Занятие 4.

Вероятность в теории чисел. Метод второго момента.

Следующий результат, принадлежащий Эрдешу, использует довольно неожиданную вероятностную конструкцию.

Назовем множество $A \subset \mathbb{Z}$ *свободным от сумм*, если не найдется трех (не обязательно различных) элементов a, b, $c \in A$ таких, что $a + b = c$.

Теорема. Каждое множество B из n ненулевых целых чисел содержит свободное от сумм подмножество A, содержащее $|A| > n/3$ чисел.

Доказательство. Выберем простое число $p = 3k + 2$ так, что все элементы множества B меньше, чем k. Докажем, что можно найти $A \subset B$ требуемой мощности свободным от сумм по модулю p (то есть $a+b-c$ не кратно p для всех a, b, c в A). Заметим, что множество $C := \{k + 1, k + 2, \dots, 2k + 1\}$ свободно от сумм по модулю p. Выберем случайно остаток $a \in \{1, 2, \dots, p - 1\}$ по модулю p согласно равномерному распределению и рассмотрим случайную величину

$$X = |a \cdot C \cap B| = \sum_{c \in C} |\{ac\} \cap B|$$

(здесь мы отождествляем B с множеством остатков по модулю p). Заметим, что математическое ожидание случайной величины $|\{ac\} \cap B|$ равно $|B|/(p-1)$, поскольку числа ac пробегают при менябщемся a по разу все ненулевые остатки по модулю p. Таким образом, $E(X) = |C| \cdot |B|/(p-1) > |B|/3$. Значит, найдется такое a, что $|a \cdot C \cap B| > |B|/3$. В качестве A можно взять $a \cdot C \cap B$.

Упражнение. Если a_i $(i = 1, 2, \dots, s)$ — целые числа, то найдется такая константа $c > 0$, что в любом конечном множестве $B \subset \mathbb{Z}$ найдется подмножество $A \subset B$ такое, что $|A| \geq c|B|$ и $\sum a_i x_i \neq 0$ для любых $x_i \in A$.

Перейдем от комбинаторной теории чисел к "настоящей".

Рассмотрим следующий вопрос: как распределено количество $\nu(k)$ различных простых делителей наугад выбранного от 1 до n натурального числа k?

Оказывается, что это количество почти для всех чисел k величина $\nu(k)$ мало отличается от $\ln \ln n$.

Точное утверждение дается следующей теоремой:

Теорема. Пусть $w(n) \to \infty$ произвольно медленно. Тогда количество тех $k \in \{1, 2, \dots, n\}$, для которых $|\nu(k) - \ln \ln n| > w(n)\sqrt{\ln \ln n}$, есть $o(n)$.

Нам понадобится следующее теоретико-числовое утверждение:

Лемма. $\sum_{p \leq n} 1/p = \ln \ln n + O(1)$, где суммирование производится по всем простым числам $p \leq n$.

Доказательство леммы. Сначала вычислим асимптотику суммы $\sum_{p\le n}\ln p/p$. Для этого разложим на простые множители число $n!$:

$$n! = \prod_{p\le n} p^{c_p},$$

где произведение берется по простым p и $c_p = [n/p]+[n/p^2]+[n/p^3]+\dots$. Имеем

$$n/p-1 \le [n/p] \le c_p \le n/p+n/p^2+\dots = n/(p-1).$$

Подставляя эти значения в выражение для $n!$ и логарифмируя, получаем

$$n\sum\frac{\ln p}{p}-\sum\ln p \le n! \le n\sum\frac{\ln p}{(p-1)} = n\sum\frac{\ln p}{p}+n\sum\frac{\ln p}{p(p-1)}$$

(суммирование везде по простым $p\le n$). Заметим, что $(n/e)^n < n! < n^n$ (левое неравенство устанавливается, например, по индукции). Кроме того, $\prod_{p\le n} p \le 4^n$. Это следует из того, что указанное произведение делит $C_n^{n/2}\cdot C_{n/2}^{n/4}\cdot C_{n/4}^{n/8}\dots$ (подробности, связанные с тем, что n может не быть степенью двойки, опускаются). Наконец, заметим, что $\sum \ln p/(p^2-p) = O(1)$. Резуюмируя получаем, что

$$S_n := \sum_{p\le n}\frac{\ln p}{p} = \frac{1}{n}\ln n! + O(1) = \ln n + O(1).$$

Выразим интересующую нас сумму $\Sigma_n = \sum_{p\le n} 1/p$ через S_1, S_2, $\dots$, S_n. Имеем

$$\Sigma_n = \sum_{p\le n} 1/p = \sum_{k=1}^{n}\frac{S_k-S_{k-1}}{\ln k} = \sum_{k=1}^{n} S_k\left(\frac{1}{\ln k}-\frac{1}{\ln(k+1)}\right)+S_n\ln(n+1).$$

Заметим, что

$$S_k\left(\frac{1}{\ln k}-\frac{1}{\ln(k+1)}\right) = (\ln(k+1)+O(1))\left(\frac{1}{\ln k}-\frac{1}{\ln(k+1)}\right).$$

Слагаемые типа $O(1)(\frac{1}{\ln k}-\frac{1}{\ln(k+1)})$ дадут в сумме $O(1)$. Также $S_n\ln(n+1) = O(1)$. Далее,

$$\ln(k+1)\left(\frac{1}{\ln k}-\frac{1}{\ln(k+1)}\right) = \frac{\ln(k+1)}{\ln k}-1 = \ln\frac{\ln(k+1)}{\ln k}+O\left(\left(\frac{\ln(k+1)}{\ln k}-1\right)^2\right),$$

мы воспользовались формулой Тейлора $x-1=\ln x+O((x-1)^2)$, $x\to 1$. Заметим, что $\ln(k+1)-\ln(k)=\ln(1+1/k)=O(1/k)$, так что $\left(\frac{\ln(k+1)}{\ln k}-1\right)^2=o(1/k^2)$. Значит, соответствующие поправки $O\left(\left(\frac{\ln(k+1)}{\ln k}-1\right)^2\right)$ также дадут в сумме $O(1)$.

Итого

$$\Sigma_n=\sum_{k=2}^{n}\ln\frac{\ln(k+1)}{\ln k}+O(1)=\ln\ln n+O(1).$$

Лемма доказана.

Доказательство теоремы.

Пусть k выбирается в множестве $\{1,2,\dots,n\}$ согласно равномерному распределению. Определим для каждого простого $p\le n$ случайную величину X_p как индикатор события $\{k \text{ кратно } p\}$. Тогда $\nu=\sum_p p\le nX_p$ и требуется доказать, что

$$\mathsf{P}\left(|\nu-\ln\ln n|>w(n)\sqrt{\ln\ln n}\right)\to 0 \qquad (1)$$

Пусть $M=n^{1/5}$ и $\nu_1=\sum_{p\le M}X_p$. Заметим, что $\nu_1\le\nu\le\nu_1+4$ (так как любое число k от 1 до n имеет не более четырех простых множителей, больших, чем M). Поэтому можно заменить в (1) ν на ν_1.

Заметим, что

$$E(\nu)=n^{-1}\sum_{p\le n}[n/p]=\sum_{p<n}1/p+O(1)=\ln\ln n+O(1)$$

в силу нашей леммы. Таким образом, можно заменить также $\ln\ln n$ на $E\nu_1=E\nu+O(1)$.

Для оценки $\mathsf{P}(|\nu_1-E\nu_1|>w(n)\sqrt{\ln\ln n})$ воспользуемся неравенством Чебышева

$$\mathsf{P}\left(|\nu_1-E\nu_1|>w(n)\sqrt{\ln\ln n}\right)\le\frac{\operatorname{Var}\nu_1}{w^2(n)\ln\ln n},$$

где Var обозначает дисперсию случайной величины.

Таким образом, достаточно доказать, что $Var\nu_1=O(\ln\ln n)$.

Для этого заметим, что

$$\operatorname{Var}\nu_1=\sum_{p\le M}\operatorname{Var}X_p+2\sum_{p<q\le M}\operatorname{Cov}(X_p,X_q),$$

где $\operatorname{Cov}(X,Y)=E(X\cdot Y)-E(X)\cdot E(Y)$ обозначает ковариацию случайных величин X и Y. Это равенство получается из формулы $E(X)=E(X^2)-(EX)^2$ путем раскрытия скобок.

Заметим, что $\operatorname{Var} X_p \leq E(X_p)^2 = E(X_p)$, так что первая сумма не превосходит $E(\nu_1) = \ln\ln n + O(1)$.

Оценим $\operatorname{Cov}(X_p, X_q)$:

$$\operatorname{Cov}(X_p, X_q) = \frac{[n/pq]}{n} - \frac{[n/p]}{n} \cdot \frac{[n/q]}{n} \leq \frac{1}{pq} - (1/p - 1/n)(1/q - 1/n) < 1/n(1/p + 1/q) \leq 2/n.$$

Поскольку количество пар различных p, q не превосходит $M^2 < n$, получаем, что сумма ковариаций есть $O(1)$.

Доказательство теоремы завершено.

Использование неравенства Чеьышева называется *методом второго момента.*

Приведем пример из теории случайных графов.

Сначала рассмотрим такую ситуацию: имеются случайные величины $X_1, X_2, \ldots, X_m$, являющиеся индикаторами некоторых событий $A_1, A_2, \ldots, A_m$ (то есть $X_i = 1$, если происходит событие A_i и $X_i = 0$ в противном случае). Положим $X = \sum X_i$. Как оценить сверху вероятность того, что $X = 0$? Воспользуемся неравенством Чебышева:

$$\mathsf{P}\{X = 0\} \leq \mathsf{P}\{|X - E(X)|^2 \geq E(X)^2\} \leq \frac{\operatorname{Var}(X)}{E(X)^2}. \tag{1}$$

Теперь оценим $\operatorname{Var}(X)$. Будем писать $i \sim j$, если $i \neq j$ и события A_i, A_j не являются независимыми. Имеем

$$\operatorname{Var}(X) = \sum \operatorname{Var}(X_i) + 2\sum_{i \sim j} \operatorname{Cov}(X_i, X_j) \leq E(X) + 2\sum_{i \sim j} E(X_i \cdot X_j). \tag{2}$$

Теперь докажем следующее типичное утверждение теории случайных графов:

Теорема. Рассмотрим случайный граф $G(n, p)$ на n вершинах, в котором каждое ребро проводится случайно независимо от других ребер с вероятностью p. Тогда при $p = o(n^{-2/3})$ вероятность наличия в графе $G(n, p)$ клики на четырех вершинах стремится к нулю, а при $p \gg n^{-2/3}$ (то есть $\lim p/n^{-2/3} = +\infty$) вероятность наличия такой клики стремится к бесконечности.

Доказательство. Пронумеруем четверки вершин нашего графа индексом i. Пусть A_i — вероятность того, что вершины i-ой четверки образуют клику, X_i — идикатор события A_i, $X = \sum X_i$. Имеем $E(X) = p^6 \cdot C_n^4$, так что при $p \ll n^{-2/3}$ имеем $E(X) = o(1)$ и вероятность того, что $X \geq 1$ (то есть вероятность наличия клики) стремится к нулю. При $p \gg n^{-2/3}$ имеем $E(X) \to +\infty$. Докажем, что $\operatorname{Var}(X) = o(E(X)^2)$ — из этого в силу (1) будет следовать требуемое. Оценим $\mathrm{Var}(X)$ как в (2). Первое слагаемое $E(X)$ есть $o(E(X)^2)$. Оценим сумму $E(X_i X_j) = \mathsf{P}(A_i \cap A_j)$ для всех пар зависимых событий. Для каждого из $O(n^4)$ событий имеется $O(n^2)$ зависимых событий, соответствующих четверкам, пересекающимся по двум вершинам, и $O(n)$ событий, соответствующих четверкам, пересекающимся по трем вершинам. В первом случае вероятность одновременного выполнения событий есть p^{11}, во втором — p^9. Итак,

$$\sum_{i \sim j} \mathsf{P}(A_i \cap A_j) = O(n^6 p^{11} + n^5 p^9) = o(n^8 p^{12}) = o(E(X)^2),$$

что завершает доказательство.