) of appearing k times in the sample, where

$$p(k) = \binom{j}{k}\binom{N-j}{n-k} \Big/ \binom{N}{n} \qquad \text{(iv)}$$

The expression (iv) is called the hypergeometric distribution (Feller 1965) and forms the basis of a close approximation by the Poisson distribution which forms the basis for the theoretical work by Pielou (1969) on the theory of abundance distributions. In fact, we may adapt the Poisson distribution directly to a theory of sampling by writing

$$\int_0^\infty g(x).e^{-rx}(rx)^k/k! \qquad \text{(v)}$$

where g(x) is the continuous version of the community distribution g and the remainder of the integrand is the Poisson distribution.

One could certainly write a parallel (discrete) equation involving the hypergeometric distribution but it is somewhat unwieldy in computations and is so closely approximated by the Poisson for reasonably large values of N that the difference between the two is statistically indiscernible. In short, apart from the usual statistical fluctuations that are normal in samples, one cannot do better than the transformation (v). This point must be stressed because the Poisson distribution has a great deal to say about how g is transformed into f.

The lognormal distribution

In particular, when k = 0, the Poisson distribution that produces the null class of species (the ones that do not appear in the sample) has a shape that bears no resemblance to the "veil line" (Preston 1948). Preston claims that all species to the left of the vertical line shown in Figure 4a will not appear in the sample. The Poisson distribution, however, subtracts species in a very different manner, namely a gradual curve that is far from vertical (Figure 4b). When the null class of species is subtracted from the lognormal distribution, what remains resembles the J-curve even less than the untruncated lognormal does.

To finish with Preston, it might also be mentioned that when the logistic distribution is subjected to the same transformation by "octaves" (essentially an exponential transformation along the abundance axis) that produces the normal

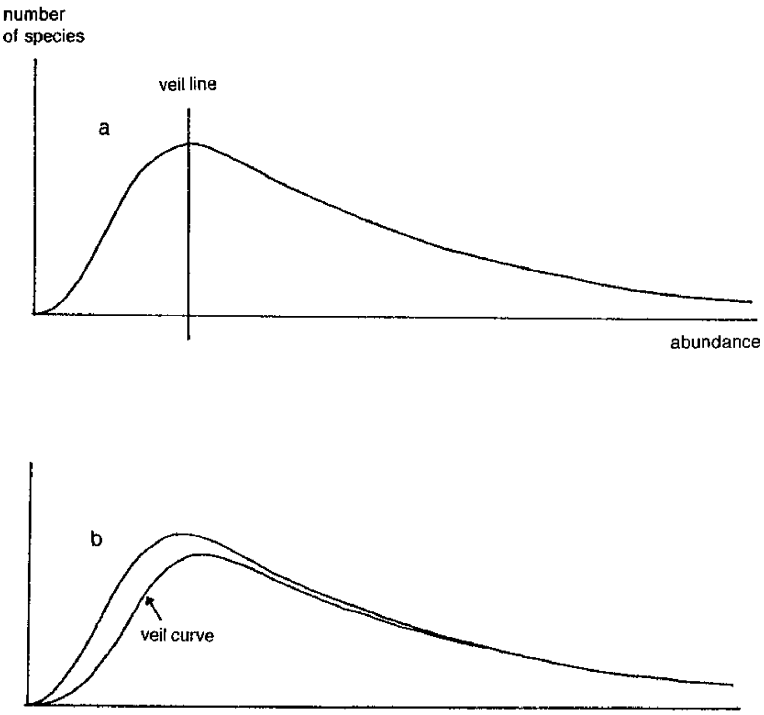


Figure 4. The effect of sampling on the lognormal distribution. In (a) all species to the left of the veil line will not appear in the sample. In (b) species to the left and above the veil curve will not appear in the sample.

distribution from the lognormal, a curve that closely resembles a truncated normal distribution also results. If the logistic distribution prevails in nature, this would readily explain the success of the lognormal when truncated. As we have just shown, however, truncation may not be considered a statistically valid operation.

The appearance of a truncated normal-looking distribution is readily explained in the presence of the logistic distribution by the cumulative effect of the subtracted δ. The exponentially transformed formula for the logistic distribution is $c(\ln 2 - 2^k\delta)$. A tail inevitably accompanies the exponential drop owing to the sampling process.

Given the relative tolerance of goodness-of-fit tests, exponential transformations of logistic distributions will readily pass as truncated normal distributions. If it should turn out that the lognormal and the related negative binomial distribution (Pielou 1969), do not have ecological relevance as distributions, it may be freely speculated that the truncation operation, along with the coincidence just outlined, have jointly operated to ensure the survival of these distributions to the present day.

The log-series distribution

A third surviving candidate as a widely-used model of abundances in natural communities is the log-series distribution of Williams (1964). This distribution is discrete and takes the following form:

$f(k) = r_k/k$,

where the series r_k converges to 0 as k approaches infinity. As early as the 1930s Williams thought he saw the inverse linear curve in the abundances of Lepidoptera caught in light traps near Cambridge. In the author's opinion, he was quite correct to do so. However, concern over the non-convergence of the inverse linear function caused him to consult a mathematical friend (Williams 1964) who suggested that the convergent series be installed in the function to force convergence. In mathematics, such a step would be called a "kludge." In short, the form of Williams' log-series distribution, while close to that which appears to prevail in real communities (Taylor et al. 1976), suffers from unwanted baggage in the form of a factor that was not derived from biological considerations, but from mathematical expediency.

From a practical point of view, the log-series distribution is (almost) statistically indistinguishable from the logistic distribution. The two functions look similar because they are similar. There is, however, a great difference between their theoretical foundations.

The biosurvey data

The same goodness-of-fit tests may be (and have been) applied to the logistic distribution, comparing it with real biosurvey data from reputable sources. At the time of this writing, some 30 such surveys have been compared to the logistic distribution via the chi-square test. All pass with extremely high values of α. In 25 of the cases $\alpha = 0.5$ and in two of them $\alpha = 0.75$.

Here, α represents the probability that a *valid* hypothesis (the empirical distribution is statistically indistinguishable from the theoretical one) will be *rejected*. In rejective mode (where the test is most useful) one wants as small an α as possible. In acceptance mode, on the other hand, one wants α values to be as large as possible, allowing the test every opportunity to reject the hypothesis, in effect. In short, the χ^2 values (which measure the difference between the empirical and theoretical frequencies) were exceptionally low.

The surveys examined so far span a large range of environments, scales, and taxa. They include temperate zone freshwater benthic microbiota, tropical rain forest butterflies, temperate montane birds, desert oases macrofauna, temperate deciduous forest microfungi, temperate freshwater fish, temperate pine forest robber flies, tropical rain forest birds, temperate deciduous forest vascular plants, temperate freshwater amphibious fungi, tropical reef fish. arctic lichen communities, temperate saline stream macroinvertebrates, boreal forest mayflies, temperate freshwater benthic invertebrates, Mediterranean montane butterflies, temperate coniferous forest birds, boreal stream macroinvertebrates, temperate marine coastal waterbirds, antarctic marine benthic macroinvertebrates, temperate coniferous forest microfungi, and others.

The fitting process may be illustrated by a data set that is entirely typical of the biosurveys examined to date. Gutierrez & Menendez (1995) sampled butterflies in the Iberian peninsula during two one-month periods in the early summer of successive years. They recorded the occurrence of 79 species of butterflies as the mean number of individuals per locality within the sampling area, but admit that another 28 species were seen but not unidentified. (The latter are presumably rare or unknown and would almost certainly weight the low end of the abundance spectrum rather than the high end.)

The mean abundance μ is 4.26 and the number of species f_a in the lowest abundance (= 1) category is 23. With these two statistics in hand, the parameters of the appropriate logistic distribution are readily computed, with nothing fancier than a hand calculator, in about ten minutes. One solves the following equation numerically:

$$(2\mu T/e)\ln(a/\varepsilon) = a^{T}\cdot\varepsilon^{1-T},$$

where $T = R/f_a$, R is the richness of the sample, f_a is the number of species of minimum abundance and a is the size of interval employed in the sample histogram. In the case at hand, the following parameter values emerge from the solution:

$$\varepsilon = 0.384$$

$$\delta = 0.0357$$

It might be asked whether the variance should not be used with the mean (instead of the number of species at the lowest abundance) as a basis for the estimation of the epsilon and delta parameters. Variance is, to say the least, highly variable in distributions with the J-curve shape, with up to 50 percent of the variance arising from the abundance of a single species! Consequently, it gives rather unstable results.

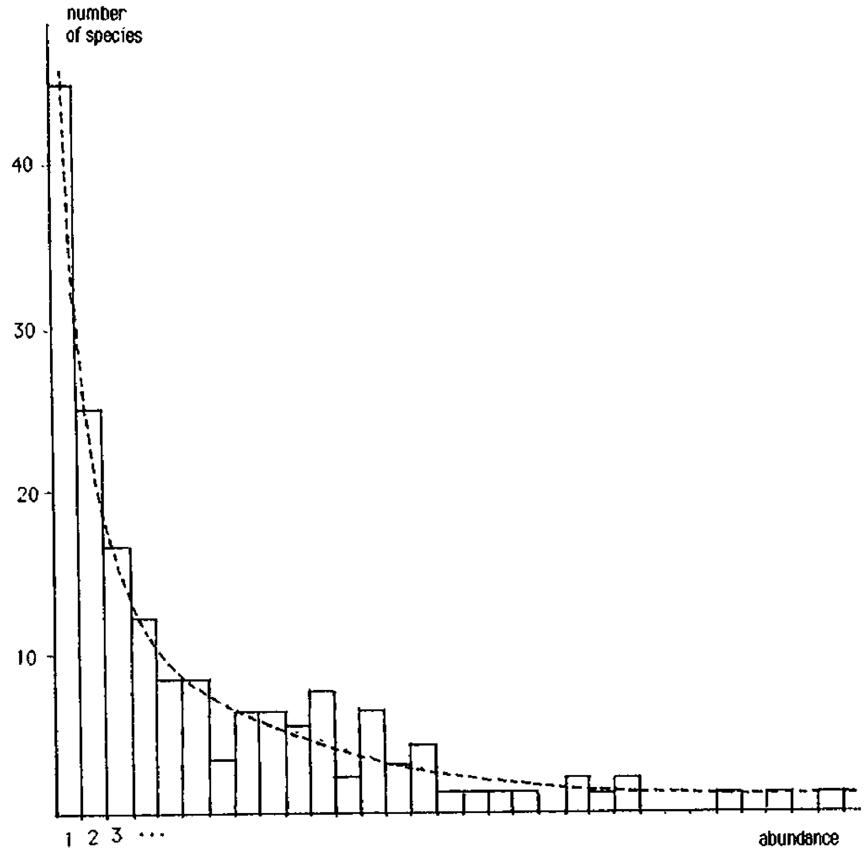


Figure 5. Fitting the logistic distribution to butterfly abundances. The histogram (data from Gutierrez & Menendez [1995]) shows the frequency of butterfly species in each abundance (number of individuals) class. The curve is the logistic distribution fitted to the data.

The butterfly community distribution has the logistic coefficient c = 0.304. Figure 5 shows the distribution that arises from this sample with the appropriate logistic distribution superimposed on it. The chi-square statistic for the fit has the value 1.94 and the test hypothesis is accepted at the alpha level of 0.5 which, as was pointed out earlier, is as low as most tables go. The critical value at this level (df = 7) is 6.346.

From the delta-value, one readily derives the logistic limit by inversion:

$\Delta = 28.01$

This means that, as far as the sample is concerned, no species in a community that follows the logistic distribution should appear more than 28 times. As it happens, one species appears 27 times in the sample and none more.

Although there are insufficient data to make any safe assertions on this point at the moment, it certainly appears as though the maximum abundances in survey after survey are "piling up" toward their logistic limits. In other words, the logistic limit exceeds the maximum abundance by a wide margin in a few of the biosurveys, by a medium margin in several of the biosurveys and by a small margin in many of them. In only three biosurveys was the logistic limit violated. However, in each case a) one species violated the limit, and b) the "aberrant" species was a species of schooling fish. The strength of the piling-up phenomenon was frankly unanticipated when this parameter was named the "logistic limit."

A more complete study of extant biosurveys will be reported in a subsequent publication.

Abundance and community dynamics

The analysis that led to expression (i) assumed only that the individuals of each species enjoyed a constant probability p of increasing (reproducing) or decreasing (dying) over some fixed interval of time. All other factors in the analysis could apply to natural communities as well as the abstract community of the MSL model. In natural communities, for example, the function f(k) is well defined in that, at any moment a definite number of species occupy each possible abundance category.

One may readily conjecture that in natural communities a similar rule holds. In other words, the two probabilities may fluctuate over a period of days, weeks, or years depending on the time scales of the species involved, but if they did not have a long-term tendency to equality they would pass, relatively quickly, either to extinction or to a position of superdominance.

Most biologists, if asked to predict how the basic MSL model would behave, might well answer that species would typically increase or decrease their abundance a little. The more knowledgeable might guess that if all the species started with the same abundance, then they would, in time, form a normal distribution of abundances. The latter would be right, up to a point. But few biologists, or mathematicians for that matter, would guess that the species would arrange themselves in a the shape of an inverse linear curve. The same game may be played with natural communities. How many people would guess, if told that all each individual in a natural community had the same probability of increase as of decrease, that the species therein would inevitably fall into such a pattern of abundances?

What may have been lacking in community studies up to this point is the notion of the equilibrium of the system as a whole, instead of the equilibrium of individual species. As we have seen, the most parsimonious assumption about living communities, that they enjoy some kind of long term equiprobable increase/decrease property, leads directly to the conclusion that their abundances will tend to take the shape of a logistic distribution.

Another way to look at the equiprobability property is to view species abundance as a random walk on the abundance axis (Dotle et al. 1984). This is not an ordinary random walk, by any means. The transition probabilities for a given species will depend on its current k-value or abundance.

The notion of species abundance as a random walk is not quite as bizarre as it might at first sight seem. Most predation events, for example, are certainly the results of random encounters. If the varying hare searches for food in this hummock instead of the one over there, for example, the lynx may not see it. The hare, in turn, may or may not find plant material that is present. If the wind changes direction slightly, the spore of a lignin-digesting fungus will miss a small bruise on a tree and not infect it. There is no organism that is not subject to this kind of random, i.e., unpredictable, event. In other words, if there is indeed a "walk" by species along the abundance axis in real habitats, it is assuredly not "random," but *effectively random.*

There is no way to predict the outcome of environmental events in detail. Moreover, there is no *apparent* overall pattern in the waxing and waning of their abundances except for after-the-fact assertions involving weather or other large-scale disturbances. It is also important to notice that this view of community abundances no longer depends explicitly on the MSL model, since only the equiprobable birth/death process has been invoked. In view of the robustness of the logistic distribution, it is possible that the complications of the model involving artificial food webs, as well as the implicit assumption of uniform spatial distribution of species may be red herrings, so to speak.

An amusing analogy makes the counter-intuitive distributions produced by the random walk model seem more natural. Imagine 200 race cars traveling about a circular track that is divided into equal-length zones with varying speed limits. If the cars start at random times and race around the track, always obeying the speed limits of successive zones, the expected number of cars to be seen in each zone will be directly proportional to the amount of time they must spend there, that is, inversely proportional to the speed limit. The equilibrium condition, that as many cars enter each zone as leave its predecessor, is trivially true.

From this point of view, species that have small populations tend to change their abundance less frequently, other things being equal, than species with large populations. The probabilities, after all, operate at the level of individuals and not at the level of species. As a direct result, more species tend to collect at the low end of the abundance spectrum than at the high end. This observation explains the general shape of the abundance distributions produced by the dynamical model and the random walk model. It remains to be seen to what extent it also applies to real communities.

Conclusions

To summarize, the possibility that the resemblance of the distributions produced by the MSL system to those found in natural communities is more than accidental has led to several hypotheses. The main hypothesis, that the distribution of abundance in natural communities is best described by the logistic distribution, has fostered the development of a new theory of the sampling process from a mathematical point of view, has necessitated the testing of a large number of natural sample distributions and, more recently, has inspired a search for evidence of random fluctuations in abundance of species in natural communities.

The state of the overall theory thus suggested by the MSL system can be summarized under several points that can all be considered hypothetical to varying degrees, some highly so, others practically established. In what follows the word "match" will mean passage of a goodness-of-fit test and a "close match" will mean passage with an unusually low chi-square statistic.

1. A dynamical model called the MSL system produces artificial abundance distributions that closely resemble natural ones.

    a) The MSL system produces the logistic distribution.

    b) The logistic distribution closely matches the vast majority of natural abundance distributions as sampled.

2. When a natural abundance distribution is sampled, its general shape is preserved by the operation. Specifically, its formula (except for changes in parameter values) is preserved by the sampling transformation (v).

    a) The species lost to a sample are Poisson-distributed in the community and follow what might be called a "veil curve" that has no resemblance to a vertical line

    b) No truncated distribution can be a candidate for the distribution of species abundances in communities.

    c) The lognormal and negative binomial distributions are usually truncated in order to succeed as models of community abundances.

3. When the exponential transformation is applied to the logistic distribution, the resulting curve closely resembles the truncated normal distribution, but is not the truncated normal distribution.

    a) Truncated normal distributions frequently match samples that are subjected to the exponential transformation. The exponentially transformed logistic distribution matches samples more frequently.

4. The logistic limit Δ equals or exceeds the maximum abundance in almost all biosurveys. Its ecological reality is demonstrated by the tendency of maximum abundances to approach but not to exceed it.

Anyone taking seriously the possibility that (almost) all natural communities follow the logistic abundance distribution will find new avenues of research before them: First, there may be a potential for linking the population statistics revealed by samples with those of the community sampled via knowledge of logistic distribution parameters.

Second, the enormous confusion that surrounds the measurement of "biodiversity" might be to some extent clarified if one makes a definition that recognizes the role of a single, dominant distribution that governs virtually all natural communities. Third, although the idea of equilibrium of single populations may have suffered in recent years, the logistic distribution (or one like it) might open the possibility of studying the equilibrium of whole communities, recognizing that while some species wax and others wane, the J-curve remains.

Finally, the theory of evolution might benefit from recognition that small populations are the rule and not the exception in most natural environments, whence speciation has a natural springboard at the low end of the species abundance distribution.

Future publications will report on progress with the hypotheses listed above. In particular, the results of analysis on an expanded list of 100 randomly selected biosurveys should appear soon.

Acknowledgement: I wish to thank David Schneider of Memorial University of Newfoundland and Michelle Bowman of the University of Western Ontario for helpful comments that clarified this paper.

References

Dewdney, A. K. 1996. Microecology in a small stream, Environmental Science Program, U. of Western Ontario, London, Canada.

Dewdney, A. K. 1997. The Abundance of Species in Natural Communities: the J-Curve, Biodiversity, and Evolution. Monograph, The Environmental Science Program, U. of Western Ontario, London, Canada.

Dotle, P. G. & Snell, J. L. 1984. Random Walks and Electrical Networks. The Carus Mathematical Monographs No. 22, The Mathematical Association of America.

Feller, W. 1968. An Introduction to Probability Theory and Its Applications (Vol 1). Wiley, New York.

Gutierrez, D. & Menendez, R. 1995. Distribution and abundance of butterflies in a mountain area in the northern Iberian peninsula. Ecography 18: 209-216.

Huston, M. A. 1994. Biological Diversity. Cambridge U. Press, Cambridge, Mass.

Judson, O. P. 1994. The rise of the individual-based model in ecology. Trends in Ecology and Evolution 9(1): 9-14.

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

May, R. M. 1964. Stability and Complexity in Model Ecosystems, Princeton U. Press, Princeton.

May R. M. 1974. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. Science 186: 645-647.

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

Preston, F. W. 1948. The commonness, and rarity, of species. Ecology 69: 254-83.

Taylor, L. R., Kempton, R. A. & Woiwood, I. P. 1976. Diversity statistics and the log-series model. J. Anim. Ecol. 45: 255-272.

Williams, C. B. 1964. Patterns in the Balance of Nature. Academic Press, London.