Случайные графы. Лекция 21.10. Эволюция случайного графа. Часть 2. Теорема о гигантской компоненте

Д. А. Шабанов

кафедра дискретной математики

21.10.2025

III. Случай $np = c$, $c > 1$

Перейдем к обсуждению случая $np = c$, где c — константа, большая единицы. Мы уже знаем (следствие 3.4), что объем унициклических компонент ограничен, их предельное распределение — пуассоновское (теорема 3.4).

III. Случай $np = c$, $c > 1$

Перейдем к обсуждению случая $np = c$, где c — константа, большая единицы. Мы уже знаем (следствие 3.4), что объем унициклических компонент ограничен, их предельное распределение — пуассоновское (теорема 3.4). Что касается остальных компонент, то, оказывается, что при $c > 1$ наблюдается перколяционный эффект — в случайном графе $G(n, p)$ появляется гигантская компонента размера порядка n, причем такая компонента единственная. Тем самым, при крайне малом увеличении вероятности появлении ребра в графе часть маленьких древесных компонент слипаются в одну большую.

III. Случай $np = c$, $c > 1$

Перейдем к обсуждению случая $np = c$, где c — константа, большая единицы. Мы уже знаем (следствие 3.4), что объем унициклических компонент ограничен, их предельное распределение — пуассоновское (теорема 3.4). Что касается остальных компонент, то, оказывается, что при $c > 1$ наблюдается перколяционный эффект — в случайном графе $G(n, p)$ появляется гигантская компонента размера порядка n, причем такая компонента единственная. Тем самым, при крайне малом увеличении вероятности появлении ребра в графе часть маленьких древесных компонент слипаются в одну большую.

Для установления данной замечательной теоремы о фазовом переходе нам понадобятся некоторые сведения относительно ветвящихся процессов Гальтона–Ватсона.

Ветвящиеся процессы

Определение

Пусть ξ — случайная величина со значениями в $\mathbb{Z}_+$, а $\{\xi_k^{(n)}, k, n \in \mathbb{N}\}$ — независимые случайные величины с тем же распределением, что и ξ. Тогда ветвящимся процессом Гальтона–Ватсона называется случайный процесс $\{X_n, n \in \mathbb{Z}_+\}$, где

$$X_0 = 1, \;\; X_n = \sum_{k=1}^{X_{n-1}} \xi_k^{(n)}, n > 0.$$

Ветвящиеся процессы

Определение

Пусть ξ *— случайная величина со значениями в* $\mathbb{Z}_+$*, а* $\{\xi_k^{(n)}, k, n \in \mathbb{N}\}$ *— независимые случайные величины с тем же распределением, что и* ξ*. Тогда ветвящимся процессом Гальтона–Ватсона называется случайный процесс* $\{X_n, n \in \mathbb{Z}_+\}$*, где*

$$X_0 = 1, \;\; X_n = \sum_{k=1}^{X_{n-1}} \xi_k^{(n)}, n > 0.$$

Смысл модели предельно прост — в начальный момент времени имеется одна частица, а затем в каждый следующий момент времени частица порождает случайное число таких же частиц и умирает. Анализ ветвящихся процессов удобно производить с помощью производящих функций.

Производящие функции

Определение

Пусть ξ *— случайная величина. Тогда ее производящей функцией называется*

$$\varphi_\xi(z) = \mathsf{E} z^\xi.$$

Производящие функции

Определение

Пусть ξ — случайная величина. Тогда ее производящей функцией называется

$$\varphi_\xi(z) = \mathsf{E} z^\xi.$$

Если ξ принимает значения в $\mathbb{Z}_+$, то ее производящая функция является степенным рядом

$$\varphi_\xi(z) = \sum_{k=0}^{\infty} z^k \mathsf{P}(\xi = k).$$

Легко видеть, что данный степенной ряд сходится абсолютно и равномерно в области $\{|z| \leqslant 1\}$, и, тем самым, внутри круга это аналитическая функция.

Вероятность вырождения

Исторически первым содержательным результатом теории ветвящихся процессов была теорема о вероятности вырождения. Обозначим через

$$q = \mathsf{P}(\exists n : X_n = 0)$$

— вероятность вырождения ветвящегося процесса Гальтона–Ватсона. Тогда имеет место следующее простое утверждение.

Вероятность вырождения

Исторически первым содержательным результатом теории ветвящихся процессов была теорема о вероятности вырождения. Обозначим через

$$q = \mathsf{P}(\exists n : X_n = 0)$$

— вероятность вырождения ветвящегося процесса Гальтона–Ватсона. Тогда имеет место следующее простое утверждение.

Утверждение (3.1)

Пусть q — вероятность вырождения ветвящегося процесса Гальтона–Ватсона с законом размножения частиц ξ. Тогда

$$q = \varphi_\xi(q).$$

Доказательство утверждения 3.1

Пусть первая частица в процессе породила k потомков. Тогда для каждого потомка вероятность того, что ветвящийся процесс, начатый с него, выродится равна q и не зависит от процессов других частиц.

Доказательство утверждения 3.1

Пусть первая частица в процессе породила k потомков. Тогда для каждого потомка вероятность того, что ветвящийся процесс, начатый с него, выродится равна q и не зависит от процессов других частиц. Тогда получаем

$$q = \mathsf{P}(\exists n : X_n = 0) = \sum_{k=0}^{\infty} q^k \mathsf{P}(\xi = k) = \varphi_\xi(q).$$

$\square$

Теорема о вероятности вырождения

Однако понятно, что корней уравнения $z = \varphi_\xi(z)$ может быть несколько. Например, 1 — всегда решение. Кто же из них будет являться вероятностью вырождения? Ответ на этот вопрос дает следующая теорема, которую мы приводим без доказательства.

Теорема о вероятности вырождения

Однако понятно, что корней уравнения $z = \varphi_\xi(z)$ может быть несколько. Например, 1 — всегда решение. Кто же из них будет являться вероятностью вырождения? Ответ на этот вопрос дает следующая теорема, которую мы приводим без доказательства.

Теорема (3.6, о вероятности вырождения)

Пусть $\{X_n, n \in \mathbb{Z}_+\}$ — ветвящийся процесс Гальтона–Ватсона с законом размножения частиц ξ. Пусть $\mathsf{P}(\xi = 1) < 1$. Тогда
1) если $\mathsf{E}\xi \leqslant 1$, то на отрезке $[0, 1]$ существует только одно решение единица уравнения $z = \varphi_\xi(z)$. В этом случае $q = 1$.
2) если $\mathsf{E}\xi > 1$, то на полуинтервале $[0, 1)$ существует только одно решение z_0 уравнения $z = \varphi_\xi(z)$. В этом случае $q = z_0 < 1$.

Теорема о вероятности вырождения

Однако понятно, что корней уравнения $z = \varphi_\xi(z)$ может быть несколько. Например, 1 — всегда решение. Кто же из них будет являться вероятностью вырождения? Ответ на этот вопрос дает следующая теорема, которую мы приводим без доказательства.

Теорема (3.6, о вероятности вырождения)

Пусть $\{X_n, n \in \mathbb{Z}_+\}$ — ветвящийся процесс Гальтона–Ватсона с законом размножения частиц ξ. Пусть $\mathsf{P}(\xi = 1) < 1$. Тогда
1) если $\mathsf{E}\xi \leqslant 1$, то на отрезке $[0, 1]$ существует только одно решение единица уравнения $z = \varphi_\xi(z)$. В этом случае $q = 1$.
2) если $\mathsf{E}\xi > 1$, то на полуинтервале $[0, 1)$ существует только одно решение z_0 уравнения $z = \varphi_\xi(z)$. В этом случае $q = z_0 < 1$.

Пример: пусть $\xi \sim \mathrm{Pois}(c)$. Тогда $\varphi_\xi(z) = e^{c(z-1)}$, и, значит,

$$q = e^{c(q-1)}.$$

Если же обозначить $\beta = 1 - q$, то получим

$$\beta + e^{-c\beta} = 1.$$

Теорема о гигантской компоненте

Теперь мы готовы сформулировать и доказать теорему о гигантской компоненте в случайном графе.

Теорема (3.7, о гигантской компоненте)

Пусть $c > 1$ *и* $np = c$. *Положим* $\beta = \beta(c)$ *— решение уравнения* $\beta + e^{-c\beta} = 1$ *из* $(0,1)$. *Тогда с вероятностью, стремящейся к 1, случайный граф* $G(n,p)$ *содержит гигантскую компоненту, чей размер при делении на* n *сходится по вероятности к* β. *При этом размер второй наибольшей компоненты не превосходит* $\frac{16c}{(c-1)^2}\ln n$.

Теорема о гигантской компоненте

Теперь мы готовы сформулировать и доказать теорему о гигантской компоненте в случайном графе.

Теорема (3.7, о гигантской компоненте)

Пусть $c > 1$ *и* $np = c$. *Положим* $\beta = \beta(c)$ *— решение уравнения* $\beta + e^{-c\beta} = 1$ *из* $(0, 1)$. *Тогда с вероятностью, стремящейся к 1, случайный граф* $G(n, p)$ *содержит гигантскую компоненту, чей размер при делении на* n *сходится по вероятности к* β. *При этом размер второй наибольшей компоненты не превосходит* $\frac{16c}{(c-1)^2}\ln n$.

Замечание

Теорема будет выполнена и при условии $np \to c > 1$ *при* $n \to +\infty$.

Доказательство теоремы. Часть 1

Обозначим $k_- = \frac{16c}{(c-1)^2}\ln n$, $k_+ = n^{2/3}$.

Доказательство теоремы. Часть 1

Обозначим $k_- = \frac{16c}{(c-1)^2} \ln n$, $k_+ = n^{2/3}$. Для каждой вершины v случайного графа $G(n,p)$ проведем ту же процедуру выявления связной компоненты, что и в теореме для $c < 1$. Но в этот раз опишем ее чуть подробнее.

Доказательство теоремы. Часть 1

Обозначим $k_- = \frac{16c}{(c-1)^2}\ln n$, $k_+ = n^{2/3}$. Для каждой вершины v случайного графа $G(n,p)$ проведем ту же процедуру выявления связной компоненты, что и в теореме для $c < 1$. Но в этот раз опишем ее чуть подробнее.

В каждый момент времени $t = 0, 1, \ldots$ множество всех вершин графа разделено на три части — S_t (рассмотренные вершины), A_t (активные) и U_t (неактивные). В начальный момент времени $A_0 = v$, S_0 — пусто, $U_0 = V(G(n,p)) \setminus v$.

Доказательство теоремы. Часть 1

Обозначим $k_- = \frac{16c}{(c-1)^2}\ln n$, $k_+ = n^{2/3}$. Для каждой вершины v случайного графа $G(n,p)$ проведем ту же процедуру выявления связной компоненты, что и в теореме для $c<1$. Но в этот раз опишем ее чуть подробнее.

В каждый момент времени $t = 0, 1, \ldots$ множество всех вершин графа разделено на три части — S_t (рассмотренные вершины), A_t (активные) и U_t (неактивные). В начальный момент времени $A_0 = v$, S_0 — пусто, $U_0 = V(G(n,p)) \setminus v$. Далее, в каждый момент времени $t \geqslant 1$

- первая активная вершина v_t из A_t переходит в рассмотренные, $S_{t+1} = S_t \cup \{v_t\}$,

Доказательство теоремы. Часть 1

Обозначим $k_- = \frac{16c}{(c-1)^2} \ln n$, $k_+ = n^{2/3}$. Для каждой вершины v случайного графа $G(n,p)$ проведем ту же процедуру выявления связной компоненты, что и в теореме для $c < 1$. Но в этот раз опишем ее чуть подробнее.

В каждый момент времени $t = 0, 1, \dots$ множество всех вершин графа разделено на три части — S_t (рассмотренные вершины), A_t (активные) и U_t (неактивные). В начальный момент времени $A_0 = v$, S_0 — пусто, $U_0 = V(G(n,p)) \setminus v$. Далее, в каждый момент времени $t \geqslant 1$

- первая активная вершина v_t из A_t переходит в рассмотренные, $S_{t+1} = S_t \cup \{v_t\}$,
- все ее соседи среди неактивных вершин (обозначим это множество через X_t) переходят в активные, $A_{t+1} = A_t \setminus \{v_t\} \cup X_t$, $U_{t+1} = U_t \setminus X_t$.

Доказательство теоремы. Часть 1

Обозначим $k_- = \frac{16c}{(c-1)^2} \ln n$, $k_+ = n^{2/3}$. Для каждой вершины v случайного графа $G(n,p)$ проведем ту же процедуру выявления связной компоненты, что и в теореме для $c < 1$. Но в этот раз опишем ее чуть подробнее.

В каждый момент времени $t = 0, 1, \ldots$ множество всех вершин графа разделено на три части — S_t (рассмотренные вершины), A_t (активные) и U_t (неактивные). В начальный момент времени $A_0 = v$, S_0 — пусто, $U_0 = V(G(n,p)) \setminus v$. Далее, в каждый момент времени $t \geqslant 1$

- первая активная вершина v_t из A_t переходит в рассмотренные, $S_{t+1} = S_t \cup \{v_t\}$,
- все ее соседи среди неактивных вершин (обозначим это множество через X_t) переходят в активные, $A_{t+1} = A_t \setminus \{v_t\} \cup X_t$, $U_{t+1} = U_t \setminus X_t$.

Процесс набора компоненты останавливается, если в момент времени t нет активных вершин или U_t пусто.

Доказательство теоремы. Часть 1

Покажем сначала, что с вероятностью, стремящейся к 1, возможны только две альтернативы:
1) процедура набора компоненты закончится ко времени k_-,
2) для любого $t \in [k_-, k_+]$ в момент времени t выполнено $|A_t| \geqslant \left(\frac{c-1}{2}\right) t$.

Доказательство теоремы. Часть 1

Покажем сначала, что с вероятностью, стремящейся к 1, возможны только две альтернативы:
1) процедура набора компоненты закончится ко времени k_-,
2) для любого $t \in [k_-, k_+]$ в момент времени t выполнено $|A_t| \geqslant \left(\frac{c-1}{2}\right) t$.

Действительно, пусть 1) и 2) не выполнены. Тогда найдется такое $t \in [k_-, k_+]$, что $|A_t| < \left(\frac{c-1}{2}\right) t$ и $|A_{t'}| > 0$ для всех $t' \leqslant t$.

Доказательство теоремы. Часть 1

Покажем сначала, что с вероятностью, стремящейся к 1, возможны только две альтернативы:
1) процедура набора компоненты закончится ко времени k_-,
2) для любого $t \in [k_-, k_+]$ в момент времени t выполнено $|A_t| \geqslant \left(\frac{c-1}{2}\right) t$.

Действительно, пусть 1) и 2) не выполнены. Тогда найдется такое $t \in [k_-, k_+]$, что $|A_t| < \left(\frac{c-1}{2}\right) t$ и $|A_{t'}| > 0$ для всех $t' \leqslant t$. Заметим, что

$$|A_t| = \sum_{k=1}^{t} |X_k| - t + 1.$$

Доказательство теоремы. Часть 1

Покажем сначала, что с вероятностью, стремящейся к 1, возможны только две альтернативы:
1) процедура набора компоненты закончится ко времени k_-,
2) для любого $t \in [k_-, k_+]$ в момент времени t выполнено $|A_t| \geqslant \left(\frac{c-1}{2}\right) t$.

Действительно, пусть 1) и 2) не выполнены. Тогда найдется такое $t \in [k_-, k_+]$, что $|A_t| < \left(\frac{c-1}{2}\right) t$ и $|A_{t'}| > 0$ для всех $t' \leqslant t$. Заметим, что

$$|A_t| = \sum_{k=1}^{t} |X_k| - t + 1.$$

Тогда

$$\mathsf{P}\left(\sum_{k=1}^{t} |X_k| < \left(\frac{c-1}{2}\right) t + t - 1\right) \leqslant \mathsf{P}\left(\sum_{k=1}^{t} |X_k| < \left(\frac{c+1}{2}\right) t\right) \leqslant$$

Доказательство теоремы. Часть 1

(т.к. на каждом шаге у нас есть не менее $n - k_+(c+1)/2$ неактивных вершин, то $|X_k| \geqslant Y_k$, где $Y_1, Y_2, \ldots$ — независимые $\mathrm{Bin}(n - k_+(c+1)/2, p)$ случайные величины)

Доказательство теоремы. Часть 1

(т.к. на каждом шаге у нас есть не менее $n - k_+(c+1)/2$ неактивных вершин, то $|X_k| \geqslant Y_k$, где $Y_1, Y_2, \ldots$ — независимые $\mathrm{Bin}(n - k_+(c+1)/2, p)$ случайные величины)

$$\leqslant \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \left(\frac{c+1}{2}\right) t\right) = \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k + \frac{c+1}{2}t - \mathsf{E}\sum_{k=1}^{t} Y_k\right) =$$

Доказательство теоремы. Часть 1

(т.к. на каждом шаге у нас есть не менее $n - k_+(c+1)/2$ неактивных вершин, то $|X_k| \geqslant Y_k$, где $Y_1, Y_2, \ldots$ — независимые $\mathrm{Bin}(n - k_+(c+1)/2, p)$ случайные величины)

$$\leqslant \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \left(\frac{c+1}{2}\right) t\right) = \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k + \frac{c+1}{2}t - \mathsf{E}\sum_{k=1}^{t} Y_k\right) =$$

$$= \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k - \frac{c-1}{2}t + \frac{c+1}{2}k_+pt\right) \leqslant$$

Доказательство теоремы. Часть 1

(т.к. на каждом шаге у нас есть не менее $n - k_+(c+1)/2$ неактивных вершин, то $|X_k| \geqslant Y_k$, где $Y_1, Y_2, \ldots$ — независимые $\mathrm{Bin}(n - k_+(c+1)/2, p)$ случайные величины)

$$\leqslant \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \left(\frac{c+1}{2}\right) t\right) = \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k + \frac{c+1}{2}t - \mathsf{E}\sum_{k=1}^{t} Y_k\right) =$$

$$= \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k - \frac{c-1}{2}t + \frac{c+1}{2}k_+pt\right) \leqslant$$

(неравенство Чернова)

$$\leqslant \exp\left\{-\left(\frac{c-1}{2}t - \frac{c+1}{2}k_+pt\right)^2 / 2\left(ct - \frac{c+1}{2}k_+pt\right)\right\} =$$

Доказательство теоремы. Часть 1

(т.к. на каждом шаге у нас есть не менее $n - k_+(c+1)/2$ неактивных вершин, то $|X_k| \geqslant Y_k$, где $Y_1, Y_2, \ldots$ — независимые $\mathrm{Bin}(n - k_+(c+1)/2, p)$ случайные величины)

$$\leqslant \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \left(\frac{c+1}{2}\right) t\right) = \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k + \frac{c+1}{2}t - \mathsf{E}\sum_{k=1}^{t} Y_k\right) =$$

$$= \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k - \frac{c-1}{2}t + \frac{c+1}{2}k_+pt\right) \leqslant$$

(неравенство Чернова)

$$\leqslant \exp\left\{-\left(\frac{c-1}{2}t - \frac{c+1}{2}k_+pt\right)^2 /2\left(ct - \frac{c+1}{2}k_+pt\right)\right\} =$$

$$= \exp\left\{-\frac{(c-1)^2}{8c}t(1+o(1))\right\} \leqslant |\text{ т.к. } t \geqslant k_-| \leqslant n^{-2(1+o(1))}.$$

Доказательство теоремы. Часть 1

(т.к. на каждом шаге у нас есть не менее $n - k_+(c+1)/2$ неактивных вершин, то $|X_k| \geqslant Y_k$, где $Y_1, Y_2, \ldots$ — независимые $\mathrm{Bin}(n - k_+(c+1)/2, p)$ случайные величины)

$$\leqslant \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \left(\frac{c+1}{2}\right) t\right) = \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k + \frac{c+1}{2}t - \mathsf{E}\sum_{k=1}^{t} Y_k\right) =$$

$$= \mathsf{P}\left(\sum_{k=1}^{t} Y_k < \mathsf{E}\sum_{k=1}^{t} Y_k - \frac{c-1}{2}t + \frac{c+1}{2}k_+pt\right) \leqslant$$

(неравенство Чернова)

$$\leqslant \exp\left\{-\left(\frac{c-1}{2}t - \frac{c+1}{2}k_+pt\right)^2 /2\left(ct - \frac{c+1}{2}k_+pt\right)\right\} =$$

$$= \exp\left\{-\frac{(c-1)^2}{8c}t(1+o(1))\right\} \leqslant |\text{ т.к. } t \geqslant k_-| \leqslant n^{-2(1+o(1))}.$$

Перебирая по всем начальным вершинам и по всем $t \leqslant k_+$, получаем, что вероятность того, что хоть для одной из них не будет выполнено 1) или 2) не превосходит $n^{-1/3(1+o(1))} = o(1)$.

Доказательство теоремы. Часть 2

Итак, мы показали, что либо размер компоненты связности не превосходит k_-, либо он имеет порядок не менее k_+.

Доказательство теоремы. Часть 2

Итак, мы показали, что либо размер компоненты связности не превосходит k_-, либо он имеет порядок не менее k_+. Предположим, что две вершины v' и v'' лежат в "больших" компонентах. С какой вероятностью эти компоненты будут разными? Покажем, что с вероятностью, стремящейся к 1, это будет одна и та же компонента.

Доказательство теоремы. Часть 2

Итак, мы показали, что либо размер компоненты связности не превосходит k_-, либо он имеет порядок не менее k_+. Предположим, что две вершины v' и v'' лежат в "больших" компонентах. С какой вероятностью эти компоненты будут разными? Покажем, что с вероятностью, стремящейся к 1, это будет одна и та же компонента.

Запустим процессы набора компоненты связности из v' и v''. После k_+ шагов процедуры в каждом процессе имеется не менее $(c-1)k_+/2$ активных вершин.

Доказательство теоремы. Часть 2

Итак, мы показали, что либо размер компоненты связности не превосходит k_-, либо он имеет порядок не менее k_+. Предположим, что две вершины v' и v'' лежат в "больших" компонентах. С какой вероятностью эти компоненты будут разными? Покажем, что с вероятностью, стремящейся к 1, это будет одна и та же компонента.

Запустим процессы набора компоненты связности из v' и v''. После k_+ шагов процедуры в каждом процессе имеется не менее $(c-1)k_+/2$ активных вершин. Тогда условная вероятность того, что между этими множествами активных вершин не будет ребер в случайном графе $G(n,p)$ не превосходит

$$(1-p)^{\left(\frac{c-1}{2}k_+\right)^2} \leqslant \exp\left\{-\frac{(c-1)^2 c}{4} n^{1/3}\right\} = o\left(\frac{1}{n^2}\right).$$

Доказательство теоремы. Часть 2

Итак, мы показали, что либо размер компоненты связности не превосходит k_-, либо он имеет порядок не менее k_+. Предположим, что две вершины v' и v'' лежат в “больших” компонентах. С какой вероятностью эти компоненты будут разными? Покажем, что с вероятностью, стремящейся к 1, это будет одна и та же компонента.

Запустим процессы набора компоненты связности из v' и v''. После k_+ шагов процедуры в каждом процессе имеется не менее $(c-1)k_+/2$ активных вершин. Тогда условная вероятность того, что между этими множествами активных вершин не будет ребер в случайном графе $G(n,p)$ не превосходит

$$(1-p)^{\left(\frac{c-1}{2}k_+\right)^2} \leqslant \exp\left\{-\frac{(c-1)^2 c}{4} n^{1/3}\right\} = o\left(\frac{1}{n^2}\right).$$

Следовательно, вероятность того, что найдутся две вершины из разных больших компонент, стремится к нулю.

Доказательство теоремы. Часть 3

Подведем промежуточные итоги.

Доказательство теоремы. Часть 3

Подведем промежуточные итоги. В случайном графе у нас есть “маленькие” вершины, лежащие в компонентах размера не более $k_- = \frac{16c}{(c-1)^2} \ln n$, а также большие вершины из единственной компоненты размера больше, чем $k_+ = n^{2/3}$.

Доказательство теоремы. Часть 3

Подведем промежуточные итоги. В случайном графе у нас есть "маленькие" вершины, лежащие в компонентах размера не более $k_- = \frac{16c}{(c-1)^2}\ln n$, а также большие вершины из единственной компоненты размера больше, чем $k_+ = n^{2/3}$.

Пусть v — некоторая вершина $G(n,p)$. Тогда из процесса набора компоненты, содержащей v, сразу следует, что вероятность того, что v — маленькая, не превосходит $\rho(n-k_-,p)$ — вероятности вырождения ветвящегося процесса Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n-k_-,p)$:

$$\mathsf{P}\left(v \text{ — маленькая}\right) \leqslant \rho(n-k_-,p).$$

Доказательство теоремы. Часть 3

Подведем промежуточные итоги. В случайном графе у нас есть “маленькие” вершины, лежащие в компонентах размера не более $k_- = \frac{16c}{(c-1)^2} \ln n$, а также большие вершины из единственной компоненты размера больше, чем $k_+ = n^{2/3}$.

Пусть v — некоторая вершина $G(n,p)$. Тогда из процесса набора компоненты, содержащей v, сразу следует, что вероятность того, что v — маленькая, не превосходит $\rho(n-k_-, p)$ — вероятности вырождения ветвящегося процесса Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n-k_-, p)$:

$$\mathsf{P}\left(v \text{ — маленькая}\right) \leqslant \rho(n-k_-, p).$$

С другой стороны, $\mathsf{P}\left(v \text{ — маленькая}\right)$ не меньше, чем вероятность того, что ветвящийся процесс Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n,p)$ выродился и всего в процессе было не более k_- вершин.

Доказательство теоремы. Часть 3

Подведем промежуточные итоги. В случайном графе у нас есть “маленькие” вершины, лежащие в компонентах размера не более $k_- = \frac{16c}{(c-1)^2}\ln n$, а также большие вершины из единственной компоненты размера больше, чем $k_+ = n^{2/3}$.

Пусть v — некоторая вершина $G(n,p)$. Тогда из процесса набора компоненты, содержащей v, сразу следует, что вероятность того, что v — маленькая, не превосходит $\rho(n-k_-,p)$ — вероятности вырождения ветвящегося процесса Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n-k_-,p)$:

$$\mathsf{P}\left(v \text{ — маленькая}\right) \leqslant \rho(n-k_-,p).$$

С другой стороны, $\mathsf{P}\left(v \text{ — маленькая}\right)$ не меньше, чем вероятность того, что ветвящийся процесс Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n,p)$ выродился и всего в процессе было не более k_- вершин. В силу того, что k_- растет с ростом n, данная вероятность асимптотически эквивалентна $\rho(n,p)$ — вероятности вырождения процесса Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n,p)$.

Доказательство теоремы. Часть 3

Подведем промежуточные итоги. В случайном графе у нас есть "маленькие" вершины, лежащие в компонентах размера не более $k_- = \frac{16c}{(c-1)^2}\ln n$, а также большие вершины из единственной компоненты размера больше, чем $k_+ = n^{2/3}$.

Пусть v — некоторая вершина $G(n,p)$. Тогда из процесса набора компоненты, содержащей v, сразу следует, что вероятность того, что v — маленькая, не превосходит $\rho(n-k_-,p)$ — вероятности вырождения ветвящегося процесса Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n-k_-,p)$:

$$\mathsf{P}\left(v \text{ — маленькая}\right) \leqslant \rho(n-k_-,p).$$

С другой стороны, $\mathsf{P}\left(v \text{ — маленькая}\right)$ не меньше, чем вероятность того, что ветвящийся процесс Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n,p)$ выродился и всего в процессе было не более k_- вершин. В силу того, что k_- растет с ростом n, данная вероятность асимптотически эквивалентна $\rho(n,p)$ — вероятности вырождения процесса Гальтона–Ватсона с законом размножения $\mathrm{Bin}(n,p)$. Таким образом,

$$\rho(n,p)(1-o(1)) \leqslant \mathsf{P}\left(v \text{ — маленькая}\right) \leqslant \rho(n-k_-,p).$$

В силу теоремы Пуассона и $\rho(n,p)$, и $\rho(n-k_-,p)$ сходятся к $1-\beta$.

Доказательство теоремы. Часть 3

Обозначим через Z_n число маленьких вершин в случайном графе $G(n, p)$.

Доказательство теоремы. Часть 3

Обозначим через Z_n число маленьких вершин в случайном графе $G(n, p)$. Тогда в силу предыдущих рассуждений получаем, что

$$\mathsf{E} Z_n = n(1 - \beta)(1 + o(1)).$$

Доказательство теоремы. Часть 3

Обозначим через Z_n число маленьких вершин в случайном графе $G(n,p)$. Тогда в силу предыдущих рассуждений получаем, что

$$\mathsf{E}Z_n = n(1-\beta)(1+o(1)).$$

При фиксированной маленькой вершине v и множества ее соседей вероятность того, что вершина w из другой связной компоненте тоже является маленькой, не превосходит $\rho(n-2k_-,p)$. Отсюда

$$\mathsf{E}Z_n(Z_n-1) \leqslant n\rho(n-k_-,p)(k_- + n\rho(n-2k_-,p)) = (\mathsf{E}Z_n)^2(1+o(1)).$$

Доказательство теоремы. Часть 3

Обозначим через Z_n число маленьких вершин в случайном графе $G(n, p)$. Тогда в силу предыдущих рассуждений получаем, что

$$\mathsf{E}Z_n = n(1-\beta)(1+o(1)).$$

При фиксированной маленькой вершине v и множества ее соседей вероятность того, что вершина w из другой связной компоненте тоже является маленькой, не превосходит $\rho(n-2k_-, p)$. Отсюда

$$\mathsf{E}Z_n(Z_n-1) \leqslant n\rho(n-k_-, p)(k_- + n\rho(n-2k_-, p)) = (\mathsf{E}Z_n)^2(1+o(1)).$$

Применяя неравенство Чебышева, для любого фиксированного $\varepsilon > 0$ имеем

$$\mathsf{P}\left(\left|\frac{Z_n - \mathsf{E}Z_n}{n}\right| \geqslant \varepsilon\right) \leqslant \frac{\mathsf{D}Z_n}{n^2\varepsilon^2} = \frac{o((\mathsf{E}Z_n)^2)}{n^2\varepsilon^2} \longrightarrow 0.$$

Завершение доказательства

В итоге, размер гигантской компоненты $\mathcal{C}_1(G(n,p))$, равный $n - Z_n$ с вероятностью, стремящейся к 1, удовлетворяет следующему предельному соотношению:

$$\frac{|\mathcal{C}_1(G(n,p))|}{n} \xrightarrow{\mathsf{P}} \beta.$$

Теорема 3.7 доказана. $\square$

ЦПТ для $|\mathcal{C}_1(G(n,p))|$

Итак, мы выяснили, что в случае $np = c > 1$ в случайном графе $G(n,p)$ имеется единственная гигантская компонента $\mathcal{C}_1(G(n,p))$ размера порядка n, причем ее асимптотический размер — это βn, где $\beta + e^{-c\beta} = 1$.

ЦПТ для $|\mathcal{C}_1(G(n,p))|$

Итак, мы выяснили, что в случае $np = c > 1$ в случайном графе $G(n,p)$ имеется единственная гигантская компонента $\mathcal{C}_1(G(n,p))$ размера порядка n, причем ее асимптотический размер — это βn, где $\beta + e^{-c\beta} = 1$. В теореме 2 был доказан закон больших чисел для размера $\mathcal{C}_1(G(n,p))$. Однако, как было установлено В.Е. Степановым в 1970 году имеет место и более сильное утверждение относительно предельного распределения $|\mathcal{C}_1(G(n,p))|$, которое можно назвать центральной предельной теоремой для размера гигантской компоненты.

ЦПТ для $|\mathcal{C}_1(G(n,p))|$

Итак, мы выяснили, что в случае $np = c > 1$ в случайном графе $G(n,p)$ имеется единственная гигантская компонента $\mathcal{C}_1(G(n,p))$ размера порядка n, причем ее асимптотический размер — это βn, где $\beta + e^{-c\beta} = 1$. В теореме 2 был доказан закон больших чисел для размера $\mathcal{C}_1(G(n,p))$. Однако, как было установлено В.Е. Степановым в 1970 году имеет место и более сильное утверждение относительно предельного распределения $|\mathcal{C}_1(G(n,p))|$, которое можно назвать центральной предельной теоремой для размера гигантской компоненты.

Теорема (3.8)

Пусть $c > 1$ и $np = c$. Положим $\beta = \beta(c) \in (0,1)$ — решение уравнения $\beta + e^{-c\beta} = 1$ и обозначим через $|\mathcal{C}_1(G(n,p))|$ — размер гигантской компоненты случайного графа $G(n,p)$. Тогда

$$\sqrt{n}\left(\frac{|\mathcal{C}_1(G(n,p))|}{n} - \beta\right) \xrightarrow{d} \mathcal{N}(0,\sigma^2),$$

где $\sigma^2 = \frac{\beta(1-\beta)}{(1-c(1-\beta))^2}$.

k-ядро случайного графа

Определение

Пусть $G = (V, E)$ *— некоторый граф, а* $k \in \mathbb{N}$*. Тогда* k*-ядром графа* G *называется такой максимальный по размеру связный индуцированный подграф* H *графа* G*, что* $\delta(H) \geqslant k$*.*

k-ядро случайного графа

Определение

Пусть $G = (V, E)$ — некоторый граф, а $k \in \mathbb{N}$. Тогда k-ядром графа G называется такой максимальный по размеру связный индуцированный подграф H графа G, что $\delta(H) \geqslant k$.

Ясно, что 1-ядро — это просто наибольшая по размеру компонента связности. Посмотрим, что происходит с k-ядром случайного графа $G(n, c/n)$ для фиксированных $k > 1$ и $c > 1$. Будем обозначать его $\mathcal{C}_k(G(n, p))$. Начнем со следующей леммы о размере $\mathcal{C}_2(G(n, p))$.

k-ядро случайного графа

Определение

Пусть $G = (V, E)$ — некоторый граф, а $k \in \mathbb{N}$. Тогда k-ядром графа G называется такой максимальный по размеру связный индуцированный подграф H графа G, что $\delta(H) \geqslant k$.

Ясно, что 1-ядро — это просто наибольшая по размеру компонента связности. Посмотрим, что происходит с k-ядром случайного графа $G(n, c/n)$ для фиксированных $k > 1$ и $c > 1$. Будем обозначать его $\mathcal{C}_k(G(n,p))$. Начнем со следующей леммы о размере $\mathcal{C}_2(G(n,p))$.

Лемма (3.5)

Пусть $c > 1$, а $\beta = \beta(c) \in (0,1)$ — решение уравнения $\beta + e^{-c\beta} = 1$. Тогда размер $\mathcal{C}_2(G(n,p))$ при делении на n сходится по вероятности к $(1 - c(1-\beta))\beta$.

Доказательство леммы 3.5

Возьмем фиксированную вершину $v \in V(G(n,p))$ и оценим вероятность того, что она принадлежит $\mathcal{C}_2(G(n,p))$.

Доказательство леммы 3.5

Возьмем фиксированную вершину $v \in V(G(n,p))$ и оценим вероятность того, что она принадлежит $\mathcal{C}_2(G(n,p))$. Рассмотрим случайный граф G_1, индуцированный $G(n,p)$ на $V(G(n,p)) \setminus \{v\}$. Ясно, что его распределение есть в точности $G(n-1,p)$. Тогда по теореме о гигантской компоненте с вероятностью, стремящейся к 1, в G_1 есть гигантская компонента $\mathcal{C}_1(G_1)$ размера $\beta n(1+o(1))$, а все остальные имеют размер $O(\ln n)$.

Доказательство леммы 3.5

Возьмем фиксированную вершину $v \in V(G(n,p))$ и оценим вероятность того, что она принадлежит $\mathcal{C}_2(G(n,p))$. Рассмотрим случайный граф G_1, индуцированный $G(n,p)$ на $V(G(n,p)) \setminus \{v\}$. Ясно, что его распределение есть в точности $G(n-1,p)$. Тогда по теореме о гигантской компоненте с вероятностью, стремящейся к 1, в G_1 есть гигантская компонента $\mathcal{C}_1(G_1)$ размера $\beta n(1+o(1))$, а все остальные имеют размер $O(\ln n)$.

Мы знаем, что уницикличеcкие компоненты занимают $O_{\mathsf{P}}(1)$ вершин. Кроме того, нам понадобится следующее важное упражнение.

Упражнение

Пусть $np \to c$, $c > 1$. Докажите, что с вероятностью, стремящейся к 1, гигантская компонента является единственной сложной компонентой в случайном графе $G(n,p)$.

Доказательство леммы 3.5

Обозначим через $\mathcal{A}$ событие, состоящее в том, что

- случайный граф G_1 содержит гигантскую компоненту $\mathcal{C}_1(G_1)$ с размером $||\mathcal{C}_1(G_1)| - \beta n| \leqslant n^{1/2} \ln n$;
- остальные компоненты G_1 — это деревья и унициклические графы размера не более $a \cdot \ln n$;
- унициклические компоненты занимают не более $\ln n$ вершин.

Доказательство леммы 3.5

Обозначим через $\mathcal{A}$ событие, состоящее в том, что

- случайный граф G_1 содержит гигантскую компоненту $\mathcal{C}_1(G_1)$ с размером $||\mathcal{C}_1(G_1)| - \beta n| \leqslant n^{1/2} \ln n$;
- остальные компоненты G_1 — это деревья и унициклические графы размера не более $a \cdot \ln n$;
- унициклические компоненты занимают не более $\ln n$ вершин.

Мы знаем, что $\mathsf{P}(\mathcal{A}) \to 1$ при $n \to +\infty$. Рассмотрим возможных соседей v. Заметим, что

$$\mathsf{P}(v \text{ имеет соседей в унициклических компонентах } G_1 | \mathcal{A}) \leqslant$$

$$\leqslant \ln n \cdot p = O\left(\frac{\ln n}{n}\right) \longrightarrow 0.$$

Доказательство леммы 3.5

Обозначим через $\mathcal{A}$ событие, состоящее в том, что

- случайный граф G_1 содержит гигантскую компоненту $\mathcal{C}_1(G_1)$ с размером $||\mathcal{C}_1(G_1)| - \beta n| \leqslant n^{1/2} \ln n$;
- остальные компоненты G_1 — это деревья и унициклические графы размера не более $a \cdot \ln n$;
- унициклические компоненты занимают не более $\ln n$ вершин.

Мы знаем, что $\mathsf{P}(\mathcal{A}) \to 1$ при $n \to +\infty$. Рассмотрим возможных соседей v. Заметим, что

$$\mathsf{P}(v \text{ имеет соседей в унициклических компонентах } G_1 | \mathcal{A}) \leqslant$$

$$\leqslant \ln n \cdot p = O\left(\frac{\ln n}{n}\right) \longrightarrow 0.$$

Далее,

$$\mathsf{P}(v \text{ имеет хотя бы двух соседей в древесной компоненте } G_1 | \mathcal{A}) \leqslant$$

$$\leqslant n \cdot (a \ln n)^2 \cdot p^2 = O\left(\frac{\ln^2 n}{n}\right) \longrightarrow 0.$$

Доказательство леммы 3.5

Тем самым, с унициклическими компонентами мы не связаны, а в древесных имеем максимум одного соседа. Значит, если v будет иметь всего лишь одного соседа в $\mathcal{C}_1(G_1)$, то она не войдет в 2-ядро $G(n,p)$. Если же соседей будет хотя бы 2, то, наоборот, v заведомо будет там содержаться в силу связности $\mathcal{C}_1(G_1)$.

Доказательство леммы 3.5

Тем самым, с унициклическими компонентами мы не связаны, а в древесных имеем максимум одного соседа. Значит, если v будет иметь всего лишь одного соседа в $\mathcal{C}_1(G_1)$, то она не войдет в 2-ядро $G(n,p)$. Если же соседей будет хотя бы 2, то, наоборот, v заведомо будет там содержаться в силу связности $\mathcal{C}_1(G_1)$.

В итоге, при G_1, удовлетворяющем $\mathcal{A}$, выполнено

$\mathsf{P}(v \in \mathcal{C}_2(G(n,p))|G_1) = o(1) + \mathsf{P}(v$ имеет хотя бы 2-х соседей в $\mathcal{C}_1(G_1)|G_1) =$

$$= o(1) + 1 - (1-p)^{|\mathcal{C}_1(G_1)|} - |\mathcal{C}_1(G_1)|p\,(1-p)^{|\mathcal{C}_1(G_1)|-1} =$$

Тем самым, с унициклическими компонентами мы не связаны, а в древесных имеем максимум одного соседа. Значит, если v будет иметь всего лишь одного соседа в $\mathcal{C}_1(G_1)$, то она не войдет в 2-ядро $G(n,p)$. Если же соседей будет хотя бы 2, то, наоборот, v заведомо будет там содержаться в силу связности $\mathcal{C}_1(G_1)$.

В итоге, при G_1, удовлетворяющем $\mathcal{A}$, выполнено

$\mathsf{P}(v \in \mathcal{C}_2(G(n,p))|G_1) = o(1) + \mathsf{P}(v$ имеет хотя бы 2-х соседей в $\mathcal{C}_1(G_1)|G_1) =$

$$= o(1) + 1 - (1-p)^{|\mathcal{C}_1(G_1)|} - |\mathcal{C}_1(G_1)|p\,(1-p)^{|\mathcal{C}_1(G_1)|-1} =$$

(подставляем $p = c/n$, $\mathcal{C}_1(G_1) \sim \beta \cdot n$)

$$= o(1) + 1 - \left(1 - \frac{c}{n}\right)^{\beta n(1+o(1))} - \beta c \left(1 - \frac{c}{n}\right)^{\beta n(1+o(1))} \longrightarrow$$

$$\longrightarrow 1 - e^{-\beta c} - \beta c \cdot e^{-\beta c}.$$

Доказательство леммы 3.5

Вспомним, что $e^{-\beta c} = 1 - \beta$. В итоге, получаем, что

$$\mathsf{P}(v \in \mathcal{C}_2(G(n,p))) = (1 - c(1-\beta))\beta(1 + o(1)),$$

откуда следует, что

$$\mathsf{E}\left|\mathcal{C}_2(G(n,p))\right| = (1 - c(1-\beta))\beta \cdot n(1 + o(1)).$$

Доказательство леммы 3.5

Вспомним, что $e^{-\beta c} = 1 - \beta$. В итоге, получаем, что

$$\mathsf{P}(v \in \mathcal{C}_2(G(n,p))) = (1 - c(1-\beta))\beta(1 + o(1)),$$

откуда следует, что

$$\mathsf{E}\,|\mathcal{C}_2(G(n,p))| = (1 - c(1-\beta))\beta \cdot n(1 + o(1)).$$

Концентрация $|\mathcal{C}_2(G(n,p))|$ вокруг своего среднего значения доказывается с помощью рассмотрения второго факториального момента $|\mathcal{C}_2(G(n,p))|$ по аналогии с размером гигантской компоненты. $\square$

k-ядро случайного графа

При $k > 2$ ситуация с k-ядром оказывается куда более нетривиальной. Оказывается, есть пороговое значение параметра c, до которого такое ядро пусто, а после него — уже линейно по числу вершину!

k-ядро случайного графа

При $k > 2$ ситуация с k-ядром оказывается куда более нетривиальной. Оказывается, есть пороговое значение параметра c, до которого такое ядро пусто, а после него — уже линейно по числу вершину!

Чтобы сформулировать результат нам понадобится ряд обозначений. Пусть $Z = Z(\lambda)$ — это пуассоновская случайная величина $\mathrm{Pois}(\lambda)$. Обозначим:

$$p_k(\lambda) = \mathsf{P}(Z \geqslant k), \quad \pi_k(\lambda) = \mathsf{P}(Z \geqslant k-1).$$

k-ядро случайного графа

При $k > 2$ ситуация с k-ядром оказывается куда более нетривиальной. Оказывается, есть пороговое значение параметра c, до которого такое ядро пусто, а после него — уже линейно по числу вершину!

Чтобы сформулировать результат нам понадобится ряд обозначений. Пусть $Z = Z(\lambda)$ — это пуассоновская случайная величина $\mathrm{Pois}(\lambda)$. Обозначим:

$$p_k(\lambda) = \mathsf{P}(Z \geqslant k), \quad \pi_k(\lambda) = \mathsf{P}(Z \geqslant k-1).$$

Положим:

$$\gamma_k = \inf_{\lambda > 0} \frac{\lambda}{\pi_k(\lambda)}.$$

Далее, для любого $c > \gamma_k$ введем $\lambda_k(c)$ как решение уравнения (относительно λ)

$$c = \frac{\lambda}{\pi_k(\lambda)}.$$

Решений у уравнения, на самом деле, два и мы берем большее в качестве $\lambda_k(c)$.

Теорема о k-ядре

Теорема (3.9)

Пусть $np = c > 1$ и $k > 2$ — натуральное число.

1. *Если $c < \gamma_k$, то с вероятностью, стремящейся к 1, k-ядро случайного графа $G(n,p)$ пусто.*
2. *Если $c > \gamma_k$, то с вероятностью, стремящейся к 1, случайный граф $G(n,p)$ содержит непустое k-ядро $\mathcal{C}_k(G(n,p))$, причем его размер удовлетворяет соотношению:*

$$\frac{|\mathcal{C}_k(G(n,p))|}{n} \xrightarrow{\mathsf{P}} p_k(\lambda_k(c)) \quad n \to +\infty.$$

IV. Случай $np \sim 1$

В предыдущих частях мы подробно изучили структуру случайного графа $G(n, p)$ при $np = c < 1$ и $np = c > 1$. Нам осталось изучить наиболее сложный случай, когда $np \sim 1$, т.е. когда мы находимся внутри фазового перехода.

IV. Случай $np \sim 1$

В предыдущих частях мы подробно изучили структуру случайного графа $G(n,p)$ при $np = c < 1$ и $np = c > 1$. Нам осталось изучить наиболее сложный случай, когда $np \sim 1$, т.е. когда мы находимся внутри фазового перехода. Здесь наиболее интересна следующая параметризация:

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}},$$

где $\lambda \in \mathbb{R}$ — фиксированная константа. Оказывается, если $\lambda = \lambda(n) \to -\infty$, то в этом случае структура $G(n,p)$ схожа со структурой в случае $c < 1$ (нет сложных компонент, все компоненты маленькие), а если $\lambda = \lambda(n) \to +\infty$, то ситуация схожа со случаем $c > 1$ (есть гигантская компонента, остальные заметно меньше). Для удобства, далее будем использовать обозначение $G_n(\lambda)$ для подобной параметрической модели случайного графа.

Граф $G_n(\lambda)$

В случайном графе $G_n(\lambda)$ введем следующие обозначения для интересующих нас характеристик:
$X(n, \ell)$ — число ℓ-компонент,
$X(n, k, \ell)$ — число ℓ-компонент на k вершинах,
$Y(n, \ell)$ — общее число вершин во всех ℓ компонентах.

Граф $G_n(\lambda)$

В случайном графе $G_n(\lambda)$ введем следующие обозначения для интересующих нас характеристик:
$X(n, \ell)$ — число ℓ-компонент,
$X(n, k, \ell)$ — число ℓ-компонент на k вершинах,
$Y(n, \ell)$ — общее число вершин во всех ℓ компонентах.

Используя обозначение $C(k, k+\ell)$ для числа связных графов на k вершинах с $k+\ell$ ребрами, легко установить следующие соотношения для этих величин.

Утверждение (3.3)

Имеют место следующие равенства:

$$X(n, \ell) = \sum_{k=1}^{n} X(n, k, \ell), \quad Y(n, \ell) = \sum_{k=1}^{n} kX(n, k, \ell),$$

$$\mathsf{E}X(n, k, \ell) = \binom{n}{k} C(k, k+\ell) p^{k+\ell} (1-p)^{\binom{k}{2} - (k+\ell) + k(n-k)}.$$

Асимптотика $\mathrm{E}X(n,k,\ell)$

Теперь найдем асимптотику математического ожидания в последней формуле. Для этого нам понадобятся следующие оценки факториальных коэффициентов. Напомним обозначение $(n)_k = n(n-1)\dots(n-k+1)$.

Асимптотика $\mathrm{E}X(n,k,\ell)$

Теперь найдем асимптотику математического ожидания в последней формуле. Для этого нам понадобятся следующие оценки факториальных коэффициентов. Напомним обозначение $(n)_k = n(n-1)\ldots(n-k+1)$.

Утверждение (3.4)

Пусть $a = a(n)$, $b = b(n)$, $n \in \mathbb{N}$ две последовательности натуральных чисел.
1) Если $a = o(b^{3/4})$, то

$$\frac{(b)_a}{b^a} = \left(1 + O\left(\frac{a}{b}\right) + O\left(\frac{a^4}{b^3}\right)\right) e^{-\frac{a^2}{2b} - \frac{a^3}{6b^2}}.$$

2) Существует такая абсолютная константа $c > 0$, что для произвольного $a \leqslant b$ выполнено

$$\frac{(b)_a}{b^a} = O\left(e^{-\frac{a^2}{2b} - \frac{a^3}{6b^2} - c\frac{a^4}{b^3}}\right).$$

Асимптотика $\mathrm{E}X(n,k,\ell)$

Теперь найдем асимптотику математического ожидания в последней формуле. Для этого нам понадобятся следующие оценки факториальных коэффициентов. Напомним обозначение $(n)_k = n(n-1)\dots(n-k+1)$.

Утверждение (3.4)

Пусть $a = a(n)$, $b = b(n)$, $n \in \mathbb{N}$ *две последовательности натуральных чисел.*
1) Если $a = o(b^{3/4})$, *то*

$$\frac{(b)_a}{b^a} = \left(1 + O\left(\frac{a}{b}\right) + O\left(\frac{a^4}{b^3}\right)\right) e^{-\frac{a^2}{2b} - \frac{a^3}{6b^2}}.$$

2) Существует такая абсолютная константа $c > 0$, *что для произвольного* $a \leqslant b$ *выполнено*

$$\frac{(b)_a}{b^a} = O\left(e^{-\frac{a^2}{2b} - \frac{a^3}{6b^2} - c\frac{a^4}{b^3}}\right).$$

Доказательство утверждения 3.4 чисто аналитическое, а потому мы оставим его без доказательства.

Асимптотика $\mathrm{E}X(n,k,\ell)$

Теперь докажем лемму о математическом ожидании случайной величины $X(n,k,\ell)$.

Асимптотика $\mathsf{E}X(n,k,\ell)$

Теперь докажем лемму о математическом ожидании случайной величины $X(n,k,\ell)$.

Лемма (3.6)

В модели $G_n(\lambda)$ при фиксированном $\ell \geqslant -1$ выполнено
1) если $k = o(n^{3/4})$, то

$$\mathsf{E}X(n,k,\ell) = \left(1 + \ell\lambda n^{-1/3} + O\left(n^{-2/3}\right) + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) \cdot$$

$$\cdot n^{-\ell} C(k,k+\ell)\frac{e^{-k}}{k!}e^{-F(x_k)},$$

2) для произвольного k

$$\mathsf{E}X(n,k,\ell) = O\left(n^{-\ell}C(k,k+\ell)\frac{e^{-k}}{k!}\right)e^{-F(x_k)-c\frac{k^4}{n^3}},$$

где $x_k = k/n^{2/3}$, $F(x) = \frac{1}{6}(x^3 - 3x^2\lambda + 3x\lambda^2) = \frac{1}{6}((x-\lambda)^3 + \lambda^3)$.

Доказательство леммы 3.6

1) Рассмотрим биномиальный коэффициент $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n^k}{k!} \cdot \frac{(n)_k}{n^k} = |\text{ утв. } 3.4| = \frac{n^k}{k!}\left(1 + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) e^{-\frac{k^2}{2n} - \frac{k^3}{6n^2}}.$$

Доказательство леммы 3.6

1) Рассмотрим биномиальный коэффициент $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n^k}{k!} \cdot \frac{(n)_k}{n^k} = |\text{ утв. } 3.4| = \frac{n^k}{k!}\left(1 + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) e^{-\frac{k^2}{2n} - \frac{k^3}{6n^2}}.$$

Далее,

$$p^{k+\ell} = n^{-k-\ell}(1+\lambda n^{-1/3})^{k+\ell} = n^{-k-\ell}(1+\ell\lambda n^{-1/3}+O(n^{-2/3}))(1+\lambda n^{-1/3})^k =$$

Доказательство леммы 3.6

1) Рассмотрим биномиальный коэффициент $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n^k}{k!}\cdot\frac{(n)_k}{n^k} = |\text{ утв. } 3.4| = \frac{n^k}{k!}\left(1+O\left(\frac{k}{n}\right)+O\left(\frac{k^4}{n^3}\right)\right)e^{-\frac{k^2}{2n}-\frac{k^3}{6n^2}}.$$

Далее,

$$p^{k+\ell} = n^{-k-\ell}(1+\lambda n^{-1/3})^{k+\ell} = n^{-k-\ell}(1+\ell\lambda n^{-1/3}+O(n^{-2/3}))(1+\lambda n^{-1/3})^k =$$

(т.к. $\ln(1+\lambda n^{-1/3}) = \lambda n^{-1/3} - \lambda^2 n^{-2/3}/2 + O(1/n)$)

$$= n^{-k-\ell}(1+\ell\lambda n^{-1/3}+O(n^{-2/3}))e^{\lambda k n^{-1/3}-\lambda^2 k n^{-2/3}/2+O(k/n)};$$

Доказательство леммы 3.6

1) Рассмотрим биномиальный коэффициент $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n^k}{k!} \cdot \frac{(n)_k}{n^k} = |\text{ утв. } 3.4| = \frac{n^k}{k!}\left(1 + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) e^{-\frac{k^2}{2n} - \frac{k^3}{6n^2}}.$$

Далее,

$$p^{k+\ell} = n^{-k-\ell}(1+\lambda n^{-1/3})^{k+\ell} = n^{-k-\ell}(1+\ell\lambda n^{-1/3}+O(n^{-2/3}))(1+\lambda n^{-1/3})^k =$$

(т.к. $\ln(1+\lambda n^{-1/3}) = \lambda n^{-1/3} - \lambda^2 n^{-2/3}/2 + O(1/n)$)

$$= n^{-k-\ell}(1+\ell\lambda n^{-1/3} + O(n^{-2/3}))e^{\lambda k n^{-1/3} - \lambda^2 k n^{-2/3}/2 + O(k/n)};$$

$$(1-p)^{\binom{k}{2} - (k+\ell) + k(n-k)} = (1-p)^{-(3k)/2-\ell}(1-p)^{kn-k^2/2} =$$

(т.к. $\ln(1-p) = -\frac{1+\lambda n^{-1/3}}{n} + O(n^{-2})$)

$$= (1+O(k/n))e^{-\frac{1+\lambda n^{-1/3}}{n}(kn-k^2/2)+O(k/n)}.$$

Доказательство леммы 3.6

Собирая все три представления вместе, получаем следующую асимптотическую формулу для $\mathsf{E}X(n,k,\ell)$:

$$\mathsf{E}X(n,k,\ell) = \binom{n}{k} C(k,k+\ell) p^{k+\ell} (1-p)^{\binom{k}{2}-(k+\ell)+k(n-k)} =$$

$$= C(k,k+\ell)\frac{n^k}{k!}\left(1+O\left(\frac{k}{n}\right)+O\left(\frac{k^4}{n^3}\right)\right)e^{-\frac{k^2}{2n}-\frac{k^3}{6n^2}}n^{-k-\ell}\cdot$$

$$\cdot(1+\ell\lambda n^{-1/3}+O(n^{-2/3}))e^{\lambda k n^{-1/3}-\lambda^2 k n^{-2/3}/2}\left(1+O\left(\frac{k}{n}\right)\right)\cdot$$

$$\cdot e^{-\frac{1+\lambda n^{-1/3}}{n}(kn-k^2/2)} =$$

Доказательство леммы 3.6

$$= \left(1 + \ell\lambda n^{-1/3} + O\left(n^{-2/3}\right) + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) \cdot$$

$$\cdot \frac{n^{-\ell}}{k!} C(k, k+\ell) e^{-k} e^{-\frac{k^3}{6n^2} - \frac{\lambda^2 n^{-2/3} k}{2} + \frac{\lambda k^2}{2n^{4/3}}} =$$

Доказательство леммы 3.6

$$= \left(1 + \ell\lambda n^{-1/3} + O\left(n^{-2/3}\right) + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) \cdot$$

$$\cdot \frac{n^{-\ell}}{k!} C(k, k+\ell) e^{-k} e^{-\frac{k^3}{6n^2} - \frac{\lambda^2 n^{-2/3} k}{2} + \frac{\lambda k^2}{2n^{4/3}}} =$$

$$= \left(1 + \ell\lambda n^{-1/3} + O\left(n^{-2/3}\right) + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) \frac{n^{-\ell}}{k!} C(k, k+\ell) e^{-k} e^{-F(x_k)}.$$

Доказательство леммы 3.6

$$= \left(1 + \ell\lambda n^{-1/3} + O\left(n^{-2/3}\right) + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) \cdot$$

$$\cdot \frac{n^{-\ell}}{k!} C(k, k+\ell) e^{-k} e^{-\frac{k^3}{6n^2} - \frac{\lambda^2 n^{-2/3} k}{2} + \frac{\lambda k^2}{2n^{4/3}}} =$$

$$= \left(1 + \ell\lambda n^{-1/3} + O\left(n^{-2/3}\right) + O\left(\frac{k}{n}\right) + O\left(\frac{k^4}{n^3}\right)\right) \frac{n^{-\ell}}{k!} C(k, k+\ell) e^{-k} e^{-F(x_k)}.$$

2) Доказывается полностью аналогично с использованием второй оценки из утверждения 3.4. □