

有序聚类法在大叶樟苗高生长 时期划分中的应用

彭方仁

(南京林业大学)

关键词 大叶樟; 有序聚类法

樟树是我国的珍贵树种,也是亚热带常绿阔叶林的代表种。樟树树姿秀丽,寿命长,四季常青,且具芳香,近年被广泛选作南方城镇及“四旁”绿化树种。四川大叶樟与本地香樟相比,具有生长期长、生长量大、耐寒性较强等优点,在江浙一带引种已初获成功。但有关它的生长特性还有待进一步研究。一年生的播种苗在年生长过程中呈现明显的节律性。通常把苗木的年生长过程划分为出苗期、生长初期、生长盛期和生长后期。但这种划分一般是以苗木的形态特征而进行直观评定,这种定性的分析往往带有很大的主观性和盲目性。本文试图通过有序样品聚类分析的方法,比较合理而客观地对四川大叶樟苗木生长时期进行定量分析。

一、试验材料和方法

(一) 试验材料

本试验在南京市花木公司所属的伊刘苗圃内进行,所用种子于1987年3月从四川引进。4月7日播种,5月初出苗,随后,及时除草分次间苗,在苗木生长期未进行追肥,仅除草松土三次并及时灌溉。

(二) 生长量测定

在试验地内随机选取20株苗木作生长量测定的固定标准株,从5月中旬至11月上旬,每隔10d定期测定苗高的生长量,定株记载苗木的生长过程。

(三) 有序样品的聚类分析

出苗期通常是指从播种到幼苗地上部分出现真叶、地下部分出现侧根时为止的时期,在本文中相当于出苗结束的日期(5月15日),对此许多学者并无太多争议。本文将着重讨论苗木生长初期与生长盛期、生长盛期与生长后期的划分。自5月下旬至10月下旬给各旬净生长量分别编号为 x_1, x_2, \dots, x_{16} ,问题就转化为如何将这16个有序样品分成三类进行聚类分析。

首先根据类直径计算公式 $D(i, j) = \sum_{k=1}^j (x_k - x_{i+j})^2$ 求出各类直径。其次,根据最小误差函数计算公式 $e[P_0(n, k)] = \min_{k < j < n} \{e[P(j-1)(k-1)] + D(j, n)\}$ 计算最小误差函数值。最后根据最小误差函数值将样品进行分类。

二、结果与分析

1. 苗高年生长进程 播种后,随着春季气温、土温的回升和播种地提供的发芽适宜条件,种子即出土萌发。萌发后,随着细胞的不断分生扩大,根、茎、叶等营养器官迅速发展,并开始积累营养物质。到7、8月份,营养器官基本成熟,此时气温较高,雨量充足,环境条件适宜,苗高的生长量增大。到了秋季,环境条件发生了变化,苗木生理机能也随之发生变化,高生长趋缓并渐停止。高生长过程表现出明显的“慢、快、慢”的生长节律。大叶樟为4月7日播种,5月5日幼苗基本出齐。5月15日开始定株观察记录,苗高生长到11月5日基本停止。从播种到出土为29 d,从出土到高生长终止为186 d。高生长定期观察结果如表1。

表1 大叶樟播种苗高生长定期观察结果

月 份	五 月		六 月		七 月			八 月			九 月			十 月		十一 月		
日 期	15	25	5	15	25	5	15	25	5	15	25	5	15	25	5	15	25	5
连续生长量 (cm)	3.28	4.31	5.43	6.67	8.52	12.77	17.50	24.18	32.87	42.42	53.86	61.35	65.30	67.06	68.02	68.64	68.75	68.82
净生长量 (cm)		1.03	1.12	1.24	1.85	4.25	4.73	6.68	8.69	9.55	11.44	7.49	3.95	1.76	0.96	0.62	0.11	0.07
累积生长量 (%)	4.77	6.26	7.89	9.69	12.38	18.56	25.43	35.14	47.76	61.40	78.76	89.15	94.89	97.44	98.84	99.74	99.90	100

2. 苗高生长时期的划分 由于苗木的生长是连续性的,以每次测定的净生长量值为一个样品,那么这些样品的次序是不能打乱的,生长期的划分必须以这些样品的自然顺序进行,为此我们采用有序样品聚类法(最优分割法)。

首先,根据下列公式计算样品间的类直径

$$D(i, j) = \sum_{k=i}^j (x_k - \bar{x}_{ij})^2 \quad (i = 1, 2 \dots 16, j = 1, 2 \dots 16, i < j)$$

如计算 $D(2, 4)$, 先计算 \bar{x}_{24} ,

$$\bar{x}_{24} = \frac{1}{3}(x_2 + x_3 + x_4) = \frac{1}{3}(1.12 + 1.24 + 1.85) = 1.40$$

那么 $D(2, 4) = (1.12 - 1.40)^2 + (1.24 - 1.40)^2 + (1.85 - 1.40)^2 = 0.31$

类直径的计算结果如表2所示。

其次计算最小误差函数值

$$e[P_0(n, k) = \min_{1 \leq j < k} \{e[P(j-1)(k-1)] + D(j, n)\}]$$

如: $e[P_0(3, 2) = \min_{2 \leq j < 3} \{D(1, j-1) + D(j, 3)\}]$

$$= \min_{2 \leq j < 3} \{D(1, 1) + D(2, 3), D(1, 2) + D(3, 3)\}$$

$$= \min \{0 + 0.001, 0.004 + 0\} = 0.001$$

表 2 类直径计算结果 $D(i, j)$

$i \backslash j$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0															
2	0.004	0														
3	0.02	0.007	0													
...
15	181.36	169.34	156.39	142.28	130.16	128.54	128.69	126.63	119.54	92.40	32.45	6.71	0.68	0.06	0	0
16	191.53	187.57	176.71	164.99	155.77	155.12	154.97	152.32	136.63	107.43	38.81	8.86	1.33	0.31	0.48	0

类似地

$$\begin{aligned}
 e[P_0(4,2)] &= \min\{D(1,1) + D(2,4), \\
 &\quad D(1,2) + D(3,4), D(1,3) \\
 &\quad + D(4,4)\} \\
 &= \min\{0.31, 0.004 + 0.19, \\
 &\quad 0.02\} \\
 &= \min\{0.31, 0.19, \\
 &\quad 0.02\} = 0.02
 \end{aligned}$$

最小误差函数值的计算结果如表 3 所示。

最后根据表 3 就可将 16 个有序样品分成三类。 $e[P_0(16,3)] = 36.60$ ，相应的 $j = 11 + 1 = 12$ ，这说明精确最优解的分类中，首先分出第三类， $G_3 = \{x_{12}, x_{13}, x_{14}, x_{15}, x_{16}\}$ ，再对余下的 11 个样品 x_1, x_2, \dots, x_{11} 进行分类。由表 3 知 $e[P_0(11,2)] = 27.74$ ，相应的 j 值为 $j = 6 + 1 = 7$ 。即前两类是 $G_1 = \{x_1, x_2, \dots, x_6\}$ 和 $G_2 = \{x_7, x_8, \dots, x_{11}\}$ 。因此，16 个有序样品分成三类的聚类分析结果为： $\{x_1, x_2, \dots, x_6\}, \{x_7, x_8, \dots, x_{11}\}, \{x_{12}, x_{13}, \dots, x_{16}\}$ ，相应的时间为第一类 5 月 15 日—7 月 15 日，第二类 7 月 25 日—9 月 5 日，第三类为 9 月 15 日—10 月 25 日。

根据以上聚类分析的结果，一年生大叶樟苗木的高生长可划分为如下四个生长时期：出苗期为 4 月 7 日—5 月 5 日；生长初期为 5 月 6 日—7 月 15 日，生长盛期为 7 月 16 日—9 月 5 日；生长后期为 9 月 6 日—11 月 5 日。这一生长期的划分与前人直观评定的结果基本相吻合。

3. 各生长期生长情况比较 从表 4 看出，生长盛期所占的时间较短，但生长量却占全年总生长量的绝大部分(64%)。有学者指出，生长期的长短不是造成苗高生长差异的主要原因，而速生期的长短及高生长率最大值的高低对苗高的生长起决定作用。生长盛期正值当地夏季的高温、干旱季节。而苗木的快速生长需要水肥能得到保证，采取适

当遮荫,及时灌溉等措施解决水、肥、热的供求矛盾,促进生长盛期的提早到来和推迟结束,提高生长盛期生长速率是苗木培育的关键。

表3 最小误差函数值 $e[p_0(n, k)]$

$k \backslash n$	2	3	$k \backslash n$	2	3	$k \backslash n$	2	3
2	0		7	3.88(4)	0.69(6)	12	47.10(6)	27.74(11)
3	0.001(2)	0	8	14.75(4)	2.71(6)	13	77.27(4)	30.14(11)
4	0.02(3)	0.004(3)	9	18.33(6)	4.25(7)	14	106.96(4)	32.02(11)
5	0.57(4)	0.02(4)	10	25.48(6)	7.85(7)	15	129.11(4)	34.45(11)
6	0.69(4)	0.14(4)	11	27.74(6)	12.18(7)	16	142.43(12)	36.60(11)

表4 各生长时期生长情况比较

生长期	指 标	日 期	时 间 (d)	净 生 长 量 (cm)	占总生长量百分数 (%)
出 苗 期		4月7日—5月5日	28	2.76	4
生 长 初 期		5月6日—7月15日	71	14.74	21
生 长 盛 期		7月16日—9月5日	52	43.85	64
生 长 后 期		9月6日—11月5日	61	7.47	11

参 考 文 献

- [1] 庄晨辉, 1987, 应用聚类分析方法划分黑荆树生长阶段, 林业勘察设计(福建), (1), 52—55.
 [2] 郭德武, 1985, 逐步聚类法在苗木分级中的应用, 辽宁林业科技, (4), 12—15.

THE APPLICATION OF SEQUENCE CLUSTER ANALYSIS ON DIVISION OF THE GROWTH STAGE OF SEEDLING HEIGHT FOR *CINNAMOMUM PLATYPHYLLUM*

Peng Fangren

(Nanjing Forestry University)

Abstract Sequence sample cluster analysis is used to conduct a research on the growth process of the annual seedling height of *Cinnamomum platyphyllum* for seed sowing. The results indicate that the growth process of the seedling height can be divided into four stages: juvenile stage (April 7—May 5), early growth stage (May 6—July 15), rapid growth stage (July 16—September 5) and late growth stage (September 6—November 5). The rapid growth stage takes shorter time, but the increment at this stage accounts for most of the total increment of the whole year. It is very important to solve the contradiction between the supply and demand of water, fertilizer, temperature in the rapid growth stage and raise its growth rate.

Key words *Cinnamomum platyphyllum*; sequence cluster analysis