

Original article

# A simple method to fit geometric series and broken stick models in community ecology and island biogeography

Simone Fattorini

*Dipartimento di Biologia Animale e dell'Uomo, Università di Roma "La Sapienza", Viale dell'Università 32, I-00185 Rome, Italy  
Present address: Via Casal dei Pazzi 49, 00156 Rome, Italy*

Received 31 October 2004; accepted 29 April 2005

Available online 06 June 2005

## Abstract

Species abundance distributions are widely used in explaining natural communities, their natural evolution and the impacts of environmental disturbance. A commonly used approach is that of rank-abundance distributions. Favored, biologically founded models are the geometric series (GS) and the broken stick (BS) model. Comparing observed abundance distributions with those predicted by models is an extremely time-consuming task. Also, using goodness-of-fit tests for frequency distributions (like Chi-square or Kolmogorov–Smirnov tests) to compare observed with expected frequencies is problematic because the best way to calculate expected frequencies may be controversial. More important, the Chi-square test may prove if an observed distribution statistically differs from a model, but does not allow the investigator to choose among competing models from which the observed distribution does not differ. Both models can be easily tested by regression analysis. In GS, if a log scale is used for abundance, the species exactly fall along a straight line. The BS distribution shows up as nearly linear when a log scale is used for the rank axis. Regression analysis is proposed here as a simpler and more efficient method to fit the GS and BS models. Also, regression analysis (1) does not suffer from assumptions related to Chi-square tests; (2) obviates the need to establish expected frequencies, and (3) offers the possibility to choose the best fit among competing models. A possible extension of abundance-rank analysis to species richness on islands is also proposed as a method to discriminate between relict and equilibrium models. Examples of application to field data are also presented. © 2005 Elsevier SAS. All rights reserved.

*Keywords:* Species abundance distributions; Rank-abundance plots; Geometric series; Broken stick; Regression analysis; Island biogeography

## 1. Introduction

The abundance distributions of species tend toward characteristic patterns, which can be represented as frequency-abundance or abundance-rank models (e.g. Magurran, 1988, 2004). In the frequency-abundance representations, the horizontal axis represents classes of abundance and the vertical axis represents number of species. In the abundance-rank representations, all the species in a sample are ranked from most abundant to least abundant. Each species has a rank, which is plotted on the horizontal axis, and an abundance, plotted on the vertical axis. Thus the abundance for the most abundant species is plotted first, then the next most common and so on until the array is completed by the rarest species of all. Plotting abundance against rank yields a curve which is by definition monotonically decreasing. Species abundance distributions appear to share similar shapes in most ecosystems,

leading to applications and generalizations ranging from the prediction of species diversity patterns (He and Legendre, 2002) to the effects of environmental variables on community structure (Detsis et al., 2000; Small and McCarthy, 2002), from the ecosystem organization (Büssenschütt and Pahl-Wostl, 1999; Harte et al., 1999; Solé et al., 2002) to the explanation of general biogeographical patterns (Hubbell, 2001).

It is therefore not surprising that theoretical ecologists have proposed a plethora of models to fit the data. The most known are: the geometric series (GS, niche preemption) model, the broken stick (BS) model (see Wilson, 1991; Magurran, 1988, 2004 for reviews), the extended BS model (Büssenschütt and Pahl-Wostl, 1999), the sequential breakage model (Kolasa and Strayer, 1988), the overlapping niche model, the negative binomial distribution, the truncated negative binomial distribution, the logseries distribution, the lognormal distribution (Pielou, 1975; Magurran, 1988, 2004), the gamma distribution (Hughes, 1986; Wilson, 1991), the first kind beta distribution, the second kind beta distribution, the Dirichlet distribution,

*E-mail address:* [Simone\\_fattorini@virgilio.it](mailto:Simone_fattorini@virgilio.it) (S. Fattorini).

bution (Engen, 1978), the two-dimensional Poisson lognormal distribution (Engen et al., 2002), and the power fraction (Tokeshi, 1996). Among them, Tokeshi's (1993, 1996) innovative models of niche apportionment may represent a notable advance in measurement techniques (Magurran, 2004), but they are still rarely applied because the numerical procedures suggested to obtain the fit are fairly complex (Mouillot and Leprêtre, 2000). Although several models have attracted the attention of theoretical ecologists, the most widely applied models are: the lognormal distribution (LN), the GS, the logarithmic series (LS), and the MacArthur BS (Magurran, 1988). The least equitable distribution of these (i.e. the distribution with lower evenness among species) is GS. A more equitable distribution is represented by BS, while intermediate curves generally approximate to a LS or a LN distribution. Theoretical ecologists still debate if LS and LN are merely statistically descriptive models, or if they possess a biological significance (e.g. Giller, 1984; Krebs, 1999; Büssenschütt and Pahl-Wostl, 1999; Dewdney, 2003). By contrast, GS and BS represent two opposite, biologically founded methods, against which empirical data may be compared with reference to explicit ecological assumptions (Giller 1984; Büssenschütt and Pahl-Wostl, 1999).

The current approach to assess if a particular observed distribution fit a model involves two steps: (1) calculating the theoretically expected frequencies according to the model, and hence (2) comparing them with the observed ones by means of some goodness-of-fit test for frequency distributions (Magurran, 1988; Moreno, 2001). This approach may involve a hard work and has some shortcomings. First, the best way to calculate expected frequencies may be controversial (e.g. Magurran, 1988; Krebs 1999) and there is no agreement about the proper test to use (Chi-square or Kolmogorov–Smirnov) (Tokeshi, 1993; Krebs, 1999; Keeley and Fotheringham, 2003). More important, the Chi-square test may prove if an observed distribution statistically differs from a model, but does not allow choosing among competing models from which the observed distribution does not differ. Thus, the choice among different models is sometimes based merely on visual inspection of plotted data (Schmiegelow et al., 1997; Siemann et al., 1999). Wilson (1991) proposed the use of algorithms which minimize the deviance in a ranked-abundance plot by replicate quadrates of plants. However, this is a time-consuming procedure difficult to extend to animals, which requires specifically written software and rather complicated designs (cf. Steel et al., 2004).

In this paper, I propose to use regression analysis as an efficient and simpler method to fit GS and BS, and I show that this approach may offer statistical tools to choose the best fit between competing models. I also suggest a possible extension of abundance-rank analysis to island biogeography.

## 2. Methods

### 2.1. Geometric series

According to the GS model, the sizes of niche hyper-volumes (the multidimensional space occupied by the

individual species belonging to a community, expressed as species abundances) are determined by some species pre-empting a relevant part of the niche space, leaving the remaining species to occupy the rest (niche preemption model) (Pielou, 1975). The first (most successful) species in the sequence pre-empt a fraction  $k$  of resource hyperspace, the second species a fraction  $k$  of hyperspace not occupied by the first, and so on. This distribution is typical of highly dominated communities and it is usually found in resource poor environments, or in those that have suffered a short time ago an environmental catastrophe, or have been colonized a short time ago (Giller, 1984). If a log scale is used for abundance, the species exactly fall along a straight line.

Both LS and LN may be related to GS. LS predicts samples or communities dominated by a few very common species, similar to the GS, but also with many rare species. LS was initially derived by Fisher et al. (1943) by considering a special case of the zero-truncated negative binomial distribution. In fact, LS is mathematically closely related to GS (May, 1975) and both can fit the same data (Magurran, 1988). When an abundance-rank representation is used, and the abundance of each species is plotted on a logarithmic scale, LS approximates a straight line, being virtually indistinguishable from GS (Taylor et al., 1976). LS provides a statistically satisfactory description of samples from a wide range of communities, but its ecological justification is unclear (May, 1975; Caswell, 1976; Dewdney, 2003). Also, in some communities middle-ranked species are relatively numerous, and there are fewer rare species than the LS distribution predicts. By the logarithmic transformation of number of individuals, the number of species follows a normal distribution. The LN distribution predicts samples or communities in which most species are of intermediate abundance. The preferred plot is the same as that used for LS. The species are grouped together into classes according to the logarithm of their abundance, and species count against class is plotted. The LN distribution forms a normal distribution instead of the monotonically decreasing distribution of a LS. However, if a community with a LN distribution is undersampled, more rare species will not be fully represented in the sample, so that in practice the left-hand tail of the distribution is not represented. With real sample data, if the mode of a lognormal curve is not revealed, it is practically impossible to distinguish from LS. The LN distribution has received most attention, but results are controversial (Krebs, 1999; Halley and Inchausti, 2002). Most mature and varied communities apparently approximate a LN pattern (May, 1975; Preston, 1980; Magurran, 1988; Brown and Nicoletto, 1991), presumably only a reflection of the Central Limit Theorem (May, 1975). A detailed account of LS and LN can be found in Krebs (1999), who also provides some methods to fit these distributions. However, with real data, both models are hardly distinguishable from the GS, and on a log-abundance vs. rank plot, LS cannot be discerned from a GS by regression analysis.

Download English Version:

<https://daneshyari.com/en/article/9444957>

Download Persian Version:

<https://daneshyari.com/article/9444957>

[Daneshyari.com](https://daneshyari.com)