Дискретная математика

том 34 выпуск 4 * 2022

УДК 519.212.2+519.179.4 DOI https://doi.org/10.4213/dm1729

О максимальном объеме дерева случайного леса

© 2022 г. Ю. Л. Павлов*

Рассматриваются леса Гальтона – Ватсона, состоящие из N корневых деревьев и n некорневых вершин. Распределение числа прямых потомков каждой частицы образующего лес критического ветвящегося процесса имеет правильно меняющийся хвост и бесконечную дисперсию. Доказана предельная теорема для максимального объема дерева, когда $N, n \to \infty$, $n/N \to \infty$. Эта теорема справедлива в значительно более широкой области изменения параметров N и n, чем это было известно ранее.

Работа выполнена при поддержке средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

Ключевые слова: лес Гальтона – Ватсона, объем дерева, степень вершины, предельные теоремы

1. Введение

В статьях [1,2] рассматривались случайные леса Гальтона – Ватсона, содержащие N корневых деревьев и n некорневых вершин. Предполагалось, что в образующих такие леса критических ветвящихся процессах распределения числа прямых потомков частиц имеют конечные математические ожидания, но бесконечные дисперсии. В последние годы методы теории ветвящихся процессов успешно используются при изучении структуры и динамики конфигурационных графов, предназначенных для моделирования современных сложных сетей коммуникаций, в частности, сети Интернет (см., например, [3]). Многочисленные наблюдения за реальными сетями показали, что в соответствующих моделях степени вершин графа можно считать независимыми одинаково распределенными случайными величинами.

Конфигурационные графы впервые были введены в [4]. В моделирующих сети конфигурационных графах принято считать, что степень любой вершины не меньше единицы и равна числу инцидентных этой вершине различимых полуребер, т. е. ребер, для которых смежные вершины еще не определены. Замечено [3,5], что в сложных сетях коммуникаций число узлов степени k при больших k пропорционально $k^{-\tau}$, где $\tau > 1$. Поэтому в моделирующих такие сети конфигурационных графах

*Место работы: Институт прикладных математических исследований Карельского научного центра РАН, e-mail: pavlov@krc.karelia.ru

обычно предполагается, что степень каждой вершины является значением случайной величины η, имеющей распределение

$$\mathbf{P}\{\eta = k\} = \frac{h(k)}{k^{\tau}}, \tag{1}$$

где $k = 1, 2, \ldots, \tau > 1$, $h(k)$ — медленно меняющаяся на бесконечности функция.

Конфигурационный граф формируется путем попарного равновероятного соединения полуребер друг с другом для образования ребер. Понятно, что сумма степеней вершин любого графа должна быть четной, поэтому при нечетной сумме в граф вводится вспомогательная вершина единичной степени или дополнительное полуребро присоединяется к равновероятно выбранной вершине. В [5] отмечается, что появление такого вспомогательного элемента не влияет на асимптотическое поведение основных характеристик графа при стремлении к бесконечности числа вершин.

При построении конфигурационного графа возможно появление кратных ребер, петель и циклов. Хотя таких объетов нет в траекториях ветвящихся процессов, известно ([3]), что существует тесная связь между конфигурационными графами и ветвящимися процессами. Обсудим эту связь, интерпретируя вершины графа как частицы процесса. Если рассмотреть некоторый подграф конфигурационного графа и выделить в нем произвольную вершину, то такому подграфу можно поставить в соответствие траекторию процесса Гальтона – Ватсона, начинающегося с одной частицы, прообразом которой служит выделенная вершина. По аналогии с ветвящимся процессом эту вершину можно назвать начальной. Естественно считать, что распределение числа прямых потомков начальной частицы совпадает с распределением (1) степени начальной вершины графа. Вершины, смежные с начальной, назовем вершинами первого поколения. Любая такая вершина соединена с начальной хотя бы одним ребром. Поэтому распределение числа прямых потомков частиц первого поколения отличается от распределения размера потомства начальной частицы. Обозначим ξ случайную величину, значения которой интерпретируются как численность прямых потомков любой частицы процесса, кроме начальной. Пусть в распределении (1) $\tau > 2$, следовательно, случайная величина η имеет конечное математическое ожидание $\mathbf{E}\eta$. В [3] показано, что в соответствующем конфигурационному графу ветвящемся процессе Гальтона – Ватсона распределение случайной величины ξ имеет вид

$$\mathbf{P}\{\xi = k\} = \frac{(k+1)\mathbf{P}\{\eta = k+1\}}{\mathbf{E}\eta}, \qquad k = 0, 1, 2, \ldots . \tag{2}$$

Известно (см., например, [3]), что если число вершин рассматриваемого подграфа конфигурационного графа ограничено или достаточно медленно стремится к бесконечности, то такой подграф с вероятностью, стремящейся к единице, является деревом, корнем которого служит начальная вершина. Этот факт дает основание говорить о локальной древовидности конфигурационного графа, состоящей в том, что в некоторой «окрестности» каждой вершины асимптотически нет кратных ребер, петель и циклов.

Если рассмотреть частицы первого поколения в качестве начальных, то можно говорить о начинающемся с нескольких частиц ветвящемся процессе, в котором (2) является распределением числа прямых потомков любой частицы. Совокупность всех траекторий такого процесса образует случайный лес Гальтона – Ватсона. В [1,2] впервые предложено использовать результаты о случайных лесах для исследования структуры конфигурационных графов. Теория случайных лесов наиболее подробно изложена в [6]. Полученные там результаты доказаны (см. также [7]) при условии конечности дисперсии числа прямых потомков частиц ветвящегося процесса. Однако в ветвящихся процессах, используемых для изучения конфигурационных графов, соответствующих реальным сетям, это условие нарушено ([3]) и, следовательно, результаты [6] неприменимы. Поэтому в [1, 2] рассматривались случайные леса в случае бесконечной дисперсии распределений числа потомков частиц.

Пусть ветвящийся процесс Гальтона – Ватсона G начинается с N частиц и число ξ прямых потомков каждой частицы имеет распределение

$$p_k = \mathbf{P}\{\xi = k\} = \frac{h(k+1)}{(k+1)^\tau}, \qquad k = 0, 1, 2, \ldots, \quad \tau \in (2, 3), \tag{3}$$

где медленно меняющаяся на бесконечности функция $h(x)$ при $x \geqslant 1$ принимает только положительные значения. Легко видеть, что распределение (3) имеет конечное математическое ожидание и бесконечную дисперсию. Далее будем считать, что процесс G является критическим: $\mathbf{E}\xi = 1$. Очевидно, что G распадается на N независимых процессов $G_1, \ldots, G_N$, каждый из которых начинается с одной частицы. Подмножество траекторий такого процесса, общее число частиц которого до вырождения равно $N + n$, вместе с индуцированным процессом распределением вероятностей на этом множестве, образует случайный лес Гальтона – Ватсона $\mathfrak{F}_{N,n}$ с N корневыми деревьями и n некорневыми вершинами.

Объемом корневого дерева называется число его вершин, включая корень. Обозначим $\nu_{\max}$ максимальный объем дерева в $\mathfrak{F}_{N,n}$. В статье [1] найдены предельные распределения $\nu_{\max}$ в трех областях стремления N и n к бесконечности: $n/N \to 0$, $0 < C_1 \leqslant n/N \leqslant C_2 < \infty$, $n/N \to \infty$. В последней области возникли наибольшие технические трудности, и соответствующую теорему удалось доказать только при небольшой скорости стремления n/N к бесконечности. Для того чтобы сформулировать возникшие при этом ограничения, введем необходимые обозначения.

Рассмотрим случайную величину $\xi(\lambda)$, имеющую распределение

$$p_k(\lambda) = \mathbf{P}\{\xi(\lambda) = k\} = \frac{\lambda^k p_k}{F(\lambda)}, \qquad k = 0, 1, 2, \ldots, \tag{4}$$

где $0 < \lambda < 1$, вероятности p_k определены в (3) и

$$F(\lambda) = \sum_{k=0}^{\infty} p_k \lambda^k. \tag{5}$$

В качестве значения параметра λ распределения (4) возьмем единственное решение уравнения

$$a = \mathbf{E}\xi(\lambda) = \frac{\lambda F'(\lambda)}{F(\lambda)} = \frac{n}{N+n}. \tag{6}$$

Пусть $N, n \to \infty$ так, что $n/N \to \infty$ и существуют такие положительные Δ, υ, где $\upsilon < 1/2$, что

$$\left(\frac{n}{N}\right)^2 = o(N^{\upsilon}(1-\lambda)^{1+\Delta}). \tag{7}$$

Обозначим $\alpha = \lambda/F(\lambda)$, $\beta = -\ln\alpha$, и пусть существует такое $\omega > 0$, что $\beta > (\ln N)^{-\omega}$. В [1] доказано, что при выполнении этих условий $\lambda \uparrow 1$ и $\nu_{\max}$ имеет экспоненциальное предельное распределение. В следующей статье [2] показано, что условие (7) можно слегка ослабить и заменить на

$$\left(\frac{n}{N}\right)^2 = O(N^{\upsilon}(1-\lambda)^{1+\Delta}). \tag{8}$$

Естественно возник вопрос о предельном распределении $\nu_{\max}$ в существенно более широкой области изменения параметров N, n, чем в [1]. В настоящей статье для этой случайной величины доказана предельная теорема, в которой условия на поведение N, n и β значительно слабее. В частности, предполагается, что существует такое ω, $0 < \omega < 1/2$, что

$$\beta > N^{-\omega}. \tag{9}$$

Эта теорема, как и в [1], доказана с использованием обобщенной схемы размещения частиц по ячейкам [8]. В рамках этой схемы устанавливается локальная сходимость к предельным законам распределений сумм вспомогательных независимых случайных величин, образующих схемы серий. Новый результат удалось получить, проводя такое доказательство иначе, чем в [1].

Статья имеет следующую структуру. В разделе 2 сформулировано основное утверждение (теорема 1). Кроме того, в нем доказана лемма 1 о том, что условия теоремы 1 на поведение N и n слабее, чем (8). В разделе 3 вводится вспомогательный ветвящийся процесс $G(\lambda)$ с распределением (4) числа прямых потомков каждой частицы, обсуждается связь между случайными лесами и обобщенной схемой размещения частиц по ячейкам и приводится лемма 2, являющаяся следствием этой связи. В лемме 3 раздела 4 рассматривается предельное поведение хвоста распределения числа частиц, существовавших в процессе $G(\lambda)$ до его вырождения, при условии, что процесс начинается с одной частицы. В разделе 5 доказаны леммы 4 и 5 об асимптотике общего числа частиц, существовавших в процессе $G(\lambda)$. В разделе 6 доказаны леммы 6 и 7 об асимптотическом поведении сумм вспомогательных случайных величин. Перечисленные леммы используются в разделе 7, где доказана теорема 1.

2. Основной результат

Обозначим

$$L = L(N, n, \delta) = \min\Big(n(1-\lambda)^{\tau-1+\delta}, \frac{N^2}{n}(1-\lambda)^{3-\tau+\delta}\Big), \tag{10}$$

где $\delta > 0$. В доказанной ниже теореме предполагается, что при $N, n \to \infty$ существует такое $\delta_0 > 0$ что $L(N, n, \delta_0) \to \infty$. Поскольку $\lambda \uparrow 1$, легко видеть, что в этом случае для всех δ, $0 < \delta \leqslant \delta_0$,

$$L(N, n, \delta) \to \infty. \tag{11}$$

Введем стремящуюся к бесконечности последовательность B_r, $r = 1, 2, \ldots$, которая при $r \to \infty$ удовлетворяет условию

$$B_r \sim (rh(B_r))^{\frac{1}{(\tau-1)}}. \tag{12}$$

Пусть $g(x)$ обозначает плотность устойчивого закона с параметром $\tau-1$ и характеристической функцией

$$f(t)=\exp\biggl\{-c|t|^{\tau-1}\biggl(1-i\frac{t}{|t|}\operatorname{tg}\frac{\pi(\tau-1)}{2}\biggr)\biggr\},$$

где

$$c=\frac{\Gamma(2-\tau)}{\tau-1}\cos\frac{\pi(\tau-1)}{2},$$

$\Gamma(x)$ — гамма-функция. Заметим, что $0<g(0)<\infty$ при $\tau\in(2,3)$ в силу одновершинности устойчивых распределений (см. [9, п. 2.7]).

Для $\nu_{\max}$ справедливо следующее утверждение.

Теорема 1. *Пусть случайный лес Гальтона – Ватсона, состоящий из N корневых деревьев и содержащий n некорневых вершин, образован ветвящимся процессом с распределением* (3) *числа прямых потомков каждой частицы, где $\tau\in(2,3)$, а медленно меняющаяся на бесконечности функция $h(x)$ выбрана так, что $\mathbf{E}\xi=1$. Пусть $N,n\to\infty$ так, что $n/N\to\infty$, выполнено условие* (9) *и существует такое $\delta_0>0$, что верно соотношение* (11). *Если последовательность $r=r(N,n)$ удовлетворяет условию*

$$\frac{Ng(0)}{r\beta e^{r\beta}B_r}\to z, \tag{13}$$

где z — фиксированное положительное число, то

$$\mathbf{P}\{\nu_{\max}\leqslant r\}\to e^{-z}.$$

Во введении отмечалось, что условие (11) слабее, чем ограничение (8). Покажем это.

Лемма 1. *Из условия* (8) *следует* (11). *Обратное неверно.*

Доказательство. Пусть выполнено (8). Тогда

$$n(1-\lambda)^{\tau-1+\delta}=n\biggl(\frac{1}{xN^{\upsilon}}\Bigl(\frac{n}{N}\Bigr)^2\biggr)^{\frac{\tau-1+\delta}{1+\Delta}},$$

где $0<x\leqslant C_3<\infty$, здесь и далее $C_3,C_4,\ldots$ означают некоторые положительные постоянные. Отсюда

$$n(1-\lambda)^{\tau-1+\delta}\geqslant C_4\Bigl(\frac{n}{N}\Bigr)^{1+2\frac{\tau-1+\delta}{1+\Delta}}N^{1-\upsilon\frac{\tau-1+\delta}{1+\Delta}}. \tag{14}$$

Очевидно, что при $0<\upsilon<1/2$, $2<\tau<3$ и достаточно малых δ, Δ

$$\upsilon^{\frac{\tau-1+\delta}{1+\Delta}}<1, \tag{15}$$

и из (14) следует, что

$$n(1-\lambda)^{\tau-1+\delta}\to\infty. \tag{16}$$

Далее, из (8) находим, что

$$\frac{N^2}{n}(1-\lambda)^{3-\tau+\delta}\geqslant C_5n^{\frac{4}{2+\upsilon}-1}(1-\lambda)^{3-\tau+\delta-\frac{2+2\Delta}{2+\upsilon}}.$$

Легко видеть, учитывая (16), что в рассматриваемых условиях последнее выражение стремится к бесконечности, отсюда и из (10) получаем (11).

Покажем теперь, что условие (11) может выполняться и в случае нарушения (8). Пусть

$$\left(\frac{n}{N}\right)^2 = xN^{\upsilon}(1-\lambda)^{1+\Delta}, \tag{17}$$

где $x \to \infty$. Тогда

$$n(1-\lambda)^{\tau-1+\delta} = n\left(\frac{1}{xN^{\upsilon}}\left(\frac{n}{N}\right)^2\right)^{\frac{\tau-1+\delta}{1+\Delta}}. \tag{18}$$

Отсюда и из (8) следует, что если x стремится к бесконечности достаточно медленно, то справедливо соотношение (16).

Из (17) вытекает, что

$$\frac{N^2}{n}(1-\lambda)^{3-\tau+\delta} = x^{-\frac{2}{2+\upsilon}} n^{\frac{2-\upsilon}{2+\upsilon}} (1-\lambda)^{3-\tau+\delta-\frac{2+2\Delta}{2+\upsilon}}.$$

Принимая во внимание (16), отсюда находим, что если $x \to \infty$, но не слишком быстро, то

$$\frac{N^2}{n}(1-\lambda)^{3-\tau+\delta} \to \infty$$

и, следовательно, (11) имеет место. Лемма 1 доказана. □

3. Случайный лес и обобщенная схема размещения

Рассмотрим ветвящийся процесс $G(\lambda)$, отличающийся от G только тем, что число прямых потомков каждой частицы процесса $G(\lambda)$ имеет распределение (4), а не (3). Так же, как и G, процесс $G(\lambda)$ распадается на N начинающихся с одной частицы независимых процессов $G_1(\lambda), ..., G_N(\lambda)$. Обозначим $\nu^{(1)}(\lambda), \ldots, \nu^{(N)}(\lambda)$ случайные величины, равные числам частиц, существовавших в процессах $G_1(\lambda), \ldots, G_N(\lambda)$ соответственно до их вырождения. Разумеется, эти случайные величины независимы и одинаково распределены. Пусть $\nu(\lambda)$ — общее число частиц, существовавших в процессе $G(\lambda)$ до его вырождения. Тогда $\nu(\lambda) = \nu^{(1)}(\lambda) + \cdots + \nu^{(N)}(\lambda)$. Обозначим также $\nu_1(\mathfrak{F}), \ldots, \nu_N(\mathfrak{F})$ случайные величины, равные объемам деревьев леса из $\mathfrak{F}_{N,n}$, корни которых занумерованы числами $1, \ldots, N$ соответственно. В [6] показано, что если натуральные числа $k_1, \ldots, k_N$ выбраны так, что $k_1 + \cdots + k_N = N + n$, то

$$\mathbf{P}\{\nu_1(\mathfrak{F}) = k_1, \ldots, \nu_N(\mathfrak{F}) = k_N\} = \mathbf{P}\{\nu^{(1)}(\lambda) = k_1, \ldots, \nu^{(N)}(\lambda) = k_N \,|\, \nu(\lambda) = N+n\}. \tag{19}$$

Это равенство означает, что два набора случайных величин $\nu_1(\mathfrak{F}), \ldots, \nu_N(\mathfrak{F})$ и $\nu^{(1)}(\lambda), \ldots, \nu^{(N)}(\lambda)$ удовлетворяют условиям обобщенной схемы размещения частиц по ячейкам [8].

Введем такие независимые случайные величины $\nu_r^{(1)}(\lambda), \ldots, \nu_r^{(N)}(\lambda)$, что

$$\mathbf{P}\{\nu_r^{(i)}(\lambda) = k\} = \mathbf{P}\{\nu^{(i)}(\lambda) = k \mid \nu^{(i)}(\lambda) \leqslant r\}, \tag{20}$$

где $i = 1, \ldots, N$, $k \leqslant r$. Кроме того, пусть

$$P_r = \mathbf{P}\{\nu_r^{(i)}(\lambda) > r\} \tag{21}$$

и $\nu_r(\lambda) = \nu_r^{(1)}(\lambda) + \dots + \nu_r^{(N)}(\lambda)$. В [6] доказано, что из (19) вытекает следующее утверждение.
Лемма 2. *При таких* n, *что* $\mathbf{P}\{\nu(\lambda) = N + n\} > 0$

$$\mathbf{P}\{\nu_{\max} \leqslant r\} = (1 - P_r)^N \frac{\mathbf{P}\{\nu_r(\lambda) = N + n\}}{\mathbf{P}\{\nu(\lambda) = N + n\}}.$$

Идея доказательства теоремы 1 основана на использовании леммы 2. Для этого в следующем разделе рассматривается асимптотика NP_r, а предельные распределения $\nu(\lambda)$ и $\nu_r(\lambda)$ найдены в разделах 5 и 6 соответственно.

4. Асимптотика NP_r

Рассмотрим предельное поведение вероятности (21).
Лемма 3. *Пусть выполнены условия теоремы* 1. *Тогда* $NP_r \to z$.

Доказательство. Подобное утверждение было доказано в лемме 5 статьи [1]. Но условия на поведение N, n и β изменились, поэтому здесь приводится полное доказательство с необходимыми изменениями.

Обозначим $\xi_1(\lambda), \xi_2(\lambda), \dots$ независимые случайные величины, распределения которых одинаковы и совпадают с (4). Пусть $\zeta_S(\lambda) = \xi_1(\lambda) + \dots + \xi_S(\lambda)$. В [8, гл. II, §1, лемма 3] показано, что при $k = 0, 1, 2, \dots$

$$\mathbf{P}\{\nu^{(1)}(\lambda) = r + k + 1\} = \frac{1}{r + k + 1} \mathbf{P}\{\zeta_{r+k+1}(\lambda) = r + k\}. \tag{22}$$

Пусть случайные величины $\xi_1, \xi_2, \dots$ независимы, их распределения одинаковы и заданы равенствами (3), а $\zeta_S = \xi_1 + \dots + \xi_S$. Из (3), (4) и (22) вытекает равенство

$$\mathbf{P}\{\nu^{(1)}(\lambda) = r + k + 1\} = \frac{\alpha^{r+k}}{(r + k + 1)F(\lambda)} \mathbf{P}\{\zeta_{r+k+1} = r + k\}, \tag{23}$$

где, напомним, $\alpha = \lambda / F(\lambda)$. В [1, соотношение (4.6)] доказано, что при $s \to \infty$

$$\mathbf{P}\{\zeta_{s+1} = s\} \sim \frac{g(0)}{B_s}. \tag{24}$$

Используя (5), (21), (23), (24), равенство $\beta = -\ln \alpha$ и соотношения $\lambda \to 1, F(\lambda) \to 1$, находим, что

$$\begin{aligned} P_r &= \frac{1}{F(\lambda)} \sum_{k=0}^{\infty} \frac{\alpha^{r+k}}{r + k + 1} \mathbf{P}\{\zeta_{r+k+1} = r + k\} = (g(0) + o(1)) \sum_{k=r}^{\infty} \frac{1}{(k+1)B_k} e^{-k\beta} \\ &\sim g(0) \int_r^{\infty} \frac{1}{xB_{[x]}} e^{-x\beta} dx \sim g(0) \int_1^{\infty} \frac{1}{yB_{[yr]}} e^{-yr\beta} dy = \frac{g(0)}{B_r} \int_1^{\infty} \frac{B_r}{yB_{[yr]}} e^{-yr\beta} dy. \end{aligned} \tag{25}$$

В силу (12) при $r \to \infty$

$$\frac{B_r}{B_{[yr]}} \to \frac{1}{y^{1/(\tau-1)}}$$

равномерно внутри любого конечного интервала $[1, K]$. Поэтому

$$\int_1^K \frac{B_r}{yB_{[yr]}} e^{-yr\beta} dy \sim \int_1^K \frac{1}{y^{\tau/(\tau-1)}} e^{-yr\beta} dy = (r\beta)^{1/(\tau-1)} \int_{r\beta}^{Kr\beta} \frac{1}{u^{\tau/(\tau-1)}} e^{-u} du. \quad (26)$$

Далее нам понадобится хорошо известное свойство медленно меняющихся функций, состоящее в том, что для любого $\delta > 0$ и достаточно больших x

$$x^{-\delta} < h(x) < x^{\delta}. \quad (27)$$

Заметим, что $r\beta \to \infty$. Действительно, если бы это было не так, то из (12), (13), (27) следовало бы, что для любого сколь угодно малого $\delta > 0$

$$N \leqslant C_6 r^{\frac{1+\delta}{\tau-1}}. \quad (28)$$

С другой стороны, справедливо соотношение

$$\frac{r}{\sqrt{n}} \to 0, \quad (29)$$

иначе, в силу (9) и (12), условие (13) не выполнялось бы. Из (10), (11) видим, что $N^2/n \to \infty$, а это противоречит (28). Согласно лемме 5 из [6, гл. II, § 2] при $x \to \infty$ и любом фиксированном d

$$\int_x^\infty y^d e^{-y} dy = x^d e^{-x}(1 + o(1)),$$

поэтому из (26) получаем, что

$$\int_1^K \frac{B_r}{yB_{[ry]}} e^{-yr\beta} dy \sim \frac{1}{r\beta} e^{-r\beta}. \quad (30)$$

Легко видеть, что

$$\int_K^\infty \frac{B_r}{yB_{[ry]}} e^{-yr\beta} dy \leqslant \frac{C_7}{K} \int_K^\infty e^{-yr\beta} dy = \frac{C_7}{Kr\beta} e^{-Kr\beta}. \quad (31)$$

Поскольку число K можно выбрать сколь угодно большим, из (25), (30), (31) находим, что

$$P_r \sim \frac{g(0)}{r\beta e^{r\beta} B_r}.$$

Отсюда и из (13) следует утверждение леммы 3. □

5. Предельное распределение $\nu(\lambda)$

Обозначим $m = \mathbf{E}\nu^{(1)}(\lambda)$, $\sigma^2 = \mathbf{D}\nu^{(1)}(\lambda)$. Пусть $\psi(t)$ обозначает характеристическую функцию случайной величины $\frac{\nu(\lambda) - Nm}{\sigma\sqrt{N}}$. Покажем, что распределение этой случайной величины слабо сходится к стандартному нормальному закону.

Лемма 4. *Пусть* $N, n \to \infty$ *так, что* $n/N \to \infty$ *и выполнено условие* (11). *Тогда* $\psi(t) \to e^{-\frac{t^2}{2}}$ *при любом фиксированном* t.

Доказательство. Обозначим $\varphi(t)$ характеристическую функцию случайной величины $\nu^{(1)}(\lambda)$. Следуя обычному методу доказательства слабой сходимости, заметим, что при достаточно малых t справедливо разложение

$$\ln \varphi(t) = itm - \frac{\sigma^2 t^2}{2} + \frac{t^3}{6} Q(t), \tag{32}$$

где

$$|Q(t)| \leqslant 2 \max_{|s| \leqslant |t|} |(\ln \varphi(s))'''_s|. \tag{33}$$

В [1] показано, что

$$m = \frac{1}{1-a}, \quad \sigma^2 = \frac{b^2}{(1-a)^3}, \tag{34}$$

где $a = \mathbf{E}\xi(\lambda)$, $b^2 = \mathbf{D}\xi(\lambda)$. Из (6) следует, что $a \to 1$, поэтому из (32)–(34) находим, что $\sigma\sqrt{N} \to \infty$ и при достаточно больших N, n и любом фиксированном t

$$\ln \psi(t) = -\frac{t^2}{2} + \frac{t^3}{6\sigma^3\sqrt{N}} Q\left(\frac{t}{\sigma\sqrt{N}}\right). \tag{35}$$

Осталось показать, что последнее слагаемое в (35) стремится к нулю.

Из оценки (4.13) статьи [2], полученной с учетом (3) и (27), следует, что

$$|(\ln \varphi(t))'''_t| \leqslant C_8 (1-\lambda)^{\tau-4-\delta} \left(\frac{n}{N}\right)^4 \max\left(1, (1-\lambda)^{\tau-2-\delta} \frac{n}{N}\right). \tag{36}$$

Кроме того, в этой же статье показано (см. (4.6)), что

$$\sigma^2 > \frac{C_9}{(1-\lambda)^{3-\tau-\delta}} \left(\frac{n}{N}\right)^3. \tag{37}$$

Отсюда и из (33), (36) находим:

$$\frac{1}{\sigma^3\sqrt{N}} \left| Q\left(\frac{t}{\sigma\sqrt{N}}\right) \right| < C_{10} \frac{\max(1, (1-\lambda)^{\tau-2-\delta} n/N)}{\sqrt{n(1-\lambda)^{\tau-1+5\delta}}}. \tag{38}$$

Отсюда и из (10), (11) получаем, что при $\delta \leqslant \delta_0/7$ правая часть (38) стремится к нулю, что и завершает доказательство леммы 4. $\square$

Покажем теперь, что на самом деле имеет место локальная сходимость распределения $\frac{\nu(\lambda)-N-n}{\sigma\sqrt{N}}$ к нормальному закону.

Лемма 5. *Пусть выполнены условия леммы* 4. *Тогда для натуральных* l *равномерно относительно* $u = \frac{l-N-n}{\sigma\sqrt{N}}$ *в любом фиксированном конечном интервале*

$$\mathbf{P}\{\nu(\lambda) = l\} = \frac{1+o(1)}{\sigma\sqrt{2\pi N}} e^{-\frac{u^2}{2}}.$$

Доказательство. Будем следовать стандартной схеме доказательства локальных предельных теорем. Представим вероятность $\mathbf{P}\{\nu(\lambda)=l\}$ по формуле обращения в виде интеграла:

$$\mathbf{P}\{\nu(\lambda)=l\}=(2\pi\sigma\sqrt{N})^{-1}\int_{-\pi\sigma\sqrt{N}}^{\pi\sigma\sqrt{N}} e^{-iut}\psi(t)\,dt. \tag{39}$$

Поскольку

$$(2\pi)^{-\frac{1}{2}}e^{-\frac{u^2}{2}}=(2\pi)^{-1}\int_{-\infty}^{\infty}\exp\{-iut-t^2/2\}\,dt, \tag{40}$$

из (39), (40) следует, что разность

$$R=2\pi\Big(\sigma\sqrt{N}\mathbf{P}\{\nu(\lambda)=l\}-(2\pi)^{-\frac{1}{2}}e^{-\frac{u^2}{2}}\Big)$$

можно представить в виде суммы $R=I_1+I_2+I_3+I_4+I_5$, где

$$\begin{gathered} I_1=\int_{-A}^{A} e^{-iut}\big(\psi(t)-e^{-\frac{t^2}{2}}\big)\,dt,\quad I_2=\int_{A<|t|\leqslant\varepsilon\sqrt{L}} e^{-iut}\psi(t)\,dt,\\ I_3=\int_{\varepsilon\sqrt{L}<|t|\leqslant\varepsilon\sigma\sqrt{N}} e^{-iut}\psi(t)\,dt,\quad I_4=\int_{\varepsilon\sigma\sqrt{N}<|t|\leqslant\pi\sigma\sqrt{N}} e^{-iut}\psi(t)\,dt,\\ I_5=-\int_{A<|t|}\exp\{-iut-t^2/2\}\,dt, \end{gathered} \tag{41}$$

а положительные постоянные A и ε будут выбраны позднее.

Разумеется, лемма будет доказана, если выбором достаточно больших N, n, A и достаточно малого ε разность R можно сделать сколь угодно малой. Для этого оценим интегралы I_1-I_5.

Очевидно, что при любом фиксированном A в силу леммы 4 справедливо соотношение $I_1\to 0$. Далее,

$$|I_5|\leqslant\int_{A<|t|} e^{-\frac{t^2}{2}}\,dt, \tag{42}$$

и этот интеграл тем меньше, чем больше A.

Оценим интеграл I_4. Хорошо известно, что при $\varepsilon<|t|\leqslant\pi$ справедливо неравенство

$$|\psi(t)|\leqslant e^{-C_{11}}, \tag{43}$$

поэтому

$$|I_4|\leqslant 2\pi\sigma\sqrt{N}e^{-C_{11}N}. \tag{44}$$

Неравенство (37) получено в [2] прямыми вычислениями с использованием (3)–(5), при этом при больших k значения $h(k+1)$ оценивались снизу с помощью (27). Повторяя эти рассуждения и применяя оценку сверху значений медленно меняющейся функции, нетрудно показать, что

$$\sigma^2<\frac{C_{12}}{(1-\lambda)^{3-\tau+\delta}}\Big(\frac{n}{N}\Big)^3. \tag{45}$$

Из (10), (11) следует, что

$$\sigma^2 < C_{13}N^3(1-\lambda)^{6-2\tau+2\delta},$$

поэтому из (44) следует, что при достаточно малых δ

$$|I_4| \leqslant C_{14}N^2(1-\lambda)^{3-\tau+\delta}e^{-C_{11}N} \to 0. \tag{46}$$

Из (10), (35) и (38) получаем, что в области интегрирования I_2 при достаточно малом ε

$$|\psi(t)| \leqslant e^{-C_{15}t^2},$$

поэтому

$$|I_2| \leqslant \int_{A<|t|} e^{-C_{15}t^2}\,dt.$$

Понятно, что последний интеграл можно сделать сколь угодно малым выбором достаточно большого A.

Осталось рассмотреть I_3. Для получения необходимой оценки воспользуемся идеей доказательства леммы 4 из [8, гл. 2, § 4]. Обозначим $F_\lambda(x)$ производящую функцию распределения (4):

$$F_\lambda(x) = \sum_{k=0}^{\infty} p_k(\lambda)x^k. \tag{47}$$

Поскольку $\lambda < 1$, функция $F_\lambda(x)$ имеет все производные и при $|x| \leqslant 1$ в окрестности единицы она может быть представлена следующим образом:

$$F_\lambda(x) = F_\lambda(1) + F'_\lambda(1)(x-1) + \frac{\widetilde{Q}(1)}{2}(x-1)^2, \tag{48}$$

где

$$|\widetilde{Q}(1)| \leqslant 2\max_{|s|\leqslant 1}|F''_\lambda(s)|. \tag{49}$$

Обозначим $U(x)$ производящую функцию распределения случайной величины $\nu^{(1)}(\lambda)$. Пусть $z = \exp\bigl\{\frac{it}{\sigma\sqrt{N}}\bigr\}$, тогда, очевидно, $U(z) = \varphi\bigl(\frac{t}{\sigma\sqrt{N}}\bigr)$. Хорошо известно (см. лемму 1 из [8, гл. 2, § 4]), что

$$U(z) = zF_\lambda(U(z)). \tag{50}$$

Очевидно, что $|z| = 1$, $|U(z)| \leqslant 1$. Положим $x = U(z)$. Тогда из (48)–(50) следует, что

$$U(z) = z\biggl(1 + F'_\lambda(z)(U(z)-1) + \frac{\widetilde{Q}(1)}{2}(U(z)-1)^2\biggr). \tag{51}$$

Решая это квадратное уравнение относительно $U(z)$ и выбирая ветвь корня так, чтобы $0 \leqslant U(z) \leqslant 1$ для действительных z, получим, что

$$U(z) = 1 - \sqrt{\frac{2(1-z)}{\widetilde{Q}(1)}}. \tag{52}$$

Учитывая, что

$$\sqrt{-2it} = \sqrt{|t|}\left(1 - i\frac{t}{|t|}\right),$$

из (52) видим, что в области интегрирования I_3

$$\left|\varphi\left(\frac{t}{\sigma\sqrt{N}}\right)\right| \leqslant \exp\left\{-C_{16}\sqrt{\frac{|t|}{\sigma\sqrt{N}|\widetilde{Q}(1)|}}\right\}. \tag{53}$$

Из (3), (4), (27) и (47) нетрудно получить, что при λ, достаточно близких к единице,

$$C_{17}\Phi(\lambda, \tau - 2 + \delta, 1) < F_\lambda^{''}(1) < C_{17}\Phi(\lambda, \tau - 2 - \delta, 1), \tag{54}$$

где $\Phi(\lambda, s, v)$ — трансцендентная функция Лерча (см. [10, п. 1.11]), имеющая вид

$$\Phi(\lambda, s, v) = \sum_{k=0}^{\infty}\frac{\lambda^k}{(k+v)^s}. \tag{55}$$

Функция (55) разлагается в ряд следующим образом [10, п. 1.11, формула (8)]):

$$\Phi(\lambda, s, v) = \frac{\Gamma(1-s)}{\lambda^v}\left(\ln\frac{1}{\lambda}\right)^{s-1} + \frac{1}{\lambda^v}\sum_{k=0}^{\infty}\zeta(s-k, \lambda)\frac{\ln^k\lambda}{k!},$$

где $|\ln\lambda| < 2\pi$, $\lambda \notin \{0, -1, -2, \dots\}$, $s \notin \{1, 2, \dots\}$, а $\zeta(s-k, \lambda)$ — значения обобщенной дзета-функции. Отсюда и из (3), (4), (54), (55) находим, что при λ, достаточно близких к единице,

$$C_{18}(1-\lambda)^{\tau-3+\delta} < F_\lambda^{''}(1) < C_{18}(1-\lambda)^{\tau-3-\delta}. \tag{56}$$

Собрав (45), (49), (53) и (56), получаем, что при достаточно малых δ

$$\left|\varphi\left(\frac{t}{\sigma\sqrt{N}}\right)\right| \leqslant \exp\left\{-C_{18}|t|^{\frac{\delta}{2}}\sqrt{\frac{NL^{\frac{1-\delta}{2}}}{n\sqrt{n}}(1-\lambda)^{\frac{9-3\tau+3\delta}{2}}}\right\}. \tag{57}$$

Предположим, что $L = n(1-\lambda)^{\tau-1+\delta}$. Тогда из (57) следует, что

$$\left|\varphi\left(\frac{t}{\sigma\sqrt{N}}\right)\right| \leqslant \exp\left\{-C_{18}|t|^{\frac{\delta}{2}}\sqrt{\frac{N}{n^{1+\frac{\delta}{2}}}(1-\lambda)^{4-\tau+2\delta}}\right\}.$$

Из этого неравенства, (10) и (11) находим, что при достаточно малых δ

$$|I_3| \leqslant \int_{\varepsilon\sqrt{L}}^{\infty} e^{-C_{19}|t|^{\frac{\delta}{2}}}\,dt \to 0. \tag{58}$$

Пусть $L = N^2(1-\lambda)^{3-\tau+\delta}/n$. Тогда, согласно (57),

$$\left|\varphi\left(\frac{t}{\sigma\sqrt{N}}\right)\right| \leqslant \exp\left\{-C_{18}|t|^{\frac{\delta}{2}}\sqrt{\frac{N^{2-\delta}}{n^{2-\frac{\delta}{2}}}(1-\lambda)^{6-2\tau+2\delta-\frac{\delta^2}{2}}}\right\},$$

откуда опять приходим к (58), что и завершает доказательство леммы 5. □

6. Предельные распределения $\nu_r(\lambda)$

Пусть $\Psi_r(t)$ означает характеристическую функцию случайной величины $\frac{\nu_r(\lambda)-N-n}{\sigma\sqrt{N}}$. Покажем, что распределение этой случайной величины слабо сходится к стандартному нормальному закону.

Лемма 6. *Пусть выполнены условия теоремы* 1. *Тогда* $\psi_r(t) \to e^{-\frac{t^2}{2}}$ *при любом фиксированном* t.

Доказательство. Будем следовать доказательству леммы 8 из [1]. Обозначим $\varphi_r(t)$ характеристическую функцию случайной величины $\nu_r^{(1)}(\lambda)$. Из (20) и (21) вытекает, что

$$\varphi_r(t) = (1-P_r)^{-1}\biggl(\varphi(t) - \sum_{k=1}^{\infty}\mathbf{P}\{\nu^{(1)}(\lambda) = r+k\}e^{it(r+k)}\biggr). \tag{59}$$

Тогда

$$\Psi_r(t) = (1-P_r)^{-N}\exp\biggl\{-\frac{i(N+n)t}{\sigma\sqrt{N}}\biggr\} \times\biggl(\varphi\biggl(\frac{t}{\sigma\sqrt{N}}\biggr) - \sum_{k=1}^{\infty}\mathbf{P}\{\nu^{(1)}(\lambda) = r+k\}\exp\biggl\{\frac{it(r+k)}{\sigma\sqrt{N}}\biggr\}\biggr)^N. \tag{60}$$

Из леммы 4 и (60) находим, что

$$\Psi_r(t) = e^{-\frac{t^2}{2}}(1-P_r)^{-N}(1+o(1)) \times\biggl(1-(1+o(1))\sum_{k=1}^{\infty}\mathbf{P}\{\nu^{(1)}(\lambda) = r+k\}\exp\biggl\{\frac{it(r+k)}{\sigma\sqrt{N}}\biggr\}\biggr)^N. \tag{61}$$

Поскольку $|e^{ix}-1| \leqslant |x|$,

$$\sum_{k=1}^{\infty}\mathbf{P}\{\nu^{(1)}(\lambda) = r+k\}\exp\biggl\{\frac{it(r+k)}{\sigma\sqrt{N}}\biggr\} = P_r + R(t), \tag{62}$$

где

$$|R(t)| \leqslant \frac{|t|}{\sigma\sqrt{N}}\sum_{k=1}^{\infty}(r+k)\mathbf{P}\{\nu^{(1)}(\lambda) = r+k\}. \tag{63}$$

Из (61)–(63) видно, что для доказательства леммы 6 достаточно установить справедливость соотношения

$$(\sigma\sqrt{N})^{-1}\sum_{k=1}^{\infty}(r+k)\mathbf{P}\{\nu^{(1)}(\lambda) = r+k\} = o(N^{-1}). \tag{64}$$

Аналогично тому, как было получено соотношение (25), с помощью (24) находим, что

$$\sum_{k=1}^{\infty}(r+k)\mathbf{P}\{\nu^{(1)}(\lambda)=r+k\}=\frac{1}{F(\lambda)}\sum_{k=1}^{\infty}\frac{r+k}{r+k+1}\alpha^{r+k}\mathbf{P}\{\zeta_{r+k+1}=r+k\}$$
$$=\frac{g(0)+o(1)}{F(\lambda)B_r}\sum_{k=1}^{\infty}e^{-(r+k)\beta}=\frac{g(0)+o(1)}{B_r}\sum_{k=r+1}^{\infty}e^{-k\beta}.$$

Поэтому

$$\sum_{k=1}^{\infty}(r+k)\mathbf{P}\{\nu^{(1)}(\lambda)=r+k\}<\frac{C_{20}}{B_r}\int_r^{\infty}e^{-y\beta}dy=\frac{C_{20}e^{-r\beta}}{\beta B_r}.$$

Таким образом, согласно (64), осталось показать, что

$$\frac{N}{\sigma\sqrt{N}\beta e^{r\beta}B_r}\to 0. \tag{65}$$

Как видно из (13), $N\leqslant C_{21}r\beta e^{r\beta}B_r$, поэтому, используя (37), получаем

$$\frac{N}{\sigma\sqrt{N}\beta e^{r\beta}B_r}<C_{22}\frac{rN(1-\lambda)^{\frac{3-\tau-\delta}{2}}}{n\sqrt{n}}.$$

Отсюда и из (29) следует (65). Лемма 6 доказана. □

В следующей лемме установлена локальная сходимость.

Лемма 7. *Пусть выполнены условия теоремы* 1. *Тогда для натуральных l равномерно относительно $u=\frac{l-N-n}{\sigma\sqrt{N}}$ в любом фиксированном конечном интервале*

$$\mathbf{P}\{\nu_r(\lambda)=l\}=\frac{1+o(1)}{\sigma\sqrt{2\pi N}}e^{-\frac{u^2}{2}}.$$

Доказательство. Лемма 7 доказывается аналогично лемме 5. По формуле обращения

$$\mathbf{P}\{\nu_r(\lambda)=l\}=\frac{1}{2\pi\sigma\sqrt{N}}\int_{-\pi\sigma\sqrt{N}}^{\pi\sigma\sqrt{N}}e^{-iut}\Psi_r(t)\,dt. \tag{66}$$

Полагая

$$R_r=2\pi\Big(\sigma\sqrt{N}\mathbf{P}\{\nu_r(\lambda)=l\}-(2\pi)^{-\frac{1}{2}}e^{-\frac{u^2}{2}}\Big)$$

и используя (40), (66), запишем равенство

$$R_r=I_{1,r}+I_{2,r}+I_{3,r}+I_{4,r}+I_5, \tag{67}$$

где интегралы $I_{1,r}-I_{4,r}$ отличаются от I_1-I_4 в (41) только заменой $\Psi(t)$ на $\Psi_r(t)$, а I_5 остается без изменений.

Оценим предельное поведение интегралов, образующих сумму (67). Из леммы 6 следует, что $I_{1,r}\to 0$ при любом фиксированном A. Для оценки I_5 используем (42). Для $|\varphi_r(t)|$ при $\varepsilon<|t|\leqslant\pi$ справедлива оценка, подобная (43), поэтому для $I_{4,r}$ можно написать оценку вида (44) и повторить рассуждения, которые приводят, как и в (46), к соотношению $I_{4,r} \to 0$. Из (59) и леммы 3 следует, что в областях интегрирования $I_{2,r}$ и $I_{3,r}$ при достаточно малом ε

$$\left|\varphi_r\left(\frac{t}{\sigma\sqrt{N}}\right)\right| \leqslant \left(1 + O\left(\frac{1}{N}\right)\right)\left|\varphi\left(\frac{t}{\sigma\sqrt{N}}\right)\right|. \tag{68}$$

Из (10), (11), (32), (33) и (38) находим, что если $A < |t| \leqslant \varepsilon\sqrt{L}$, то

$$\left|\varphi\left(\frac{t}{\sigma\sqrt{N}}\right)\right| \leqslant e^{-C_{23}\frac{t^2}{N}},$$

поэтому (68) влечет за собой оценку

$$|I_{2,r}| \leqslant C_{24}\int_A^\infty e^{-C_{23}t^2}\,dt,$$

а последний интеграл можно сделать сколь угодно малым, выбирая достаточно большое A. Наконец, интеграл $I_{3,r}$ также оценивается с помощью (57) и (68), что дает оценку, аналогичную (58). Лемма 7 доказана. □

7. Доказательство теоремы

Теорема 1 доказывается с помощью леммы 2. Из лемм 5 и 7 следует, что

$$\frac{\mathbf{P}\{\nu_r(\lambda) = N + n\}}{\mathbf{P}\{\nu(\lambda) = N + n\}} \to 1, \tag{69}$$

а из леммы 3 получаем, что

$$(1 - P_r)^N \to e^{-z}.$$

Отсюда, из (69) и леммы 2 следует утверждение теоремы.

Список литературы

1. Павлов Ю. Л., “Максимальное дерево случайного леса в конфигурационном графе”, *Матем. сб.*, **212**:9 (2021), 146–163.
2. Павлов Ю. Л., Чеплюкова И. А., “Объемы деревьев случайного леса и конфигурационные графы”, *Труды МИАН*, **316** (2022), 298–315.
3. Hofstad R., *Random Graphs and Complex Networks*, Vol.1, Cambridge University Press, Cambridge, 2017, xvi+321 c., Vol.2, https://www.win.tue.nl/∼rhofstad/NotesRGCNII.pdf.
4. Bollobas B., “A probabilistic proof of an asymptotic formula for the number of labelled regular graphs”, *Eur. J. Comb.*, **1**:4 (1980), 311–316.
5. Reittu H., Norros I., “On the power-law random graph model of massive data networks”, *Performance Evalation*, **55**:1-2 (2004), 3–23.
6. Павлов Ю. Л., *Случайные леса*, Петрозаводск: Карельский научный центр РАН, 1996, 259 с.
7. Казимиров Н. И., Павлов Ю. Л., “Одно замечание о лесах Гальтона – Ватсона”, *Дискретная математика*, **12**:1 (2000), 47–59.
8. Колчин В. Ф., *Случайные отображения*, М.: Наука, 1984, 207 с.
9. Золотарев В. М., *Одномерные устойчивые распределения*, М.: Наука, 1983, 304 с.
10. Бейтман Г., Эрдейи А., *Высшие трансцендентные функции. Гипергеометрическая функция. Функции Лежандра*, М.: Наука, 1965, 296 с.

Статья поступила 05.06.2022.