西北林学院学报 2009，24(6)：1～5
Journal of Northwest Forestry University

北京山区天然栎林直径分布的研究

张雄清， 雷渊才*

（中国林业科学研究院 资源信息研究所，北京 100091）

摘 要：以偏度、峰度为指标研究了北京山区栎类天然林直径分布特征，运用 Weibull 分布、Normal 分布、Lognormal 分布和 Gamma 分布 4 种概率密度函数研究了栎树林分的直径分布规律，并利用 Kolmogorov-Smirnov(KS)检验法来检验 4 种概率分布的拟合效果，结果表明：栎树林分中等稍偏小胸径的林木占大多数，Gamma 分布拟合效果最好，并建立该分布的参数预估方程。

关键词：栎树；直径分布；偏度；峰度

中图分类号：S792.180.6　　**文献标识码**：A　　**文章编号**：1001-7461(2009)06-0001-05

Studies on Diameter Distribution of Oak Plants in Beijing

ZHANG Xiong-qing, LEI Yuan-cai

(*Institute of Forest Resources Information Techniques, CAF, Beijing* 100091, *China*)

Abstract: Some characters of diameter distribution of oak plants in Beijing were studied by using indices of skewness and kurtosis. The stand diameter distributions in the oak plants were simulated by four probability density functions: Weibull distribution, Normal distribution, Lognormal distribution and Gamma distribution. The four distribution functions were tested through Kolmogorov-Smirnov test. The results showed that the majorities of oak plants were medium and small diameter at breast height trees and Gamma distribution simulates much better than that of the others. In the end, the relations between Gamma function parameters and stand variables were established.

Key words: oak plants; diameter distribution; skewness; kurtosis

林分直径结构是最基本的林分结构，它不仅是林分树高、断面积和材积等结构的基础，而且也是估算林分材种出材量、指导抚育间伐、掌握林木枯损过程、确定合理轮伐期以及评定生产力的基础。精确的林分直径结构预测能为科学地进行人工林的定向培育和经营提供可靠的理论依据[1-3]。

栎类质地坚硬、耐腐蚀、强度大、抗冲击、具有广泛的用途和很高的经济价值[4]。国内许多学者对栎类林分做过相关研究，如王利等[5]研究了麻栎树高与胸径的相关关系，马增旺等[6]对 Weibull 分布的参数估计在蒙古栎天然林直径分布研究中的应用进行了研究，赵浩彦等[7]对北京栓皮栎林分的干形质量做了初步研究，张文辉等[8]研究了北京地区东灵山辽东栎林优势林木种群的直径结构等。但是，未曾见到国内学者对北京山区天然栎林的直径分布结构变化的研究。因此，笔者利用 Weibull 分布、Normal 分布、Lognormal 分布、Gamma 分布 4 种分布函数，模拟北京山区天然栎林直径分布，并选择出较优的直径分布模型，进而预测栎类林分直径分布的变化趋势。对研究栎类林分的直径分布规律、了解林分的结构特征、估计栎类林分蓄积量、预测林分生长和进行密度调控有重要意义，为北京山区栎类天然林的经营管理提供重要依据。

1 材料与方法

1.1 研究地概况

北京山区位于北京市北部和西部，北部山地为军都山，属于燕山山脉；西部山区称西山，属于太行

收稿日期：2008-12-11　修回日期：2009-04-13
基金项目："十一五"国家科技支撑计划重点项目(2006BAD23B0202-1)
作者简介：张雄清，男，在读硕士，研究方向为森林模型。
*通讯作者：雷渊才，男，博士生导师。Email：yclei@caf.ac.cn。

山脉的小五台山余脉，海拔1 000 m以上，最高峰东灵山海拔2 303 m，也是华北地区较高的山峰。

北京山区属于暖温带半湿润—半干旱季风气候。年平均气温2～8℃，年≥0℃的积温2 300～3 600℃；年日照时数2 600 h；年降水量500 mm左右，多集中在植物生长的旺季6～8月，约为全年降水量的75%，对植物生长发育十分有利；山区的无霜期随地势的变化有所不同，但通常不超过160 d，而东灵山顶峰5月份仍有积雪。

北京山区土壤以山地棕壤、亚高山草甸土和褐土为主，一般呈微酸性反应。成土母质主要有花岗岩、砂岩、砾岩、安山岩等。

1.2 数据收集与整理

数据来源于2001年北京市林业调查设计院一类调查结果。其中，栎类天然林固定样地78个，每个样地面积为0.067 hm^2。样地主要调查因子有胸径、方位角、林分年龄、平均树高、郁闭度、地位级、水平距、坡向、坡位、坡度、海拔高度、土层厚度等。考虑到分布模型的检验，从78个样地中随机抽取63个样地用于拟合直径分布，剩下的15个样地作为模型的检验。表1为栎类林分因子统计数据。

表1 林分及林木特征因子

Table 1 Descriptive statistics of stand and tree variables

项目	林分特征因子(78块)				林木特征因子(6 678株)	
	平均直径 /cm	林分年龄 /a	优势木平均高 /m	林分密度 /(株·hm^{-2})	胸高断面积 /(m^2·株$^{-1}$)	胸高直径 /cm
均值	10.85	31.58	7.02	1 141	0.011 340	10.99
标准差	4.14	15.33	2.60	564	0.011 192	4.89
最小值	6.40	5.00	0.80	388	0.001 963	5.00
最大值	27.00	70.00	16.00	2 716	0.128 825	40.50

1.3 方法

1.3.1 *偏度和峰度* 偏度、峰度是描述林分直径分布特征的主要指标。偏度和峰度的计算公式为：

$$偏度：SK=\mu^3/\rho^3 \quad (1)$$

$$峰度：ST=\mu^4/\rho^4-4 \quad (2)$$

式中：ρ为标准差，μ^3、μ^4分别为三阶、四阶中心矩。

可以看出，偏度(SK)表示非对称分布的偏斜方向与偏斜程度，当$SK>0$，表示左偏；$SK<0$，表示右偏；SK的绝对值愈大，表明偏斜程度愈大。峰度(ST)表示分布的尖峭或平坦程度，$ST>0$，表示尖峭；$ST<0$，表示平坦，正态分布的峰度为0；ST绝对值愈大，表明与正态分布的差别愈大[9]。

1.3.2 *林分直径分布模型的选择* 选择广泛使用的Weibull分布、Normal分布、Lognormal分布和Gamma分布函数描述北京山区栎类天然林林分的直径分布，并对这4种分布使用检验指标进行模拟比较，最后选择出适合于该地区栎类林分的直径分布模型。选好模型后，利用林分的特征因子和林分直径分布的主要指标，建立该模型的参数预估方程，并对这些因子进行共线性检验。

(1)Weibull分布。Weibull分布的概率函数密度为：

$$f(a;a,b,c)=\frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1}\exp\left[-\left(\frac{x-a}{b}\right)^{c}\right] \quad (a\leqslant x\leqslant\infty) \quad (3)$$

对该函数取积分后，得到累积分布函数：

$$F(x;a,b,c)=1-\exp\left[-\left(\frac{x-a}{b}\right)^{c}\right] \quad (4)$$

式中：x为林木实测直径值，a为位置参数，b为尺度参数，c为形状参数。

当$c<1$时，该曲线为反"J"型分布函数；$1<c<3.6$时，为具有正偏的山状曲线的密度函数；$c=3.6$时，近似于正态分布；$c>3.6$时，分布曲线由正偏逐渐移向左偏；$c=2$时，为该分布的特殊情况，即Rayleigh分布；当$c\to\infty$时，变为单点分布。

在(3)式中，当$a=0$，则为两参数Weibull分布密度函数：

$$f(x;b,c)=\frac{c}{b}\left(\frac{x}{b}\right)^{c-1}\exp\left[-\left(\frac{x}{b}\right)^{c}\right] \quad (5)$$

Weibull分布函数是描述林分直径分布的有效模型，近年来被广泛应用[10-13]。

(2)Normal(正态)分布。Normal分布的概率密度函数为：

$$f(x)=\frac{1}{\sqrt{2\pi}c}\exp\left[-\left(\frac{(x-b)^2}{2c^2}\right)\right] \quad (6)$$

式中：b为正态分布的数学期望；c为其标准差。正态曲线是一条对称型山状。b决定正态分布的位置，c决定正态分布的分散大小；c越小，分布越集中，正态曲线呈高而瘦；c越大，分布越分散，正态曲线呈矮而胖[14-15]。

Normal分布是一种常见的分布，许多树种的直径分布都近似Normal分布。

(3)Lognormal(对数正态)分布。Lognormal分布的概率密度函数为：

$$f(x)=\frac{1}{\sqrt{2\pi}cx}\exp\left(\frac{-(\ln(x)-b)^2}{2c^2}\right) \quad (x>0) \quad (7)$$

式中：b 为随机变量 $\ln(x)$ 的平均数，c 为随机变量 $\ln(x)$ 的标准差。

Lognormal 分布为偏态概率分布，有的林分直径分布为此形状[16-17]。

(4) Gamma 分布。Gamma 分布的概率密度函数为：

$$f(x)=\frac{b^c}{\Gamma(c)}x^{(c-1)}\exp(-bx)\quad x\geqslant 0, b>0, p>0 \tag{8}$$

式中：b、c 为参数，其中，$b>0$，称为尺度参数，$c>0$，称为形状参数。$c>1$ 时，该密度函数是单峰，峰值位于 $x=(c-1)/b$；$1<c\leqslant 2$ 时，密度函数曲线先上凸，后下凹；当 $c>2$ 时，其密度函数曲线是先下凸，中间上凸，最后又下凸；当 $c=1$ 时，为指数分布；当 $c=n/2$，$b=1/2$ 时，为 χ^2 分布。

Gamma 分布是适应性比较强的概率分布[18-19]。

1.3.3　*分布模型的检验标准*　采用 Weibull 分布、Normal 分布、Lognormal 分布和 Gamma 分布对各样地的直径结构进行分布拟合，并利用 Kolmogorov-Smirnov(KS) 检验来选择拟合最好的分布模型，KS 统计值越小，表示该分布的拟合效果越好[20]。

$$D=\underset{-<x<\infty}{Sup}|F^*(x)-F_0(x)| \tag{9}$$

式中：D 为 KS 统计值；Sup 为最大值；$F^*(x)$ 为林分中各样地直径分布的理论概率；$F_0(x)$ 为林分中各样地直径分布的实际概率。

1.3.4　*模型的发展*　在选择了分布模型后，为了进一步预测和分析北京山区栎类天然林林分的直径分布规律，有必要对模型的参数进行预估。在建立模型的参数预估方程时，用林分年龄、平均高、平均胸径、密度的单一因子来估计的回归模型，往往不能提供充足的信息。但用年龄、林分平均高、平均胸径、密度等多因子来拟合时，首先要有效诊断所用变量间是否存在多重共线关系及其严重程度，这是保证最小二乘法可得到最优解及参数预估精度和稳定性的必须解决的问题。多重共线性的存在，将会使参数预估方程的建立出现以下几个问题：(1)各个解释变量对被解释变量的影响很难精确鉴别；(2)由于存在多重共线性，模型回归参数估计量的方差会很大，这将使预估的参数值不准确；(3)模型回归参数的估计量对删除或增加少量的观测值以及删除一个不显著的解释变量都可能非常敏感。

本文采用方差膨胀因子(VIF, variance inflation factor) 作为共线性的检验指标。

$$VIF=1/(1-R_j^2) \tag{10}$$

式中：R_j^2 表示第 j 个自变量对其他自变量进行回归得到的判定系数 R^2。

VIF 值越大，说明自变量 j 与其他自变量间存在共线性的可能性越大。通常，当 $VIF>10$ 时，认为变量 x_j 与其他变量之间存在多重共线性[19,21]。建立预估方程后，利用未参与建模的样地对该预估方程进行适应性检验。参数估计及模型检验均利用 SAS 软件完成[22]。

2　结果与分析

2.1　偏度、峰度分析

从表 2 可知，偏度系数正值的样地数占总样地数的 92.78%，平均值为 0.824 0。从数理统计上分析，这些样地直径分布偏斜程度不大，曲线顶峰左偏现象明显，而右偏不明显。从栎类林分特征上看，林分胸径差异不大，且中等稍偏小胸径的林木占大多数。峰度为正值的样地数占 60.82%，平均值为 0.910 2，表明大多数样地的栎类林分的直径分布曲线比标准正态分布曲线略为陡峭，但总的来说曲线比较平坦，直径分布离散程度小而集中。从栎类林分的生长规律上看，与偏度分析结果一致。这种直径分布特征主要是由于研究的对象起源是人工林，同时还与林分密度、平均直径以及年龄因素综合作用有关[23-24]。

表 2　偏度、峰度统计表($n=68$)①

Table 2　Statistics of skewness and kurtosis

偏度平均值	偏度(α)	峰度平均值	峰度(β)
0.8240	92.78%	0.9102	60.82%

①α 为偏度为正值的样地数所占比例；β 为峰度为正值的样地数所占比例。

2.2　林分年龄、平均直径与偏度和峰度的相关性

经查表，样本数为 63，在 0.01 的显著水平，临界值 $F_{(1,61)}=7.03$。分析发现，林分年龄与偏度存在较显著的相关性，而与峰度相关性不显著；平均直径与偏度、峰度的相关性较显著。林分年龄和平均直径与偏度之间的相关性均较显著，而且趋势一致(表 3)，说明随着林分年龄的增加或平均直径的增大，林分直径分布的偏度趋于增加。这也反映了随着林分的生长，强烈的竞争最终将导致大量林木的自然稀疏。高密度的林分促进了林分竞争，也使林分自然稀疏的发生时间来的更早。自然稀疏一旦发生，首先淘汰的是严重被压的林木，人工林内主要为小径阶的林木，于是，林分算术平均直径会显著提高，而直径分布最高峰所对应的径阶值还会保持不变或移动较小，这样必然导致林分直径分布曲线向左偏。

2.3　直径分布函数的选择及检验

这 4 种分布中，为了选择拟合最好的分布模型，

采用 Kolmogorov-Smirnov(KS)检验法进行选择，其显著水平为 0.01(表 4)。

由表 4 知，Gamma 分布模型 KS 统计值的平均值和标准差分别是 0.102 8 和 0.032 7，其余 3 种分布依次是 Lognormal(0.105 6、0.035 0)、Weibull(0.114 9，0.034 7)、Normal(0.116 2、0.036 8)，因此，根据 KS 平均值的大小，得出 4 种分布拟合的效果排序为：Gamma＞Lognormal＞ Weibull＞Normal。陈新美等[25]、陈昌雄等[18]利用这 4 种分布对左偏的阔叶林直径分布进行了拟合，并通过 χ^2 检验得出一致结果。由此得出 Gamma 分布对于左偏的阔叶林拟合效果较好。

表 3 栎类林分年龄、平均直径与直径分布特征数的相关分析①

Table 3 Correlation analysis between stand variables and diameter distribution characteristic factors

	回归常数	回归系数	相关系数	F 检验统计值
林分年龄与偏度	38.511	−6.289	0.274	7.708**
林分年龄与峰度	34.234	−0.994	0.157	2.389
平均直径与偏度	12.716	−2.080	0.468	26.595**
平均直径与峰度	11.311	−0.338	0.275	7.780**

① ** 表示在 0.01 水平上显著。

表 4 各分布参数统计

Table 4 Parameters and KS statistic of different diameter distribution models

方法	b			c			KS 统计值			
	最小值	最大值	标准差	最小值	最大值	标准差	最小值	最大值	标准差	平均值
Normal	6.07	24.02	2.91	1.05	7.79	1.50	0.049 8	0.206 3	0.036 8	0.116 2
Lognormal	1.80	3.20	0.25	0.14	0.54	0.09	0.049 3	0.242 7	0.035 0	0.105 6
Weibull	6.50	15.00	3.29	1.90	7.90	1.14	0.057 7	0.192 4	0.034 7	0.114 9
Gamma	0.14	4.49	0.89	1.87	42.20	7.72	0.046 0	0.231 5	0.032 7	0.102 8

从理论上来讲，Weibull 分布具有较大的灵活性，对直径分布适应性较强，可以拟合不同偏度、峰度的单峰山状曲线，又可以拟合倒“J”型曲线[26-27]，但在实际应用中由于参数估计需要用到较复杂的程序，在计算过程中可能出现迭代发散等异常情况而失败[28]，同时，Weibull 分布也不适合拟合幼中龄林的直径分布[29]。Normal 分布曲线变化小，只拟合林分发育过程某一阶段的直径分布，有一定局限性[13]。Lognormal 分布适应于幼龄林的直径分布[26]。因此，采用 Gamma 分布来确定北京山区栎类天然林林分直径分布规律更为适宜。

2.4 参数预估方程的建立

为了进一步预测和分析北京山区栎类天然林林分的直径分布规律，利用近似 Gamma 分布的 63 块样地的参数(b,c)的求解值分别与林分特征因子、直径分布的主要指标等建立多元线性回归方程，即可得到 Gamma 分布参数的预估方程。

根据表 5，VIF 的取值范围是 1.044～5.200，数值都不大，从而可以认为偏度、峰度、胸径、树高、年龄、密度之间不存在共线性；另外，由 t 检验可知，只有偏度、峰度、平均胸径的伴随概率接近于 0，其他都大于 0.05。因此，利用偏度、峰度、平均胸径 3 个变量建立参数 b,c 的预估方程：

$$b=-1.716+0.887SK-0.193ST+0.219D,\quad R=0.795 \tag{11}$$

$$c=40.54-16.17SK-3.65ST-2.04D,\quad R=0.566 \tag{12}$$

查表得 $R_{0.01}=0.320$，式(11)和式(12)的 R 值均大于 $R_{0.01}$，说明复相关系数显著。因此，利用上述回归方程把 15 块检验样地的平均胸径、偏度、峰度代入式(11)和式(12)中，可得到 Gamma 分布的参数估计值，再用 Gamma 分布函数进行拟合(表 6)。

表 5 参数 b,c 回归分析表①

Table 5 Regression analysis of parameter b,c

自变量	b		c		方差膨胀因子(VIF)
	t 统计值	伴随概率	t 统计值	伴随概率	
偏度	3.578	0.000	−6.455	0.000	5.200
峰度	−3.099	0.000	5.711	0.000	4.349
平均胸径	3.573	0.000	4.629	0.000	2.105
平均树高	1.005	0.318	−0.433	0.666	1.334
平均年龄	0.918	0.361	−1.566	0.124	1.512
林分密度	−1.463	0.147	−1.456	0.149	1.044

①显著水平 $\alpha=0.05$。

表 6 Gamma 分布拟合效果

Table 6 Results of gamma function simulation

标准地号	KS 统计值	p 值	标准地号	KS 统计值	p 值
3	0.112 568	＞0.250	686	0.0 820 279	＞0.250
11	0.139 083	＞0.250	811	0.1 154 488	＞0.250
37	0.103 823	＞0.250	875	0.3 507 884	＜0.001
60	0.085 515	＞0.250	932	0.1 343 427	＞0.250
111	0.350 788	＜0.001	1 699	0.1 325 963	＞0.250
254	0.129 533 7	＞0.250	2 543	0.1 245 975	＞0.250
288	0.101 189	＞0.250	2 720	0.1 178 563	＞0.250
478	0.424 212	＜0.001			

由表 6 知，经过 KS 检验后，有 4 块验样地不符合 Gamma 分布(p＜0.001)，符合率为73.33%，表明根据林分平均胸径、偏度、峰度来确定 Gamma 分布的参数估计值拟合该地区的栎类林分效果较好。

3 结论

北京山区栎类天然林林分直径分布形态属于单

峰型，基本上是左偏，直径分布的离散程度小而集中，从林分特征上反映出中等稍偏小胸径的林木占大多数。栎类林分年龄与偏度相关显著，平均直径与偏度、峰度呈较显著的正相关。

北京山区栎类天然林分直径分布用 Gamma 分布拟合效果最好。

Gamma 分布对于左偏的阔叶林直径分布拟合效果较好。

经过参数的回归分析，可以用林分平均胸径、偏度、峰度预估 Gamma 分布的参数，进而预测栎类林分直径结构的变化，这对北京山区栎类天然林的经营管理具有理论和实践意义。

参考文献：

[1] 田松岩，宋国华，宋存彦，等. 东北林区天然林林分结构及林分生产力的研究[J]. 林业科技，2005：30(4)：21-22.

[2] 孟宪宇. 测树学[M]. 北京：中国林业出版社，1996：66-67.

[3] 张建国，段爱国. 理论生长方程与直径结构模型的研究[M]. 北京：科学出版社，2004：88-97.

[4] 刘振西. 栎树资源及其开发利用[J]. 自然资源，1995，17(2)：56-59.

[5] 王利，丰震，白世红，等. 麻栎树高与胸径相关关系的研究[J]. 山东林业科技，2005(4)：33-34.

[6] 马增旺，李宗领，王林，等. Weibull 分布的参数估计在蒙古栎天然林直径分布研究中的应用[J]. 河北林业科技，1996 (5)：23-16.

[7] 赵浩彦，陆元昌. 北京栓皮栎林分的干形质量初步研究[J]. 林业资源管理，2008，2(1)：44-48.

[8] 张文辉，赵则海，孙海芹，等. 东灵山辽东栎林优势林木种群直径结构的研究[J]. 植物研究，2002，22(1)：84-90.

[9] 李荣伟，覃志刚，潘攀，等. 杜仲人工林林分直径分布研究[J]. 四川林业科技，2000，21(2)：1-6.

[10] 孟宪宇. 使用 Weibull 分布对人工油松林直径分布的研究[J]. 北京林学院学报，1985，7(1)：30-40.

[11] 孟宪宇. 使用 Weibull 分布对树高分布和直径分布的研究[J]. 北京林业大学学报，1988，10(1)：40-47.

[12] LIU C M, ZHANG S Y, LEI Y C. Evaluation of three methods for predicting diameter distributions of black spruce (*Picea mariana*) plantations in central Canada[J]. Can. J. For. Res., 2004, 34: 2424-2432.

[13] JAN M E. Characterisation of diameter distribution using the weibull function: method of moments [J]. Eur J Forest Res, 2006, 125: 427-439.

[14] 茆诗松. 概率论与数理统计[M]. 北京：中国统计出版社，2003：109-111.

[15] 张文，高天雷. 马尾松林分直径结构的研究[J]. 四川林勘设计，2001(2)：26-30.

[16] 陈瑞杰，盖新敏，翁怀锋，等. 突脉青冈天然林主要种群直径分布结构特征的研究[J]. 福建林业科技，2005，32(2)：27-30.

[17] BLISS C I, REINKER K, A LOG A. Normal approach to diameter distributions in even-aged stands[J]. For. Sci., 1964, 10: 350-360.

[18] 陈昌雄，黄宝龙，林立法. 南平市延平区天然阔叶林直径分布的研究[J]. 福建林业科技，2004，31(1)：1-4，9.

[19] NELSON T C. Diameter distribution and growth of loblolly pine [J]. For. Sci., 1964, 10(1): 105-114.

[20] ZHANG S Y, LEI Y C, JIANG Z H. Modelling the relationship of tree-level product value with tree characteristics in black spruce [J]. The Forestry Chronicle, 2006 (9): 390-699.

[21] 付小勇. 云南松林分生长模型研究[D]. 昆明：西南林学院，2007.

[22] 高惠璇，耿直，李贵斌，等. SAS 系统——SAS/STAT 软件使用手册[M]. 北京：中国统计出版社，1997.

[23] 刘金福. 格氏栲天然林主要种群直径分布结构研究[J]. 福建林学院学报，2001，21(4)：325-328.

[24] 王鹏. 马尾松人工混交林直径分布规律的研究[J]. 江西林业科技，2005(4)：6-7.

[25] 陈新美，张会儒. 柞树林直径分布模拟研究[J]. 林业资源管理，2008(1)：39-43.

[26] BAILEY R L, DELL T R. Quantifying diameter distributions with the Weibull distribution [J]. For Sci, 1973, 19: 97-104.

[27] NAGEL J, BIGING G S. Schazung der Weibull function zur Generierung con durchmesserverteilungen [J]. Allg Forst-uJ-Ztg, 1995, 166(9-10): 185-189.

[28] 惠淑荣. 日本落叶松直径结构规律的研究[J]. 沈阳农业大学学报，1993，24(4)：365-369.

[29] NANANG D M. Suitability of the Normal, Log-normal and Weibull distribution for fitting diameter distributions of neem plantations in Northern Ghana[J]. For Ecol Management, 1998, 103(1): 1-7.