available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel

Frequency distribution models for spatial patterns of vegetation abundance

Jun Chen[a], *Masae Shiyomi*[a,b,*], *Yoshimichi Hori*[a], *Yasuo Yamamura*[a]

[a] *Ibaraki University, Faculty of Science, Bunkyo 2-1-1, Mito 310-8512, Japan*

[b] *The University of the Air, Ibaraki Learning Center, Bunkyo 2-1-1, Mito 310-0056, Japan*

ARTICLE INFO

Article history:
Received 6 June 2007
Received in revised form
18 September 2007
Accepted 24 September 2007

Keywords:
Beta-binomial distribution
Beta distribution
Biomass
Cover
Gamma distribution
Negative binomial distribution
Number of individuals
Number of binary occurrences

ABSTRACT

In a plant community, measurements of the number of plant individuals, the number of binary occurrences, plant cover, and biomass per unit area have frequently been used to explain species composition and spatial variation. Studies have shown that frequency distribution can be expressed in several ways: the number of individuals can be expressed using the negative binomial distribution, the number of binary occurrences can be expressed using the beta-binomial distribution, plant cover can be expressed by the beta distribution, and biomass by the gamma distribution. In this study, we have mathematically clarified the relationships between these distributions and their biological relevance. We have also defined a spatial heterogeneity index for each of the above four methods of measurement. For each of these four distribution patterns, several-fitted examples of plant populations or communities, obtained from grassland surveys, are provided.

1. Introduction

Frequency distributions of plant density, cover, biomass per unit area, and height, as measures for expressing biological abundance and biological dominance of vegetation, have been used to describe species composition and spatial patterns of vegetation in the studies of plant communities. Relative to plant populations, counting numbers of individuals in animal populations is easy because individuals can generally be recognized. Animal ecologists have studied population processes, and used them to predict and control populations by comparing spatial patterns formed by the number of individuals or comparing differences in densities between populations (Ito and Murai, 1977; Southwood and Henderson, 2000). Few plant ecologists have used individual counts, because the definition or identification of an individual plant is often difficult; instead, biomass and cover measurements are used. Harper (1977) published *Population Biology of Plants* and proposed counting individual plant parts, such as stocks, stems, leaves, and seeds, to obtain information on plant populations. Today, many population ecologists count individual plants to establish natural units, and such methods have become popular.

In this paper, we refer to biological quantities such as biomass, cover, number of individuals, and number of binary occurrences per unit area as "vegetation abundance". For frequency distribution per unit area of vegetation abundance, various statistical/probabilistic models have been derived and

* *Corresponding author at*: Ibaraki University, Faculty of Science, Bunkyo 2-1-1, Mito 310-8512, Japan. Tel.: +81 20 228 0683.
E-mail address: shiyomi@mx.ibaraki.ac.jp (M. Shiyomi).

0304-3800/$ – see front matter

doi:10.1016/j.ecolmodel.2007.09.017

utilized. Examples of such frequency distribution models include: (1) the number of individuals per unit area, expressed using the negative binomial distribution (Fisher, 1941) and the Poisson distribution (e.g., Greig-Smith, 1983); (2) the number of plant occurrences in binary counts (presence/absence) per unit area, expressed as the beta-binomial distribution (Skellam, 1948) and the binomial distribution (e.g., Greig-Smith, 1983); (3) biomass per unit area, expressed as the gamma distribution (May, 1973; Shiyomi et al., 1984); (4) plant cover per unit area, expressed as the beta distribution (Chen et al., 2006). Models (1) and (2) are probability distributions with a discrete variate for count data, and models (3) and (4) are density functions with a continuous variate. These models are based on some simplified assumptions about biological processes. In this paper, we discuss the biological and mathematical relationships between the above four frequency distribution models, and examine: (1) biological processes in derivations of the models; (2) systematic relationships between the models; (3) actual examples of frequency distributions of vegetation abundance obtained from field surveys.

Spatial heterogeneity, e.g., patchy/uniform, coarse/dense, and aggregated/random, is an important property of organisms in communities. Various measures of heterogeneity have been developed in population ecology (e.g., David and Moore, 1954; Morisita, 1959; Lloyd, 1967; Iwao and Kuno, 1971), based on the number of individuals in a population. However, in many surveys of spatial distribution, collected data may indicate presence/absence at a sampling site, cover area, or biomass per unit area. Shiyomi and Yoshimura (2000) derived, to examine populations, a basic formula that can be modified for use with any of the above plant measurements, and the formula is developed in the present study.

2. Frequency distribution models representing spatial patterns

2.1. *From the binomial distribution to the beta-binomial distribution (BBD)*

Suppose that N quadrats, each of which is divided into n smaller cells of equal area, are set on random sites in grassland. Assume that individuals of a plant species (i.e., species A) grow at random on sites in the grassland, and the probability that species A is found in a cell is p; then, the probability $P(i)$ that species A is found in i of the n cells is given by the following equation:

$$P(i)=\binom{n}{i}p^i(1-p)^{n-i} \quad (i=0,1,2,\ldots,n;\ \text{binomial distribution}) \tag{1}$$

Eq. (1) is also described by the following series:

$$P(0)=(1-p)^n$$

$$\vdots$$

$$P(i)=P(i-1)\frac{n-i+1}{i}\frac{p}{1-p} \quad (i=1,2,\ldots,n)$$

This form is convenient for calculating series of $P(i)$ for $i=0, 1,\ldots, n$. Here, $P(0)+P(1)+\cdots+P(n)=1$. The mean μ and variance σ^2 for Eq. (1) are given as follows

$$\mu=np \tag{2}$$

$$\sigma^2=np(1-p) \tag{3}$$

In many cases, species A distributes not randomly but patchily among nN cells in an actual plant community, i.e., at a high density in some places and a low density in other places within the grassland. In such case, it is possible to assume that p in the binomial distribution (Eq. (1)) follows a beta distribution. Thus, we derive the probability $P(i)$ that we find species A in i of n cells in a quadrat as follows

$$\begin{aligned}P(i)&=\int_0^1\frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}\binom{n}{i}p^i(1-p)^{n-i}\,dp\\&=\frac{1}{B(\alpha,\beta)}\binom{n}{i}\int_0^1 p^{\alpha+i-1}(1-p)^{\beta+n-i-1}\,dp\\&=\binom{n}{i}\frac{B(\alpha+i,\beta+n-i)}{B(\alpha,\beta)}\end{aligned}$$

$$(i=0,1,2,\ldots,n;\ \text{beta-binomial distribution or BBD}) \tag{4}$$

where $B(\alpha,\beta)=\Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)=\int_0^1 p^{\alpha-1}(1-p)^{\beta-1}\,dp$ is the beta function, α and β are parameters that determine the shapes of the beta function and beta distribution, and $P(0)+P(1)+\cdots+P(n)=1$. Eq. (4) is also expressed by the following series:

$$P(0)=\frac{\beta(\beta+1)\cdots(\beta+n-1)}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+n-1)}$$

$$\vdots$$

$$P(i)=\frac{P(i-1)(n-i+1)(\alpha+i-1)}{(\beta+n-i)i} \quad (i=1,2,3,\ldots,n)$$

The mean μ and variance σ^2 are

$$\mu=\frac{\alpha n}{\alpha+\beta} \tag{5}$$

$$\sigma^2=\frac{\alpha\beta n(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{6}$$

The basic formula defining the spatial heterogeneity I is given by the following equation (Shiyomi and Yoshimura, 2000):

$$I=\frac{\sigma^2-\mu(1-(\mu/n))}{\mu^2}$$

The spatial heterogeneity is determined by the following criteria: (1) if $I>0$, higher heterogeneity than would be expected in the random pattern, and a larger value indicates a high degree in heterogeneity; (2) if $I=0$, random pattern and (3) if $I<0$, lower heterogeneity than would be expected in the random pattern, and a large negative value indicates a low degree of heterogeneity. Even if some occurrences of species A are removed from random cells, I-value does not change