УДК 519.175.4

РАСПРЕДЕЛЕНИЕ МАКСИМАЛЬНЫХ K-СТЕПЕНЕЙ БИНОМИАЛЬНОГО СЛУЧАЙНОГО ГРАФА

И.В. Родионов

Институт проблем управления им. В.А. Трапезникова РАН

Россия, 117997, Москва, Профсоюзная ул., 65

E-mail: vecsell@gmail.com

М.Е. Жуковский

Московский физико-технический институт (государственный университет)

Россия, 141701, Московская область, г. Долгопрудный, Институтский пер., 9

E-mail: zhukmax@gmail.com

Ключевые слова: случайный граф, k-степень, распределение Гумбеля, неравенство Янсона.

Аннотация: В настоящей работе мы доказываем, что для максимального числа Δ_n общих соседей k вершин случайного графа $G(n,p)$ найдутся такие функции a_n, σ_n, что $\frac{\Delta_n - a_n}{\sigma_n}$ сходится по распределению к случайной величине, имеющей стандартное распределение Гумбеля.

В 1980 г. [1] Б. Боллобаш исследовал асимптотическое поведение максимальной степени Δ_n биномиального случайного графа $G(n,p)$ ([2–5]) для постоянного $p \in (0,1)$. Основной результат этой статьи следующий. Для $n \in \mathbb{N}$, $n \geqslant 2$, положим

$$a_n = pn + \sqrt{2p(1-p)n\ln n}\left(1 - \frac{\ln\ln n}{4\ln n} - \frac{\ln(2\sqrt{\pi})}{2\ln n}\right), \tag{1}$$

$$\sigma_n = \sqrt{\frac{p(1-p)n}{2\ln n}}. \tag{2}$$

Тогда $\frac{\Delta_n - a_n}{\sigma_n}$ сходится по распределению к случайной величине, имеющей стандартное распределение Гумбеля с функцией распределения $F(x) = \exp(-e^{-x})$. Кроме того, в этой статье Боллобаш получил аналогичный результат для m-ой степени (здесь степени упорядочены в порядке невозрастания) Δ_n^m (в частности, $\Delta_n^1 = \Delta_n$): для любого $y \in \mathbb{R}$

$$\mathsf{P}\left(\frac{\Delta_n^m - a_n}{\sigma_n} \leqslant y\right) \to e^{-e^{-y}} \sum_{k=0}^{m-1} \frac{e^{-ky}}{k!} \text{ при } n \to \infty.$$

Заметим, что степени ξ_i вершин $G(n,p)$ имеют биномиальное распределение с параметрами $n-1, p$. Для независимых биномиальных случайных величин распределение максимума изучено в работе [6]. А именно доказано, что если случайные величины $\xi_1, \ldots, \xi_n$ независимы и имеют биномиальное распределение с параметрами $N(n) = \omega([\ln n]^3)$ и $p = \mathrm{const}$, то их максимум D_n подчиняется следующему асимптотическому закону: для каждого $y \in \mathbb{R}$

$$(3) \qquad \lim_{n\to\infty} \mathsf{P}\left(D_n \leqslant pN + \sqrt{2p(1-p)N\ln n}\left[1 - \frac{\ln\ln n}{4\ln n} - \frac{\ln(2\sqrt{\pi})}{2\ln n} + \frac{y}{2\ln n}\right]\right) \to e^{-e^{-y}}.$$

Легко заметить, что при $N = n-1$ приведенные здесь функции a_n и σ_n совпадают с соответствующими функциями в теореме Боллобаша (простая подстановка дает несколько другие функции, но сходимость сохранится и для “оригинальных” функций из случая независимых величин).

Результат для независимых случайных величин непосредственно следует из неравенства, оценивающего вероятность большого уклонения биномиальной случайной величины от ее математического ожидания и некоторых свойств стандартного нормального распределения. В случае степеней случайного графа Боллобаш применил метод моментов для получения результата. А именно, он рассмотрел случайную величину X, равную количеству вершин со степенями, превосходящими $y\sigma_n + a_n$, где a_n и σ_n определены в (1) и (2). Результат следует из того, что для каждого $k \in \mathbb{N}$ k-ый факториальный момент величины X сходится к k-му факториальному моменту пуассоновской случайной величины с параметром e^{-y}. Очевидно, такой же подход может быть использован и для независимых биномиальных случайных величин, и для независимых одинаково распределенных величин с другими распределениями при некоторых условиях (см., например, [7], Chapter 2).

Обратимся к новым результатам. Пусть k — произвольное натуральное число. Пусть, кроме того, $\Delta_{k,n}$ — максимальное количество общих соседей k вершин в $G(n,p)$ (как мы заметили выше, случай $k=1$ уже изучен Боллобашем, так как $\Delta_{1,n} = \Delta_n$). Оказывается, что при $k \geqslant 2$ метод моментов не работает, так как дисперсия вспомогательной случайной величины стремится к бесконечности уже, например, при $k=2$ и $p > 1/2$. Введем исследуемые объекты более формально, прежде чем сформулировать наш основной результат. Пусть $v_1, \ldots, v_k \in \{1, \ldots, n\}$ — различные вершины $G(n,p)$, $N_n(v_1, \ldots, v_k)$ — множество их общих соседей. Рассмотрим последовательность $\Delta^1_{k,n}, \Delta^2_{k,n}, \ldots$ мощностей этих множеств, упорядоченных по невозрастанию.

Теорема 1. *Пусть* $y \in \mathbb{R}$*,*

$$a_{k,n} = np^k + \sqrt{2kp^k(1-p^k)n\ln n}\left(1 - \frac{\ln[k!]}{2k\ln n} - \frac{\ln[4\pi k\ln n]}{4k\ln n}\right),$$

$$\sigma_{k,n} = \sqrt{\frac{p^k(1-p^k)n}{2k\ln n}}.$$

Тогда

$$\mathsf{P}\left(\frac{\Delta^m_{k,n} - a_{k,n}}{\sigma_{k,n}} \leqslant y\right) \to e^{-e^{-y}} \sum_{i=0}^{m-1} \frac{e^{-yi}}{i!} \quad \textit{при } n \to \infty.$$

Заметим, что $\Delta_{k,n} = \Delta^1_{k,n}$ является максимальным членомпоследовательности $\binom{n}{k}$ биномиальных случайных величин с параметрами $n-k, p^k$. Поэтому для $N = \binom{n}{k}$ наш результат не дублирует утверждение (3) (в отличие от результата Боллобаша). Проиллюстрируем различие в зависимостях между случайными величинами в последовательностях с помощью сравнения случаев $k = 1$ и $k = 2$. Пусть $\xi_{N,p}$ — биномиальная случайная величина с параметрами N, p. Пусть, кроме того, X^k_n — количество k-вершинных множеств с более чем $b_{k,n}(y) := a_{k,n} + y\sigma_{k,n}$ общими соседями. Тогда

$$\mathsf{E}X^1_n = n\mathsf{P}\left(\xi_{n-1,p} > b_{1,n}(y)\right) \to e^{-y}, \quad \mathsf{E}X^1_n(X^1_n - 1) =$$
$$n(n-1)\left(p\left[\mathsf{P}\left(\xi_{n-1,p} > b_{1,n}(y) - 1\right)\right]^2 + (1-p)\left[\mathsf{P}\left(\xi_{n-1,p} > b_{1,n}(y)\right)\right]^2\right) \to e^{-2y}$$

(см. [1, 8]). При $p > 1/2$, используя аналогичную технику, можно доказать, что

$$\mathsf{E}X^2_n = \binom{n}{2}\mathsf{P}\left(\xi_{n-2,p^2} > b_{2,n}(y)\right) \to e^{-y},$$

$$\mathsf{E}X^2_n(X^2_n - 1) > n(n-1)(n-2)\sum_{i=b_{2,n}(y)+1}^{n-3}\mathsf{P}\left(\xi_{n-3,p} = i\right)\left[\mathsf{P}\left(\xi_{i,p} > b_{2,n}(y)\right)\right]^2 \to \infty$$

в отличие от первого случая.

Обсудим схему доказательства теоремы 1 для случая $m = 1$. Рассмотрим случайную величину X^k_n, равную количеству k-вершинных множеств с более чем $b_{k,n}(y) := a_{k,n} + y\sigma_{k,n}$ общими соседями. Положим $\overline{n} := \{1, \ldots, n\}$. Для следующих множеств $U = \{u_1, \ldots, u_k\} \in \binom{\overline{n}}{k}$ и $W \subset \overline{n} \setminus U$ рассмотрим событие $B_{U,W} = \{W = N_n(U)\}$. Положим $B_U = \bigvee_{W \subset \overline{n} \setminus U:\, |W| > b_{k,n}} B_{U,W}$. Тогда в силу следствия из теоремы Муавра–Лапласа (см. (3) в [8]) и асимптотики $1 - \Phi(x) \sim \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}$ (см. (1′) в [8]), где $\Phi(x) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt$, выполнено

$$\mathsf{P}(B_U) = \sum_{i > b_{k,n}}\binom{n-k}{i}(p^k)^i(1-p^k)^{n-k-i} \sim 1 - \Phi\left(\frac{b - (n-k)p^k}{\sqrt{(n-k)p^k(1-p^k)}}\right) =$$
$$1 - \Phi\left(\left[\sqrt{2k\ln n}\left(1 - \frac{\ln[k!]}{2k\ln n} - \frac{\ln[4\pi k\ln n]}{4k\ln n}\right) + \frac{y}{\sqrt{2k\ln n}}\right](1 + O(n^{-1}))\right)$$
$$\sim \frac{k!}{n^k}e^{-y}.$$

Причина, по которой $\mathsf{D}X^k_n$ может стремиться к бесконечности, состоит в том, что основной вклад в дисперсию дают множества, собственные подмножества которых имеют слишком много общих соседей (настолько много, что вероятность подобного вклада стремится к нулю). В этой связи мы рассмотрели случайную величину $\tilde{X}^k_n$, равную количеству таких множеств $U_k := \{u_1, \ldots, u_k\}$, что для каждого $\ell \in \{1, \ldots, k-1\}$ и любых различных $i_1, \ldots, i_\ell \in \{1, \ldots, k\}$

$$|N_n(u_{i_1}, \ldots, u_{i_\ell})| \leqslant \Gamma_\ell \text{ и } |N_n(u_1, \ldots, u_k)| > b_{k,n},$$

где $\Gamma_\ell = np^\ell + \sqrt{2\ell}\sqrt{np^\ell(1-p^\ell)\ln n}$. Несмотря на то, что это определение позволяет ограничить дисперсию, оценить все факториальные моменты такой случайной величины непросто. К счастью, нам удалось доказать, что асимптотического равенства

$\mathsf{E}\tilde{X}_n^k(\tilde{X}_n^k - 1) \sim (\mathsf{E}X_n^k)^2$ достаточно для получения нашего результата. Это наблюдение следует из доказанного нами варианта неравенства Янсона, о котором речь пойдет ниже. Следует отметить, что результат Боллобаша напрямую следовал бы из неравенства Янсона для произвольных возрастающих свойств ([9], Theorem 1, Inequality (3)), если бы для различных вершин u, v соседи вершины u не зависели бы от соседей вершины v. В таком случае, достаточно бы было найти асимптотику первого момента для нахождения асимптотики $\mathsf{P}(X_n^1 = 0)$. К сожалению, упомянутые объекты зависимы. Тем не менее наш вариант неравенства Янсона применим и в этом случае, то есть из него и того, что $\mathsf{E}X_n^1 \sim e^{-y}$, мгновенно следует результат Боллобаша.

Сформулируем это ключевое неравенство. Пусть $U = \{u_1, \ldots, u_k\} \in \binom{\overline{n}}{k}$, а множества $W_1, \ldots, W_k \subset \overline{n} \setminus U$. Положим

$$B_{U,W_1,\ldots,W_k} = \{W_1 = N_n(u_1), \ldots, W_k = N_n(u_k)\}.$$

Обозначим

$$\lambda = \sum_{U,W} \mathsf{P}(B_{U,W}), \quad \tilde{\lambda} = \sum_{U,W_1,\ldots,W_k} \mathsf{P}(B_{U,W_1,\ldots,W_k}),$$

$$\tilde{\Delta} = \sum_{U_1,U_2} \mathsf{P}(|N_n(U_1 \cap U_2)| \leqslant \Gamma_{|U_1 \cap U_2|}, |N_n(U_1)| > b_{k,n}, |N_n(U_2)| > b_{k,n}),$$

где первые два суммирования ведутся по всем $U \in \binom{\overline{n}}{k}$ и $W, W_1, \ldots, W_k \subset \overline{n} \setminus U$ таким, что $|W| > b_{k,n}$ и для каждого $\ell \in \{1, \ldots, k-1\}$ и всех различных $i_1, \ldots, i_\ell \in \{1, \ldots, k\}$, $|W_{i_1} \cap \ldots \cap W_{i_\ell}| \leqslant \Gamma_\ell$, а третье суммирование — по таким $(U_1, U_2) \in \binom{\binom{\overline{n}}{k}}{2}$, что множество $U_1 \cap U_2$ не пусто.

Лемма 1. *Справедливы оценки*

$$\exp\left[-\lambda + o(1)\right] \leqslant \mathsf{P}(X = 0) \leqslant \mathsf{P}(\tilde{X} = 0) \leqslant$$
$$\exp\left[-(1 + o(\max\{1, e^{\lambda}\}))\left(\tilde{\lambda} - e^{\lambda}\frac{\tilde{\Delta}}{2}\right)\right].$$

Разумеется, утверждение теоремы следует из этой леммы и технического утверждения, сформулированного ниже.

Лемма 2. *При* $n \to \infty$

$$\tilde{\lambda} \sim \lambda \sim e^{-y}, \quad \tilde{\Delta} \to 0.$$

Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (19-01-00090a).

Список литературы

1. Bollobás B. The distribution of the maximum degree of a random graph // Discrete Mathematics. 1980. Vol. 32. P. 201-203.

2. Жуковский М.Е., Райгородский А.М. Случайные графы: модели и асимптотические характеристики // Успехи математических наук. 2015. Т. 70, № 1. С. 35-88.
3. Alon N., Spencer J.H. The Probabilistic Method. Third Edition, John Wiley & Sons, 2008.
4. Bollobás B. Random Graphs / 2nd Edition, Cambridge University Press, 2001.
5. Janson S., Łuczak T., Rucinski A. Random Graphs. New York: Wiley, 2000.
6. Nadarajah S., Mitov K. Asymptotics of Maxima of Discrete Random Variables // Extremes. 2002. Vol. 5. P. 287-294.
7. Leadbetter M.R., Lindgren G., Rootzén H., Extremes and Related Properties of Random Sequences and Processes. Springer, 1983.
8. Bollobás B. Degree sequences of random graphs // Discrete Mathematics. 1981. Vol. 33. P. 1-19.
9. Riordan O., Warnke L. The Janson inequalities for general up-sets // Random Structures & Algorithms. 2015. Vol. 46, No. 2. P. 391-395.