

均匀递归树的分支结构*

冯群强** 苏 淳 胡治水

(中国科学技术大学统计与金融系, 合肥 230026)

摘要 研究均匀递归树的分支结构中的有关问题. 用独立和的方法得出了在大小为 n 的均匀递归树上分支数目 η_n 的分布律, 建立了 η_n 的强大数律, 中心极限定理和重对数律; 证明了 η_n 和顶点 n 的深度 ξ_n 是同分布的; 得出了大小为 m 的分支数 $\zeta_{n,m}$ 的分布律, 并且证明了 $\zeta_{n,m}$ 的极限分布就是参数 $\lambda = 1/m$ 的 Poisson 分布, 给出了各种分支数目的联合分布及其极限分布; 还研究了大小为 n 的均匀递归树上最大分支的大小, 证明了在 $n \rightarrow \infty$ 时, 它几乎必然趋于无穷.

关键词 均匀递归树 分支 深度 分布律 极限定理

1 引言

不含圈的简单连通图称为树^[1]. 大小为 n 的递归树是一类按照随机连接方式得到的树. 递归树的构造过程如下^[2]: 将顶点集合记为 $\{1, 2, \dots, n\}$, 并且给定一组分布律 $\{p_{k,i}, i = 1, 2, \dots, k\}, k = 1, 2, \dots, n-1$, 即

$$p_{k,i} \geq 0, \quad \sum_{i=1}^k p_{k,i} = 1, \quad k = 1, 2, \dots, n-1.$$

第 1 步, 将所有顶点放在一个平面上; 第 2 步将顶点 2 与顶点 1 连结; 第 3 步将顶点 3 与集合 $\{1, 2\}$ 中的一个顶点连结, 连结的概率分别是 p_{21}, p_{22} ; 一般地, 第 $k+1$ 步将顶点 $k+1$ 与集合 $\{1, 2, \dots, k\}$ 中的一个顶点连结, 连结的概率分别为 $p_{k,i}, i = 1, 2, \dots, k$. 经过 n 步之后, 就得到了以顶点 1 为根的一个根树, 称之为递归树. 如果

$$p_{k,i} = 1/k, \quad i = 1, 2, \dots, k, \quad k = 1, 2, \dots, n-1,$$

即在每一步上都将新的顶点以相等的概率与已有的顶点连接, 那么所得到的递归树就称为均匀递归树, 记作 T_n . 由于造树过程中的每一步都有多种不同的可能

2003-10-17 收稿, 2005-01-28 收修改稿

* 国家自然科学基金 (批准号: 10371117) 和中国科学技术大学高水平大学建设基金资助项目

** E-mail: fengqq@mail.ustc.edu.cn

SCIENCE IN CHINA Ser. A Mathematics

连法, 所以可得到 $(n-1)!$ 种不同的树, 即 $|T_n| = (n-1)!$, 并且每种均匀递归树出现的可能性都是相等的. 递归树与许多应用问题都有联系, 如传染病传染规律研究 [3], 家谱研究 [4], 金字塔式结构研究 [5], 等等. 下面给出一个为均匀递归树的传染病模型.

假设某个群体中一共有 n 个人先后感染了某种传染病 (例如严重急性呼吸综合征), 其中只有 1 个人去过疫区, 从而他是这个群体的感染源头. 第 2 个人一定是被他传染的; 由于不清楚传染机制, 所以假定第 3 个人分别以概率 $1/2$ 被前两个人所传染; 一般地, 假定第 k 个人以概率 $1/(k-1)$ 被前 $(k-1)$ 个人所传染, $k = 2, 3, \dots, n$. 如果以顶点 k 表示第 k 个人, 对于 $k \geq 2$, 将顶点 k 与传染给他疾病的人所对应的顶点之间连一条边, 那么就得到了一个均匀递归树. 由此看来, 研究均匀递归树有利于从一定程度上弄清严重急性呼吸综合征之类的传染病的传染规律.

在 T_n 上, 用 D_j 表示第 j 代顶点的集合. 以 D_1 中的点作为根点的子树称为 T_n 的分支, 分支也是递归树 [6]. 显然 T_n 的分支数就是 D_1 中的顶点数目 $|D_1|$, 将之记作 η_n . 大小为 m ($1 \leq m \leq n-1$) 的分支称为 m -分支, m -分支的数目记作 $\zeta_{n,m}$. 显然有 $\eta_n = \sum_{i=1}^{n-1} \zeta_{n,i}$. 此外, 当顶点 $k \in D_j$ 时, 称顶点 k 的深度为 j , 以 ξ_k 记顶点 k 的深度.

不少作者都对深度问题作过讨论, 例如, Szymański 给出了顶点 n 的深度 ξ_n 的分布 [7], Devroye 证明了 ξ_n 的中心极限定理 [8], Mahmoud 进一步研究了 ξ_n 和 $\sum_{k=1}^n \xi_k$ 的极限性状 [9,10], Meir 和 Moon 还研究了各代子点数目分布律 [6].

不难想像, T_n 中的分支的结构问题是均匀递归树性质研究中的重要方面, 但是却未见有文献加以讨论, 本文的目的在于讨论其中的一些问题. 在第 2 节中, 充分利用顶点 $2, 3, \dots, n-1$ 属于 D_1 的事件的相互独立性, 非常方便地得到了 D_1 中的顶点数目 η_n 的强大数律、中心极限定理和重对数律, 并且直接推出了 η_n 的分布律; 通过第 3 节的讨论, 证明了 η_n 和顶点 n 的深度 ξ_n 的同分布性; 在第 4 节中, 不但给出了 m -分支的数目 $\zeta_{n,m}$ 的分布和极限分布, 而且得到了各种分支数目的联合分布, 算出了该联合分布的数学期望和协方差阵, 同时还证明了分支数目的渐近独立性, 为进一步探讨可列无穷大均匀递归树的性质铺平了道路; 最后, 在第 5 节中, 研究了 T_n 上的最大分支的大小 ν_n , 证明了在 $n \rightarrow \infty$ 时, ν_n 几乎必然趋于无穷.

2 均匀递归树的分支数目

本节讨论 η_n 的性质, 并顺便给出 η_n 的分布律. 虽然文献 [6] 的结果中包含了 η_n 的分布律, 但是其推导办法很繁琐, 而且未对 η_n 的性质作深入的讨论.

记 $X_j = I(j+1 \in D_1)$, 其中 $I(A)$ 是事件 A 的示性函数. 由造树过程知, 顶点 j 是否属于 D_1 , 只取决于造树过程的第 j 步, 所以 $X_1, X_2, \dots, X_{n-1}$ 是相互独

立的 Bernoulli 随机变量, 并且

$$\eta_n = \sum_{j=1}^{n-1} X_j = 1 + \sum_{j=2}^{n-1} X_j. \quad (2.1)$$

由造树规则知

$$P(X_j = 1) = P(j+1 \in D_1) = 1/j, \quad j = 1, \dots, n-1,$$

所以

$$EX_j = \frac{1}{j}, \quad \text{Var} X_j = \frac{j-1}{j^2},$$

$$E|X_j - EX_j|^3 \leq EX_j^3 = 1/j, \quad j = 1, \dots, n-1.$$

记 $\log x = \ln \max\{e, x\}$, 则当 $n \rightarrow \infty$ 时, 有

$$E\eta_n = \sum_{j=1}^{n-1} EX_j = \sum_{j=1}^{n-1} \frac{1}{j} = \log n + O(1), \quad (2.2)$$

$$B_n := \text{Var}\eta_n = \sum_{j=1}^{n-1} \text{Var} X_j = \sum_{j=1}^{n-1} \frac{j-1}{j^2} = \log n + O(1), \quad (2.3)$$

$$G_n := \sum_{j=1}^{n-1} E|X_j - EX_j|^3 \leq C \sum_{j=1}^{n-1} EX_j^3 = C \sum_{j=1}^{n-1} \frac{1}{j} = C \log n + O(1). \quad (2.4)$$

由此不难得出 η_n 的有关极限定理. 首先给出 η_n 的 Marcinkiewicz 强大数律:

定理 2.1 对任意 $1 \leq p < 2$, 当 $n \rightarrow \infty$ 时, 都有

$$\frac{\eta_n - \log n}{\log^{1/p} n} \rightarrow 0 \quad \text{a.s.} \quad (2.5)$$

特别地,

$$\eta_n / \log n \rightarrow 1 \quad \text{a.s.} \quad (2.6)$$

证 由表达式 (2.1) 知, η_n 是相互独立的随机变量的和, 且上述计算表明

$$\sum_{n=1}^{\infty} \frac{\text{Var} X_n}{\log^{2/p} n} \leq \sum_{n=1}^{\infty} \frac{1}{n \log^{2/p} n} < \infty.$$

所以由文献 [11] 中定理 6.6 知

$$\frac{\eta_n - E\eta_n}{\log^{1/p} n} = \frac{\sum_{j=1}^{n-1} (X_j - EX_j)}{\log^{1/p} n} \rightarrow 0 \quad \text{a.s.}$$

又显然有

$$\frac{E\eta_n - \log n}{\log^{1/p} n} \rightarrow 0, \quad \frac{E\eta_n}{\log n} \rightarrow 1,$$

所以定理的结论成立.

下面证明 η_n 的中心极限定理:

定理 2.2

$$\frac{\eta_n - \log n}{\sqrt{\log n}} \xrightarrow{d} N(0, 1).$$

证 由于当 $n \rightarrow \infty$ 时, 有

$$\frac{1}{B_n^{3/2}} \sum_{j=1}^{n-1} E|X_j - EX_j|^3 \leq C(\log n)^{-1/2} \rightarrow 0,$$

所以由 Lyapunov 条件知

$$\frac{\eta_n - E\eta_n}{B_n} \xrightarrow{d} N(0, 1).$$

再由 (2.2) 和 (2.3) 式, 即知定理结论成立.

容易证明 η_n 的重对数律:

定理 2.3

$$\limsup_{n \rightarrow \infty} \frac{\eta_n - \log n}{\sqrt{2 \log n \log \log \log n}} = 1 \quad \text{a.s.},$$

$$\liminf_{n \rightarrow \infty} \frac{\eta_n - \log n}{\sqrt{2 \log n \log \log \log n}} = -1 \quad \text{a.s.}$$

证 由表达式 (2.1) 知, η_n 是一致有界的相互独立的随机变量的和, 并且当 $n \rightarrow \infty$ 时, 有

$$E\eta_n = \log n + O(1), \quad B_n = \log n + O(1), \quad \log \log B_n \sim \log \log \log n,$$

故由著名的 Kolmogorov 重对数律 (参阅文献 [11] 的定理 7.1) 立得本定理.

利用表达式 (2.1), 很容易写出 η_n 的分布律. 以带下标的 m 表示正整数, 记

$$\beta_{n,0} = 1, \quad \beta_{n,k} = \sum_{1 \leq m_1 < \dots < m_k \leq n-2} m_1 \cdots m_k, \quad k = 1, 2, \dots, n-2, \quad (2.7)$$

其中, 当取定 $k \in \{1, 2, \dots, n-2\}$ 时, 上式右端表示对一切不同的正整数 $1 \leq m_1 < \dots < m_k \leq n-2$ 的乘积求和.

定理 2.4 当 $n \geq 2$ 时, 有

$$P(\eta_n = k) = \frac{\beta_{n,n-1-k}}{(n-1)!}, \quad k = 1, \dots, n-1. \quad (2.8)$$

证 由于 η_n 是 $n-1$ 个相互独立的 Bernoulli 随机变量 $X_1, X_2, \dots, X_{n-1}$ 的和, 所以事件 $(\eta_n = k)$ 发生, 当且仅当 $X_1, X_2, \dots, X_{n-1}$ 中有 k 个等于 1, 其余的等于 0, 故有

$$\begin{aligned} P(\eta_n = k) &= \sum_{1 \leq j_1 < \dots < j_k \leq n-1} \left(\prod_{i=1}^k \frac{1}{j_i} \prod_{j \notin \{j_1, \dots, j_k\}} \left(1 - \frac{1}{j}\right) \right) \\ &= \frac{1}{(n-1)!} \sum_{1 \leq j_1 < \dots < j_k \leq n-1} \prod_{j \notin \{j_1, \dots, j_k\}} (j-1) \\ &= \frac{1}{(n-1)!} \sum_{1 \leq m_1 < \dots < m_{n-1-k} \leq n-2} m_1 \cdots m_{n-1-k} = \frac{\beta_{n,n-1-k}}{(n-1)!}. \end{aligned}$$

3 顶点 n 的深度

正如前面所言, 许多文献都研究过均匀递归树 T_n 上顶点 n 的深度 ξ_n , 这种研究具有其应用背景: 在选取最佳方案时, 由于方案的比较涉及到多个指标, 而利用计算机一般只可以对两个方案的优劣进行比较, 现在经过比较, 已经筛选出 $n-1$ 个比原有方案更好的方案, 要从它们中进一步找出最好的方案来 (假定其中的 1 号方案的各项指标都优于其他方案). 于是我们以相等的概率 $1/(n-1)$ 从这 $n-1$ 个方案中取出一个方案 j_1 , 将其余 $n-2$ 个方案逐一同它比较, 如果它们都比它差, 那么它就是 1 号方案, 此时过程结束; 否则我们就可以找出比它更好的 j_1-1 个方案. 于是第 2 步, 再以相等的概率 $1/(j_1-1)$ 从这 j_1-1 个方案中取出一个方案 j_2 , 如果经过逐一比较, 发现它最好, 那么它就是 1 号方案; 否则又将继续, 直至到达方案 1 为止. 另一方面, 如果反观顶点 n 的选择自己祖先的过程, 则可看出: 第 1 步, 它以相等的概率 $1/(n-1)$ 从已有的 $n-1$ 个顶点中选择自己的父点 j_1 , 如果 $j_1=1$, 则过程结束; 否则, 顶点 j_1 又以相等的概率 $1/(j_1-1)$ 从已有的 j_1-1 个顶点中选择自己的父点 j_2 , 如果 $j_2=1$, 则过程结束; 否则又将继续, 直至到达顶点 1 为止. 所以在方案选择问题中, 所需要的选择步数就刚好是均匀递归树上顶点 n 的深度 ξ_n .

我们来证明在 T_n 上, ξ_n 与分支数目 η_n 同分布. 为此, 给出 ξ_n 的一种刻画. 在均匀递归树 T_n 上存在着一条由根点到达顶点 n 的最短路径 Z_n , Z_n 上除根点之外的其余顶点的数目就是顶点 n 的深度 ξ_n . 显然 $n \in Z_n$, 因此有

$$\xi_n = 1 + \sum_{j=2}^{n-1} I(j \in Z_n). \quad (3.1)$$

先来讨论 $I(j \in Z_n)$ 的分布律. 显然 $P(n-1 \in Z_n) = 1/(n-1)$, 这是因为只有当顶点 n 以其为父点时, 才有 $n-1 \in Z_n$. 而 $P(n-2 \in Z_n) = 1/(n-2)$, 这是因为或者顶点 n 以其为父点, 或者顶点 n 以顶点 $n-1$ 为父点, 而顶点 $n-1$ 以其为父点时, 有 $n-2 \in Z_n$, 所以

$$P(n-2 \in Z_n) = \frac{1}{n-1} + \frac{1}{n-1} \cdot \frac{1}{n-2} = \frac{1}{n-2},$$

由此并结合归纳法, 即知

$$P(i \in Z_n) = 1/i, \quad i = 2, \dots, n-1. \quad (3.2)$$

事实上, 假设对某个 $2 \leq i \leq n-2$, (3.2) 式中的结论对 $i < j \leq n-1$ 都已经成立, 我们来证明该结论对 i 也成立. 以 $A_{i,j}$ 表示顶点 j 是顶点 i 的子点的事件, 易知

$$I(i \in Z_n) = \sum_{j=i+1}^n I(j \in Z_n, A_{i,j}),$$

由于事件 $A_{i,j}$ 只与造树过程的第 j 步有关, 而事件 $(j \in Z_n)$ 只与第 $j+1, \dots, n$

步有关, 所以事件 $A_{i,j}$ 与 $(j \in Z_n)$ 相互独立. 因此,

$$\begin{aligned} P(i \in Z_n) &= EI(i \in Z_n) \\ &= \sum_{j=i+1}^n EI(j \in Z_n, A_{i,j}) \\ &= \sum_{j=i+1}^n P(j \in Z_n, A_{i,j}) \\ &= \sum_{j=i+1}^n P(j \in Z_n) P(A_{i,j}) \\ &= \sum_{j=i+1}^{n-1} \frac{1}{j(j-1)} + \frac{1}{n-1} \\ &= 1/i. \end{aligned}$$

再来证明随机变量 $I(2 \in Z_n), \dots, I(n-1 \in Z_n)$ 相互独立. 易知, 对任意 $2 \leq k \leq n-2$ 和 $2 \leq j_k < \dots < j_2 < j_1 \leq n-1$, 事件

$$(I(j_i \in Z_n) = 1, I(j \in Z_n) = 0, j \neq j_i, i = 1, 2, \dots, k)$$

就是顶点 n 以顶点 j_1 为父点, 顶点 j_i 以顶点 j_{i+1} 为父点的事件, $i = 1, \dots, k-1$. 记 $j_0 = n$, 则由父点的选择过程知

$$\begin{aligned} &P(I(j_i \in Z_n) = 1, I(j \in Z_n) = 0, j \neq j_i; i = 1, 2, \dots, k) \\ &= \prod_{i=0}^k \frac{1}{j_i - 1} \\ &= \frac{1}{(n-1)!} \prod_{j \notin \{j_1, \dots, j_k\}} (j-1) \\ &= \prod_{i=1}^k \frac{1}{j_i} \prod_{j \notin \{j_1, \dots, j_k\}} \frac{j-1}{j} \\ &= \prod_{i=1}^k P(I(j_i \in Z_n) = 1) \prod_{j \notin \{j_1, \dots, j_k\}} P(I(j_i \in Z_n) = 0). \end{aligned}$$

这就证明了 $I(2 \in Z_n), \dots, I(n-1 \in Z_n)$ 是相互独立的 Bernoulli 随机变量.

将表达式 (3.1) 与 η_n 的表达式 (2.1) 相比较 (注意在 (2.1) 式中 $X_1 = I(2 \in D_1) = 1$), 发现它们完全相同, 所以 ξ_n 与 η_n 有相同的分布. 故而将定理 2.1~2.3 中的 η_n 换成 ξ_n 后仍然成立. 虽然 Devroye^[8] 和 Mahmoud^[10] 中证明过深度 ξ_n 的中心极限定理和大数律, 但是他们的证明方法不仅复杂得多, 而且只得到了弱大数律.

4 m -分支的数目

本节中研究各种不同大小的分支数目的分布及其极限分布, 并且给出所有

不同分支数目的联合分布以及联合极限分布.

4.1 $\zeta_{n,1}$ 的分布律和极限分布

在各种分支当中, 1-分支亦即顶点 1 的子叶占有特殊的地位, 所以首先单独对它进行讨论, 并且运用了与其他分支不同的讨论方法, 虽然 1-分支的性质可以看成下一小节内容的特殊情形.

先来研究 $\zeta_{n,1}$ 的概率递推公式. 利用全概率公式, 容易得出

$$\begin{aligned} & P(\zeta_{n+1,1} = k) \\ &= P(\zeta_{n,1} = k)P(\zeta_{n+1,1} = k \mid \zeta_{n,1} = k) \\ &\quad + P(\zeta_{n,1} = k-1)P(\zeta_{n+1,1} = k \mid \zeta_{n,1} = k-1) \\ &\quad + P(\zeta_{n,1} = k+1)P(\zeta_{n+1,1} = k \mid \zeta_{n,1} = k+1) \\ &= \frac{n-k-1}{n}P(\zeta_{n,1} = k) + \frac{1}{n}P(\zeta_{n,1} = k-1) + \frac{k+1}{n}P(\zeta_{n,1} = k+1), \end{aligned}$$

亦即

$$\begin{aligned} & nP(\zeta_{n+1,1} = k) - (n-1)P(\zeta_{n,1} = k) \\ &= P(\zeta_{n,1} = k-1) + (k+1)P(\zeta_{n,1} = k+1) - kP(\zeta_{n,1} = k). \end{aligned}$$

在上式中以 j 取代 n , 并对 j 由 1 到 n 求和, 得

$$\begin{aligned} & nP(\zeta_{n+1,1} = k) \\ &= \sum_{j=1}^n P(\zeta_j = k-1) + \sum_{j=1}^n (k+1)P(\zeta_j = k+1) - \sum_{j=1}^n kP(\zeta_j = k). \end{aligned} \quad (4.1)$$

利用递归式 (4.1), 可以证明如下结论:

定理 4.1 对任意 $n \in \mathbb{N}$, 都有

$$P(\zeta_{n,1} = k) = a_{n-k}/k!, \quad k = 0, 1, \dots, n-1, \quad (4.2)$$

其中

$$a_i = \sum_{j=0}^{i-1} \frac{(-1)^j}{j!}, \quad i \geq 1, \quad (4.3)$$

并且当 $n \rightarrow \infty$ 时, $\zeta_{n,1}$ 的极限分布是参数为 $\lambda = 1$ 的 Poisson 分布.

证 容易验证数列 $\{a_i\}$ 满足如下递归式:

$$a_1 = 1, \quad a_2 = 0, \quad a_j = \frac{1}{j-1} \sum_{i=1}^{j-2} a_i, \quad j \geq 2, \quad (4.4)$$

并且易知 (4.2) 式对 $n = 1, 2$ 成立.

假设 (4.2) 式对 n ($n \geq 2$) 成立, 我们来证明它对 $n+1$ 也成立. 由均匀递归树的结构立知

$$P(\zeta_{n+1,1} = n-1) = 0, \quad P(\zeta_{n+1,1} = n) = 1/n!.$$

而由归纳假设、(4.1) 和 (4.4) 式, 可知对任意 $k \in \{0, 1, \dots, n-2\}$, 都有

$$\begin{aligned} nP(\zeta_{n+1,1} = k) &= \sum_{j=1}^n P(\zeta_j = k-1) + \sum_{j=1}^n (k+1)P(\zeta_j = k+1) - \sum_{j=1}^n kP(\zeta_j = k) \\ &= \sum_{j=k}^n \frac{a_{j-(k-1)}}{(k-1)!} + \sum_{j=k+2}^n (k+1) \cdot \frac{a_{j-k-1}}{(k+1)!} - \sum_{j=k+1}^n k \cdot \frac{a_{j-k}}{k!} \\ &= \frac{a_{n-(k-1)}}{(k-1)!} + \frac{1}{k!} \sum_{j=1}^{n-k-1} a_j \\ &= \frac{(k+n-k)a_{n-(k-1)}}{k!} \\ &= \frac{na_{n+1-k}}{k!}, \end{aligned}$$

所以 (4.2) 式对任意 $n \in \mathbb{N}$ 都成立.

由 (4.2) 式立得 $\zeta_{n,1}$ 的极限分布为 $\lambda = 1$ 的 Poisson 分布.

此外, 由 (4.2) 式容易算得:

推论 4.1 对任意自然数 $n \geq 3$, 都有 $E\zeta_{n,1} = 1$, $\text{Var}\zeta_{n,1} = 1$.

4.2 一般情形

以 $[x]$ 表示不超过 x 的最大整数. 显然, 对任意 $m \in \{1, \dots, n-1\}$, 都有

$$P(\zeta_{n,m} \geq 0) = 1, \quad P\left(\zeta_{n,m} > \left[\frac{n-1}{m}\right]\right) = 0. \quad (4.5)$$

定理 4.2 在大小为 n 的均匀递归树上, m -分支的数目 $\zeta_{n,m}$ 的分布律是

$$P(\zeta_{n,m} = k) = \frac{1}{m^k k!} \sum_{i=0}^{\left[\frac{n-1}{m}\right]-k} \frac{(-1)^i}{i! m^i}, \quad k = 0, 1, \dots, \left[\frac{n-1}{m}\right]. \quad (4.6)$$

特别地, 当 $m > (n-1)/2$ 时, $\zeta_{n,m}$ 是 Bernoulli 随机变量, 有

$$P(\zeta_{n,m} = 1) = 1 - P(\zeta_{n,m} = 0) = 1/m.$$

证 从集合 $\{2, 3, \dots, n\}$ 中选 i 个 m 元子集, 将它们作成 i 个 m -分支. 虽然分支只能与顶点 1 相连, 但是每个 m -分支上的顶点却都有 $(m-1)!$ 种不同连法, 再让其余 $n-mi-1$ 个顶点按递归树的构造法则任意连接, 那么, 所有不同的连接方式数目为

$$\begin{aligned} &\binom{n-1}{m} \cdots \binom{n-m(i-1)-1}{m} ((m-1)!)^i (n-mi-1)! / i! \\ &= \frac{(n-1)!}{m^i i!}, \quad 1 \leq i \leq \left[\frac{n-1}{m}\right]. \end{aligned} \quad (4.7)$$

另一方面, 由 (4.5) 式易知

$$\sum_{j=0}^{\left[\frac{n-1}{m}\right]} P(\zeta_{n,m} = j) = 1,$$

在 (4.7) 式中令 $i = 1$, 则可得到

$$\frac{\binom{n-1}{m} (m-1)! (n-m-1)!}{(n-1)!} = \frac{1}{m},$$

注意到恰含有 j 个 m -分支的递归树都被计算了 j 次, 以及 $|T_n| = (n-1)!$, 所以上式左端就是 $\sum_{j=1}^{\lfloor \frac{n-1}{m} \rfloor} \binom{j}{1} P(\zeta_{n,m} = j)$, 由此即得

$$\sum_{j=1}^{\lfloor \frac{n-1}{m} \rfloor} \binom{j}{1} P(\zeta_{n,m} = j) = \frac{1}{m}, \quad (4.8)$$

类似可知

$$\sum_{j=2}^{\lfloor \frac{n-1}{m} \rfloor} \binom{j}{2} P(\zeta_{n,m} = j) = \frac{1}{2m^2}, \quad (4.9)$$

...

$$\sum_{j=i}^{\lfloor \frac{n-1}{m} \rfloor} \binom{j}{i} P(\zeta_{n,m} = j) = \frac{1}{m^i i!},$$

...

$$P\left(\zeta_{n,m} = \left\lfloor \frac{n-1}{m} \right\rfloor - 1\right) + \left\lfloor \frac{n-1}{m} \right\rfloor P\left(\zeta_{n,m} = \left\lfloor \frac{n-1}{m} \right\rfloor\right) = \frac{1}{m^{\lfloor \frac{n-1}{m} \rfloor - 1} (\lfloor \frac{n-1}{m} \rfloor - 1)!},$$

$$P\left(\zeta_{n,m} = \left\lfloor \frac{n-1}{m} \right\rfloor\right) = \frac{1}{m^{\lfloor \frac{n-1}{m} \rfloor} \lfloor \frac{n-1}{m} \rfloor!}.$$

上述最后一个式子就是 $\zeta_{n,m} = \lfloor \frac{n-1}{m} \rfloor$ 的概率, 将其代入倒数第 2 个式子, 即可得到 $\zeta_{n,m} = \lfloor \frac{n-1}{m} \rfloor - 1$ 的概率, 所以由上述 $\lfloor \frac{n-1}{m} \rfloor + 1$ 个式子, 由下往上逐个反解回去, 即得 (4.6) 式.

当 $m > \frac{n-1}{2}$ 时, 有 $\lfloor \frac{n-1}{m} \rfloor = 1$, 所以 $\zeta_{n,m}$ 只能取 0 和 1 两个值. 并且由 (4.6) 式易知

$$P(\zeta_{n,m} = 1) = 1/m.$$

根据 (4.8) 和 (4.9) 式, 可以得出 $\zeta_{n,m}$ 的期望与方差:

推论 4.2 (1) 对任意 $n \geq 2$, $\zeta_{n,m}$ 的期望为

$$E\zeta_{n,m} = 1/m, \quad m = 1, \dots, n-1;$$

(2) 对任意 $n \geq 3$, $\zeta_{n,m}$ 的方差为

$$\text{Var}\zeta_{n,m} = \begin{cases} \frac{1}{m}, & 1 \leq m \leq \frac{n-1}{2}, \\ \frac{m-1}{m^2}, & \frac{n-1}{2} < m \leq n-1. \end{cases} \quad (4.10)$$

证 由 (4.8) 式立知 $E\zeta_{n,m} = 1/m$, 而由 (4.9) 式可知, 当 $1 \leq m \leq (n-1)/2$ 时,

$$E\zeta_{n,m}^2 - E\zeta_{n,m} = 1/m^2,$$

故有

$$\text{Var}\zeta_{n,m} = E\zeta_{n,m}^2 - (E\zeta_{n,m})^2 = 1/m.$$

当 $(n-1)/2 < m \leq n-1$ 时, 结论显然.

由定理 4.2, 容易得出

定理 4.3 对任意 $m \in \mathbb{N}$, 当 $n \rightarrow \infty$ 时, $\zeta_{n,m}$ 的极限分布是参数为 $\lambda = 1/m$ 的 Poisson 分布, 即

$$\lim_{n \rightarrow \infty} P(\zeta_{n,m} = k) = e^{-1/m} \frac{1}{m^k k!}, \quad k = 0, 1, \dots, \quad m = 1, 2, \dots \quad (4.11)$$

4.3 $\zeta_{n,m}$ 的联合分布

显然, 只有研究各种不同分支的数目 $(\zeta_{n,1}, \zeta_{n,2}, \dots, \zeta_{n,n-1})$ 的联合分布, 才能真正弄清楚均匀递归树 T_n 上的集合 D_1 的整体结构情况.

定理 4.4 均匀递归树 T_n 上各种不同分支的数目 $(\zeta_{n,1}, \zeta_{n,2}, \dots, \zeta_{n,n-1})$ 的联合分布是

$$P(\zeta_{n,1} = x_1, \zeta_{n,2} = x_2, \dots, \zeta_{n,n-1} = x_{n-1}) = \prod_{m=1}^{n-1} \frac{1}{m^{x_m} x_m!}, \quad (4.12)$$

其中 $x_1, \dots, x_{n-1}$ 是满足条件 $\sum_{i=1}^{n-1} i x_i = n-1$ 的任意一组非负整数.

证 我们来计算事件 $(\zeta_{n,1} = x_1, \dots, \zeta_{n,n-1} = x_{n-1})$ 中的基本事件数目. 首先考察 $n-1$ 个顶点的分组情况 (分在同一个组中的顶点即属于同一个分支). 由于递归树上的分支的编号方式是自动形成的, 即按照其根点的号码的大小编号, 所以由无编号分组模式知: 一共有

$$\frac{(n-1)!}{(1!)^{x_1} (2!)^{x_2} \dots ((n-1)!)^{x_{n-1}}} \cdot \frac{1}{x_1! x_2! \dots x_{n-1}!}$$

种不同的分组方式 (上式中的第一个分式是有编号分组方式数目, 而 $x_1! x_2! \dots x_{n-1}!$ 则是编号方式数目). 又由于分支也是递归树, 故一个 m -分支上的顶点有 $(m-1)!$ 种不同的连接方式, 所以事件 $(\zeta_{n,1} = x_1, \dots, \zeta_{n,n-1} = x_{n-1})$ 中的基本事件数目为分组方式数目与连接方式数目的乘积, 即

$$\begin{aligned} & \frac{(n-1)!}{(1!)^{x_1} (2!)^{x_2} \dots ((n-1)!)^{x_{n-1}}} \cdot \frac{(0!)^{x_1} (1!)^{x_2} \dots ((n-2)!)^{x_{n-1}}}{x_1! x_2! \dots x_{n-1}!} \\ &= \frac{(n-1)!}{1^{x_1} 2^{x_2} \dots (n-1)^{x_{n-1}} x_1! x_2! \dots x_{n-1}!}. \end{aligned}$$

由于各个基本事件都以相等的概率 $1/(n-1)!$ 发生, 所以

$$\begin{aligned} & P(\zeta_{n,1} = x_1, \zeta_{n,2} = x_2, \dots, \zeta_{n,n-1} = x_{n-1}) \\ &= \frac{1}{(n-1)!} \cdot \frac{(n-1)!}{1^{x_1} 2^{x_2} \dots (n-1)^{x_{n-1}} x_1! x_2! \dots x_{n-1}!} \\ &= \prod_{m=1}^{n-1} \frac{1}{m^{x_m} x_m!}. \end{aligned}$$

由 4.2 小节的讨论, 已经知道随机向量 $(\zeta_{n,1}, \zeta_{n,2}, \dots, \zeta_{n,n-1})$ 的数学期望, 即

$$E\zeta_{n,m} = 1/m, \quad m = 1, \dots, n-1, \quad (4.13)$$

现在来给出它的协方差阵.

定理 4.5 对任意 $1 \leq k < l \leq n-1$, 当 $k+l \leq n-1$ 时, 有

$$\text{Cov}(\zeta_{n,k}, \zeta_{n,l}) = 0; \quad (4.14)$$

当 $k+l > n-1$ 时, 有

$$\text{Cov}(\zeta_{n,k}, \zeta_{n,l}) = -\frac{1}{kl}. \quad (4.15)$$

证 当 $1 \leq k < l \leq n-1$ 且 $k+l > n-1$ 时, 显然有

$$P(\zeta_{n,k} = i, \zeta_{n,l} = j) = 0, \quad i, j > 0,$$

因此

$$E\zeta_{n,k}\zeta_{n,l} = 0, \quad \text{Cov}(\zeta_{n,k}, \zeta_{n,l}) = -E\zeta_{n,k}E\zeta_{n,l} = -\frac{1}{kl}.$$

对 $1 \leq k < l \leq n-1$ 且 $k+l \leq n-1$, 当 $i, j > 0$, $ik+jl \leq n-1$ 时, 恰有 i 个 k -分支和 j 个 l -分支的均匀递归树的数目为 $(n-1)! \cdot P(\zeta_{n,k} = i, \zeta_{n,l} = j)$.

另一方面, 设 A 和 B 分别为集合 $\{2, 3, \dots, n\}$ 的两个不相交的 k 元和 l 元子集, 则包含有以 A 和 B 中的顶点为 k -分支和 l -分支的均匀递归树的数目是 $(k-1)!(l-1)!(n-k-l-1)!$. 如果考虑子集 A 和 B 的不同选法, 将选法数目与 $(k-1)!(l-1)!(n-k-l-1)!$ 相乘, 得到

$$M := \binom{n-1}{k} \cdot \binom{n-k-1}{l} (k-1)! (l-1)! (n-k-l-1)! = \frac{(n-1)!}{kl}.$$

显然, 在 M 中每个恰有 i 个 k -分支和 j 个 l -分支的均匀递归树都分别被计算了 i 次和 j 次, 所以

$$\sum_{(i,j): i,j>0, ik+jl \leq n-1} ij((n-1)! \cdot P(\zeta_{n,k} = i, \zeta_{n,l} = j)) = M = \frac{(n-1)!}{kl},$$

亦即

$$\sum_{i,j} ijP(\zeta_{n,k} = i, \zeta_{n,l} = j) = \frac{1}{kl},$$

由此和 (4.13) 式, 得知

$$E\zeta_{n,k}\zeta_{n,l} = \frac{1}{kl} = E\zeta_{n,k}E\zeta_{n,l}, \quad 1 \leq k < l \leq n-1, \quad k+l \leq n-1,$$

故知此时 $\text{Cov}(\zeta_{n,k}, \zeta_{n,l}) = 0$.

综合上述结果, 可知

推论 4.3 随机向量 $(\zeta_{n,1}, \zeta_{n,2}, \dots, \zeta_{n,n-1})$ 的协方差阵为 $B_n = (b_{ij})_{1 \leq i,j \leq n-1}$, 其中

$$b_{ii} = \begin{cases} \frac{1}{i}, & 1 \leq i \leq \frac{n-1}{2}, \\ \frac{i-1}{i^2}, & \frac{n-1}{2} < i \leq n-1, \end{cases}$$

$$b_{ij} = \begin{cases} 0, & i \neq j, i+j \leq n-1, \\ -\frac{1}{ij}, & i \neq j, i+j > n-1. \end{cases}$$

4.4 $\zeta_{n,m}$ 联合分布的极限分布

我们来研究 $\zeta_{n,m}$ 联合分布的极限, 先证明一个引理:

引理 4.1 对任意 $m \in \mathbb{N}$, 当 $n \rightarrow \infty$ 时, $P(\zeta_{n,1} = 0, \dots, \zeta_{n,m} = 0)$ 的极限都存在.

证 定理 4.3 表明结论已经对 $m = 1$ 成立. 我们来看 $m = 2$ 的情形. 记

$$P(\zeta_{n,1} = 0, \zeta_{n,2} = 0) = a_n, \quad P(\zeta_{n,1} = 0) = b_n.$$

由定理 4.4 知

$$b_n = P(\zeta_{n,1} = 0) = \sum_{2x_2+3x_3+\dots+(n-1)x_{n-1}=n-1} \prod_{m=2}^{n-1-i} \frac{1}{m^{x_m} x_m!},$$

$$a_n = P(\zeta_{n,1} = 0, \zeta_{n,2} = 0) = \sum_{3x_3+\dots+(n-1)x_{n-1}=n-1} \prod_{m=2}^{n-1-i} \frac{1}{m^{x_m} x_m!},$$

因此

$$b_n = \sum_{j=0}^{[(n-1)/2]} \frac{1}{2^j j!} \sum_{3x_3+\dots+(n-1)x_{n-1}=n-2j-1} \prod_{m=3}^{n-2j-1} \frac{1}{m^{x_m} x_m!}$$

$$= \sum_{j=0}^{[(n-1)/2]} \frac{1}{2^j j!} a_{n-2j}. \quad (4.16)$$

假设 $n \rightarrow \infty$ 时, a_n 的极限不存在, 则存在 $\alpha, \beta \in [0, 1]$, 使得

$$\liminf_{n \rightarrow \infty} a_n = \alpha, \quad \limsup_{n \rightarrow \infty} a_n = \beta.$$

注意到 $16 - 9\sqrt{e} > \sqrt{256} - \sqrt{81 \times 3} > 0$, 令 $\delta = \frac{\beta-\alpha}{4} > 0$, $\delta_1 = \left(4 - \frac{9\sqrt{e}}{4}\right)\delta > 0$. 由于 $\lim_{n \rightarrow \infty} b_n = e^{-1}$, 所以对任意取定的 $0 < \varepsilon < (11\sqrt{e} - 18)\delta$ (注意 $11\sqrt{e} - 18 > \sqrt{121 \times 2.7} - \sqrt{324} > \sqrt{326} - \sqrt{324} > 0$), 只要 n_0 充分大, 就对任意 $n_2 > n_1 > n_0$, 都有

$$|b_{n_1} - b_{n_2}| < \varepsilon. \quad (4.17)$$

而且只要 n_0 充分大, 就有

$$\alpha - \delta < a_n < \beta + \delta, \quad n \geq n_0, \quad (4.18)$$

$$\sum_{j=n+1}^{\infty} \frac{1}{2^j j!} < \delta_1, \quad n \geq n_0. \quad (4.19)$$

取定 n_0 充分大, 使得 (4.17)~(4.19) 式都成立.

当 n 充分大时, 可将 (4.16) 式写为

$$b_n = a_n + \sum_{j=1}^{n_0} \frac{1}{2^j j!} a_{n-2j} + \sum_{j=n_0+1}^{[(n-1)/2]} \frac{1}{2^j j!} a_{n-2j}. \quad (4.20)$$

由于 $\limsup_{n \rightarrow \infty} a_n = \beta$, $\liminf_{n \rightarrow \infty} a_n = \alpha$, 所以存在 $n_2 > n_1 > 3n_0$, 使得

$$a_{n_1} > \beta - \delta_1, \quad a_{n_2} < \alpha + \delta_1.$$

由于 $n_2 - 2n_0 > n_1 - 2n_0 > n_0$, 所以 (4.18) 式对一切 $n = n_1 - 2j$, $n = n_2 - 2j$, $j \in \{1, 2, \dots, n_0\}$ 都成立. 由此, 并结合 (4.19) 和 (4.20) 式, 就有

$$\begin{aligned} b_{n_1} &= a_{n_1} + \sum_{j=1}^{n_0} \frac{1}{2^j j!} a_{n_1-2j} + \sum_{j=n_0+1}^{[(n-1)/2]} \frac{1}{2^j j!} a_{n_1-2j} \\ &> (\beta - \delta_1) + \sum_{j=1}^{n_0} \frac{1}{2^j j!} (\alpha - \delta) + \sum_{j=n_0+1}^{[(n-1)/2]} \frac{1}{2^j j!} a_{n_1-2j} \\ &> (\beta - \delta_1) + (\sqrt{e} - 1)(\alpha - \delta) - \sum_{j=n_0+1}^{\infty} \frac{1}{2^j j!} \\ &> (\beta - \delta_1) + (\sqrt{e} - 1)(\alpha - \delta) - \delta_1, \\ b_{n_2} &= a_{n_2} + \sum_{j=1}^{n_0} \frac{1}{2^j j!} a_{n_2-2j} + \sum_{j=n_0+1}^{[(n-1)/2]} \frac{1}{2^j j!} a_{n_2-2j} \\ &< (\alpha + \delta_1) + \sum_{j=1}^{n_0} \frac{1}{2^j j!} (\beta + \delta) + \sum_{j=n_0+1}^{[(n-1)/2]} \frac{1}{2^j j!} a_{n_2-2j} \\ &< (\alpha + \delta_1) + (\sqrt{e} - 1)(\beta + \delta) + \sum_{j=n_0+1}^{\infty} \frac{1}{2^j j!} \\ &< (\alpha + \delta_1) + (\sqrt{e} - 1)(\beta + \delta) + \delta_1. \end{aligned}$$

将上述二式相减, 得到

$$\begin{aligned} b_{n_1} - b_{n_2} &> (2 - \sqrt{e})(\beta - \alpha) - 2(\sqrt{e} - 1)\delta - 4\delta_1 \\ &= 4\delta - 2(\sqrt{e} - 1)\delta - (16 - 9\sqrt{e})\delta \\ &= (11\sqrt{e} - 18)\delta, \end{aligned}$$

此与 (4.17) 式相矛盾. 所以当 $n \rightarrow \infty$ 时, $P(\zeta_{n,1} = 0, \zeta_{n,2} = 0)$ 的极限存在.

再运用归纳法, 即可证得引理 4.1 的结论对一切 $m \in \mathbb{N}$ 都成立, 归纳过渡的过程与 $m = 1 \Rightarrow m = 2$ 的过程类似.

下面给出本小节的主要结果:

定理 4.6 均匀递归树上的分支数目渐近独立, 并且对任意 $m \in \mathbb{N}$, 任意有序非负整数组 $\{x_1, \dots, x_m\}$, 都有

$$\lim_{n \rightarrow \infty} P(\zeta_{n,1} = x_1, \dots, \zeta_{n,m} = x_m) = \prod_{j=1}^m \lim_{n \rightarrow \infty} P(\zeta_{n,j} = x_j) = \prod_{j=1}^m e^{-1/j} \frac{1}{j^{x_j} x_j!}.$$

证 本定理的证明包括两个部分: (1) 极限 $\lim_{n \rightarrow \infty} P(\zeta_{n,1} = x_1, \dots, \zeta_{n,m} = x_m)$ 存在; (2) 对任意 $m \in \mathbb{N}$, 任意有序非负整数组 $\{x_1, \dots, x_m\}$, 都有

$$\lim_{n \rightarrow \infty} P(\zeta_{n,1} = x_1, \dots, \zeta_{n,m} = x_m) = \prod_{j=1}^m e^{-1/j} \frac{1}{j^{x_j} x_j!}. \quad (4.21)$$

定理 4.3 已经表明, 对任意 $j \in \mathbb{N}$, 任意非负整数 i , 都有

$$\lim_{n \rightarrow \infty} P(\zeta_{n,j} = i) = e^{-1/j} \frac{1}{j^i i!},$$

所以结论已经对 $m = 1$ 成立. 特别地,

$$\lim_{n \rightarrow \infty} P(\zeta_{n,1} = 0) = e^{-1}. \quad (4.22)$$

我们来看 $m = 2$ 的情形. 由 (4.12) 式知

$$\begin{aligned} P(\zeta_{n,1} = i, \zeta_{n,2} = j) &= \frac{1}{i! \cdot 2^j j!} \sum_{3x_3 + \dots + (n-1-i)x_{n-1-i} = n-1-i-2j} \prod_{m=3}^{n-1-i} \frac{1}{m^{x_m} x_m!} \\ &= \frac{1}{i! 2^j j!} P(\zeta_{n-i-2j,1} = 0, \zeta_{n-i-2j,2} = 0), \end{aligned}$$

所以由引理 4.1 知极限

$$\lim_{n \rightarrow \infty} P(\zeta_{n,1} = i, \zeta_{n,2} = j) = \frac{1}{i! 2^j j!} \lim_{n \rightarrow \infty} P(\zeta_{n-i-2j,1} = 0, \zeta_{n-i-2j,2} = 0)$$

存在. 关键是证明

$$\lim_{n \rightarrow \infty} P(\zeta_{n,1} = 0, \zeta_{n,2} = 0) = e^{-1-\frac{1}{2}}, \quad (4.23)$$

为此, 记

$$\lim_{n \rightarrow \infty} P(\zeta_{n,1} = 0, \zeta_{n,2} = 0) = P(\zeta_1 = 0, \zeta_2 = 0) =: p, \quad 0 \leq p \leq 1. \quad (4.24)$$

注意

$$P(\zeta_{n,1} = 0, \zeta_{n,2} = 0) = \sum_{3x_3 + \dots + (n-1-i)x_{n-1-i} = n-1} \prod_{m=2}^{n-1} \frac{1}{m^{x_m} x_m!}.$$

另一方面, (4.22) 式就是

$$\begin{aligned} &\lim_{n \rightarrow \infty} \sum_{2x_2 + 3x_3 + \dots + (n-1-i)x_{n-1-i} = n-1} \prod_{m=2}^{n-1} \frac{1}{m^{x_m} x_m!} \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{[(n-1)/2]} \frac{1}{2^j j!} \sum_{3x_3 + \dots + (n-1-2j)x_{n-1-2j} = n-1-2j} \prod_{m=3}^{n-1-2j} \frac{1}{m^{x_m} x_m!} \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{[(n-1)/2]} \frac{1}{2^j j!} P(\zeta_{n-2j,1} = 0, \zeta_{n-2j,2} = 0) \\ &= e^{-1}, \end{aligned}$$

亦即

$$\lim_{n \rightarrow \infty} e \sum_{j=0}^{[(n-1)/2]} \frac{1}{2^j j!} P(\zeta_{n-2j,1} = 0, \zeta_{n-2j,2} = 0) = 1. \quad (4.25)$$

上式表明, 对充分大的正整数 n_0 , 当 n 远远大于 n_0 时, 有

$$\begin{aligned} & e \sum_{j=0}^{n_0} \frac{1}{2^j j!} P(\zeta_{n-2j,1} = 0, \zeta_{n-2j,2} = 0) \\ & \leq e \sum_{j=0}^{[(n-1)/2]} \frac{1}{2^j j!} P(\zeta_{n-2j,1} = 0, \zeta_{n-2j,2} = 0) \\ & \leq e \sum_{j=0}^{n_0} \frac{1}{2^j j!} P(\zeta_{n-2j,1} = 0, \zeta_{n-2j,2} = 0) + e \sum_{j=n_0+1}^{\infty} \frac{1}{2^j j!}. \end{aligned}$$

令 $n \rightarrow \infty$, 由 (4.24) 和 (4.25) 式, 得

$$ep \sum_{j=0}^{n_0} \frac{1}{2^j j!} \leq 1 \leq ep \sum_{j=0}^{n_0} \frac{1}{2^j j!} + e \sum_{j=n_0+1}^{\infty} \frac{1}{2^j j!}.$$

再在上式中令 $n_0 \rightarrow \infty$, 即得 $e^{1+\frac{1}{2}}p = 1$. (4.23) 式获证. 从而定理的结论对 $m = 2$ 成立.

对于 $m \geq 3$, 可利用归纳法, 而由 $m-1$ 向 m 过渡的过程与由 $m=1$ 推 $m=2$ 的过程类似, 故略.

5 均匀递归树的最大分支

在均匀递归树 T_n 上以 ν_n 表示它的最大分支的大小, 即

$$\nu_n = \max\{m : \zeta_{n,m} \geq 1\}.$$

在传染病模型中, ν_n 表示以顶点 1 的某个子点为源头, 所传染的最多人数.

命题 5.1

$$\lim_{n \rightarrow \infty} P\left(\nu_n > \frac{n-1}{2}\right) = \ln 2.$$

证 由于

$$\left(\nu_n > \frac{n-1}{2}\right) = \bigcup_{m=[(n-1)/2]+1}^{n-1} (\nu_n = m),$$

而事件 $(\nu_n = [(n-1)/2] + 1), (\nu_n = [(n-1)/2] + 2), \dots, (\nu_n = n-1)$ 中任意两个都不能同时发生, 所以当 $n \rightarrow \infty$ 时, 有

$$P\left(\nu_n > \frac{n-1}{2}\right) = \sum_{m=[(n-1)/2]+1}^{n-1} P(\nu_n = m) = \sum_{m=[(n-1)/2]+1}^{n-1} \frac{1}{m} \rightarrow \ln 2.$$

注 1 这个结论显示, 大于 $[(n-1)/2]$ 的分支存在的概率很大, 因从而从一定意义上解释了严重急性呼吸综合征流行中存在超级传播者的事实.

不仅如此, 我们还可以证明随 $n \rightarrow \infty$, 最大分支的大小 ν_n 几乎必然趋于无穷. 易知

$$\nu_n \leq \nu_{n+1} \leq \nu_n + 1,$$

因此, ν_n 随 n 增加而非降, 故知 ν_n 的极限存在.

定理 5.1

$$\lim_{n \rightarrow \infty} \nu_n = \infty \quad \text{a.s.}$$

证 由于 ν_n 是最大分支的大小, 而 η_n 是分支数目, 所以 $\eta_n \cdot \nu_n \geq n - 1$. 由此并结合 η_n 的强大数律 (2.6) 式, 即得

$$\liminf_{n \rightarrow \infty} \frac{\log n \cdot \nu_n}{n} \geq \lim_{n \rightarrow \infty} \frac{\log n}{\eta_n} = 1 \quad \text{a.s.}$$

既然 $\lim_{n \rightarrow \infty} \frac{\log n}{n} = 0$, 所以 $\lim_{n \rightarrow \infty} \nu_n = \infty \quad \text{a.s.}$

参 考 文 献

- 1 徐俊明. 图论及其应用. 合肥: 中国科学技术大学出版社, 1998
- 2 Bergeron F, Flajolet P, Salvy B. Varieties of increasing trees. In: Proc 17th Coll Trees in Algebra and Programming. Lecture Notes Comput Sci, Vol 581. Berlin: Springer, 1992. 24~48
- 3 Moon J W. The distance between nodes in recursive trees. London Mathematics Society Lecture Notes Series, 1974, 13: 125~132
- 4 Najock D, Heyde C. On the number of the terminal vertices in certain random trees with an application to stemma construction in philology. J Appl Prob, 1982, 19: 675~680
- 5 Gastwirth J. A probability model of a pyramid scheme. Amer Statist, 1977, 31: 79~82
- 6 Meir A, Moon J W. On the altitude of nodes in random trees. Canad J Math, 1978, 30: 997~1015
- 7 Szymański J. On the maximum degree and the height of a random recursive tree. In: Karoński M, Ruciński A, eds. Random Graphs. North-Holland: John Wiley & Sons, 1990. 313~324
- 8 Devroye L. Applications of the theory of records in the study of random trees. Acta Informatica, 1988, 26: 123~130
- 9 Mahmoud H M. Distances in random plane-oriented recursive trees. J Comput Appl Math, 1992, 41: 237~245
- 10 Mahmoud H M. Limiting distributions for path lengths in recursive trees. Probab Engrg Inform Sci, 1991, 5: 53~59
- 11 Petrov V V. Limit Theorems of Probability Theory. Oxford: Clarendon Press, 1995