Московский физико-технический институт
Физтех-школа прикладной математики и информатики

ДИСКРЕТНЫЙ АНАЛИЗ

III СЕМЕСТР

Лектор: *Андрей Михайлович Райгородский*

Автор: Максимов Даниил

Проект на Github

осень 2022

Содержание

1 Продолжение теории графов

1.1 Асимптотика числа унициклических графов

Утверждение 1.1. *Для числа унициклических графов $U(n)$ имеет место следующая асимптотика:*

$$U(n) \sim \sqrt{\frac{\pi}{8}} \cdot n^{n-\frac{1}{2}}$$

Доказательство. Перепишем сумму в более удобном виде:

$$U(n) = \frac{1}{2}\sum_{r=3}^{n} C_n^r \cdot r! \cdot n^{n-1-r} = \frac{1}{2}n^{n-1}\sum_{r=3}^{n}\left(n^{-r} \cdot \prod_{t=0}^{r-1}(n-t)\right) =$$
$$\frac{1}{2}n^{n-1}\sum_{r=3}^{n}\prod_{t=0}^{r-1}\left(1-\frac{t}{n}\right) = \frac{1}{2}n^{n-1}\sum_{r=3}^{n}\exp\left(\sum_{t=0}^{r-1}\ln\left(1-\frac{t}{n}\right)\right)$$

Отдельно преобразуем внутреннюю сумму (при помощи формулы Тейлора с остаточным членом в форме O-большого):

$$\sum_{t=0}^{r-1}\ln\left(1-\frac{t}{n}\right) = \sum_{t=0}^{r-1}\left(-\frac{t}{n}+O\left(\frac{t^2}{n^2}\right)\right) = -\frac{r(r-1)}{2n}+O\left(\frac{r^3}{n^2}\right)$$

Отметим 1 интересный факт о полученном выражении: если $r = r(n)$ и $r^3 = o(n^2)$ (то есть $r = o(n^{2/3})$), то

$$-\frac{r(r-1)}{2n}+O\left(\frac{r^3}{n^2}\right) = -\frac{r(r-1)}{2n}+o(1) \sim -\frac{r(r-1)}{2n}$$

Теперь, мы применим классический трюк: разделим сумму на 2 и попробуем оценить каждую часть по отдельности

$$\sum_{r=3}^{n}\exp\left(-\frac{r(r-1)}{2n}+O\left(\frac{r^3}{n^2}\right)\right) = \underbrace{\sum_{r=3}^{l(n)}\exp\left(-\frac{r(r-1)}{2n}+O\left(\frac{r^3}{n^2}\right)\right)}_{S_1} +$$
$$\underbrace{\sum_{r=l(n)+1}^{n}\exp\left(-\frac{r(r-1)}{2n}+O\left(\frac{r^3}{n^2}\right)\right)}_{S_2}$$

▷ Попробуем выяснить, при каких $l(n)$ S_2 будет зануляться:

$$S_2 = \sum_{r=l(n)+1}^{n} \exp\left(\sum_{t=0}^{r-1} \ln\left(1 - \frac{t}{n}\right)\right) \leqslant$$

$$\sum_{r=l(n)+1}^{n} \exp\left(-\sum_{t=0}^{r-1} \frac{t}{n}\right) = \sum_{r=l(n)+1}^{n} \exp\left(-\frac{r(r-1)}{2n}\right) \leqslant$$

$$n \cdot \exp\left(-\frac{(l(n)+1)l(n)}{2n}\right)$$

В переходе от первой строки ко второй мы используем тот факт, что для $-1 < x < 0 \Rightarrow \ln(1+x) < x$. Если $l(n) = \Omega(n^{1/2})$, то экспонента будет стремиться к нулю, утаскивая S_2 за собой.

▷ Теперь посмотрим, что мы могли бы сделать с S_1, если учитывать условие из S_2. Всё, что мы можем сделать - потребовать $l(n) = o(n^{2/3})$ и применить уже известный факт:

$$S_1 = \sum_{r=3}^{l(n)} \exp\left(-\frac{r(r-1)}{2n} + O\left(\frac{r^3}{n^2}\right)\right) \sim \sum_{r=3}^{l(n)} \exp\left(-\frac{r(r-1)}{2n}\right) \sim \sum_{r=3}^{l(n)} \exp\left(-\frac{r^2}{2n}\right)$$

Для определенности, мы можем взять $l(n) = \lfloor n^{0.6} \rfloor$, или, например, $l(n) = \lfloor n^{\pi/5} \rfloor$. Получившуюся сумму очень хотелось бы посчитать так:

$$\sum_{r=3}^{l(n)} \exp\left(-\frac{r^2}{2n}\right) = \sum_{r=0}^{\infty} \exp\left(-\frac{r^2}{2n}\right) - \sum_{r=0}^{2} \exp\left(-\frac{r^2}{2n}\right) - \sum_{r=l(n)+1}^{\infty} \exp\left(-\frac{r^2}{2n}\right)$$

Но возникает вопрос: а почему эта запись корректна? Почему ряд сходится? Мы воспользуемся интегральным признаком Коши и его следствием:

$$\sum_{r=0}^{\infty} \exp\left(-\frac{r^2}{2n}\right) \sim \int_0^{\infty} e^{-\frac{r^2}{2n}} dr$$

Мы свели задачу к известному интегралу Гаусса, обладающего конкретным значением (значит, ряды сходятся):

$$\int_0^{\infty} e^{-\frac{r^2}{2n}} dr = \frac{\sqrt{2\pi n}}{2} = \sqrt{\frac{\pi n}{2}}$$

Со второй суммой всё совсем просто:

$$\lim_{n\to\infty} \sum_{r=0}^{2} e^{-\frac{r^2}{2n}} = 1 + 1 + 1 = 3$$

Для последней придётся поисследовать ряд. Например, во сколько раз уменьшается следующее слагаемое по сравнению с предыдущим?

$$e^{-\frac{(r+1)^2}{2n}} / e^{-\frac{r^2}{2n}} = e^{-\frac{2r+1}{2n}} = e^{-\frac{r}{n} - \frac{1}{2n}} < e^{-\frac{r}{n}}$$

Если $r > n^2$, то $e^{-r/n} < e^{-n}$ - зависимость от r пропадает, и мы можем получить оценку сверху на сумму. Значит, повторим старую стратегию:

$$\sum_{r=l(n)+1}^{\infty} \exp\left(-\frac{r^2}{2n}\right) = \underbrace{\sum_{r=l(n)+1}^{n^2} \exp\left(-\frac{r^2}{2n}\right)}_{S_1'} + \underbrace{\sum_{r=n^2+1}^{\infty} \exp\left(-\frac{r^2}{2n}\right)}_{S_2'}$$

– Для S_1' делаем действия, аналогичные для S_1 (напоминаю, есть ограничения $n^{1/2} < l(n) < n^{2/3}$):

$$S_1' \leqslant \sum_{r=l(n)+1}^{n^2} \exp\left(-\frac{(l(n)+1)^2}{2n}\right) \leqslant n^2 \cdot \exp\left(-\frac{l^2(n)}{2n}\right) \xrightarrow[n\to\infty]{} 0$$

– Для S_2' теперь пишем оценку исходя из упомянутого выше факта:

$$S_2' \leqslant \sum_{r=n^2+1}^{\infty} \exp\left(-\frac{(n^2+1)^2}{2n}\right) \cdot (e^{-n})^{r-n^2-1} = \exp\left(-\frac{(n^2+1)^2}{2n}\right) \cdot (1+e^{-n}+e^{-2n}+\ldots) =$$

$$\exp\left(-\frac{(n^2+1)^2}{2n}\right) \cdot \frac{1}{1-e^{-n}} \xrightarrow[n\to\infty]{} 0$$

Итого, мы нашли эквивалентность для S_1:

$$S_1 \sim \sqrt{\frac{\pi n}{2}} - 3 \sim \sqrt{\frac{\pi n}{2}}$$

В общей формуле это выливается в следующее:

$$U(n) = \frac{1}{2} n^{n-1}(S_1 + S_2) \sim \frac{1}{2} n^{n-1} \sqrt{\frac{\pi}{2}} n^{\frac{1}{2}} = \sqrt{\frac{\pi}{8}} n^{n-\frac{1}{2}}$$

$\square$

1.2 Асимптотика биномиальных коэффициентов

Лемма 1.1. *Имеют место следующие 2 утверждения:*

1. $C_n^k \leqslant \frac{n^k}{k!}$
2. $C_n^k = \frac{n^k}{k!} \exp\left(-\frac{k(k-1)}{2n} + O(k^3/n^2)\right)$

Доказательство. Распишем биномиальный коэффициент, как мы это сделали в доказательстве теоремы:

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\ldots(n-(k-1))}{k!} = \frac{n^k}{k!} \prod_{t=0}^{k-1} \left(1 - \frac{t}{n}\right) \leqslant \frac{n^k}{k!}$$

Дрля второго надо расписать натуральный логарифм через формулу Тейлора с О-большим:

$$C_n^k = \frac{n^k}{k!}\exp\left(\sum_{t=0}^{k-1}\ln\left(1-\frac{t}{n}\right)\right) = \frac{n^k}{k!}\exp\left(\sum_{t=0}^{k-1}\left(-\frac{t}{n}+O\left(\frac{t^2}{n^2}\right)\right)\right) = \frac{n^k}{k!}\exp\left(-\frac{k(k-1)}{2n}+O\left(\frac{k^3}{n^2}\right)\right)$$

□

Следствие.

▷ Если $k = k(n) = o(\sqrt{n})$, то имеет место эквивалентность:

$$C_n^k \sim \frac{n^k}{k!}$$

▷ Если $k^3 = o(n^2)$, то имеет место другая эквивалентность:

$$C_n^k \sim \frac{n^k}{k!}\cdot e^{-\frac{k(k-1)}{2n}}$$

Замечание. На практике очень полезно знать, как ведут себя биномиальные коэффициенты при $k \sim an$, $a \in (0;1)$. Например, это играет важную роль в кодах, исправляющих ошибки.

Теорема 1.1. *(Формула Стирлинга, без доказательства) Для факториала верна следующая эквивалентность:*

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

Теорема 1.2. *Если $k = \lfloor an \rfloor$, где $a \in (0;1)$, то для биномиальных коэффициентов справедлива следующая эквивалентность:*

$$C_n^{\lfloor an \rfloor} \sim (a^{-a}(1-a)^{-(1-a)} + o(1))^n$$

Доказательство. Начнём аккуратно расписывать биномиальный коэффициент через формулу Стирлинга:

$$C_n^{\lfloor an \rfloor} = \frac{n!}{\lfloor an \rfloor!(n-\lfloor an \rfloor)!} \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot \frac{1}{\sqrt{2\pi\lfloor an \rfloor}}\left(\frac{e}{\lfloor an \rfloor}\right)^{\lfloor an \rfloor} \cdot \frac{1}{\sqrt{2\pi(n-\lfloor an \rfloor)}}\left(\frac{e}{n-\lfloor an \rfloor}\right)^{n-\lfloor an \rfloor} = n^n \underbrace{\frac{\sqrt{2\pi n}}{\sqrt{2\pi\lfloor an \rfloor}\cdot\sqrt{2\pi(n-\lfloor an \rfloor)}}}_{\Pi(n)} \cdot \frac{1}{(\lfloor an \rfloor)^{\lfloor an \rfloor}} \cdot \frac{1}{(n-\lfloor an \rfloor)^{n-\lfloor an \rfloor}}$$

Если бы могли как-то избавиться от округлений, то можно было бы сократить n^n. А что, идея хорошая: давайте распишем округление как $\lfloor an \rfloor = an - \varepsilon_n$, $\varepsilon_n \in [0; 1)$. Тогда

$$C_n^{\lfloor an \rfloor} \sim n^n \Pi(n) \frac{1}{(an-\varepsilon_n)^{an-\varepsilon_n}} \cdot \frac{1}{(n - an + \varepsilon_n)^{n-an+\varepsilon_n}} =$$
$$\Pi(n) \frac{n^n}{(an)^{an-\varepsilon_n} \cdot (n-an)^{n-an+\varepsilon_n}} \cdot \frac{1}{\left(1 - \frac{\varepsilon_n}{an}\right)^{an-\varepsilon_n} \cdot \left(1 + \frac{\varepsilon_n}{n-an}\right)^{n-an+\varepsilon_n}} =$$
$$\Pi(n)Q(n) \cdot \frac{n^n}{(an)^{an}(n-an)^{n-an}} = \Pi(n)Q(n) \cdot \left(a^{-a}(1-a)^{-(1-a)}\right)^n$$

Остаётся доказать, что 2 коэффициента спереди не вносят вклад в эквивалентность. Для этой цели, вернёмся к исходной формуле и покажем такую равносильность (для простоты записи обозначим $A = a^{-a}(1-a)^{-(1-a)}$):

$$C_n^{\lfloor an \rfloor} \sim (A + o(1))^n \iff \ln C_n^{\lfloor an \rfloor} \sim n \ln(A)$$

- ▷ ($\Rightarrow$) Обозначим $o(1) = \delta \to 0$. Тогда

$$\ln C_n^{\lfloor an \rfloor} = n \ln(A + \delta) = n \ln A + n \ln(1 + \delta/A)$$

Из условия $\ln(1 + \delta/A)$ тоже стремится к нулю. Стало быть

$$\lim_{n\to\infty} \frac{n \ln A + n \ln(1 + \delta/A)}{n \ln A} = 1$$

- ▷ ($\Leftarrow$) Условие означает, что

$$\lim_{n\to\infty} \frac{\ln C_n^{\lfloor an \rfloor}}{n \ln(A)} = 1 \Leftrightarrow \ln C_n^{\lfloor an \rfloor} = n \ln(A) + o(1)$$

Не придумал

Значит, нужно проверить такую эквивалентность:

$$\ln C_n^{\lfloor an \rfloor} \sim \ln(\Pi(n)Q(n)A^n) \sim n \ln(A)$$

Действительно, $\ln(\Pi(n)Q(n)A^n) = \ln(\Pi(n)) + \ln(Q(n)) + n\ln(A)$

- ▷ То, что мы обозначили за $\Pi(n)$, является $O(n)$. Стало быть, $\ln(\Pi(n)) = O(\ln n) = o(n)$
- ▷ Аккуратно разберёмся с каждым сомножителем, входящим в $Q(n)$:

$$\ln\left(\frac{1}{(an)^{-\varepsilon_n}}\right) = \varepsilon_n \ln(an) = o(n); \quad \ln\left(\frac{1}{(n-an)^{\varepsilon_n}}\right) = -\varepsilon_n \cdot \ln(n - an) = o(n)$$

И другие 2:

$$\left(1 - \frac{\varepsilon_n}{an}\right)^{\varepsilon_n - an} \to 1 \cdot e^{-\varepsilon_n} = o(n); \quad \left(1 - \frac{\varepsilon_n}{n-an}\right)^{n-an+\varepsilon_n} \to e^{-\varepsilon_n} \cdot 1 = o(n)$$

- ▷ $\ln(A^n) = n\ln(A)$. Тут, думаю, пояснений не надо.

Таким образом
$$\lim_{n\to\infty} \frac{\ln(\Pi(n)) + \ln(Q(n)) + n\ln(A)}{n\ln(A)} = 1$$

□

Следствие. Если проделать всё то же самое, но при этом положить $k = n$ и избежать этим округлений, то получится такая формула:

$$C_{2n}^n \sim \frac{2^{2n}}{\sqrt{\pi n}}$$

Упражнение. *Если $k = k(n) \sim an$, $a \in (0;1)$, то*

$$C_n^k = (A + o(1))^n$$

где $A = a^{-a}(1-a)^{-(1-a)}$

Упражнение. *Если $\sum_{i=1}^n k_i = n$ и $\forall i \in \{1,\ldots,n\}$ $k_i \sim a_i n$, то верна формула*

$$P(k_1,\ldots,k_s) = (a_1^{-a_1}\cdot\ldots\cdot a_s^{-a_s} + o(1))^n$$

1.3 Небольшое напоминание

Замечание. Когда мы говорим слово *граф*, то мы подразумеваем простой граф (без петель, без ориентации и без кратных рёбер).

Напоминание. Пусть $G = (V,E)$ — произвольный граф. Тогда подмножество $W \subseteq V$ называется *независимым*, если выполнено условие:

$$\forall x, y \in W \ (x,y) \notin E$$

Определение 1.1. Пусть $G = (V,E)$ — произвольный граф. Тогда подмножество $W \subseteq V$ называется *кликой*, если верно следующее условие:

$$\forall x, y \in W \ (x,y) \in E$$

Замечание. Иными словами, клика — это индуцированный полный подграф в G.

Замечание. Мы могли бы сказать, что независимое множество W является *антикликой*, но в серьёзной науке так говорить не принято, хоть в олимпиадном движении это понятие и популярно.

Напоминание. *Числом независимости* графа $G = (V,E)$ называется максимальный размер среди всех его независимых множеств:

$$\alpha(G) = \max_{W\subseteq G\text{ — независимое}} |W|$$

Определение 1.2. *Кликовым числом* графа $G = (V,E)$ называется максимальный размер клики в этом графе:
$$w(G) = \max_{W\subseteq G\text{ — клика}} |W|$$

Определение 1.3. *Хроматическим числом* графа $G = (V, E)$ называется минимальное число цветов, которыми можно покрасить вершины графа так, что у любого ребра концы будут разноцветными.

Для формального описания, введём множество функций-раскрасок:

$$\mathcal{C} = \{\chi\colon V \to \mathbb{N} \bigm| \forall (x, y) \in E\ \chi(x) \neq \chi(y)\}$$

Тогда хроматическое число можно записать так:

$$\chi(G) = \min_{\chi \in \mathcal{C}} |\chi(V)|$$

Замечание. Мы используем букву χ из-за ассоциации с латинским словом chroma, обозначающим цвет.

Упражнение. *Верны следующие простые утверждения:*

1. $\chi(G) \geqslant w(G)$
2. $\chi(G) \geqslant |V|/\alpha(G)$
3. $\chi(G) \leqslant \Delta(G) + 1$, *где* $\Delta(G)$ *- это максимальная степень вершины в графе*

1.4 Вероятности, связанные с характеристическими числами графов

Теорема 1.3. *Почти любой граф* G *на* n *вершинах таков, что у него одновременно* $w(G) < 2\log_2 n$ *и* $\alpha(G) < 2\log_2 n$

Замечание. *Почти любой* нужно понимать как существование следующего предела:

$$\frac{\#G \text{ на } n \text{ вершинах с нужными свойствами}}{2^{C_n^2}} \xrightarrow[n\to\infty]{} 1$$

Доказательство. Введём классическое вероятностное пространство. Положим Ω_n — множество всех графов на n вершинах (без изоморфизма). Тогда, хотелось бы доказать следующий факт:

$$P(w(G) < 2\log_2 n) \xrightarrow[n\to\infty]{} 1$$

Это эквивалентно стремлению вероятности противоположного события к нулю:

$$P(w(G) \geqslant 2\log_2 n) \xrightarrow[n\to\infty]{} 0$$

Зафиксируем некоторое $k = \lfloor 2\log_2 n \rfloor$ и занумеруем все наборы из k вершин как $A_1, \ldots, A_{C_n^k}$. Обозначим $\mathcal{A}_i$ - событие, когда вершины A_i в выпавшем графе G образуют клику. $\mathcal{A} = \bigcup_{i=1}^{C_n^k} \mathcal{A}_i$ — событие, что будет хотя бы одна клика размера k. По свойству вероятностной меры

$$P(\mathcal{A}) \leqslant \sum_{i=1}^{C_n^k} P(\mathcal{A}_i)$$

Что есть $P(\mathcal{A}_i)$? Это надо поделить число графов, в которых A_i образует клику, на число всех графов. То есть

$$P(\mathcal{A}) \leqslant \sum_{i=1}^{C_n^k} \frac{2^{C_n^2 - C_k^2}}{2^{C_n^2}} = \sum_{i=1}^{C_n^k} \left(\frac{1}{2}\right)^{C_k^2} = C_n^k \cdot 2^{-C_k^2} \leqslant \frac{n^k}{k!} \cdot 2^{-C_k^2} = \frac{2^{k\log_2 n - k^2/2 + k/2}}{k!} \leqslant$$
$$\frac{1}{k!} \cdot 2^{\log_2^2 n + \log_2 n - \frac{(2\log_2 n - 1)^2}{2}} \leqslant \frac{2^{3\log_2 n}}{k!} \leqslant \frac{2^{1.5(k+1)}}{k!} \to 0$$

Абсолютно аналогично можно доказать вероятность для числа независимости графа:

$$P(\alpha(G) < 2\log_2 n) \xrightarrow[n\to\infty]{} 1$$

Теперь поясним, почему вероятность выполнения этих условий одновременно тоже будет стремиться к единице. Пусть $\mathcal{B}$ — это событие, что будет хотя бы одно независимое множество размера k. Тогда

$$P(\mathcal{A} \cup \mathcal{B}) \leqslant P(\mathcal{A}) + P(\mathcal{B}) \to 0$$

Значит, $P(\overline{\mathcal{A} \cup \mathcal{B}}) \to 1$, что в точности является нужным событием. $\square$

Пример. Зафиксируем $n = k^2$ и рассмотрим граф, состоящий из k клик, в каждой из которых по k элементов. Несложно понять, что в таком графе $w(G) = \sqrt{n} = \alpha(G)$

Теорема 1.4. *(Теорема Тура́на) Для любого графа $G = (V, E)$ такого, что $|V| = n$ и $\alpha(G) = \alpha$, имеет место оценка на число рёбер:*

$$|E| \geqslant n \cdot \left\lfloor \frac{n}{\alpha} \right\rfloor - \alpha \frac{\left\lfloor \frac{n}{\alpha} \right\rfloor \left(\left\lfloor \frac{n}{\alpha} \right\rfloor + 1 \right)}{2}$$

Доказательство. Пусть $A \subset V$ - это независимое подмножество вершин, реализующее число независимости (то есть $|A| = \alpha(G) = \alpha$). Тогда совершенно очевиден следующий факт:

$$\forall x \in V \backslash A\ \exists y \in A \ \big|\ (x, y) \in E$$

Следовательно, у нас уже есть не менее $n - \alpha$ рёбер. Повторим рассуждение для $V \backslash A$ (в нём выделяем A': $|A'| \leqslant \alpha$ и так далее). Тогда, мы получим ещё $\geqslant (n - \alpha) - \alpha = n - 2\alpha$ рёбер. Продолжая спуск в разности множеств, получим следующую оценку на $|E|$:

$$|E| \geqslant (n - \alpha) + (n - 2\alpha) + \ldots + \left(n - \left\lfloor \frac{n}{\alpha} \right\rfloor \alpha\right) = n \cdot \left\lfloor \frac{n}{\alpha} \right\rfloor - \alpha \frac{\left\lfloor \frac{n}{\alpha} \right\rfloor \left(\left\lfloor \frac{n}{\alpha} \right\rfloor + 1 \right)}{2}$$

$\square$

Замечание. «Мощь» теоремы заключается в том, что оценка неулучшаемая. Это показано ниже.

Пример. Пусть $\alpha \mid n$, тогда оценка теоремы Турана принимает вид $|E| \geqslant \frac{n^2}{2\alpha} - \frac{n}{2}$.

Сюда бы картиночку со второй лекции осени 2022, где-то 1:02:00

Посмотрим такой граф, для которого будет достигнуто равенство. Разобьём n вершин на α равновеликих клик. Тогда понятно, что $\alpha(G)=\alpha$, но при этом число рёбер

$$|E|=\alpha\cdot C^2_{n/\alpha}=\alpha\cdot\frac{\frac{n}{\alpha}\left(\frac{n}{\alpha}-1\right)}{2}=\frac{n^2}{2\alpha}-\frac{n}{2}$$

1.5 Дистанционные графы

Определение 1.4. Граф $G=(V,E)$ называется *дистанционным*, если $V\subset\mathbb{R}^n$, а $E=\{\{\vec{x},\vec{y}\}\colon|\vec{x}-\vec{y}|=a\}$, $a>0$.

Пример. Мы уже сталкивались с дистанционными графами в прошлом году. Примером послужит граф, к которому применялась теорема Эрдёша-Хватала:

- $V=\{\vec{x}\in\mathbb{R}^n\colon\forall i\ x_i\in\{0,1\},\ x_1+\ldots+x_n=3\}$
- $E=\{\{\vec{x},\vec{y}\}\colon(\vec{x},\vec{y})=1\}$

Теорема 1.5. *Пусть G — дистанционный граф на плоскости, $|V|=4n$, $\alpha(G)\leqslant n$. Тогда $|E|\geqslant 7n$.*

Доказательство. Для начала, повторим рассуждения из теоремы Турана. Выделим $A\subseteq V$, реализующее число независимости. Из него мы получим $\geqslant 3n$ рёбер в G. Посмотрим на индуцированный при помощи $V\backslash A$ граф и проделаем то же самое. Это даст нам $\geqslant 2n$ ребёр, ну и последний раз даст $\geqslant n$ рёбер. Суммарно это $|E|\geqslant 6n$, что не дотягивает до утверждения теоремы. Мы сможем доказать теорему, если мы из каких-то соображений улучшим оценку рёбер на первом шаге до $\geqslant 4n$.

Для этого разобьём $V\backslash A$ на 2 непересекающихся множества:

$$V\backslash A=V_1\sqcup V_2$$

где $V_1=\{x\in V\backslash A\colon\exists!y\in A,\ (x,y)\in E\}$

Утверждение 1.2. *Если $|V_2|\geqslant n$, то на первом шаге мы найдём $\geqslant 4n$ рёбер*

Доказательство. Действительно, V_2 даст $\geqslant 2|V_2|$ рёбер, тогда как от V_1 будет ровно $4n-\alpha(G)-|V_2|\geqslant 3n-|V_2|$ рёбер. В сумме это $3n+|V_2|$ рёбер и утверждение стало тривиальным. □

Лемма 1.2. *При всех условиях теоремы, $|V_2|\geqslant n$*

Доказательство. Предположим противное: $|V_2|\leqslant n-1$. В таком случае, $|V_1|\geqslant 2n+1$. Более того, тогда должна существовать вершина $y\in A$, соединенная ребрами с тремя из V_1 (это следует по принципу Дирихле):

$$\Rightarrow\exists y\in A\ \Big|\ \exists x_1,x_2,x_3\in V_1\quad\forall i\in\{1,2,3\}\ (x_i,y)\in E$$

Ну и катарсис доказательства: если между x_1,x_2 нет ребра, то $(A\backslash\{y\})\cup\{x_1,x_2\}$ будет бо́льшим независимым множеством, чем A. Аналогично для остальных пар иксов, а стало быть мы получаем клику на вершинах $\{y,x_1,x_2,x_3\}$ в дистанционном графе на плоскости. Несложно убедиться, что такой существовать в принципе не может. □

□

Замечание. Теорема не является лучшим результатом. Примерно теми же методами доказывается оценка $|E| \geqslant \frac{26}{3}n$.

Следствие. (теоремы Турана) Если $G_n = (V_n, E_n)$, $|V_n| = n$, а также $\alpha(G_n) = \alpha_n = o(n)$, то имеет место следующее неравенство:

$$|E_n| \geqslant (1 + o(1))\frac{n^2}{2\alpha}$$

Доказательство. Несложно увидеть следующую эквивалентность:

$$\left\lfloor \frac{n}{\alpha} \right\rfloor \sim \frac{n}{\alpha}$$

Если подставить эту эквивалентность для формулы из теоремы Турана, то получим данное следствие. □

Теорема 1.6. *Пусть $G_s = (V_s, E_s)$ — последовательность дистанционных графов таких, что $|V_s| = n(s)$, $\alpha(G_s) = \alpha(s)$, $V_s \subset \mathbb{R}^{d(s)}$. Пусть также $d(s) \cdot \alpha(s) = o(n(s))$. В таком случае*

$$|E_s| \geqslant (1 + o(1))\frac{n^2(s)}{\alpha(s)}$$

Замечание. Пафос теоремы состоит в том, что для дистанционных графов мы улучшаем следствие теоремы Турана аж в 2 раза.

Лемма 1.3. *(без доказательства) Если K_l — это клика на l вершинах, то в $\mathbb{R}^n$ всегда существует K_{n+1} и не существует K_{n+2}.*

Доказательство. (теоремы) Будем повторять идею доказательства предыдущей теоремы, но уже ориентируясь на другую лемму:

Лемма 1.4. *При условиях теоремы, $|V_1| \leqslant d(s) \cdot \alpha(s)$*

Доказательство. Предположим, что $|V_1| \geqslant d(s) \cdot \alpha(s) + 1$. Тогда

$$\exists y \in A, x_1, \ldots, x_{d+1} \in V_1 \;\Big|\; \forall i \in \{1, \ldots, d+1\} \;\; (x_i, y) \in E$$

Ровно тем же методом показываем, что вершины $\{y, x_1, \ldots, x_{d+1}\}$ образуют клику. По лемме о кликах, такой в нашем пространстве не существует. □

Следовательно, мы найдём $\geqslant d(s)\cdot\alpha(s) + 2(n(s) - \alpha(s) - d(s)\cdot\alpha(s))$ рёбер. Если преобразовать данное выражение, то получим $\geqslant 2n(s) - d(s)\alpha(s) - 2\alpha(s)$ рёбер. Что поменяется в этой формуле, когда мы перейдём к $V \backslash A$? Вместо $n(s)$ будет $n(s) - \alpha(s)$, и после раскрытия скобок получится то же выражение, только $4\alpha(s)$ в конце. Аналогично происходит на остальных итерациях, а потому всего их будет k штук:

$$k = \left\lfloor \frac{2n(s) - d(s)\alpha(s)}{2\alpha(s)} \right\rfloor \sim \frac{2n(s)}{2\alpha(s)} = \frac{n(s)}{\alpha(s)}$$

если обозначить $a(s) = 2n(s) - d(s)\alpha(s)$, то

$$|E_s| \geqslant \sum_{t=1}^{k} (a(s) - t \cdot 2\alpha(s)) = ka(s) - 2\alpha(s) \cdot \frac{k(k-1)}{2} \sim ka(s) - k^2\alpha(s) \sim$$
$$a(s) \cdot \frac{n(s)}{\alpha(s)} - \alpha(s) \cdot \left(\frac{a(s)}{2\alpha(s)}\right)^2 \sim \frac{2n^2(s)}{\alpha(s)} - \frac{4n^2(s)}{4\alpha(s)} = \frac{n^2(s)}{\alpha(s)}$$

□

Пример. Вернёмся к графу, который появлялся в теореме Эрдёша-Хватала:

▷ $V = \{\vec{x} = (x_1, \ldots, x_d)\colon x_i \in \{0, 1\},\ \sum_{i=1}^{d} x_i = 3\}$

▷ $E = \{\{\vec{x}, \vec{y}\}\colon |\vec{x} - \vec{y}| = 2\}$

Напомним 3 факта об этом графе:

1. $|V| = n = C_d^3$
2. $\alpha(G) = \begin{cases} d,\ d \equiv 0 \pmod 4 \\ d-1,\ d \equiv 1 \pmod 4 \\ d-2, \text{ иначе} \end{cases}$

Положим за номер графа в последовательности размерность пространства, в котором он находится. Тогда $n \sim d^3/6$, $\alpha(G) \sim d$, а также $\alpha \cdot d \sim d^2 = o(n)$. Стало быть, доказанная теорема верна для этой последовательности, и тут мы улучшаем оценку в 2 раза по сравнению со следствием теоремы Турана.

1.6 Жадные алгоритмы для оценки характеристических чисел случайных графов

Замечание. Это, конечно, замечательно обсуждать абстрактную математику, но и приложения забывать нельзя. В частности, хотелось бы иметь алгоритмы для вычисления характеристических чисел произвольного графа G (то есть его $\chi(G)$, $\alpha(G)$ и $\mu(G)$).

Задачи о нахождении таких чисел лежат в классе NP-полных задач, то есть мы не можем сделать их за полиномиальное от числа вершин время.

Жадный алгоритм для подсчёта хроматического числа

Задача. *Пусть дан граф $G = (V, E)$, $V = \{1, \ldots, n\}$. Нас просят дать оценку на хроматическое число графа $\chi(G)$ за полиномиальное от n время работы.*

Решение. Попробуем исполнить следующий алгоритм: Идём по номерам V от 1 по возрастанию. Для каждой следующей вершины будем находить цвет, в который её нужно покрасить, чтобы не испортить текущую раскраску. В качестве цвета текущей вершины положим минимальное натуральное число, которое не появится при перечислении цветов уже покрашенных вершин, соединённых с этой.

Замечание. Оценку на хроматическое число графа, которую можно получить данным алгоритмом, обозначим за $\chi_g(G)$ (g - greedy).

Следствие. Описанный алгоритм можно также использовать для оценки $\alpha_g(G)$ и $w_g(G)$:

- ▷ Чтобы получить оценку на число независимости графа, просто возьмём максимальный остов одного цвета.
- ▷ $w_g(G) = \alpha_g(G')$, где G' — это инвертированный граф (рёбра, которых не было в G, добавили, а старые наоборот, убрали).

Теорема 1.7. *Если ввести (Ω, F, P) — вероятностное пространство случайных графов ($|\Omega| = 2^{C_n^2}$, $F = 2^\Omega$, P — равномерная вероятность), то для описанных жадных алгоритмов и их оценок имеют место следующие утверждения:*

- ▷ $\forall \varepsilon > 0 \;\; P\left(G\colon \frac{\alpha(G)}{\alpha_g(G)} \leqslant 2 + \varepsilon\right) \xrightarrow[n\to\infty]{} 1$
- ▷ $\forall \varepsilon > 0 \;\; P\left(G\colon \frac{\chi_g(G)}{\chi(G)} \leqslant 2 + \varepsilon\right) \xrightarrow[n\to\infty]{} 1$
- ▷ $\forall \varepsilon > 0 \;\; P\left(G\colon \frac{w(G)}{w_g(G)} \leqslant 2 + \varepsilon\right) \xrightarrow[n\to\infty]{} 1$

Замечание. Для жадных алгоритмов $\varepsilon > 0$ из вероятности убрать нельзя. Однако, существуют и такие детерминированные алгоритмы, для которых это сделать возможно.

Доказательство. Проведём доказательство только для числа независимости. Мы уже знаем, что $\lim_{n\to\infty} P(\alpha(G) < 2\log_2 n) = 1$. Если мы как-то докажем, что

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P(\alpha_g(G) \geqslant (1-\varepsilon)\log_2 n) = 1$$

То этого будет достаточно. Действительно, обратим внимание на 2 факта:

1. Если $A_n \subseteq B_n$ и $\lim_{n\to\infty} P(A_n) = 1$, то $\lim_{n\to\infty} P(B_n) = 1$ — просто из свойства $P(A_n) \leqslant P(B_n) \leqslant 1$
2. Что из себя представляет пересечение свойств $\alpha(G) < 2\log_2 n$ и $\alpha_g(G) \geqslant (1-\varepsilon)\log_2 n$?

   $$\frac{\alpha(G)}{\alpha_g(G)} \leqslant \frac{2\log_2 n}{(1-\varepsilon)\log_2 n} = \frac{2}{1-\varepsilon} = 2 + \varepsilon'$$

   Понятно, что из предела к 1 при ε мы вполне законно говорим и о пределе для ε'.

Чтобы доказать достаточно неравенство, рассмотрим его дополнение:

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P(\underbrace{\alpha_g(G) < (1-\varepsilon)\log_2 n}_{\mathcal{A}}) = 0$$

Теперь покажем некоторое событие $\mathcal{B} \subseteq \mathcal{A}$, потому что из $\mathcal{A} \Rightarrow \mathcal{B}$. Это позволит оценить вероятность $\mathcal{A}$ сверху:

$$\mathcal{B} = \begin{cases} \exists a_1, \ldots, a_m \in \mathbb{N}_0\colon \forall i \; a_i < (1-\varepsilon)\log_2 n \\ \exists C_1, \ldots, C_m \subseteq V\colon \forall i \; |C_i| = a_i \\ \forall i \neq j \quad C_i \cap C_j = \varnothing \\ \forall x \notin (C_1 \cup \ldots \cup C_m) \; \forall i \; \exists y \in C_i\colon (x, y_i) \in E \end{cases}$$

Теперь объясним, что такое $\mathcal{B}$ и каково m. Что утверждает $\mathcal{A}$ в нашем алгоритме? А то, что он не смог найти одноцветное множество размера $\geqslant (1-\varepsilon)\log_2 n$. Значит, $\chi_g(G)$ лежит в полуинтервале $(n/(1-\varepsilon)\log_2 n; n]$. Положим $m = \lfloor n/\big(2(1-\varepsilon)\log_2 n\big)\rfloor$ и выберем произвольные m одноцветных множеств, найденных алгоритмом. Тогда, они удовлетворяют всем условиям выше, и, более того, верно следующее:

$$\sum_{i=1}^{m}|C_i| = \sum_{i=1}^{m}a_i < m(1-\varepsilon)\log_2 n \leqslant \frac{n}{2}$$

Теперь, начнём восстанавливать вероятность $P(\mathcal{B})$ по цепочке снизу-вверх:

- Пусть $x \notin (C_1 \cup \ldots \cup C_m)$ и $i \in \{1, \ldots, m\}$ уже зафиксированы, и мы хотим найти вероятность оставшегося:

$$P(\exists y \in C_i\colon (x, y_i) \in E) = 1 - P(\forall y \in C_i\ (x, y_i) \notin E) = 1 - \frac{1}{2^{a_i}}$$

- Теперь позволим i быть произвольным. Так как $\forall i \neq j\ C_i \cap C_j = \varnothing$, то выбор рёбер с разными множествами вообще никак не влияет друг на друга (иначе говоря, события предыдущего пункта независимы):

$$P(\forall i\ \exists y \in C_i\colon (x, y_i) \in E) = \prod_{i=1}^{m}\left(1 - \frac{1}{2^{a_i}}\right) < \left(1 - \frac{1}{2^{(1-\varepsilon)\log_2 n}}\right)^m = \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^m$$

- Теперь мы добавляем варьирование x. Вероятность тогда тоже будет считаться через произведение, ибо снова имеется независимость событий:

$$P(\forall x \notin (C_1 \cup \ldots \cup C_m)\ \forall i\ \exists y \in C_i\colon (x, y_i) \in E) \leqslant \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^{m\cdot\frac{n}{2}} = e^{\frac{mn}{2}\cdot\ln\left(1 - \frac{1}{n^{1-\varepsilon}}\right)} \leqslant$$
$$\exp\left(-\frac{mn}{2}\cdot\frac{1}{n^{1-\varepsilon}}\right) = \exp\left(-\frac{mn^{\varepsilon}}{2}\right)$$

  Для достаточно больших n мы можем сделать оценку на m:

$$m = \left\lfloor\frac{n}{2(1-\varepsilon)\log_2 n}\right\rfloor \geqslant \frac{n}{2\log_2 n}$$

  Из этого следует последняя оценка на вероятность с x:

$$P(\forall x \notin (C_1 \cup \ldots \cup C_m)\ \forall i\ \exists y \in C_i\colon (x, y_i) \in E) \leqslant \exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n}\right)$$

- Теперь высвободим строчку с $\exists C_1, \ldots, C_m \subseteq V \ldots$ Вероятность будем считать через дробление на отдельно взятые C_i (то есть для данного графа G мы перебираем все возможные C_i во вложенных суммах). Понятно, что вероятности всех «плохих» наборов подмножеств занулятся (ибо будет ложное утверждение в событии) и останутся только те, что подходят условиям на C_i. Итого:

$$P(\exists C_1, \ldots, C_m \subseteq V \ldots) < \exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n}\right) \sum_{\substack{C_1,\ldots,C_m\subseteq V:\\ \forall i\ |C_i|=a_i\\ \forall i\neq j\ C_i\cap C_j=\varnothing}} 1 =$$

$$\exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n}\right) \cdot C_n^{a_1} \cdot C_{n-a_1}^{a_2} \cdot \ldots \cdot C_{n-a_1-\ldots-a_{m-1}}^{a_m} <$$

$$\exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n}\right) \cdot n^{a_1+\ldots+a_m} \leqslant \exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n} + \frac{n\ln n}{2}\right)$$

▷ Ну и наконец-то мы приступаем к оценке $P(\mathcal{B})$:

$$P(\mathcal{B}) \leqslant \sum_{a_1=1}^{(1-\varepsilon)\log_2 n} \cdots \sum_{a_m=1}^{(1-\varepsilon)\ln n} \exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n} + \frac{n\ln n}{2}\right) <$$

$$(\log_2 n)^m \cdot \exp\left(-\frac{n^{1+\varepsilon}}{4\log_2 n} + \frac{n\ln n}{2}\right) \leqslant \exp\left(\frac{n\ln(\log_2 n)}{2(1-\varepsilon)\log_2 n} - \frac{n^{1+\varepsilon}}{4\log_2 n} + \frac{n\ln n}{2}\right) \xrightarrow[n\to\infty]{} 0$$

□

Гипотеза. *Не существует полиномиального алгоритма A, для которого имеет место предел:*

$$P\left(\frac{\alpha(G)}{\alpha_A(G)} < 2\right) \xrightarrow[n\to\infty]{} 1$$

Теорема 1.8. *(Кучéра, без доказательства)* $\forall \varepsilon > 0\ \forall \delta > 0\ \exists \{G_n\}_{n=1}^{\infty} \mid |V_n| = n$ *верно, что*

$$\frac{\#\text{перестановок множества вершин, при которых } \frac{\alpha(G_n)}{\alpha_g(G_n) \geqslant n^{1-\varepsilon}}}{n!} > 1 - \delta$$

Замечание. Говоря человеческим языком, доля перестановок вершин, на которых жадный алгоритм ошибается с заданной наперёд точностью, стремится к единице. То есть предыдущая теорема, несмотря на покрытие *почти всех* графов, оставляет за собой ещё очень много других, на которых жадный алгоритм будет работать плохо.

1.7 Модель Эрдёша-Рéньи случайных графов

Замечание. Также эту модель называют *биномиальной*. Она одна из самых старых и наиболее изученных.

Определение 1.5. *Моделью Эрдёша-Рéньи $G(n,p)$, $p \in (0;1)$* называется вероятностное пространство (Ω, F, P), где

▷ $\Omega = \{G = (\{1,\ldots,n\}, E)\} \Rightarrow |\Omega| = 2^{C_n^2}$

▷ $F = 2^{\Omega}$

▷ $\forall G = (\{1,\ldots,n\}, E) \in \Omega \quad P(G) := P(\{G\}) = p^{|E|} \cdot (1-p)^{C_n^2 - |E|}$

Замечание. Понятно, что p символизирует вероятность того, что в граф будет взято конкретное ребро.

Замечание. Важно отметить, что $G(n, p)$ используется за обозначение как самой модели Эрдёша-Ре́ньи, так и для обозначения случайного графа в этой модели.

Теорема 1.9. *Пусть в $G(n, p)$, где $p = p(n)$, причём $np(n) \to 0, n \to \infty$. Тогда*

$$P(\text{в } G(n,p) \text{ есть } \triangle) = P(X \geqslant 1) \xrightarrow[n\to\infty]{} 0$$

где $X(G)$ — это случайная величина, равная числу треугольников в графе G.

Доказательство. Воспользуемся классическим приёмом теории вероятностей: разобьём нашу случайную величину на сумму индикаторов:

$$X = X_1 + \ldots + X_{C_n^3}$$

где X_i — это индикатор того, что i-я тройка оказалась подграфом G. Тогда

$$\mathbb{E}X = \sum_{i=1}^{C_n^3} \mathbb{E}X_i = C_n^3 \cdot p^3 \sim \frac{(np)^3}{6} \xrightarrow[n\to\infty]{} 0$$

По неравенству Маркова отсюда моментально следует, что и $P(X \geqslant 1) \to 0$ при $n \to \infty$. □

Определение 1.6. Если в модели (Ω, F, P) вероятность события A зависит от параметра $n \in \mathbb{N}$, причём

$$\exists \lim_{n\to\infty} P(A) = 1$$

то говорят, что *событие A случится асимптотически почти наверное.*

Теорема 1.10. *Пусть в $G(n, p)$, где $p = p(n)$, причём $np(n) \to \infty, n \to \infty$. Тогда асимптотически почти наверное в $G(n, p)$ есть треугольники.*

Доказательство. Пусть $X(G)$ = число треугольников в G. Тогда $P(X \geqslant 1) = 1 - P(X \leqslant 0) = 1 - P(X = 0)$. Провернём такой цыганский фокус:

$$P(X \geqslant 1) = 1 - P(X \leqslant 0) = 1 - P(-X \geqslant 0) = 1 - P(\mathbb{E}X - X \geqslant \mathbb{E}X)$$

Теперь, мы можем в 2 действия получить оценку неравенством Чебышёва:

$$P(\mathbb{E}X - X \geqslant \mathbb{E}X) \leqslant P(|\mathbb{E}X - X| \geqslant \mathbb{E}X) \leqslant \frac{DX}{(\mathbb{E}X)^2} \Longrightarrow P(X \geqslant 1) \geqslant 1 - \frac{DX}{(\mathbb{E}X)^2}$$

При этом $DX = \mathbb{E}X^2 - (\mathbb{E}X)^2$, иначе говоря

$$P(X \geqslant 1) \geqslant 2 - \frac{\mathbb{E}X^2}{(\mathbb{E}X)^2}$$

Распишем матожидание квадрата:

$$\mathbb{E}X^2 = \mathbb{E}(X_1 + \ldots + X_{C_n^3})^2 = \mathbb{E}\left(X_1^2 + \ldots + X_{C_n^3}^2 + \sum_{i \neq j} X_i X_j\right) =$$
$$\underbrace{\mathbb{E}(X_1 + \ldots + X_{C_n^3})}_{\mathbb{E}X} + \sum_{i \neq j} \mathbb{E}(X_i \cdot X_j)$$

Чтобы посчитать сумму, нужно понять, что есть 3 возможных ситуации:

1. Тройки i, j пересекаются по двум элементам. Для суммы таких i, j получится выражение $(C_n^3 \cdot C_3^2 \cdot C_{n-3}^1) \cdot p^5$
2. Тройки i, j пересекаются по одному элементу. Получаем $(C_n^3 \cdot C_3^1 \cdot C_{n-3}^2) \cdot p^6$
3. Тройки i, j не пересекаются. Тогда $(C_n^3 \cdot C_{n-3}^3) \cdot p^6$

Теперь подставим всё в дробь:

$$\frac{\mathbb{E}X^2}{(\mathbb{E}X)^2} = \frac{\mathbb{E}X + 3p^5 C_n^3 C_{n-3}^1 + 3p^6 C_n^3 C_{n-3}^2 + p^6 C_n^3 C_{n-3}^3}{(C_n^3)^2 p^6} =$$
$$\frac{1}{\mathbb{E}X} + \frac{3C_{n-3}^1}{pC_n^3} + \frac{3C_{n-3}^2}{C_n^3} + \frac{C_{n-3}^3}{C_n^3} \xrightarrow[n\to\infty]{} 1$$

Этого нам и достаточно. □

Теорема 1.11. *(без доказательства) Пусть в $G(n, p)$, где $p = p(n)$, причём $np \to c > 0, n \to \infty$. Если $X(G)$ — это число треугольников в графе, то*

$$P(X = 0) \xrightarrow[n\to\infty]{} e^{-\frac{c^3}{6}}$$

Замечание. Можно заметить, что $\mathbb{E}X = C_n^3 p^3 \sim \frac{(np)^3}{6} \to \frac{c^3}{6}$, и поэтому в показателе экспоненты стоит ничто иное, как предел матожидания числа треугольников.

Определение 1.7. *Обхватом* (на англ. *girth*) графа G называется величина $g(G)$, равная длине минимального цикла в этом графе.

Утверждение 1.3. *Чтобы лучше прочувствовать определение, заметим такое интересное свойство:*

$$g(G) > 3 \Leftrightarrow \text{в } G \text{ нет треугольников}$$

Доказательство. Действительно, $g(G) > 3$ означает, что в графе не существует цикла длины 3, коим является треугольник. □

Теорема 1.12. *(1957, Эрдеш) Имеет место следующее утверждение:*

$$\forall k, l \in \mathbb{N}\ \exists G = (V, E) \mid \chi(G) > k,\ g(G) > l$$

Доказательство. $G(n, p)$, $p = p(n) = n^{\Theta - 1}$, $\Theta = \frac{1}{2l}$.

$X_l := X_l(G)$ — количество простых циклов длины $\leqslant l$ в G. Посмотрим матожидание:

$$\mathbb{E}X_l = \sum_{r=3}^{l} C_n^r \frac{(r-1)!}{2} p^r < \sum_{r=3}^{l} \frac{n^r}{r!} \cdot r! \cdot p^r = \sum_{r=3}^{l} (np)^r < l \cdot n^{\Theta l} = l \cdot n^{1/2}$$

По неравенству Маркова:

$$P\left(X_l \geqslant \frac{n}{2}\right) \leqslant \frac{\mathbb{E}X_l}{n/2} < \frac{l\sqrt{n}}{n/2} \xrightarrow[n\to\infty]{} 0$$

Отсюда возникает следствие:

$$\exists n_1 \mid \forall n \geqslant n_1 \; P\left(X_l < \frac{n}{2}\right) > \frac{1}{2}$$

Тут был замысел, около 16й минуты 7й лекции

$Y_x := Y_x(G)$ — количество независимых множеств размера x в G. Положим $x = \left\lceil \frac{3\ln n}{p} \right\rceil \sim \frac{3\ln n}{p}$. Оценим матожидание:

$$\mathbb{E}Y_x = C_n^x (1-p)^{C_x^2} \leqslant n^x \cdot e^{-pC_x^2} = \exp\left(x \ln n - (1+o(1)) p \frac{x^2}{2}\right)$$

Посмотрим асимптотику штуки в скобочках:

$$x\left(\ln n - (1+o(1))\frac{px}{2}\right) \sim x\left(\ln n - (1+o(1))\frac{3\ln n}{2}\right) \xrightarrow[n\to\infty]{} -\infty$$

Снова применим неравенство Маркова:

$$P(\alpha(G) \geqslant x) = P(Y_x \geqslant 1) \leqslant \frac{\mathbb{E}Y_x}{1} \to 0$$

Стало быть, $\exists n_2 \mid \forall n \geqslant n_2 \; P(\alpha(G) < x) > 1/2$. Теперь можно рассмотреть $n > \max\{n_1, n_2\}$, а потому

$$\exists G \mid X_l(G) < \frac{n}{2} \wedge \alpha(G) < x$$

Получим индуцированный граф G' из G путём удаления вершин и инциндентных им рёбер. Добьёмся того, что $X_l(G') = 0$, то есть $g(G') > l$. Чтобы убрать все простые циклы (по свойству G), нам потребуется удалить не более $n/2$ вершин. Иначе говоря, $|V(G')| > n - n/2 = n/2$. При этом, мы не могли улучшить число независимости почему?. Отсюда

$$\chi(G') > \frac{n}{2x} \sim \frac{np}{2 \cdot 3\ln n} = \frac{n^{\Theta}}{6\ln n} \to +\infty$$

По итогу $\exists n_3 \mid \forall n \geqslant n_3 \; \chi(G') > k$. □

Замечание. Теорема указывает на существование и количество таких графов (их большинство), а вот уже конструкцию никакую не даёт. Первая конструкция была придумана Ловасом Ласло примерно через 10 лет после публикации Эрдёша.

Утверждение 1.4. *Имеют место следующие утверждения:*

1. *Если $p(n) = o(1/n^2)$ и мы обозначим за $X(G)$ — число рёбер в G, то асимптотически почти наверное $X = 0$ и $\chi(G) = 1$*

2. *Пусть $p(n) = w(1/n^2)$ (по определению это значит, что $pn^2 \to +\infty$), а также $p = o(1/n)$. Тогда асимптотически почти наверное $\chi(G) = 2$.*

3. *Пусть $p = c/n$, $c < 1$. Тогда асимптотически почти наверное $\chi(G) = 3$. Более того, существует $\Psi = \Psi(n) = o(n)$ и асимптотически почти наверное $n - \Psi(n)$ вершин графа $G(n,p)$ принадлежат древесным комптонентам (то есть компонентам-деревьям), а $\Psi(n)$ вершин принадлежат унициклическим компонентам.*

Доказательство.

1. Посмотрим на матожидание X:

$$\mathbb{E}X = C_n^2 \cdot p \sim \frac{pn^2}{2} = o(1)$$

Тогда, мы можем оценить вероятность наличия рёбер:

$$P(X \geqslant 1) \leqslant \mathbb{E}X \to 0$$

Почему $\chi(G) = 1$ асимптотически почти наверное?

2. Оцени наличие рёбер в графе через известный трюк:

$$P(X \geqslant 1) \geqslant 1 - \frac{DX}{(\mathbb{E}X)^2}$$

где $DX = C_n^2 p(1-p)$. Стало быть

$$\frac{DX}{(\mathbb{E}X)^2} = \frac{C_n^2 p(1-p)}{(C_n^2 p)^2} = \frac{1-p}{p \cdot C_n^2} \xrightarrow[n\to\infty]{} 0$$

□

Теорема 1.13. *(80-е годы прошлого века, Бела Боллобаш) Пусть $p = n^{-\alpha}$, $\alpha \in (5/6; 1)$. Тогда существует функция $u(n,p)$ такое, что для $G(n,p)$ асимптотически почти наверное $\chi(G) \in \{u, u+1, u+2, u+3\}$.*

Упражнение. *$u(n,p) \to +\infty$ всегда.*

Замечание. Сейчас теорема улучшена до ограничений $\alpha \in (1/2; 1)$ и результата $\chi(G) \in \{u, u+1\}$.

Теорема 1.14. *(80-е годы прошлого века, Бела Боллобаш) Пусть $p = 1/2$. Тогда существует функция $\varphi(n) = o(n/\ln n)$ такая, что асимтотически почти наверное имеет место неравенство:*

$$\frac{n}{2\log_2 n} - \varphi(n) \leqslant \chi(G) \leqslant \frac{n}{2\log_2 n} + \varphi(n)$$

Замечание. Разница с предыдущей теоремой в том, что конкретно для $p = 1/2$ отрезок значений $\chi(G)$ будет сужаться. При этом заменить $\varphi(n)$ на константу нельзя, недавно было доказано, что $\varphi(n) = \Omega(\sqrt[4]{n})$.

1.7.1 Случайные блуждания

Теорема 1.15. *(Неравенство большого уклонения) Пусть $\xi_1, \ldots, \xi_n$ — одинаковые независимые случайные величины, $P\{\xi_i = 1\} = \frac{1}{2} = P\{\xi_i = -1\}$. Тогда утверждается, что*

$$\forall a > 0 \quad P(\xi_1 + \ldots + \xi_n \geqslant a) \leqslant e^{-\frac{a^2}{2n}}$$

Доказательство. Пусть $\lambda > 0$. Тогда имеют место следующие равенства:

$$P(\xi_1 + \ldots + \xi_n \geqslant a) = P(\lambda(\xi_1 + \ldots + \xi_n) \geqslant \lambda a) = P(e^{\lambda(\xi_1+\ldots+\xi_n)} \geqslant e^{\lambda a})$$

По неравенству Маркова почему можно применить?

$$P(e^{\lambda(\xi_1+\ldots+\xi_n)} \geqslant e^{\lambda a}) \leqslant e^{-\lambda a} \cdot \mathbb{E}(e^{\lambda(\xi_1+\ldots+\xi_n)}) = e^{-\lambda a} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda \xi_i}) = e^{-\lambda a} \left(\frac{e^{\lambda} + e^{-\lambda}}{2} \right)^n =$$
$$e^{-\lambda a} \left(\sum_{l=0}^{\infty} \frac{\lambda^{2l}}{(2l)!} \right)^n \leqslant e^{-\lambda a} \left(\sum_{l=0}^{\infty} \frac{\lambda^{2l}}{l! \cdot 2^l} \right)^n = \exp\left(-\lambda a + \frac{\lambda^2}{2} n \right)$$

Лучшая оценка будет при $\lambda = \frac{a}{n}$. Если подставить, то

$$P(\xi_1 + \ldots + \xi_n \geqslant a) \leqslant \exp\left(-\frac{a^2}{2n} \right)$$

□

Определение 1.8. Пусть дано вероятностное пространство $G(n, p)$. Случайная величина f называется *липшицевой по рёбрам*, если $|f(G) - f(G')| \leqslant 1$, коль скоро графы G и G' отличаются на 1 ребро.

Определение 1.9. Пусть дано вероятностное пространство $G(n, p)$. Случайная величина f называется *липшицевой по вершинам*, если $|f(G) - f(G')| \leqslant 1$, коль скоро графы G и G' отличаются в окрестности ровно одной вершины (то есть множества рёбер у какой-то вершины в графах G и G' различаются)

Замечание. Число рёбер графа является липшицевым по рёбрам, а хроматическое число графа — липшицево по вершинам.

Теорема 1.16. *(Неравенство Азумы, без доказательства) Пусть дано вероятностное пространство $G(n, p)$. Тогда:*

- ▷ *Если f — липшицева по рёбрам случайная величина, то $P(f - \mathbb{E}f \geqslant a) \leqslant \exp\left(-\frac{a^2}{2C_n^2}\right)$*
- ▷ *Если f — липшицева по вершинам случайная величина, то $P(f - \mathbb{E}f \geqslant a) \leqslant \exp\left(-\frac{a^2}{2(n-1)}\right)$*

Доказательство теоремы Боллобаша

Замечание. Далее и до конца доказательства мы живём в вероятностном пространстве $G(n, p)$, где $p = n^{-\alpha}$, $\alpha \in (5/6; 1)$

Лемма 1.5. *Утверждается, что*

$$\exists n_0 \mid \forall n \geqslant n_0 \quad P(\forall S \subset V\colon |S| \leqslant \sqrt{n}\ln n,\ \chi(G|_S) \leqslant 3) \geqslant 1 - \frac{1}{\ln n}$$

Доказательство. Надо показать, что

$$P(\exists S \subset V\colon |S| \leqslant \sqrt{n}\ln n,\ \chi(G|_S) > 3) = \frac{1}{\ln n}$$

Для этого мы усилим свойства нашего события (равенство ибо события тривиально эквивалентны), а затем воспользуемся неравенством для вероятности объединения: Пояснить равенство подробнее

$$\begin{aligned}
&P(\exists S \subset V\colon |S| \leqslant \sqrt{n}\ln n,\ \chi(G|_S) > 3) = \\
&\qquad P(\exists S \subset V\colon 4 \leqslant |S| \leqslant \sqrt{n}\ln n,\ \chi(G|_S) > 3,\ \forall x \in S\ \chi(G|_{S\setminus\{x\}})) \leqslant \\
&\qquad \sum_{s=4}^{\sqrt{n}\ln n} \sum_{\substack{S\subset V\\|S|=s}} P\underbrace{(\chi(G|_s) > 3 \wedge \forall x \in S\ \chi(G|_{S\setminus\{x\}}) \leqslant 3)}_{\Rightarrow |E(G|_S)|\geqslant 3s/2} \leqslant \\
&\sum_{s=4}^{\sqrt{n}\ln n} \sum_{\substack{S\subset V\\|S|=s}} C_{C_s^2}^{3s/2} \cdot p^{3s/2} = \sum_{s=4}^{\sqrt{n}\ln n} C_n^s C_{C_s^2}^{3s/2} p^{3s/2} \leqslant \sum_{s=4}^{\sqrt{n}\ln n} \left(\frac{ne}{s}\right)^s \left(\frac{C_s^2 \cdot e}{3s/2}\right)^{3s/2} p^{3s/2} < \\
&\qquad \sum_{s=4}^{\sqrt{n}\ln n} \left(\frac{ne}{s}\right)^s p^{3s/2} s^{3s/2} = \sum_{s=4}^{\sqrt{n}\ln n} \left(\frac{ne}{s} \cdot s^{3/2} p^{3/2}\right)^s \leqslant \sum_{s=4}^{\sqrt{n}\ln n} \left(n^{5/4}\sqrt{\ln n} \cdot e \cdot n^{-3\alpha/2}\right)^s
\end{aligned}$$

Настал час, когда выстрелит условие $p = n^{-\alpha}$, $\alpha \in (5/6; 1)$. Действительно, при таких α мы можем сказать, что $n^{5/4 - 3\alpha 2} = n^{-\beta}$, $\beta > 0$. Так как моном растёт быстрее корня из логарифма, то

$$\exists n_0 \mid \forall n \geqslant n_0 \quad \sum_{s=4}^{\sqrt{n}\ln n} \left(n^{5/4-3\alpha/2}\sqrt{\ln n} \cdot e\right)^s \leqslant \sum_{s=4}^{\sqrt{n}\ln n} n^{-\beta/2\cdot s}$$

Дописать неравенство и вывод □

Доказательство. (теоремы Боллобаша) Зафиксируем $\alpha \in (5/6; 1)$, по лемме найдём n_0 и рассмотрим $n \geqslant n_0$. Функцию $u(n, \alpha)$ надо взять таким образом:

$$u := u(n, \alpha) = \min\left\{t\colon P(\chi(G) \leqslant t) \geqslant 1 - \frac{1}{\ln n}\right\}$$

Тогда имеет место 2 неравенства:

- ▷ $P(\chi(G) \leqslant u - 1) < \frac{1}{\ln n}$, ибо u минимальное
- ▷ $P(\chi(G) > u) > 1 - \frac{1}{\ln n}$ — просто отрицание неравенства для u

Введём $Y(G)$ как минимальное число вершин, без которых граф G правильно красится не более чем u цветами:

$$Y(G) = \min\{k \in \mathbb{N}\colon \exists S \subseteq V, |S| = k \wedge \chi(G|_{V\setminus S}) \leqslant u\}$$

Если мы докажем, что $P\{Y \leqslant \sqrt{n}\ln n\} = 1$, то мы получим теорему. Заметим, что $Y(G)$ является липшицевой по вершинам случайной величиной пояснить. Стало быть, мы можем воспользоваться неравенством Азумы: положим $a = \sqrt{2(n-1)\ln\ln n}$. Предположим, что $\mathbb{E}Y \geqslant a$, но тогда

$$\frac{1}{\ln n} \leqslant P(\chi(G) \leqslant u) = P(Y \leqslant 0) \leqslant \underbrace{P(Y \leqslant \mathbb{E}Y - a)}_{P(Y-\mathbb{E}Y \leqslant -a)} \leqslant \exp\left(-\frac{a^2}{2(n-1)}\right) < \frac{1}{\ln n}$$

получили противоречие. Значит $\mathbb{E}Y < a$. Тогда

$$P(Y \geqslant 2a) \leqslant P(Y \geqslant \mathbb{E}Y + a) < \frac{1}{\ln n}$$

Дописать □

Доказательство. (доказательство теоремы 1.11 поправить ссылку на теорему) Мы уже знаем, что если $p = 1/2$, то асимптотически почти наверное $\alpha(G) \leqslant 2\log_2 n$, а для хроматического числа это даёт оценку $\chi(G) \geqslant n/(2\log_2 n)$ (этим мы уже доказали часть теоремы, оценка снизу).

Введём параметры Переписать. Сделать выбор интуитивным. $m = \lfloor n/\ln^2 n \rfloor$ и $f_m(k) = \mathbb{E}X_k$, где X_k - случайная величина, означающая число независимых множеств на k вершинах в графе $G(m; 1/2)$. Тогда $f_m(k) = \mathbb{E}X_k = C_m^k \cdot 2^{-C_k^2}$. Посмотрим, что происходит при разных k:

$$\begin{aligned}
&f_m(1) = C_m^1 \cdot 2^0 = m \xrightarrow[n\to\infty]{} +\infty \\
&f_m(2) = C_m^2 \cdot 2^{-1} \xrightarrow[n\to\infty]{} +\infty \\
&\vdots \\
&f_m(2\log_2 m) \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

Прокомментировать последний предел и существование некоторой границы, до которой стремимся в $+\infty$, а после в 0. Введём $k_0(m)$ таким образом:

$$k_0(m) = t \mid f_m(t) < 1,\ f_m(t-1) \geqslant 1$$

Тогда имеет место факт, что $k_0(m) \sim 2\log_2 m$, $m \to \infty$. □

10я лекция появится тут, когда я осознаю и поправлю всё предыдущее. Пока времени и сил нетути.