

Ripley's 指数的一个新变形—— $G(d)$

葛宏立¹, 周元中², 汤孟平¹, 丁丽霞¹

(1. 浙江林学院环境科技学院, 临安 311300; 2. 浙江省临安市林业局, 临安 311300)

摘要:在对 Ripley's 指数的物理背景进行分析的基础上提出了一个新的变形—— $G(d)$ 指数。原始 Ripley's 指数 $K(d)$ 是一个半径为 d (d 为尺度) 的圆的面积估计量, 是有量纲的, 且 $K(d)$ 随着 d 的增大迅速增大, 应用起来不太方便。Ripley's 指数的变形 $L(d)$ 将 $K(d)$ 换算成 d 的估计量, 再减去 d , 使得 $L(d)$ 在随机分布的假定下有数学期望 0, 这使得 $L(d)$ 的应用比 $K(d)$ 要方便。但 $L(d)$ 还是一个具有长度单位的量, 其上下包迹线呈明显的喇叭形状, 对于应用还不是十分方便。提出的 $G(d)$ 指数是一个比值, 无量纲, 数学期望也是 0。文章给出了 4 个例子, 这些例子说明, 它保持了 $L(d)$ 指数区分分布类型的能力, 同时具有稳定这个良好特性, 到了一定尺度以后, 上下包迹线趋向常数。进一步分析得知, 这些常数与单位面积的个体密度相关, 呈对数关系, 其相关系数 r^2 达到 0.9 左右。这样在实际应用中, 包迹线只要模拟到稳定点即可, 余下部分可通过回归公式计算, 从而节省计算工作量。

关键词:种群; 空间分布格局; Ripley 指数

文章编号: 1000-0933(2008)04-1491-07 中图分类号: Q143 文献标识码: A

A modified Ripley's index

GE Hong-Li¹, ZHOU Yuan-Zhong², TANG Meng-Ping¹, DING Li-Xia¹

1 School of Environmental Sciences and Technology, Zhejiang Forestry University, Lin'an 311300, China

2 Lin'an City Forestry Bureau, Lin'an 311300, China

Acta Ecologica Sinica, 2008, 28(4): 1491 ~ 1497.

Abstract: This paper proposes a new index ($G(d)$) based on the Ripley's. The original Ripley's index $K(d)$ (here d is a distance scale) is an area estimate of a circle with radius d under the assumption that individuals have a Poisson distribution. The fact that an area estimate becomes larger when d gets larger makes it inconvenient for practical applications. The modified Ripley's index $L(d)$ transforms $K(d)$ to an estimate of d and then d is subtracted from the estimate. The difference has an expectation of zero under the assumption of Poisson distribution. Although the use of $L(d)$ is more convenient than $K(d)$, the $L(d)$ is still a number with a distance measure and not very convenient for use because its upper and lower envelopes look like a bell-mouthed form.

The $G(d)$ index proposed in this paper is a ratio with an expectation of zero and with no dimension and is defined as

$$G(d) = K(d) / (\pi d^2) - 1$$

Three examples with simulations and one real application show that $G(d)$ has the property of stability, which previous

基金项目: 国家 863 资助项目 (2006AA12Z109); 国家自然科学基金资助项目 (30471390); 浙江省自然科学基金资助项目 (Y305261); 国家自然科学基金青年基金资助项目 (60403046)

收稿日期: 2007-08-07; **修订日期:** 2008-01-07

作者简介: 葛宏立 (1960 ~), 男, 浙江诸暨市人, 博士, 教授, 主要从事森林资源监测的科研与教学. E-mail: jhghlxl@sina.com

Foundation item: The project was financially supported by National High Technology Research and Development Program of China (863 Program) (No. 2006AA12Z109); National Natural Science Foundation of China (No. 30471390); Natural Science Foundation of Zhejiang Province (No. Y305261); National Natural Science Foundation of China (The Youth Fund) (No. 60403046)

Received date: 2007-08-07; **Accepted date:** 2008-01-07

Biography: GE Hong-Li, Ph. D., Professor, mainly engaged in forest resources monitoring. E-mail: jhghlxl@sina.com

Ripley's indices (i. e., $L(d)$ and $K(d)$) do not have, and is able to distinguish different spatial patterns. When the scale d reaches a certain value, such as between 2 and 6 in most cases, the envelopes become constants. Further analyses show that there is a relationship between the constants and the density of individuals (number of individuals per 100 m²). The relationship can be expressed as equations, which the stable values of upper - (SVU) and lower - (SVL) envelopes are as follows:

$$SVU = 0.152549 - 0.0396694 \ln(\text{density}), r^2 = 0.919$$

$$SVL = -0.175449 + 0.0610485 \ln(\text{density}), r^2 = 0.883$$

The results provide a possibility that it is not necessary to calculate the envelopes beyond the stable point if $G(d)$ is employed.

Key Words: plant community; spatial pattern; Ripley index

种群的空间格局分析是研究种群特征、种群间相互作用以及种群与环境关系的重要手段,在生态学中一直是研究的热点之一。种群的空间分布格局一般分为随机分布、聚集分布和规则分布(规则分布有时被称作均匀分布)3种形式。通常认为,这些分布类型的确定与样地(或称样方)大小关系十分密切,也就是说种群的分布类型与空间尺度有重要关系^[1]。Ripley's 指数是一种常用的种群格局分析指数。它以植物种的个体在空间中的平面坐标为基本数据,每个个体都被视为二维空间的一个点,所有个体组成空间分布的点图(Point map)。以此点图为基础进行格局分析,叫做“点格局分析”(Point pattern analysis)^[2]。Ripley's 指数可以分析各种尺度下的种群格局和种间关系^[3]以及这种关系随时间的变化^[4]。Ripley's 指数不仅在生态学中得到广泛应用,在其它方面也有应用,如在医学上进行疾病发病点的空间格局分析^[5]。本文对 Ripley's 指数及由 Besag 提出的一个变形^[6,7]进行深入分析,在此基础上提出一个新的指数(一个变形) $G(d)$ 。新指数 $G(d)$ 和 Ripley's 指数的区别在于:在随机分布的假定下,新指数的方差当尺度达到一定大小后,基本趋于稳定,即上下包迹线趋于平行,而 Ripley's 指数的方差随着尺度的增加一直增加,上下包迹线呈喇叭状。这样,新指数更稳定、更便于应用。

1 Ripley's 指数的物理背景

如图 1 所示,矩形表示一个样地,样地中每个点表示一个植物个体。以个体 i 为圆心, d 为半径画圆,记该圆覆盖的除中心个体 i 之外的个体数为 $a_i(d)$ 。在随机分布的假定下, $a_i(d)$ 的数学期望为:

$$E(a_i(d)) = n \frac{\pi d^2}{A} \quad (1)$$

式中, n 为样地内的总个体数, A 为样地面积。这样的圆可以画 n 个。如果圆的一部分跨越了样地边界,则 $a_i(d)$ 需要进行调整。有的根据圆周长落在样地内的比例进行调整^[6-12],有的根据圆面积落在样地内的比例进行调整^[13],以及整个样地整体校正^[14]。 n 个 a_i 的均值的期望与(1)式一样:

$$E\left(\frac{\sum_{i=1}^n a_i(d)}{n}\right) = n \frac{\pi d^2}{A} \quad (2)$$

根据上式,有:

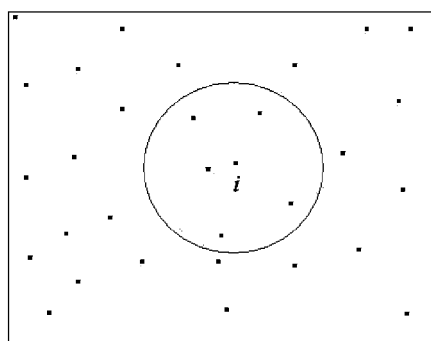


图 1 样地和个体

Fig. 1 A plot and plant individuals

$$\pi d^2 = \frac{A}{n} E \left(\frac{\sum_{i=1}^n a_i(d)}{n} \right) \quad (3)$$

如果个体在平面空间上的分布不服从随机分布,则(3)式并不成立。离随机分布越远,(3)式两边相差越大。所以可以用(3)式右边部分作为衡量一个分布与随机分布相差程度的一个指标。这就是 Ripley's 指数的基本概念^[2,9,15~18]。由于数学期望在实际中一般不容易取得,只能由样本平均数作为估计值。这样,Ripley's 指数定义为:

$$K(d) = \frac{A}{n} \frac{\sum_{i=1}^n a_i(d)}{n} = A \sum_{i=1}^n \sum_{j=1}^n \frac{\delta_{ij}(d)}{n^2} \quad (4)$$

其中

$$\delta_{ij}(d) = \begin{cases} 1 & \text{如果 } d_{ij} \leq d \\ 0 & \text{如果 } d_{ij} > d \end{cases} \quad i, j = 1, \dots, n; i \neq j$$

式中, d_{ij} 为个体*i*与个体*j*之间的距离。

由于 $K(d)$ 的实际含义是以 d 为半径的圆面积的估计值,所以 $K(d)$ 随着 d 的增大而迅速增大。这给应用带来不便。Besag^[6]提出用 $L(d)$ 取代 $K(d)$:

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d \quad (5)$$

在随机分布假设下, $K(d)$ 的期望是 πd^2 ,所以 $L(d)$ 的期望为0。在实际应用时,根据 $L(d)$ 值的大小,大于0时为聚集分布,等于0时为随机分布,小于0时为规则分布。

对于样地边缘影响,除了上面提到的用周长、面积等校正的方法外,还有缓冲带的方法,即缓冲带内的个体只计数但不作为中心个体*i*,缓冲带的宽度与 d 相等。设非缓冲带内的个体有 m 个(显然 $m \leq n$),根据上面的推导,显然(4)式变为:

$$K(d) = \frac{A}{n} \frac{\sum_{i=1}^m a_i(d)}{m} = A \sum_{i=1}^m \sum_{j=1}^n \frac{\delta_{ij}(d)}{nm} \quad (6)$$

按比例调整与缓冲带两类边缘处理方法各有缺点。按比例调整,扩大了边缘个体的影响;缓冲带方法刚好相反,削弱了边缘个体的影响。

2 Ripley's 指数的一个新变形

用(5)式表示的 Ripley's 指数,虽然与用(4)式表示的相比,使用方便,便于解释,也变得稳定。但其还是一个具有量纲的数值,实际是一个长度值。虽然其数学期望不随 d 的变化而变化(在随机分布假设下),但其方差是随着 d 的增大而增大的。这就是随机分布的 Ripley's 指数的上下两条包迹线呈现喇叭状的原因。这种情况仍然导致应用上的不方便。下面构造一个无量纲的指数。它定义为:

$$G(d) = \frac{K(d)}{\pi d^2} - 1 \quad (7)$$

等式右边第一部分是一个比值。显然在随机分布假设下, $G(d)$ 的数学期望为0。对其的分布解释则完全与 $L(d)$ 相同。

3 一些模拟例子和一个实际例子

下面用例子来说明 $G(d)$ 与 $L(d)$ 的异同。边缘处理采用缓冲带方法,即用公式(6)计算 $K(d)$ 。用蒙特卡罗法计算上下包迹线(Envelopes),即置信区间。假定种群是随机分布,用随机模型产生一组与例子数据个体数相同的随机点,然后对一系列 d 值计算指数。重复进行这一过程直到达到事先确定的次数。指数的最大值和最小值构成上下包迹线。重复次数对95%的置信水平应为20次,99%的置信水平就为100次^[9],

本文为 20 次,即 95% 的置信水平。

3.1 模拟规则数据

数据以 $5\text{m} \times 5\text{m}$ 的正方形的形式排列,样地大小为 $105\text{m} \times 105\text{m}$,21 行 21 列共 441 个数据(图 2(a))。相应的无量纲 Ripley's 指数 $G(d)$ 和原指数 $L(d)$ 见图 3(a-1)和(a-2)。两种指数的分布格局结论基本相同。当 $d < 5\text{m}$ 时,两个指数均明显小于 0,是规则分布。当 $d = 5$ 时都明显大于 0,是聚集分布。此后则以小于 0 为主,当 $d \geq 20\text{m}$,两种指数均在上下包迹线之间,即趋向随机分布。规则分布当 d 增大时趋向随机分布,这一点是与常识相符的。但当 d 不很大时,指数也有可能明显大于 0,呈聚集分布。对于此例,当 $d < 5\text{m}$ 时,所有的 $\delta_{ij}(d)$ 都为 0,当 $d = 5\text{m}$ 时,所有的 $\delta_{ij}(d)$ 都为 4,指数出现明显的跳跃。 $G(d)$ 和 $L(d)$ 的区别在于: $G(d)$ 在 $d > 7\text{m}$ 时,上下包迹线趋向常数,基本不再变化,上包迹线在 0.091 左右,下包迹线在 -0.062 左右。而 $L(d)$ 则呈喇叭状,且跳动较大。

3.2 模拟聚集数据

共模拟 738 个个体,样地大小为 $120\text{m} \times 100\text{m}$,见图 2(b),聚集成团状(接近圆形)。 d 为 1m 时,倾向于规则分布。 d 在 $3 \sim 11\text{m}$ 之间为明显的聚集分布,其中间值 $6 \sim 7\text{m}$ 大约为团状分布的半径。 d 大于 13m 后,倾向于规则分布。 $G(d)$ 在 $d > 5\text{m}$ 时,上包迹线稳定在 0.076 左右,下包迹线稳定在 -0.039 左右。 $G(d)$ 和 $L(d)$ 见图 3(b-1)和(b-2)。

3.3 模拟随机数据

共模拟 441 个个体,样地大小为 $105\text{m} \times 105\text{m}$,见图 2(c)。两种指数均一直落在上下包迹线之间,即不论什么尺度,均呈现随机分布。自从 7m 开始, $G(d)$ 的上包迹线稳定在 0.091 左右,下包迹线稳定在 -0.062 左

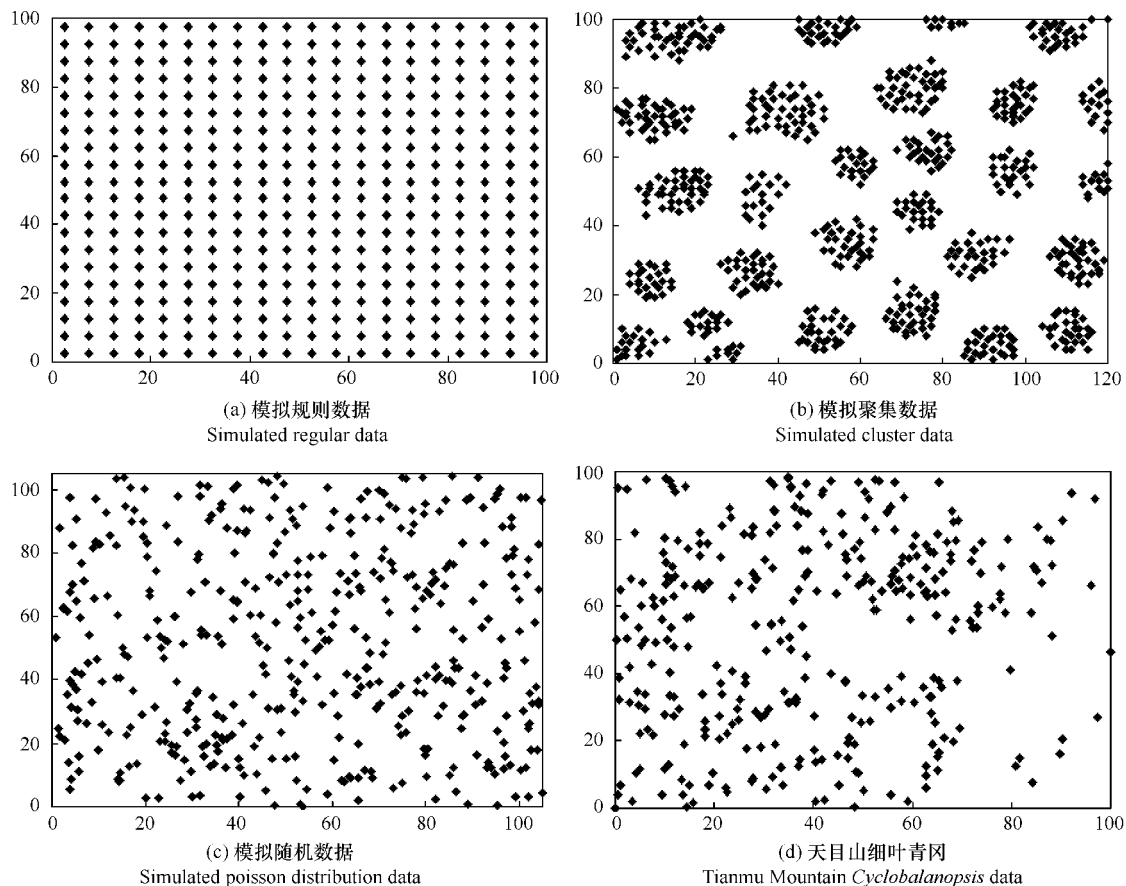


图 2 原始数据

Fig. 2 Experimental data

右。 $G(d)$ 和 $L(d)$ 见图3(c-1)和(c-2)。

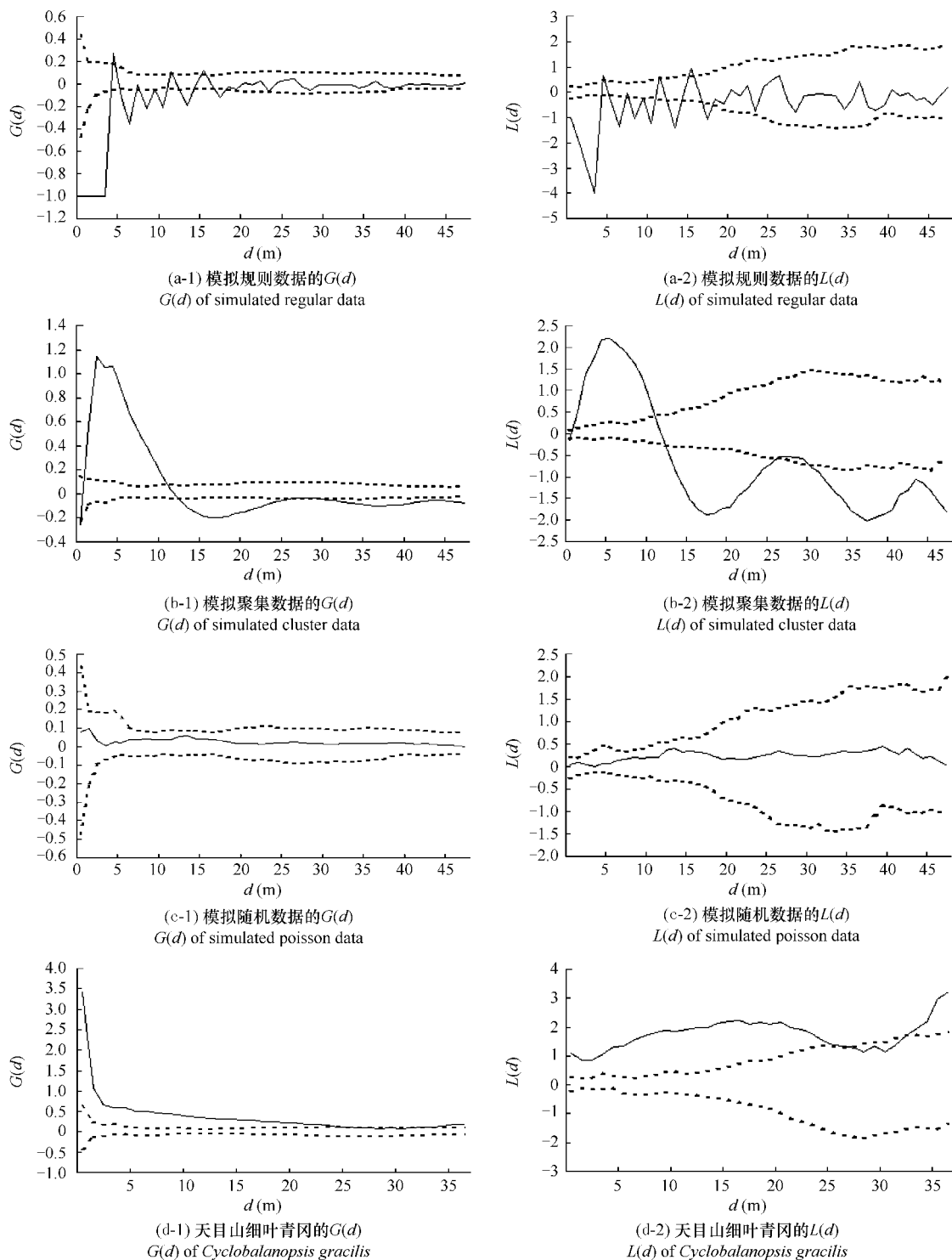


图3 两种指数比较

Fig. 3 Comparison of $G(d)$ and $L(d)$

3.4 天目山细叶青冈(*Cyclobalanopsis gracilis*)数据

共403个个体(图2(d))。该样地大小为100m×100m,于2005年设置于浙江省天目山自然保护区内。保护区大地坐标为:东经119°23'47"~119°28'27",北纬30°18'30"~30°24'55"。样地海拔580~680m。样地内

胸径 5cm 以上的木本植物共有 72 个种,1603 个个体。计算结果表明,在小尺度时,细叶青冈聚集分布特征突出,其中一个原因是细叶青冈在胸高 1.3m 以下分叉较多,测量和记载时按分叉进行。一直到 $d = 20\text{m}$ 也表现为聚集分布,但这时分叉不是主要原因。此后向随机分布靠近。自从 7m 开始, $G(d)$ 的上包迹线稳定在 0.091 左右,下包迹线稳定在 -0.065 左右。 $G(d)$ 和 $L(d)$ 见图 3(d-1)和(d-2)。

4 包迹线稳定值与个体密度之间的关系模拟计算与分析

上面的例子表明, $G(d)$ 指数当尺度达到一定大小时,随机分布的上下包迹线趋向稳定。下面通过模拟计算来分析包迹线的稳定值与单位面积上个体密度的关系。以 $105\text{m} \times 105\text{m}$ 模拟样地为例进行计算,密度分别取 100m^2 1 个个体到 10 个个体等 10 个等级,每个等级模拟 20 次(95% 的置信水平)。模拟计算结果分别见表 1 和图 4。

表 1 包迹线稳定值与个体密度的模拟计算结果
Table 1 Simulation result of envelope stable values and densities

个体密度(个体数/100m ²) Density(individuals/100 m ²)	上包迹线稳定值 Stable values of upper envelope (SVU)	下包迹线稳定值 Stable values of lower envelope (SVL)	稳定点(m) Stable point
1	0.166793	-0.202623	5
2	0.120057	-0.113000	6
3	0.098715	-0.103275	5
4	0.090542	-0.061580	7
5	0.081091	-0.079170	4
6	0.077554	-0.061286	3
7	0.070821	-0.056380	4
8	0.079878	-0.062976	3
9	0.072156	-0.048072	2
10	0.068697	-0.044028	3

很明显,包迹线的稳定值与个体密度之间存在一定关系。用对数曲线对这种关系进行拟合。上包迹线的拟合结果为:

$$SVU = 0.152549 - 0.0396694 \ln(\text{density}), r^2 = 0.919$$

式中,SVU 表示上包迹线的稳定值(Stable value of upper envelope),density 为密度,即 100m^2 内的平均个体数。

下包迹线的拟合结果为:

$$SVL = -0.175449 + 0.0610485 \ln(\text{density}), r^2 = 0.883$$

从表 1 还可看出,包迹线开始稳定的尺度大致随着密度的增加而提前,一般在 6、7 之前就已经稳定。利用上面的回归方程可以算出不同密度的稳定后的与尺度无关的上下包迹线值,用于判定一个分布是否属于聚集、随机或规则分布。这样,在实际应用中只要模拟稳定点之前的包迹线,而可以省去模拟稳定点以后的包迹线的大量计算工作。

5 结论

本文对 Ripley's 指数的物理背景重新进行了分析,在此基础上提出了 Ripley's 指数的一个新的变形——

$G(d)$ 指数。最初的 Ripley's 指数 $K(d)$ 是一个半径为 d 的圆面积估计量,随着 d 的增大迅速增大,应用起来不方便。Ripley's 指数的变形 $L(d)$ 将 $K(d)$ 换算成 d 的估计量,再减去 d ,使得 $L(d)$ 在随机分布的假定下有数学期望 0,这使得 $L(d)$ 的应用比 $K(d)$ 要方便。但 $L(d)$ 毕竟是一个具有长度单位量纲的量,其上下包迹线

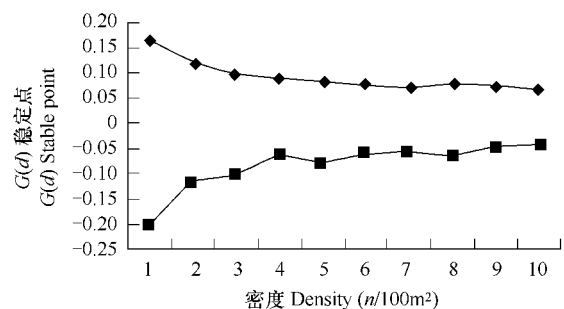


图 4 包迹线稳定值与个体密度的关系

Fig. 4 Relationship between envelope stable values and densities

呈明显的喇叭形状,对于应用还不是十分方便。本文提出的 $G(d)$ 指数是一个比值,无量纲,数学期望也是 0。文章给出的 4 个例子说明,它保持了 $L(d)$ 指数区分分布类型的能力,同时稳定,即到了一定尺度以后,上下包迹线趋向常数。这些常数与单位面积的个体密度相关,呈对数关系,其相关系数 r^2 达到 0.9 左右。

References:

- [1] Zhang J T. Analysis of spatial point pattern for plant species. *Acta Phytocologica Sinica*, 1998, 22 (4) : 344 — 349.
- [2] Diggle P J. Statistical analysis of spatial point patterns. New York: Academic Press, 1983. 148.
- [3] Joao L F B, Douglas A M. Modeling the spatial structure of tropical forests. *Forest Ecology and Management*, 1998, 110: 293 — 314.
- [4] Annett W. Fifty year record of change in tree spatial patterns within a mixed deciduous forest. *Forest Ecology and Management*, 2005, 215: 212 — 223.
- [5] Chi W X, Wang J F, Li X H, *et al.* Spatial Point Pattern Analysis of Birth Defects. *J Environ Health*, 2007, 24(4): 238 — 240.
- [6] Moeur M. Characterizing spatial patterns of trees using stem-mapped data. *Forest Science*, 1993, 39 (4) : 756 — 775.
- [7] Tang M P, Tang S Z, Lei X D, *et al.* Edge correction of Ripley's $K(d)$ function on population spatial pattern analysis. *Acta Ecologica Sinica*, 2003, 23(8): 1533 — 1538.
- [8] Hanus M, Hann D W and Marshall D D. Reconstructing the spatial pattern of trees from routine stand examination measurements. *Forest Science*, 1998, 44 (1) : 125 — 133.
- [9] Ripley B D. Modelling spatial pattern. *Journal of the Royal statistical Society, Series B*, 1977, 39: 172 — 212.
- [10] Michael L W, Arthur G. The spatial characteristics of stand structure in *Pinus torreyana*. *Plant Ecology*, 1999, 143: 153 — 170.
- [11] Zhang C Y, Zhao X H, Wang X Y, *et al.* Spatial pattern and canopy structure of Korean pine roadleaved forests in Changbaishan Mountains. *Journal of Beijing Forestry University*, 2006, 28 (Supp. 2): 45 — 51.
- [12] Tang M P, Zhou G M, Shi Y J, *et al.* Study of dominant plant populations and their spatial patterns in evergreen broad-leaved forest in Tianmu Mountain, China. *Journal of Plant Ecology*, 2006, 30 (5) : 743 — 752.
- [13] Besag J E. Comments on Ripley's paper. *Journal of the Royal Statistical Society, Series B*, 1977, 39: 193 — 195.
- [14] Ward J S, Ferrandino F J. New derivation reduces bias and increases power of Ripley's L index. *Ecol. Model*, 1999, 116: 225 — 236.
- [15] Ripley, B. D. *Spatial statistics*. New York: Wiley, 1981. 252.
- [16] Szwagrzyk J. Small-scale spatial patterns of trees in a mixed *Pinus sylvestris*-*Fagus sylvatica* forest. *Forest Ecology and Management*, 1990, 51: 301 — 315.
- [17] Chen J Q, Gay A B. Forest structure in space: a case study of an old growth spruce-fir forest in Changbaishan Natural Reserve, China. *Forest Ecology and Management*, 1999, 120: 219 — 233.
- [18] Youngblood A, Timothy M, Kent C. Stand structure in eastside old-growth ponderosa pine forests of Oregon and northern California. *Forest Ecology and Management*, 2004, 199: 191 — 217.

参考文献:

- [1] 张金屯. 植物种群空间分布的点格局分析. *植物生态学报*, 1998, 22 (4) : 344 ~ 349.
- [5] 迟文学,王劲峰,李新虎,等. 出生缺陷的空间点格局分析. *环境与健康杂志*, 2007, 24(4): 238 ~ 240.
- [7] 汤孟平,唐守正,雷相东,等. Ripley's $K(d)$ 函数分析种群空间分布格局的边缘校正. *生态学报*, 2003, 23(8): 1533 ~ 1538.
- [11] 张春雨,赵秀海,王新怡,等. 长白山自然保护区红松阔叶林空间格局研究. *北京林业大学学报*, 2006, 28(增2): 45 ~ 51.
- [12] 汤孟平,周国模,施拥军,等. 天目山常绿阔叶林优势种群及其空间分布格局. *植物生态学报*, 2006, 30 (5) : 743 ~ 752.