Труды Петрозаводского государственного университета
Серия “Математика” Выпуск 4, 1997

УДК 519.2

ПРЕДЕЛЬНЫЕ РАСПРЕДЕЛЕНИЯ НАИБОЛЬШИХ ДЕРЕВЬЕВ В СЛУЧАЙНОМ ЛЕСЕ

И. А. Чеплюкова

В статье рассматриваются случайные леса, состоящие из N корневых и n некорневых вершин, с достаточно общими ограничениями на классы деревьев. Для таких лесов получены предельные распределения наибольших деревьев для некоторых зон изменения параметров N и n.

Рассмотрим класс лесов $\mathcal{F}_{N,n}$, состоящих из N корневых деревьев с n некорневыми помеченными вершинами. В [1] показано, что такому множеству можно поставить в соответствие множество траекторий, начинающихся с N частиц ветвящегося процесса Гальтона—Ватсона G. Каждой конкретной траектории $G_{N,n}$ процесса G соответствует класс лесов $\Delta_{\mathcal{F}}(G_{N,n})$, отличающихся друг от друга только перенумерацией вершин. Рассмотрим такое множество $\mathcal{F}_{N,n}$, для которого выполняется соотношение

$$\mathsf{P}\{G = G_{N,n} | \nu = N + n\} = \frac{|\Delta_{\mathcal{F}}(G_{N,n})|}{|\mathcal{F}_{N,n}|},$$

где $|\Delta_{\mathcal{F}}(G_{N,n})|$ и $|\mathcal{F}_{N,n}|$ — число лесов в множествах $\Delta_{\mathcal{F}}(G_{N,n})$ и$\mathcal{F}_{N,n}$ соответственно, ν — общее число частиц, существовавших в процессе G до его вырождения. Обозначим $\nu_1(\mathcal{F}), \ldots, \nu_N(\mathcal{F})$ случайные величины, равные объемам деревьев из $\mathcal{F}_{N,n}$, имеющих корневые вершины с номерами $1, \ldots, N$.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект 97-01-00065.

Введем вспомогательную случайную величину ξ, имеющую дискретное распределение

$$\mathsf{P}\{\xi = k\} = p_k, \qquad k = 0, 1, 2, \ldots, \tag{1}$$

с максимальным шагом d и производящей функцией

$$F(z) = \sum_{k=0}^{\infty} p_k z^k, \tag{2}$$

где множество значений ξ, имеющих ненулевую вероятность, содержит нуль и не совпадает с множеством $\{0,1\}$. Пусть $M\xi = 1$, $D\xi = B$ и $F'''(1) < \infty$. Пусть число прямых потомков одной частицы ветвящегося процесса G имеет распределение

$$p_k(\lambda) = \lambda^k p_k / F(\lambda), \qquad k = 0, 1, 2, \ldots, \tag{3}$$

где $0 < \lambda \leq 1$.

Введем случайные величины $\nu^{(1)}, \ldots, \nu^{(N)}$, равные числу частиц, существовавших соответственно в процессах $G^{(1)}, \ldots, G^{(N)}$ до их вырождения. Как показано в [1], для всех n, удовлетворяющих условию $\mathsf{P}\{\nu = N + n\} > 0$, справедливо соотношение

$$\begin{aligned} &\mathsf{P}\{\nu_1(\mathcal{F}) = k_1, \ldots, \nu_N(\mathcal{F}) = k_N\} = \\ &= \mathsf{P}\{\nu^{(1)} = k_1, \ldots, \nu^{(N)} = k_N | \nu = N + n\}. \end{aligned} \tag{4}$$

Пусть $\nu_{(1)}(\mathcal{F}), \ldots, \nu_{(N)}(\mathcal{F})$ — вариационный ряд объемов деревьев в лесе, полученный расположением $\nu_1(\mathcal{F}), \ldots, \nu_N(\mathcal{F})$ в неубывающем порядке. В [1] получены предельные распределения максимального объема дерева $\nu_{(N)}(\mathcal{F})$ при $N, n \to \infty$ и всех соотношениях между N и n. В настоящей работе получены предельные распределения $\nu_{(N-j)}(\mathcal{F})$ при $N, n \to \infty$ так, что $n/N^2 \to 0, j = 0, 1, 2, \ldots$

Введем вспомогательные независимые одинаково распределенные случайные величины $\nu_r^{(1)}, \ldots, \nu_r^{(N)}$ и $\tilde{\nu}_r^{(1)}, \ldots, \tilde{\nu}_r^{(N)}$, для которых

$$\mathsf{P}\{\nu_r^{(i)} = k\} = \mathsf{P}\{\nu^{(i)} = k | \nu^{(i)} \leq r + 1\}, \tag{5}$$

$$\mathsf{P}\{\tilde{\nu}_r^{(i)} = k\} = \mathsf{P}\{\nu^{(i)} = k | \nu^{(i)} > r + 1\}, \ i = 1, \ldots, N, k = 1, 2, \ldots \tag{6}$$

Обозначим также

$$\nu_{r,j} = \nu_r^{(1)} + \cdots + \nu_r^{(N-j)} + \tilde{\nu}_r^{(N-j+1)} + \cdots + \tilde{\nu}_r^{(N)}, \quad j = 0, 1, 2, \ldots,$$

$$\mathsf{P}_r = \mathsf{P}\{\nu^{(1)} > r+1\},$$

тогда

$$\mathsf{P}_r = \sum_{k=1}^{\infty} \mathsf{P}\{\nu^{(1)} = r + kd + 1\}. \tag{7}$$

Аналогично лемме 1.2.2 из книги [1] из (4) и соотношения

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) \le r+1\} =$$
$$= \sum_{i=0}^{j} \binom{N}{i} \mathsf{P}\{\nu_{(1)}(\mathcal{F}) \le r+1, \ldots, \nu_{(N-i)}(\mathcal{F}) \le r+1,$$
$$\nu_{(N-i+1)}(\mathcal{F}) > r+1, \ldots, \nu_{(N)}(\mathcal{F}) > r+1\},$$

нетрудно получить следующую лемму.

Лемма 1. *При таких n, для которых $\mathsf{P}\{\nu = N+n\} > 0$, справедливо равенство*

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) \le r+1\} =$$
$$= \textstyle\sum_{i=0}^{j} (1-\mathsf{P}_r)^{N-i} \mathsf{P}_r^i \frac{N(N-1)\cdots(N-i+1)}{i!} \; \frac{\mathsf{P}\{\nu_{r,i}=n+N\}}{\mathsf{P}\{\nu=n+N\}}.$$

Лемма 1 показывает, что для получения предельных распределений случайных величин $\nu_{(N-j)}(\mathcal{F})$ достаточно изучить асимптотику $\mathsf{P}\{\nu_{r,i} = N+n\}$, поскольку результаты исследования асимптотического поведения бинома $(1-\mathsf{P}_r)^{N-i}$ и вероятностей P_r и $\mathsf{P}\{\nu = n+N\}$ получены Павловым Ю. Л. в книге [1]. С помощью этих результатов будут доказаны следующие теоремы.

Теорема 1. *Пусть $N, n \to \infty$ так, что n пробегает значения, кратные d, $n/N \to 0$, а параметр $\lambda = \lambda(N,n)$ определяется соотношением*

$$\frac{\lambda F'(\lambda)}{F(\lambda)} = \frac{n}{N+n}. \tag{8}$$

Пусть $N\mathsf{P}\{\nu^{(1)} = r+1\} \to \infty$, $N\mathsf{P}\{\nu^{(1)} = r+s+1\} \to \gamma$, где γ — некоторая неотрицательная постоянная, а натуральные r и s кратны d и удовлетворяют одному из условий:

1) $r \to \infty, s = d$

2) r — *фиксировано*, $r \ge m+1, \mathsf{P}\{\nu^{(1)} = r+s+1\} > 0, \mathsf{P}\{\nu^{(1)} =$

$= r + i + 1\} = 0$, *где* m — *наименьшее целое положительное число, такое, что* $p_m > 0$ *и* $0 < i < s$.
Тогда

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) = r + 1\} = e^{-\gamma} \sum_{i=0}^{j} \frac{\gamma^i}{i!},$$

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) = r + s + 1\} = 1 - e^{-\gamma} \sum_{i=0}^{j} \frac{\gamma^i}{i!}.$$

Теорема 2. *Пусть* $N, n \to \infty$ *так, что* n *пробегает значения, кратные* d, $n/N \to b$, *где* b — *некоторая положительная постоянная. Пусть параметр* $\lambda = \lambda(N, n)$ *определяется соотношением (8),*

$$\alpha = (\lambda_b / F(\lambda_b))^d, \tag{9}$$

где λ_b — *решение уравнения* $\dfrac{\lambda F'(\lambda)}{F(\lambda)} = \dfrac{b}{b+1}$. *Если* $r = r(N, n)$ *пробегает значения, кратные* d, *и такие, что*

$$\frac{N}{F(\lambda)} \left(\frac{\lambda}{F(\lambda)} \right)^r \frac{d}{r^{3/2}\sqrt{2\pi B}} \to \gamma, \tag{10}$$

где γ — *положительная постоянная, то для* $k = 0, \pm 1, \pm 2, \ldots$

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) \le r + kd + 1\} \to \exp\left\{ -\gamma \frac{\alpha^{k+1}}{1-\alpha} \right\} \sum_{i=0}^{j} \left(\frac{\gamma \alpha^{k+1}}{1-\alpha} \right)^i \frac{1}{i!}.$$

Теорема 3. *Пусть* $N, n \to \infty$ *так, что* n *пробегает значения, кратные* d, $n/N \to \infty, n/N^2 \to 0$. *Пусть параметр* $\lambda = \lambda(N, n)$ *определяется соотношением (8). Тогда*

$$\mathsf{P}\{\beta \nu_{(N-j)}(\mathcal{F}) - u \le z\} \to e^{-e^{-z}} \sum_{i=0}^{j} \frac{e^{-iz}}{i!},$$

где

$$\beta = \beta(\lambda) = -\ln(\lambda / F(\lambda)), \tag{11}$$

а $u = u(\lambda)$ *выбрано так, что*

$$N \beta^{1/2} u^{-3/2} e^{-u} = \sqrt{2\pi B}. \tag{12}$$

Ниже доказываются леммы 2 и 3 о предельном поведении суммы $\nu_{r,t}$, а затем с их помощью будут доказаны теоремы 1–3.

Пусть $\varphi_{r,t}(u)$ означает характеристическую функцию случайной величины $(\nu_{r,t} - N - n)/(\sigma\sqrt{N})$, где $t = 1, \ldots, j$. Тогда справедливо следующее утверждение.

Лемма 2. *Пусть $N, n \to \infty, n/N^2 \to 0$, параметр $\lambda = \lambda(N, n)$ определяется соотношением (8), а $r = r(N, n)$ пробегает значения, кратные d, и такие, что $N\mathsf{P}_r \to \gamma$, где γ — некоторая положительная постоянная. Тогда равномерно относительно u в любом конечном интервале*

$$\varphi_{r,t}(u) \to \exp\{-u^2/2\}.$$

Доказательство. Лемма 2 доказывается аналогично лемме 2.4.1 книги [1]. Обозначим через $\varphi_{\tilde{\nu}_r^{(1)}}(u), \varphi_{\nu_r^{(1)}}(u), \varphi(u)$ характеристические функции случайных величин $\tilde{\nu}_r^{(1)}, \nu_r^{(1)}, \nu^{(1)}$ соответственно. Нетрудно получить, что

$$\varphi_{\tilde{\nu}_r^{(1)}}(u) = \mathsf{P}_r^{-1} \sum_{k=1}^{\infty} \mathsf{P}\{\nu^{(1)} = r + kd + 1\} \exp\{iu(r + kd + 1)\},$$

$$\varphi_{\nu_r^{(1)}}(u) = (1 - \mathsf{P}_r)^{-1} \Big(\varphi(u) -$$

$$- \sum_{k=1}^{\infty} \mathsf{P}\{\nu^{(1)} = r + kd + 1\} \exp\{iu(r + kd + 1)\}\Big).$$

Тогда

$$\varphi_{r,t}(u) = \frac{1}{(1 - \mathsf{P}_r)^{N-t}} \Big(\varphi\Big(\frac{u}{\sigma\sqrt{N}}\Big) -$$

$$- \sum_{k=1}^{\infty} \mathsf{P}\{\nu^{(1)} = r + kd + 1\} \exp\{iu\frac{r + kd + 1}{\sigma\sqrt{N}}\}\Big)^{N-t} \times$$

$$\times \mathsf{P}_r^{-t} \Big(\sum_{k=1}^{\infty} \mathsf{P}\{\nu^{(1)} = r + kd + 1\} \exp\{iu(r + kd + 1)/(\sigma\sqrt{N})\}\Big)^{t} \times$$

$$\times \exp\{-iu(N + n)/(\sigma\sqrt{N})\}.$$

Согласно лемме 2.4.1 книги [1], при выполнении условий леммы справедливо соотношение

$$\varphi_{r,t}(u) = (1 + o(1))e^{-u^2/2}\mathsf{P}_r^{-1} \times$$

$$\times\left(\sum_{k=1}^{\infty}\mathsf{P}\{\nu^{(1)}=r+kd+1\}\exp\{iu(r+kd+1)/(\sigma\sqrt{N})\}\right)^{t}. \quad (13)$$

Нетрудно показать, что

$$\sum_{k=1}^{\infty}\mathsf{P}\{\nu^{(1)}=r+kd+1\}\exp\{iu(r+kd+1)/(\sigma\sqrt{N})\}=\mathsf{P}_r+Q(u), \quad (14)$$

где

$$|Q(u)|\le|u|(\sigma\sqrt{N})^{-1}\sum_{k=1}^{\infty}(r+kd+1)\mathsf{P}\{\nu^{(1)}=r+kd+1\}. \quad (15)$$

В [1] показано, что

$$(\sigma\sqrt{N})^{-1}\sum_{k=1}^{\infty}(r+kd+1)\mathsf{P}\{\nu^{(1)}=r+kd+1\}=o(N^{-1}) \quad (16)$$

Тогда из (13)–(16) следует утверждение леммы 2.

Лемма 2 утверждает, что распределение случайной величины $\tilde{\nu}_{r,t}$ слабо сходится к нормальному закону с параметрами $(N+n,\sigma\sqrt{N})$. Докажем локальное сближение этих распределений.

Лемма 3. *Пусть выполнены условия леммы 2 и, кроме того, пусть $N\lambda^{m+l}\to\infty$, где l — натуральное, не кратное m, удовлетворяющее условию $p_{m+l}>0$, если такого l нет, то $l=0$. Тогда для целых неотрицательных h, кратных d, равномерно относительно $(h-n)/(\sigma\sqrt{N})$ в любом конечном интервале*

$$\mathsf{P}\{\nu_{r,t}=N+h\}=\frac{d(1+o(1))}{\sigma\sqrt{2\pi N}}\exp\left\{-\frac{(h-n)^2}{2\sigma^2N}\right\}, \quad t=1,\ldots,j.$$

Доказательство. Используя формулу обращения, представим вероятность $\mathsf{P}\{\nu_{r,t}=N+h\}$ в виде интеграла

$$\mathsf{P}\{\nu_{r,t}=N+h\}=\frac{d}{2\pi\sigma\sqrt{N}}\int\limits_{-\pi\sigma\sqrt{N}/d}^{\pi\sigma\sqrt{N}/d}\exp\{-iuz\}\varphi_{r,t}(u)du,$$

где $z = (h-n)/(\sigma\sqrt{N})$. Известно, что

$$\frac{1}{\sqrt{2\pi}}e^{-z^2/2} = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-ixz-x^2/2}dx.$$

Рассмотрим разность

$$R = 2\pi\left(\frac{\sigma\sqrt{N}}{d}\mathsf{P}\{\nu_{r,t} = N+h\} - \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z^2}{2}\right\}\right).$$

Разность R можно представить в виде суммы четырех интегралов: $R = I_1 + I_2 + I_3 + I_4$, где

$$I_1 = \int\limits_{-A}^{A} e^{-iuz}(\varphi_{r,t}(u) - e^{-u^2/2})du,$$

$$I_2 = \int\limits_{A<|u|\le\epsilon\sigma\sqrt{N}} e^{-iuz}\varphi_{r,t}(u)du,$$

$$I_3 = \int\limits_{\epsilon\sigma\sqrt{N}<|u|\le d^{-1}\pi\sigma\sqrt{N}} e^{-iuz}\varphi_{r,t}(u)du,$$

$$I_4 = -\int\limits_{A<|u|} \exp\{-iuz - u^2/2\}du,$$

A и ϵ - положительные постоянные.

В силу леммы 2 интеграл $I_1 \to 0$. Оценки интегралов $I_2 - I_4$ аналогичны оценкам интегралов $I_2 - I_4$ при доказательстве леммы 2.4.2 книги [1], где показано, что $I_2 \to 0, I_3 \to 0, I_4 \to 0$, что и доказывает лемму 3.

Докажем теперь теоремы 1–3. Если выполнены условия теоремы 1, то, согласно лемме 2.2.2 книги [1],

$$N\mathsf{P}_{r-d} \to \infty,\ N\mathsf{P}_{r+k} \to \gamma, \qquad k = 0, d, 2d, \dots, s-d. \tag{17}$$

Тогда

$$(1 - \mathsf{P}_{r-d})^{N-i} \to 0, \tag{18}$$

$$(1 - \mathsf{P}_{r+k})^{N-i} \to e^{-\gamma}, \qquad i = 0, 1, \ldots, j. \tag{19}$$

При доказательстве теоремы 2.1.1 книги [1] показано, что

$$\mathsf{P}_{r+s} = \mathsf{P}\{\nu^{(1)} = r + s + l + 1\}(1 + o(1))$$

и

$$N\mathsf{P}\{\nu^{(1)} = r + s + l + 1\} \to 0.$$

Тогда

$$(1 - \mathsf{P}_{r+s})^{N-i} \to 1. \tag{20}$$

Из лемм 1, 3, леммы 2.3.2 книги [1] и соотношений (17)–(20) следует, что

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) \leq r - d + 1\} \to 0,$$

$$\mathsf{P}\{\nu_{(N-j)}(\mathcal{F}) \leq r + k + 1\} \to e^{-\gamma} + \gamma e^{-\gamma} + \gamma^2 e^{-\gamma}/2! + \cdots + \gamma^j e^{-\gamma}/j!,$$

что и доказывает теорему 1.

Если выполнены условия теоремы 2, то, согласно лемме 2.2.3 книги [1], имеем для фиксированных целых k

$$N\mathsf{P}_{r+kd} \to \gamma\alpha^{k+1}(1 - \alpha)^{-1}.$$

Тогда

$$(1 - \mathsf{P}_{r+kd})^{N-i} \to \exp\{-\gamma\alpha^{k+1}(1 - \alpha)^{-1}\}, i = 0, \ldots, j.$$

Отсюда из лемм 1, 3 и леммы 2.3.2 книги [1] следует утверждение теоремы 2. □

Пусть выполнены условия теоремы 3. Согласно лемме 2.2.5 книги [1], при r, пробегающих кратные d значения так, что $r = (u + z)/\beta + O(1), z$ — фиксированное число, справедливо

$$N\mathsf{P}_r \to e^{-z}.$$

Тогда

$$(1 - \mathsf{P}_r)^{N-i} \to e^{-e^{-z}}.$$

Отсюда из лемм 1, 3 и леммы 2.3.2 книги [1] получаем утверждение теоремы 3. □

Résumé

A random forests consisting of N rooted trees and n non-rooted vertices with sufficient common restrictions to the classes of trees are considered. For this random forests the limit distributions of the biggest trees have been got for various domains of variations N and n.

Литература

[1] Павлов Ю. Л. Случайные леса/ КНЦ РАН. Петрозаводск, 1996.