Species abundance data from fouling communities conform to the gamma distribution

Günter H. Schmidt[1] & David J. Garbutt[2]

Departments of [1] Pure and Applied Zoology, and [2] Applied Statistics, University of Reading, Whiteknights, Reading RG6 2AJ, United Kingdom

ABSTRACT: The gamma distribution successfully fitted 128 of 136 sets of species abundance data (% cover) from a fouling study. In contrast to the more widely used log-normal, which always has an expected mode, the great flexibility of the gamma allows it to assymptote to the frequency (y) axis and therefore to satisfactorily describe the deeply concave shape of the observed abundance curves.

INTRODUCTION

Diversity statistics based upon 2 resource-apportioning models – log-normal and log-series distributions – have recently been criticised for their lack of correspondence to species abundance distributions of marine invertebrate communities, as neither the log-normal nor the log-series satisfactority describe the deeply concave shape of the dominance-diversity curves of many species assemblages (Hughes 1984). Hughes (1984) introduced a multiple-parameter model which simulates recognized features of community dynamics and yields curves similar in shape to those derived from data of natural communities, but the practical applications of the model are less apparent. In particular, when information is available only on numbers of individuals, or percentage of substratum covered, the advantage of the currently employed theoretical distributions for summarising species frequency data is apparent (Ross 1980, Vandermeer 1981). The log-series distribution is a discrete distribution and applicable when counts of numbers of species with a given number of individuals are considered. A continuous distribution, however, is the appropriate model when species abundance data is recorded in continuous form, e.g. area covered (%) by sessile fouling organisms, as in the present study. The log-normal is the most frequently used continuous distribution in benthic ecology (Hughes 1984), although alternative models are available, e.g. the gamma distribution (Ross 1980).

A theoretical basis for various descriptive uses of gamma distributions in statistical ecology, including the use as species abundance distributions, is given by Dennis & Patil (1984). An advantage of the gamma over the log-normal is that it is flexible enough to describe a variety of data sets (Fig. 1; McCullagh & Nelder 1983). The form may be unimodel and skewed to the right (Fig. 1c & 1d), or asymptote to 0 (Fig. 1a). In contrast, the log-normal always has a mode between 0 and ∞ (McCullagh & Nelder 1983). Hence, if species frequency curves are highly concave (Hughes 1984) they may approximate to gamma rather than log-normal. Gamma is a widely used model in community ecology (Ross 1980, Dennis & Patil 1984); in the present paper it
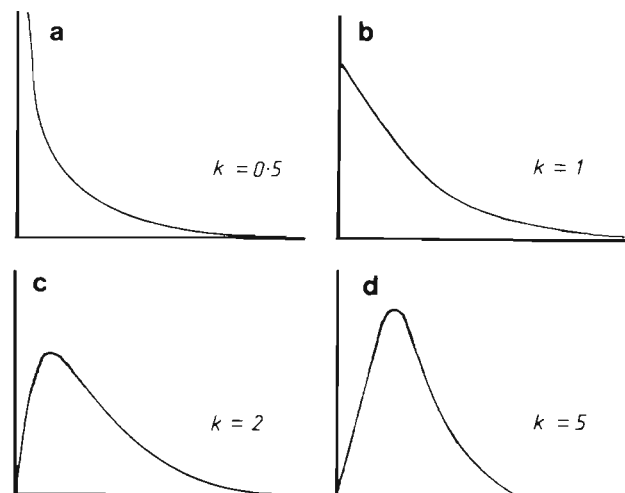

Fig. 1. The gamma distribution is highly flexible (forms a, b, c, d); the parameter k determines the shape of distribution (the mean in a to d is kept constant at 1). All distributions are positively skewed. A mode with zero at the origin is found only if k > 1. (After McCullagh & Nelder 1983)

is fitted to data collected from a fouling study (cf. Schmidt & Warner 1984).

MATERIALS AND METHODS

Black perspex panels, 0.25×0.25 m square and 5 mm thick, were attached to steel frames and sub-merged at a depth of ca 1 m in Langstone Harbour, Hampshire, UK. Fifty panels (25 vertical, 25 horizontal) were submerged on 12 Dec 1979 ('Winter series'), and 16 (8 vertical, 8 horizontal) on 1 Aug 1980 ('Summer series'); they were collected at monthly intervals until 16 Dec 1982 (for dates see Table 1). A further 23 panels were employed for 'caging experiments' (Schmidt &

Table 1. Frequency abundance data (% cover) were fitted to gamma distributions and the goodness of fit tested by means of χ^2. Degrees of freedom given in brackets. Panels were submerged 12 Dec 1979 and collected on dates shown. The table shows a representative sample from a total set of 136 (available from the first author on request)

| Date | Vertical panels | | Horizontal panels | | | |
| | | | Lower | | Upper | |
	Gamma fit	No. of species	Gamma fit	No. of species	Gamma fit	No. of species
1980						
May 1			0.39 (2)	6		
30	3.21 (4)	14	1.11 (5)	13		
	2.20 (4)	12				
Jun 27	1.02 (5)	15	5.64 (4)	12	0.38 (2)	6
	3.02 (6)	14				
Jul 25	7.53 (4)	16	11.99 (6)	15	0.73 (2)	10
	7.61 (4)	16				
Aug 22	8.66 (5)	17	6.32 (5)	17	2.29 (3)	9
	6.44 (2)*	17				
Oct 2	7.12 (7)	26	6.92 (6)	19	1.29 (3)	10
	3.35 (7)	24				
Oct 31	8.96 (4)	24	13.38 (7)	26	6.81 (4)	10
	7.56 (7)	23				
Nov 28	4.20 (3)	22	8.25 (8)	29	6.39 (3)	15
	4.07 (6)	23				
1981						
Jan 14	0.41 (4)	22	1.01 (5)	24	1.03 (1)	12
	9.57 (6)	21				
Feb 10	5.26 (5)	25	14.35 (7)*	23	2.43 (3)	13
	2.77 (6)	23				
Mar 11	2.21 (5)	21	5.94 (5)	20	1.27 (2)	12
	2.79 (6)	22				
Apr 8	7.85 (6)	21	1.86 (5)	18	3.43 (4)	16
	6.63 (5)	19				
May 5	5.39 (4)	19	5.56 (5)	18	2.58 (3)	9
	3.16 (4)	19				
Jun 3	14.83 (7)*	19	5.49 (4)	15	2.14 (4)	8
	8.03 (4)	17				
Jul 1	8.82 (5)	24	8.84 (6)	23	2.64 (4)	15
	7.60 (6)	19				
Jul 30	4.94 (7)	24	6.09 (5)	18	1.98 (5)	10
	0.80 (7)	17				
Aug 27	4.40 (6)	24	5.04 (7)	25	2.34 (5)	11
	8.63 (6)	24				
Oct 7	6.34 (7)	22	3.06 (7)	23	1.98 (5)	16
	1.53 (4)	23				
Nov 10	15.48 (8)	22	7.26 (6)	21	1.56 (5)	13
	18.91 (9)*	20				
Dec 16	13.01 (8)	21	9.67 (6)	23	4.46 (6)	21
	9.82 (8)	18				
	9.56 (6)	20				
	6.79 (7)	19				

* Lack of fit at the 5 % level
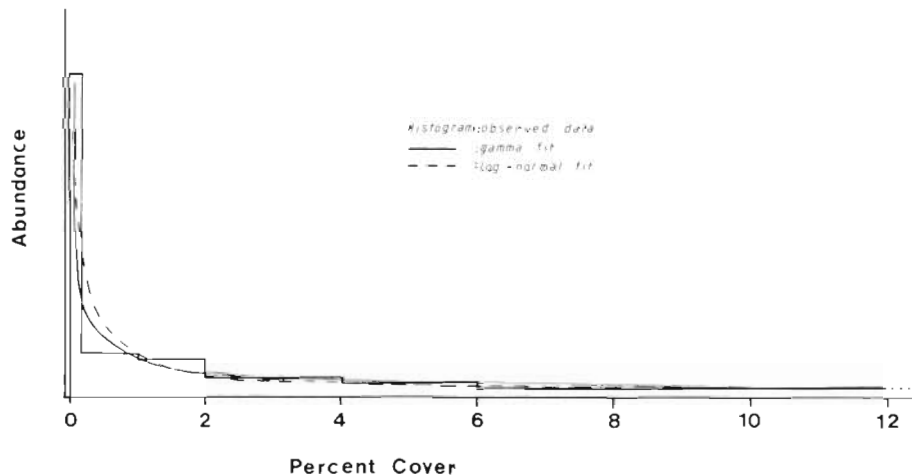
Fig. 2. Species abundance distribution for a fouling community which had developed on a 25 × 25 cm panel over a period of 5.5 mo. Gamma and log-normal were fitted to the data using the method of maximum likelihood and tested for goodness of fit by means of χ^2. The distributions were similar over much of their range but differed in the extreme left part: only gamma can assymptote to the frequency (y) axis and follow the deeply concave shape of the observed data while log-normal has an expected mode. Gamma therefore yields a better χ^2 value than log-normal (even though both fit statistically as the mode of the log-normal remains undetected by the sampling method employed): χ^2 values – gamma: 1.1 on 4 d.f.; log-normal: 3.68 on 4 d.f.

Warner 1984). The percentage area covered by each sessile species was determined by the line intersection method (Schmidt & Warner 1984). When a species was present, but not sampled by any intersection point, it was arbitrarily assigned a cover value of 0.1 %. The species abundance data thus obtained were fitted to gamma (and log-normal for the data shown in Fig. 2) and tested for goodness of fit (χ^2) using the method of maximum likelihood (Ross 1980). 'Winter' panels collected prior to May 1980 and some horizontal upper surfaces carried too few species for meaningful statistical analysis.

RESULTS AND DISCUSSION

A total of 136 data sets were fitted to gamma; 128 of them statistically conformed to the distribution (tested at the 5 % level, Table 1). A graphical representation of 1 set of data, with expected values also for log-normal distribution, is given in Fig. 2. The concave shape of the frequency curve is due to the relatively large number of species which occupied less than 1 % of the panel area. The mode of the fitted log-normal in the first octave is not detected by our method as the < 0.15 octave cannot be subdivided any further. Consequently, only the right hand side of the expected lognormal distribution is exposed, while the left side remains unsampled. This is a recognised property of sampling from a log-normal distribution when the sample size is small (Vandermeer 1981). However, although both models fit the data statistically, only the gamma which does not assume a mode that is un-

detected can assymptote on the frequency (y) axis (Fig. 1), following the deeply concave form of the observed species frequency data. It therefore yields a better χ^2 value than log-normal. Gamma may also provide a valid model for other species abundance distributions which are highly concave and lack a mode (Hughes 1984). While the theoretical basis for the gamma may be complex (Dennis & Patil 1984, Hughes 1984), its application in diversity statistics is of practical value, forming a basis for comparison of different communities or experimental treatments; diversity indices may be calculated from the fitted gamma distributions (Ross 1980, Schmidt & Warner 1984) and further analysed by ANOVA (Kempton & Taylor 1976). The present analysis was restricted to data given in continuous form; for counts of numbers of individuals the discrete negative binomial, which includes the logseries as a special case and is characterised by similar flexibility as the gamma (Ross 1980), may provide a satisfactory model.

LITERATURE CITED

Dennis, B., Patil, G. P. (1984). The gamma distribution and weighted multimodel gamma distributions as models of population abundance. Math. Biosc. 68: 187–121

Hughes, R. G. (1984). A model of the structure and dynamics of benthic marine invertebrate communities. Mar. Ecol. Prog. Ser. 15: 1–11

Kempton, R. A., Taylor, L. R. (1976). Models and statistics for species diversity. Nature, Lond. 262: 818–820

McCullagh, D., Nelder, J. A. (1983). Generalised linear models. Chapman & Hall, London

Ross, G. J. S. (1980). MLP (Maximum Likelihood Program). Rothamstead Experimental Station, Harpenden, U.K.

Schmidt, G. H., Warner, G. F. (1984). Effects of caging on the development of a sessile epifaunal community. Mar. Ecol. Prog. Ser. 15: 251–263

Vandermeer, J. (1981). Elementary mathematical ecology. John Wiley & Sons, Chichester