Math-Net.Ru

Общероссийский математический портал

Ю. Л. Павлов, Е. А. Лосева, Предельные распределения
максимального объема дерева в случайном рекурсивном
лесе, *Дискрет. матем.*, 2002, том 14, выпуск 1, 60–74

DOI: <https://doi.org/10.4213/dm230>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 52.87.193.239

26 апреля 2023 г., 13:49:59

УДК 519.2

Предельные распределения максимального объема дерева в случайном рекурсивном лесе

© 2002 г. Ю. Л. Павлов, Е. А. Лосева

Рассматривается множество лесов, состоящих из N рекурсивных некорневых деревьев и n вершин, на котором задано равномерное распределение вероятностей. Получено полное описание предельного поведения максимального объема дерева в лесе при $n \rightarrow \infty$.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант 00-01-00233.

Предельное поведение максимального объема дерева в случайном лесе впервые изучалось в статьях [1, 2]. В этих работах рассматривалось множество лесов, состоящих из N корневых деревьев и n занумерованных некорневых вершин при равномерном распределении вероятностей на этом множестве. Основным методом исследования в [1, 2] была обобщенная схема размещения частиц по ячейкам [3]. Аналогичная задача для случайного леса из некорневых деревьев решена в работе [4]. В книге [5] доказаны предельные теоремы для максимального объема дерева в случайном лесе, генерируемом ветвящимся процессом Гальтона–Ватсона, начинающимся с N частиц, при условии, что общее число частиц в процессе равно $N + n$. Распределение вероятностей на множестве таких лесов индуцируется распределением числа потомков частиц ветвящегося процесса. Такие леса состоят из деревьев, принадлежащих так называемому просто генерируемому семейству корневых деревьев [6].

В последнее время значительное внимание уделяется изучению рекурсивных деревьев (см., например, [7]). Дерево T_n с n занумерованными вершинами называется рекурсивным, если $n = 1$ или если T_n построено путем последовательного присоединения j -й вершины к одной из первых $j - 1$ вершин, $2 \leq j \leq n$.

Ниже рассматриваются леса, состоящие из рекурсивных деревьев, такие леса назовем рекурсивными. Мы будем считать, что эти леса состоят из N занумерованных некорневых деревьев и n вершин. Это значит, что каждое дерево такого леса содержит хотя бы одну из n вершин и лес строится так, что очередная, j -я вершина либо попадает в одно из деревьев, в которых еще нет вершин (то есть становится первой вершиной этого дерева), либо присоединяется к одной из первых $j - 1$ вершин леса. Понятно, что в таком лесе $n \geq N$.

Обозначим $F_{N,n}$ множество рассматриваемых лесов и введем на этом множестве равномерное распределение вероятностей. В [6] показано, что рекурсивные деревья не относятся к классу просто генерируемых деревьев. Отсюда следует, что

для их изучения нельзя применить методы теории ветвящихся процессов, как это осуществлено в [5]. Поэтому для получения теорем о максимальном объеме дерева в рекурсивном лесе ниже применяется обобщенная схема размещения.

Пусть $\eta_1, \dots, \eta_N$ — случайные величины, равные числу вершин в деревьях с номерами $1, \dots, N$ леса из $F_{N,n}$, то есть объемам соответствующих деревьев. Нетрудно видеть, что число рекурсивных деревьев с k вершинами равно $(k-1)!$. Обозначим $L(N, n)$ число лесов в $F_{N,n}$. Легко проверить, что

$$L(N, n) = n! \sum \frac{1}{k_1 k_2 \dots k_N},$$

где суммирование проводится по всем целым положительным $k_1, \dots, k_N$ таким, что $k_1 + \dots + k_N = n$. Отсюда следует, что для таких $k_1, \dots, k_N$

$$P\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \frac{n!}{k_1 k_2 \dots k_N L(N, n)}. \quad (1)$$

Введем независимые одинаково распределенные случайные величины $\xi_1, \dots, \xi_N$, для которых

$$P\{\xi_1 = k\} = -\frac{\lambda^k}{k \ln(1 - \lambda)}, \quad k = 1, 2, \dots, \quad (2)$$

где $0 < \lambda < 1$. Такое распределение носит название распределения логарифмического ряда. Из (2) следует, что

$$P\{\xi_1 = k_1, \dots, \xi_N = k_N \mid \xi_1 + \dots + \xi_N = n\} = \frac{n!}{k_1 k_2 \dots k_N L(N, n)}.$$

Отсюда и из (1) получаем равенство

$$P\{\eta_1 = k_1, \dots, \eta_N = k_N\} = P\{\xi_1 = k_1, \dots, \xi_N = k_N \mid \xi_1 + \dots + \xi_N = n\}. \quad (3)$$

Это равенство показывает, что наборы случайных величин $(\eta_1, \dots, \eta_N)$ и $(\xi_1, \dots, \xi_N)$ образуют обобщенную схему размещения [3].

Обозначим $\eta_{(N)}$ максимальный объем дерева в лесе из $F_{N,n}$, то есть

$$\eta_{(N)} = \max_{1 \leq i \leq N} \eta_i.$$

Введем независимые одинаково распределенные случайные величины $\xi_1^{(r)}, \dots, \xi_N^{(r)}$, для которых

$$P\{\xi_1^{(r)} = k\} = P\{\xi_1 = k \mid k \leq r\}, \quad k = 1, 2, \dots, r. \quad (4)$$

Обозначим также

$$\zeta_N = \xi_1 + \dots + \xi_N, \quad \zeta_N^{(r)} = \xi_1^{(r)} + \dots + \xi_N^{(r)}.$$

В [3] показано, что из (3) вытекает следующая лемма.

Лемма 1. *Справедливо равенство*

$$P\{\eta_{(N)} \leq r\} = (1 - P_r)^N \frac{P\{\zeta_N^{(r)} = n\}}{P\{\zeta_N = n\}},$$

где $P_r = P\{\xi_1 > r\}$.

Из леммы 1 следует, что изучение поведения случайной величины $\eta_{(N)}$ может быть сведено к исследованию сумм независимых слагаемых ζ_N , $\zeta_N^{(r)}$ и бинома $(1 - P_r)^N$. Ниже, в леммах 2–18, будет изучено асимптотическое поведение этих величин при $n \rightarrow \infty$ и всех регулярных изменениях параметра N , а затем с помощью леммы 1 будут доказаны предельные теоремы для $\eta_{(N)}$. Понятно, что если $n = N$, то все деревья имеют единичный объем, поэтому ниже предполагается, что $n > N$.

Заметим, что лемма 1 справедлива при любых значениях λ , $0 < \lambda < 1$. В дальнейшем будут использованы три значения параметра λ распределения (2). Если $N \rightarrow \infty$, $n = N + k$, где $k = 1, 2, \dots$ фиксировано, то будем полагать, что $\lambda = 1/N$. Если $N, n \rightarrow \infty$ так, что $n - N \rightarrow \infty$, $N/\ln n \rightarrow \infty$, то в качестве λ берется единственное решение уравнения

$$-\frac{\lambda}{(1-\lambda)\ln(1-\lambda)} = \frac{n}{N}, \quad (5)$$

а если $N/\ln n = O(1)$, то $\lambda = 1 - n^{-1}$.

Теорема 1. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow 1$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что

$$N\lambda^{r-1}/r \rightarrow \infty, \quad N\lambda^r/(r+1) \rightarrow \gamma,$$

где γ — некоторая неотрицательная постоянная. Тогда

$$\mathbf{P}\{\eta_{(N)} = r\} \rightarrow e^{-\gamma}, \quad \mathbf{P}\{\eta_{(N)} = r+1\} \rightarrow 1 - e^{-\gamma}.$$

Теорема 2. Пусть $N, n \rightarrow \infty$ и существуют постоянные C_1 и C_2 такие, что $1 < C_1 \leq n/N \leq C_2$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что

$$-\frac{\lambda^r N}{r(1-\lambda)\ln(1-\lambda)} \rightarrow \gamma,$$

где γ — некоторая положительная постоянная. Тогда при любом фиксированном $k = 0, \pm 1, \pm 2, \dots$

$$\mathbf{P}\{\eta_{(N)} \leq r+k\} = \exp\{-\gamma\lambda^{k+1}\}(1 + o(1)).$$

Теорема 3. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что $n\lambda^r/r \rightarrow e^{-z}$, где z — некоторая постоянная. Тогда

$$\mathbf{P}\{\eta_{(N)} \leq r\} \rightarrow e^{e^{-z}}.$$

Теорема 4. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow \alpha$, где α — некоторая положительная постоянная. Тогда

$$\mathbf{P}\{\eta_{(N)}/n \leq z\} \rightarrow \sum_{s=0}^{\infty} \frac{(-\alpha)^s}{s!} J_s(z),$$

где $J_0(z) = 1$,

$$J_s(z) = \int_{X(s,z)} \frac{(1-x_1-\dots-x_s)^{\alpha-1}}{x_1 \dots x_s} dx_1 \dots dx_s,$$

$X(s, z) = \{x_i \geq z, i = 1, \dots, s, x_1 + \dots + x_s \leq 1\}$, $s = 1, 2, \dots$

Теорема 5. Пусть $n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow 0$. Тогда для любого положительного z

$$\mathbf{P}\{-N(\ln(1 - \eta_{(N)}/n))/\ln n \leq z\} \rightarrow 1 - e^{-z}.$$

Для доказательства теорем 1–5 рассмотрим сначала предельное поведение величины NP_r .

Лемма 2. Пусть $N \rightarrow \infty$, $\lambda = 1/N$. Тогда $NP_1 \rightarrow 1/2$, $NP_2 \rightarrow 0$.

Доказательство. Поскольку

$$P_r = \mathbf{P}\{\xi_1 > r\} = \sum_{s=1}^{\infty} \mathbf{P}\{\xi_1 = r + s\},$$

из (2) находим, что

$$\begin{aligned} NP_r &= N\mathbf{P}\{\xi_1 = r + 1\} \sum_{s=1}^{\infty} \mathbf{P}\{\xi_1 = r + s\} / \mathbf{P}\{\xi_1 = r + 1\} \\ &= N\mathbf{P}\{\xi_1 = r + 1\} \left(1 + \sum_{s=1}^{\infty} \lambda^s (r + 1) / (r + s + 1) \right). \end{aligned} \quad (6)$$

Поскольку $\lambda \rightarrow 0$, из (6) находим, что

$$NP_r = N\mathbf{P}\{\xi_1 = r + 1\}(1 + o(1)), \quad (7)$$

и утверждение леммы следует из (2), если положить $r = 1, 2$.

Лемма 3. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow 1$, $n - N \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что $N\lambda^r / (r + 1) \rightarrow \gamma$, где γ — некоторая неотрицательная постоянная. Тогда

$$NP_{r-1} \rightarrow \infty, \quad NP_r \rightarrow \gamma, \quad NP_{r+1} \rightarrow 0.$$

Доказательство. Из (5) следует, что $\lambda \rightarrow 0$, поэтому лемма вытекает из (2) и (7).

Лемма 4. Пусть $N, n \rightarrow \infty$ и существуют постоянные C_1 и C_2 такие, что $1 < C_1 \leq n/N \leq C_2$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что

$$-\frac{\lambda^r N}{r(1 - \lambda) \ln(1 - \lambda)} \rightarrow \gamma,$$

где γ — некоторая положительная постоянная. Тогда при любом фиксированном целом k

$$NP_{r+k} = \gamma \lambda^{k+1} (1 + o(1)).$$

Доказательство. По аналогии с (6) находим, что

$$NP_{r+k} = N\mathbf{P}\{\xi_1 = r + k + 1\} \sum_{s=1}^{\infty} \lambda^{s-1} (r + k + 1) / (r + k + s). \quad (8)$$

Из (5) следует, что в условиях леммы справедливы неравенства $0 < C_3 \leq \lambda \leq C_4 < 1$. Учитывая, что $r \rightarrow \infty$, отсюда получаем, что

$$\sum_{s=1}^{\infty} \lambda^{s-1} (r+k+1)/(r+k+s) = (1-\lambda)^{-1} (1+o(1)),$$

поэтому лемма следует из (2) и (8).

Лемма 5. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что $n\lambda^r/r \rightarrow e^{-z}$, где z — некоторая постоянная. Тогда $NP_r \rightarrow e^{-z}$.

Доказательство. В силу (2)

$$P_r = -\frac{1}{\ln(1-\lambda)} \sum_{s=r+1}^{\infty} \frac{\lambda^s}{s}.$$

Заменяя суммирование интегрированием, отсюда получаем, что

$$P_r = -\frac{1+o(1)}{\ln(1-\lambda)} \int_{-r \ln \lambda}^{\infty} y^{-1} e^{-y} dy. \quad (9)$$

Заметим, что $-r \ln \lambda \rightarrow \infty$. Действительно, если это не так, то из условия $n\lambda^r/r \rightarrow e^{-z}$ следует, что $n = O(r)$. Из (5) нетрудно получить, что

$$\lambda = 1 - \frac{N}{n \ln(n/N)} (1+o(1)), \quad (10)$$

поэтому приходим к соотношению $N/\ln(n/N) = O(1)$, что противоречит условию $N/\ln n \rightarrow \infty$. Из соотношений $-r \ln \lambda \rightarrow \infty$, $\ln \lambda = -(1-\lambda)(1+o(1))$ и (9) получаем, что

$$NP_r = -\frac{N\lambda^r(1+o(1))}{r(1-\lambda) \ln(1-\lambda)}.$$

Это равенство, (5) и условие выбора r приводят к утверждению леммы 5.

Из (9) очевидным образом получаем следующие утверждения, справедливые при $\lambda = 1 - n^{-1}$.

Лемма 6. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow \alpha$, а $r = zn$, где α, z — некоторые положительные постоянные. Тогда

$$NP_r \rightarrow \alpha \int_z^{\infty} y^{-1} e^{-y} dy.$$

Лемма 7. Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow 0$, $r = n - x$, где $x = o(n)$. Тогда $NP_r \rightarrow 0$.

Предельное поведение суммы ζ_N во всех зонах изменения параметров N и n исследовано в книге [8]. Приведем соответствующие результаты.

Лемма 8. Если $N \rightarrow \infty$, $\lambda = 1/N$, то для любого фиксированного $k = 0, 1, \dots$

$$\mathbf{P}\{\zeta_N - N = k\} \rightarrow \frac{1}{2^k k!} e^{-1/2}.$$

Лемма 9. Если $N, n \rightarrow \infty$ так, что $n/N \rightarrow 1$, $n - N \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, то

$$\mathbf{P}\{\zeta_N = k\} = \frac{1}{\sqrt{2\pi(n-N)}} e^{-u^2/2} (1 + o(1))$$

равномерно относительно целых k , для которых $u = (k - n)/\sqrt{n - N}$ лежит в любом фиксированном конечном интервале.

Лемма 10. Если $N, n \rightarrow \infty$, $n/N \geq C_1 > 1$, $N/\ln n \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$ и

$$\sigma^2 = -\frac{\lambda \ln(1 - \lambda) + \lambda^2}{(1 - \lambda)^2 \ln^2(1 - \lambda)}, \quad (11)$$

то

$$\mathbf{P}\{\zeta_N = k\} = \frac{1}{\sigma \sqrt{2\pi N}} e^{-u^2/2} (1 + o(1))$$

равномерно относительно целых k , для которых $u = (k - n)/(\sigma \sqrt{N})$ лежит в любом фиксированном конечном интервале.

Лемма 11. Если $n \rightarrow \infty$, $N = \alpha \ln n + o(\ln n)$, где α — постоянная, $0 < \alpha < \infty$, $\lambda = 1 - n^{-1}$, то

$$\mathbf{P}\{\zeta_N = k\} = \frac{1}{n\Gamma(\alpha)} u^{\alpha-1} e^{-u} (1 + o(1))$$

равномерно относительно целых k , для которых $u = k/n$ лежит в любом фиксированном интервале вида $0 < u_0 \leq u \leq u_1$, где u_0, u_1 — постоянные.

Лемма 12. Если $n \rightarrow \infty$ и $N/\ln n \rightarrow 0$, то

$$\mathbf{P}\{\zeta_N = n\} = \frac{e^{-1}}{n\Gamma(N/\ln n)} (1 + o(1)).$$

Рассмотрим теперь сумму $\zeta_N^{(r)}$. Обозначим $\varphi(t)$ и $\varphi_r(t)$ характеристические функции случайных величин ξ_1 и $\xi_1^{(r)}$ соответственно. Нетрудно видеть, что

$$\varphi(t) = \frac{\ln(1 - \lambda e^{it})}{\ln(1 - \lambda)}, \quad (12)$$

$$\varphi_r(t) = (1 - P_r)^{-1} \left(\varphi(t) - \sum_{s=r+1}^{\infty} e^{its} \mathbf{P}\{\xi_1 = s\} \right). \quad (13)$$

Лемма 13. Если $N \rightarrow \infty$, $\lambda = 1/N$, то для любого фиксированного $k = 0, 1, \dots$

$$\mathbf{P}\{\zeta_N^{(1)} = N\} \rightarrow 1, \quad \mathbf{P}\{\zeta_N^{(2)} - N = k\} \rightarrow \frac{1}{2^k k!} e^{-1/2}.$$

Доказательство. Подставляя в (13) $r = 1, 2$, нетрудно получить, что

$$\varphi_1^N(t) = e^{itN}, \quad \varphi_2^N(t) = e^{itN} e^{(e^{it}-1)/2} (1 + o(1)),$$

что и доказывает лемму 13.

Используя (4), нетрудно показать, что

$$\begin{aligned} m &= \mathbf{E}\xi_1 = -\frac{\lambda}{(1-\lambda)\ln(1-\lambda)}, \\ m_r &= \mathbf{E}\xi_1^{(r)} = (1-P_r)^{-1} \left(m - \sum_{s=r+1}^{\infty} s\mathbf{P}\{\xi_1 = s\} \right), \\ \sigma_r^2 &= \mathbf{D}\xi_1^{(r)} = (1-P_r)^{-1} \left(\sigma^2 + m^2 - \sum_{s=r+1}^{\infty} s^2\mathbf{P}\{\xi_1 = s\} \right) - m_r^2. \end{aligned} \quad (14)$$

Пусть $\psi_r(t)$ означает характеристическую функцию суммы $(\zeta_N^{(r)} - Nm_r)/(\sigma_r\sqrt{N})$.

Лемма 14. Пусть $N, n \rightarrow \infty$ так, что $n/N \geq C_1 > 1$, $N/\ln n \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что $NP_r \rightarrow \gamma$, где γ — некоторая положительная постоянная. Тогда равномерно по t в любом фиксированном конечном интервале

$$\psi_r(t) \rightarrow e^{-t^2/2}.$$

Это утверждение сохраняет силу при $r \geq 2$, $n/N \rightarrow 1$, $n - N \rightarrow \infty$, если $NP_{r-1} \rightarrow \infty$, $NP_r \rightarrow \gamma$, где γ — некоторая неотрицательная постоянная. Кроме того, в этом случае $\psi_{r+1}(t) \rightarrow e^{-t^2/2}$, а для $r \geq 3$ и $\psi_{r-1}(t) \rightarrow e^{-t^2/2}$.

Доказательство. Нетрудно видеть, что

$$\begin{aligned} \psi_r(t) &= \exp \left\{ -\frac{iNtm_r}{\sigma_r\sqrt{N}} \right\} (1-P_r)^{-N} \\ &\quad \times \left(\varphi \left(\frac{t}{\sigma_r\sqrt{N}} \right) - \sum_{s=r+1}^{\infty} \exp \left\{ \frac{its}{\sigma_r\sqrt{N}} \right\} \mathbf{P}\{\xi_1 = s\} \right)^N. \end{aligned} \quad (15)$$

С помощью (2) находим, что

$$\begin{aligned} \sum_{s=r+1}^{\infty} s\mathbf{P}\{\xi_1 = s\} &= -\frac{\lambda^{r+1}}{(1-\lambda)\ln(1-\lambda)}, \\ \sum_{s=r+1}^{\infty} s^2\mathbf{P}\{\xi_1 = s\} &= -\frac{\lambda^{r+1}}{(1-\lambda)^2\ln(1-\lambda)}. \end{aligned} \quad (16)$$

Используя (11), (14) и (16), нетрудно показать, что

$$\sigma_r\sqrt{N} \rightarrow \infty, \quad (17)$$

и это соотношение остается верным в случае $n/N \rightarrow 1$ при замене r на $r+1$, а также при замене r на $r-1$, если $r > 2$.

Очевидно, что

$$\ln \psi_r(t) = -im_r t N / (\sigma_r \sqrt{N}) - N \ln \varphi_r(t / (\sigma_r \sqrt{N})). \quad (18)$$

При достаточно малых t справедливо разложение

$$\ln \varphi_r(t) = 1 + im_r t - \frac{\sigma_r^2 t^2}{2} + \frac{t^3}{3} Q_r(t), \quad (19)$$

где

$$|Q_r(t)| \leq 2 \max_{|\tau| \leq |t|} |(\ln \varphi_r(\tau))'''|. \quad (20)$$

Из (17)–(19) следует, что при любом фиксированном t

$$\ln \psi_r(t) = -\frac{t^2}{2} + \frac{t^3}{6\sigma_r^3 \sqrt{N}} Q_r\left(\frac{t}{\sigma_r \sqrt{N}}\right). \quad (21)$$

Для доказательства леммы достаточно показать, что второе слагаемое в правой части равенства (21) стремится к нулю, в том числе и при замене r на $r+1$ и на $r-1$ для $r > 2$ в случае $n/N \rightarrow 1$. Это нетрудно сделать, используя (11), (12), (14), (16), (20) и рассматривая отдельно случаи $n/N \rightarrow 1$, $1 < C_1 \leq n/N \leq C_2 < \infty$, $n/N \rightarrow \infty$.

Из леммы 14 следует, что распределение $\zeta_N^{(r)}$ слабо сходится к нормальному закону. Покажем, что на самом деле имеет место локальное сближение этих распределений.

Лемма 15. Пусть $N, n \rightarrow \infty$ так, что $n/N \geq C_1 > 1$, $N/\ln n \rightarrow \infty$, λ — единственное решение уравнения (5) из интервала $(0, 1)$, а r выбрано так, что $NP_r \rightarrow \gamma$, где γ — некоторая положительная постоянная. Тогда равномерно относительно целых k , для которых $(k - Nm_r)/(\sigma_r \sqrt{N})$ лежит в любом фиксированном конечном интервале,

$$\sigma_r \sqrt{N} \mathbf{P}\{\zeta_N^{(r)} = k\} - \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(k - Nm_r)^2}{2\sigma_r^2 N}\right\} \rightarrow 0.$$

Это утверждение сохраняет силу при $r \geq 2$, $n/N \rightarrow 1$, $n - N \rightarrow \infty$, если $NP_{r-1} \rightarrow \infty$, $NP_r \rightarrow \gamma$, где γ — некоторая неотрицательная постоянная. Кроме того, в этом случае утверждение леммы сохраняет силу для суммы $\zeta_N^{(r+1)}$, а при $r \geq 3$ и для $\zeta_N^{(r-1)}$.

Доказательство. Следуя доказательству классической локальной теоремы, представим вероятность $\mathbf{P}\{\zeta_N^{(r)} = k\}$ по формуле обращения в виде интеграла

$$\mathbf{P}\{\zeta_N^{(r)} = k\} = \frac{1}{2\pi\sigma_r \sqrt{N}} \int_{-\pi\sigma_r \sqrt{N}}^{\pi\sigma_r \sqrt{N}} e^{-itk/(\sigma_r \sqrt{N})} \varphi_r^N\left(\frac{t}{\sigma_r \sqrt{N}}\right) dt.$$

Используя равенство

$$\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(k - Nm_r)^2}{2\sigma_r^2 N}\right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left\{-\frac{it(k - Nm_r)}{\sigma_r \sqrt{N}} - \frac{t^2}{2}\right\} dt,$$

разность

$$R = 2\pi \left(\sigma_r \sqrt{N} \mathbf{P} \{ \zeta_N^{(r)} = k \} - \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{(k - Nm_r)^2}{2\sigma_r^2 N} \right\} \right)$$

можно представить в виде суммы интегралов

$$R = I_1 + I_2 + I_3 + I_4,$$

где

$$\begin{aligned} I_1 &= \int_{-A}^A \exp \left\{ -\frac{it(k - Nm_r)}{\sigma_r \sqrt{N}} \right\} (\psi_r(t) - e^{-t^2/2}) dt, \\ I_2 &= \int_{A < |t| \leq \varepsilon \sigma_r \sqrt{N}} \exp \left\{ -\frac{it(k - Nm_r)}{\sigma_r \sqrt{N}} \right\} \psi_r(t) dt, \\ I_3 &= \int_{\varepsilon \sigma_r \sqrt{N} < |t| \leq \pi \sigma_r \sqrt{N}} \exp \left\{ -\frac{it(k - Nm_r)}{\sigma_r \sqrt{N}} \right\} \psi_r(t) dt, \\ I_4 &= - \int_{|t| > A} \exp \left\{ -\frac{it(k - Nm_r)}{\sigma_r \sqrt{N}} - \frac{t^2}{2} \right\} dt, \end{aligned}$$

положительные постоянные A и ε будут выбраны позднее. Лемма будет доказана, если мы покажем, что выбором достаточно больших N, n разность R можно сделать сколь угодно малой. Для этого мы рассмотрим отдельно интегралы I_1 – I_4 и покажем, что каждый из них может быть сделан сколь угодно малым.

При выполнении условий леммы 15 справедлива лемма 14, поэтому, как нетрудно показать, $I_1 \rightarrow 0$ при любом фиксированном A . Далее,

$$|I_4| \leq \int_{|t| > A} e^{-t^2/2} dt,$$

следовательно, интеграл I_4 можно сделать сколь угодно малым выбором достаточно большого A .

Рассмотрим I_3 . Пусть $1 < C_1 \leq n/N \leq C_2 < \infty$. Заметим, что максимальный шаг распределения ξ_1 равен единице, характеристическая функция $\varphi(t)$ непрерывно зависит от λ и $0 < C_3 \leq \lambda \leq C_4 < 1$. Отсюда следует, что при $\varepsilon < |t| \leq \pi$ справедливо неравенство $|\varphi(t)| \leq e^{-C_5}$, здесь и далее символы $C_5, C_6, \dots$ означают некоторые положительные постоянные. В случае $n/N \rightarrow \infty$ мы можем, следуя доказательству леммы 4.2.3 из [8], записать равенство

$$\varphi(t) = \frac{1}{\ln(1 - \lambda)} (\ln(1 - \lambda - it) + \ln(1 + \psi_1(t) + \psi_2(t))),$$

где

$$\begin{aligned} \psi_1(t) &= \frac{1 + it - e^{it}}{1 - \lambda - it}, & \psi_2(t) &= \frac{e^{it} - 1}{1 - \lambda - it}, \\ |\psi_1(t)| &\leq |t|/2, & |\psi_2(t)| &\leq 1 - \lambda. \end{aligned}$$

Из этих соотношений легко получить, что при $\varepsilon < |t| \leq \pi$ для величины $|\varphi(t)|$ также справедлива указанная выше оценка. Отсюда и из (15) находим, что

$$|\psi_r(t)| \leq C_6 e^{-C_5 N}, \quad (22)$$

поэтому, используя (11), (14) и (16), получаем, что

$$|I_3| \leq C_7 \sigma_r \sqrt{N} e^{-C_5 N} \leq C_8 (1 - \lambda)^{-1} \sqrt{-N / \ln(1 - \lambda)} e^{-C_5 N}. \quad (23)$$

Понятно, что если $1 < C_1 \leq n/N \leq C_2$, то последнее выражение стремится к нулю. Пусть $n/N \rightarrow \infty$. Из (10) и (22) следует, что $|I_3| = o(ne^{-C_5 N})$ и соотношение $I_3 \rightarrow 0$ вытекает из условия $N / \ln n \rightarrow \infty$. Пусть $n/N \rightarrow 1$. Тогда из (12) следует, что при $\varepsilon < |t| \leq \pi$ справедливо неравенство $|\varphi(t)| < e^{-C_9 \lambda}$. Учитывая, что в этом случае $P_r = O(\lambda^r / r)$, из (15) находим, что

$$|\psi_r(t)| \leq C_{10} e^{-C_9 \lambda N}.$$

Отсюда и из легко проверяемого неравенства $\sigma_r^2 > C_{11} \lambda$ (см. (11), (14), (16)), получаем, что

$$|I_3| \leq C_{12} \lambda N e^{-C_9 \lambda N},$$

а последнее выражение стремится к нулю в силу соотношения $\lambda N \rightarrow \infty$. Нетрудно проверить, что эти рассуждения остаются в силе и при указанных в условиях леммы заменах r на $r + 1$, $r - 1$.

Нам осталось рассмотреть интеграл I_2 . Пусть сначала $n/N \leq C_2$. Тогда из (11), (14), (18)–(21) нетрудно получить, что в области интегрирования I_2

$$\ln \psi_r(t) = (-t^2/2)(1 + \delta(t, N, n)), \quad (24)$$

где величину $\delta(t, N, n)$ можно сделать сколь угодно малой выбором достаточно больших N, n и достаточно малого ε . Тогда

$$|I_2| < \int_{|t| > A} e^{-t^2/4} dt, \quad (25)$$

и интеграл I_2 можно сделать сколь угодно малым выбором достаточно большого A . В случае $n/N \rightarrow 1$, $n - N \rightarrow \infty$ аналогично проверяется возможность указанной в лемме замены r .

Пусть $n/N \rightarrow \infty$. Представим интеграл I_2 в виде суммы

$$I_2 = I'_2 + I''_2 + I'''_2,$$

где области интегрирования I'_2, I''_2, I'''_2 соответственно равны

$$\begin{aligned} \{t: A < |t| \leq \varepsilon \sqrt{N / \ln n}\}, \\ \{t: \varepsilon \sqrt{N / \ln n} < |t| \leq N / \ln n\}, \\ \{t: N / \ln n < |t| \leq \varepsilon \sigma_r \sqrt{N}\}. \end{aligned}$$

Заметим, что область интегрирования I'''_2 может оказаться пустой, в этом случае верхняя граница интегрирования I''_2 равна $\varepsilon \sigma_r \sqrt{N}$. С помощью (11), (14), (18)–(21) нетрудно получить, что в области интегрирования I'_2 справедливы соотношения (24) и (25). Нетрудно видеть, что в области интегрирования I''_2

$$\left| \operatorname{Re} \left(\frac{1 - \lambda \exp\{it/(\sigma_r \sqrt{N})\}}{1 - \lambda} \right) / \ln(1 - \lambda) \right| < -\frac{C_{13}}{\ln(1 - \lambda)}. \quad (26)$$

Поскольку

$$\varphi^N \left(\frac{t}{\sigma_r \sqrt{N}} \right) = \left(1 + \ln \left(\frac{1 - \lambda \exp\{it/(\sigma_r \sqrt{N})\}}{1 - \lambda} \right) / \ln(1 - \lambda) \right)^N, \quad (27)$$

из (15) и (26) находим, что

$$|I_2''| < -C_{14} \frac{N}{\ln(1 - \lambda)} \exp \left\{ -\frac{C_{15} N}{\ln(1 - \lambda)} \right\},$$

и остается заметить, что из условия $N/\ln n \rightarrow \infty$ следует, что $-N/\ln(1 - \lambda) \rightarrow \infty$. Для оценки I_2''' нетрудно получить, что в соответствующей области изменения t

$$\left| \operatorname{Re} \left(\frac{1 - \lambda \exp\{it/(\sigma_r \sqrt{N})\}}{1 - \lambda} \right) / \ln(1 - \lambda) \right| \leq -C_{13} \frac{\ln(|t|/((1 - \lambda)\sigma_r \sqrt{N}))}{\ln(1 - \lambda)},$$

поэтому, учитывая (27), получаем, что

$$\begin{aligned} |I_2'''| &< C_{16} \int_{-N/\ln(1-\lambda)}^{\varepsilon \sigma_r \sqrt{N}} \left(\frac{t}{\sigma_r \sqrt{N}} \right)^{C_{17} N / \ln(1-\lambda)} dt \\ &\leq C_{18} \sqrt{-\frac{N}{\ln(1-\lambda)}} \int_{\sqrt{-N/\ln(1-\lambda)}}^{\varepsilon/(1-\lambda)} y^{C_{17} N / \ln(1-\lambda)} dy. \end{aligned}$$

Опять используя соотношение $-N/\ln(1 - \lambda) \rightarrow \infty$ находим, что $I_2''' \rightarrow 0$. Лемма 15 доказана.

Лемма 16. Пусть $\lambda = 1 - n^{-1}$, $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, $N/\ln n \rightarrow \alpha$, $r = zn$, где α, z — некоторые положительные постоянные. Тогда равномерно по t в любом фиксированном конечном интервале

$$\varphi_r^N(t/n) = \frac{e^{-E(z,t)}}{(1 - it)^\alpha e^{-E(z,0)}} + o(1),$$

где

$$E(z, t) = \alpha \int_z^\infty y^{-1} e^{-y(1-it)} dy. \quad (28)$$

Доказательство. Из (2) и (13) следует, что при любом фиксированном t

$$\begin{aligned} \varphi_r^N \left(\frac{t}{n} \right) &= \varphi^N \left(\frac{t}{n} \right) (1 - P_r)^{-N} \\ &\times \left(1 - \frac{1 + o(1)}{\ln n} \varphi^{-1} \left(\frac{t}{n} \right) \sum_{s=r+1}^\infty \frac{1}{s} \left(1 - \frac{1}{n} \right)^s e^{its/n} \right)^N. \end{aligned} \quad (29)$$

Согласно лемме 4.2.1 из [8]

$$\varphi^N(t/n) \rightarrow (1 - it)^{-\alpha}. \quad (30)$$

При доказательстве леммы 1.8.8 из [3] показано, что

$$\sum_{s=r+1}^{\infty} \frac{1}{s} \left(1 - \frac{1}{n}\right)^s e^{its/n} = \int_z^{\infty} y^{-1} e^{-y(1-it)} dy + o(1).$$

Утверждение леммы 16 теперь нетрудно получить из (29), (30) и леммы 6.

Лемма 17. Пусть $\lambda = 1 - n^{-1}$, $N, n \rightarrow \infty$ так, что $N/\ln n \rightarrow \alpha$, $r = zn$, где α, z — некоторые положительные постоянные. Тогда равномерно относительно $y \geq y_0 > 0$

$$n\mathbf{P}\{\zeta_N^{(r)}/n = y\} = \frac{e^{E(z,0)-y}}{\Gamma(\alpha)} \sum_{s=0}^{\infty} \frac{(-\alpha)^s}{s!} I_s(y, z),$$

где

$$\begin{aligned} I_0(y, z) &= y^{\alpha-1}, \\ I_s(y, z) &= \int_{X(s,y,z)} \frac{(y - x_1 - \dots - x_s)^{\alpha-1}}{x_1 \dots x_s} dx_1 \dots dx_s, \\ X(s, y, z) &= \{x_i \geq z, i = 1, \dots, s, x_1 + \dots + x_s \leq y\}, \quad s = 1, 2, \dots \end{aligned}$$

Доказательство. Найдем плотность распределения, соответствующую пределу характеристической функции случайной величины $\zeta_N^{(r)}/n$, полученному в лемме 16. Для этого разложим $e^{-E(z,t)}$ в ряд по степеням $E(z,t)$ и представим эту характеристическую функцию в виде

$$e^{E(z,0)} \sum_{s=0}^{\infty} \frac{(-1)^s}{s!} \left(\frac{E(z,t)}{E(z,0)} \right)^s \frac{(E(z,0))^s}{(1-it)^\alpha}.$$

Функция $E(z,t)/E(z,0)$ является характеристической функцией распределения с плотностью

$$\alpha(xE(z,0)e^x)^{-1}, \quad x \geq z.$$

Учитывая, что характеристической функции $(1-it)^{-\alpha}$ соответствует гамма-распределение с параметрами $(\alpha, 1)$, получаем, что характеристической функции

$$(E(z,t)/E(z,0))^s (1-it)^{-\alpha}$$

соответствует плотность распределения, равная

$$\frac{e^{-y}}{\Gamma(\alpha)} \left(\frac{\alpha}{E(z,0)} \right)^s \int_{X(s,y,z)} \frac{(y - x_1 - \dots - x_s)^{\alpha-1}}{x_1 \dots x_s} dx_1 \dots dx_s,$$

поэтому $\varphi_r^N(t/n)$ сходится к характеристической функции распределения, имеющего плотность

$$g(y) = \frac{e^{E(z,0)-y}}{\Gamma(\alpha)} \sum_{s=0}^{\infty} \frac{(-\alpha)^s}{s!} I_s(y, s).$$

По формуле обращения вероятность $\mathbf{P}\{\zeta_N^{(r)}/n = y\}$ представим в виде

$$\mathbf{P}\{\zeta_N^{(r)}/n = y\} = (2\pi n)^{-1} \int_{-\pi n}^{\pi n} e^{-ity} \varphi_r^N(t/n) dt. \quad (31)$$

Поскольку

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \frac{e^{-E(z,t)}}{(1-it)^{\alpha} e^{-E(z,0)}} dt,$$

из (31) следует, что

$$2\pi n \mathbf{P}\{\zeta_N^{(r)} \mid n = y\} - 2\pi n g(y) = I_1 + I_2 + I_3 + I_4,$$

где

$$\begin{aligned} I_1 &= \int_{-A}^A e^{-ity} \left(\varphi_r^N \left(\frac{t}{n} \right) - \frac{e^{-E(z,t)}}{(1-it)^{\alpha} e^{-E(z,0)}} \right) dt, \\ I_2 &= \int_{A < |t| \leq \varepsilon n} e^{-ity} \varphi_r^N(t/n) dt, \\ I_3 &= \int_{\varepsilon n < |t| \leq \pi n} e^{-ity} \varphi_r^N(t/n) dt, \\ I_4 &= - \int_{A < |t|} e^{-ity} \frac{e^{-E(z,t)}}{(1-it)^{\alpha} e^{-E(z,0)}} dt, \end{aligned}$$

положительные постоянные A и ε будут выбраны ниже. Для доказательства леммы достаточно показать, что каждый из интегралов I_1 – I_4 может быть сделан сколь угодно малым.

Из леммы 16 следует, что при любом фиксированном A интеграл I_1 стремится к нулю. Для оценки I_4 проведем интегрирование по частям. Тогда

$$I_4 = -e^{E(z,0)} \left(-\frac{e^{-ity} e^{-E(z,t)}}{iy(1-it)^{\alpha}} \Big|_A^{\infty} + \frac{\alpha}{y} \int_A^{\infty} e^{-ity} \frac{e^{-E(z,t)}}{(1-it)^{\alpha+1}} (1 - e^{-z(1+it)}) dt \right).$$

Отсюда и из (28) следует, что существует положительная постоянная C_1 такая, что

$$|I_4| \leq \frac{C_1}{y} \left(\frac{1}{(1+A^2)^{\alpha/2}} + \int_A^{\infty} \frac{dt}{(1+t^2)^{(\alpha+1)/2}} \right),$$

и I_4 можно сделать сколь угодно малым выбором достаточно большого A .

Из (29) и леммы 16 легко видеть, что при достаточно больших N , n найдется такая постоянная $C_2 > 0$, что

$$|\varphi_r^N(t/n)| \leq C_2 |\varphi^N(t/n)|.$$

Используя это неравенство, нетрудно оценить интегралы I_2 и I_3 аналогично оценкам одноименных интегралов, найденным при доказательстве теоремы 4.2.1 в книге [8].

Лемма 18. Пусть $\lambda = 1 - n^{-1}$, $N/\ln n \rightarrow 0$, $r = n - n \exp\{-(z \ln n)/N\}$, где z — некоторая положительная постоянная. Тогда

$$\mathbf{P}\{\zeta_N^{(r)} = n\} = \frac{e^{-1}(1 - e^{-z})(1 + o(1))}{n\Gamma(N/\ln n)}.$$

Доказательство. Нетрудно видеть, учитывая (4), что

$$\mathbf{P}\{\zeta_N^{(r)} = n\} = (1 - P_r)^{-N} \mathbf{P}\{\zeta_N = n, \xi_i \leq r, i = 1, \dots, N\}.$$

Отсюда и из леммы 7 получаем, что

$$\mathbf{P}\{\zeta_N^{(r)} = n\} = (\mathbf{P}\{\zeta_N = n\} - N\mathbf{P}\{\zeta_N = n, \xi_N > r\})(1 + o(1)). \quad (32)$$

Из (2) следует, что

$$\mathbf{P}\{\zeta_N = n, \xi_N > r\} = \frac{e^{-1}(1 + o(1))}{(\ln n)^N} \sum \frac{1}{k_1 \dots k_N},$$

где суммирование проводится по всем $k_1, \dots, k_N$ таким, что $k_1 + \dots + k_N = n$, $k_N > r$, $k_1, \dots, k_{N-1} \geq 1$. Тогда

$$\mathbf{P}\{\zeta_N = n, \xi_N > r\} = \frac{e^{-1}(1 + o(1))}{n(\ln n)^N} \sum \frac{1}{k_1 \dots k_{N-1}},$$

в последней сумме $k_1 + \dots + k_{N-1} < ne^{-(z \ln n)/N}$. Нетрудно проверить (например, заменяя суммирование интегрированием), что последняя сумма равна $(\ln n)^{N-1} e^{-z}$, поэтому

$$\mathbf{P}\{\zeta_N = n, \xi_N > r\} = \frac{e^{-z-1}(1 + o(1))}{n \ln n}.$$

Отсюда, из (32), леммы 12 и соотношения $\Gamma(N/\ln n) = ((\ln n)/N)(1 + o(1))$ получаем утверждение леммы 18.

Теперь мы можем доказать теоремы 1–5. Пусть выполнены условия теоремы 1. Нетрудно видеть, что в этом случае

$$\lambda = \frac{2(n - N)}{N}(1 + o(1)). \quad (33)$$

Пусть сначала $n = N + k$, где $k = 1, 2, \dots$. Тогда из (33) находим, что в условиях теоремы $r = 2$, $\gamma = 0$. Как мы видели выше, утверждение леммы 1 верно для любого λ , $0 < \lambda < 1$, поэтому, используя леммы 1, 2, 8 и 13 при $\lambda = 1/N$, находим, что

$$\mathbf{P}\{\eta_{(N)} \leq 1\} \rightarrow 0, \quad \mathbf{P}\{\eta_{(N)} \leq 2\} \rightarrow 1,$$

следовательно, для этого случая теорема 1 верна. Пусть теперь $n - N \rightarrow \infty$. Используя (2), (11), (14), (16) и (33), нетрудно получить соотношения

$$\sigma_r^2 N = (n - N)(1 + o(1)), \quad (n - Nm_r)^2 = o(\sigma_r^2 N).$$

Отсюда и из лемм 3, 9, 15 находим, что

$$(1 - P_r)^N \rightarrow e^{-\gamma}, \quad (1 - P_{r+1})^N \rightarrow 1, \\ \mathbf{P}\{\zeta_N^{(s)} = n\} / \mathbf{P}\{\zeta_N = n\} \rightarrow 1, \quad s = r, r + 1,$$

а если $r \geq 3$, то, кроме того,

$$(1 - P_{r-1})^N \rightarrow 0, \quad \mathbf{P}\{\zeta_N^{(r-1)} = n\} / \mathbf{P}\{\zeta_N = n\} \rightarrow 1.$$

Из этих соотношений и леммы 1 следует, что

$$\mathbf{P}\{\eta_{(N)} \leq r\} \rightarrow e^{-\gamma}, \quad \mathbf{P}\{\eta_{(N)} \leq r+1\} \rightarrow 1.$$

Отсюда и из очевидного условия $\mathbf{P}\{\eta_{(N)} \leq 1\} = 0$ следует утверждение теоремы 1.

Для доказательства теоремы 2 заметим, что из (5), (11), (14) и (16) легко получить соотношения $\sigma_r = \sigma(1 + o(1))$, $(n - Nm_r)^2 = o(\sigma_r^2 N)$, поэтому теорема 2 следует из лемм 1, 4, 10 и 15. Аналогично, для доказательства теоремы 3 достаточно применить леммы 1, 5, 10 и 15. Теорема 4 легко следует из лемм 1, 6, 11 и 17, а теорема 5 из лемм 1, 7, 12 и 18.

Список литературы

1. Павлов Ю. Л., Асимптотическое распределение максимального объема дерева в случайном лесе. *Теория вероятностей и ее применения* (1977) **22**, №3, 523–533.
2. Павлов Ю. Л., Один случай предельного распределения максимального объема дерева в случайном лесе. *Матем. заметки* (1979) **25**, №5, 751–760.
3. Колчин В. Ф., *Случайные отображения*. Наука, Москва, 1984.
4. Бритиков В. Е., Предельные теоремы для максимального объема дерева в случайном лесе из некорневых деревьев. *Вероятностные задачи дискретной математики*. МИЭМ, Москва, 1987, 84–91.
5. Pavlov Yu. L., *Random forests*. VSP, Utrecht, 2000.
6. Meir A., Moon J. W., On the altitude of nodes in random trees. *Canad. J. Math.* (1978) **30**, №3, 997–1015.
7. Shapiro A., A generalized distribution model for random recursive trees. *Acta Informatica* (1997) **34**, 211–216.
8. Колчин В. Ф., *Случайные графы*. Наука, Москва, 2000.

Статья поступила 18.09.2001.