4.3. Теорема о распределении простых чисел. Обозначим через $\pi(x)$ количество простых чисел, не превосходящих x. Из бесконечности множества простых чисел вытекает, что $\pi(x) \to \infty$, $x \to \infty$. Уже долгое время самые сильные математики задаются вопросом: а какие другие свойства имеет функция $\pi(x)$? В частности, с какой скоростью $\pi(x)$ стремится к бесконечности? Понятно, что $\pi(x)$ меньше, чем x, но насколько меньше? Другими словами, какая часть натуральных чисел являются простыми?

Следующий результат называется *теоремой о распределении простых чисел*:

$$\pi(x) \sim \frac{x}{\ln x}, \qquad x \to \infty. \tag{12}$$

Доказать этот результат удалось только в конце XIX столетия сразу двум математикам Жаку Адамару и Шарлю Валле-Пуссену, которые работали независимо один от другого.

4.4. Гипотеза Харди. Два числа p и $p+2$ называются *близнецами*, если они оба являются простыми. Например, близнецами являются 3 и 5, 5 и 7, 11 и 13. Вопрос о том является ли множество пар чисел-близнецов бесконечным остается без ответа до сих пор, хотя специалисты склоняются к гипотезе о его бесконечности. Приведем (не строгое) обоснование этой гипотезы, которое использует идею рандомизации.

Согласно (12), простых чисел в множестве $\{2, 3, \ldots, x\}$ примерно в $\ln x$ раз меньше, чем натуральных. Отсюда мы "делаем" вывод, что натуральное число x с вероятностью $\frac{\pi(x)}{x} \sim \frac{1}{\ln x}$ является простым, т.е.

$$\mathsf{P}(x \in \mathbb{P}) = \frac{\pi(x)}{x} \sim \frac{1}{\ln x}. \tag{13}$$

Аналогично $\mathsf{P}(x+2 \in \mathbb{P}) \sim 1/\ln(x+2) \sim 1/\ln(x)$. Если бы события $x \in \mathbb{P}$ и $x+2 \in \mathbb{P}$ были независимы, то

$$\mathsf{P}(x \in \mathbb{P}, x+2 \in \mathbb{P}) = \mathsf{P}(x \in \mathbb{P})\mathsf{P}(x+2 \in \mathbb{P}) \sim \frac{1}{\ln^2(x)}.$$

С другой стороны, обозначим через $\pi_2(x)$ количество близнецов, не превосходящих x. Тогда

$$\mathsf{P}(x \in \mathbb{P}, x+2 \in \mathbb{P}) = \frac{\pi_2(x)}{x}$$

Сравнивая это равенство с предыдущим соотношением, заключаем, что $\pi_2(x) \sim x/\ln^2(x)$.

Полученный результат подсказывает какой является асимптотика $\pi_2(x)$. Однако события $x \in \mathbb{P}$ и $x+2 \in \mathbb{P}$ не являются независимыми, [6] поэтому наше рассуждение необходимо подправить. Согласно формуле умножения вероятности

$$\begin{aligned} &\mathsf{P}(x \in \mathbb{P}, x+2 \in \mathbb{P}) \\ &\qquad = \mathsf{P}(x \in \mathbb{P})\mathsf{P}\left(x+2 \in \mathbb{P} / x \in \mathbb{P}\right) \\ &\qquad = \mathsf{P}(x \in \mathbb{P})\mathsf{P}(x+2 \in \mathbb{P}) \cdot \frac{\mathsf{P}\left(x+2 \in \mathbb{P} / x \in \mathbb{P}\right)}{\mathsf{P}(x+2 \in \mathbb{P})}. \end{aligned} \tag{14}$$

Асимптотика первых двух сомножителей уже найдена выше. Для “вычисления” третьего сомножителя введем события $D_{p,y}$ при $p \in \mathbb{P}$, $y \in \mathbf{N}$:

$$D_{p,y} \overset{\text{def}}{=} \{p \ \texttt{не является делителем}\ y\} = \{p \nmid y\}.$$

Вычислим вероятность $\mathsf{P}(D_{p,x+2})$. Для этого разобьем натуральные числа на классы остатков от деления на p (классы конгруэнтности):

$$\mathbb{N} = C_0 \cup C_1 \cup \cdots \cup C_{p-1}, \quad \text{где} \quad C_k = \{z \in \mathbf{N} : z \mod p = k\}.$$

Так как событию $D_{p,x+2}$ благоприятствуют классы $C_1, \dots,$ C_{p-1} среди возможных $C_0, C_1, \dots, C_{p-1}$, то

$$\mathsf{P}(D_{p,x+2}) = \frac{p-1}{p}.$$

Отметим следующее свойство: *если $p \neq 2$, то остаток от деления $x+2$ на p отличается от остатка от деления x на p.* Действительно, если бы остатки были равны, то $x = \alpha p + \beta$ и $x + 2 = \gamma p + \beta$ для некоторых целых неотрицательных чисел α, β и γ. Отсюда $2 = (\gamma - \alpha)p$, что невозможно в силу $p \neq 2$.

Теперь “вычислим” вероятность $\mathsf{P}\left(D_{p,x+2} \,/ D_{p,x}\right)$. Интерпретируя условную вероятность как безусловную вероятность на суженом пространстве $D_{p,x}$, получаем из предыдущего замечания, что при $p > 2$

$$\mathsf{P}\left(D_{p,x+2} \,/ D_{p,x}\right) = \frac{|D_{p,x+2} \cap D_{p,x}|}{|D_{p,x}|} = \frac{p-2}{p-1},$$

поскольку $|D_{p,x+2} \cap D_{p,x}| = p - 2$, $|D_{p,x}| = p - 1$. Если же $p = 2$, то $\mathsf{P}\left(D_{p,x+2} \,/ D_{p,x}\right) = 1$, а $\mathsf{P}(D_{p,x}) = \frac{1}{2}$, так как имеются только два класса конгруєнтности, а благоприятствует только один.

Теперь можно “подсчитать” третий сомножитель в равенстве (14). Прежде всего,

$$\{x + 2 \in \mathbb{P}\} = \bigcap_{p \in \mathbb{P}} D_{p,x+2}.$$

“Поскольку” события $D_{p,x+2}$, $p \in \mathbb{P}$, независимы, [7] то

$$\mathsf{P}(x + 2 \in \mathbb{P}) = \prod_{p \in \mathbb{P}} \mathsf{P}(D_{p,x+2}) = \prod_{p=1}^{\infty} \frac{p-1}{p}.$$

“Аналогично”, события $D_{p,x} \cap D_{p,x+2}$, $p \in \mathbb{P}$, независимы ⑧
и поэтому

$$\mathsf{P}\left(x+2 \in \mathbb{P} \,/\, x \in \mathbb{P}\right) = \prod_{p \in \mathbb{P}} \mathsf{P}\left(D_{p,x+2} \,/\, D_{p,x}\right) = \prod_{p=2}^{\infty} \frac{p-2}{p-1}$$

(см. задачу 8). Окончательно

$$\frac{\mathsf{P}\left(x+2 \in \mathbb{P} \,/\, x \in \mathbb{P}\right)}{\mathsf{P}(x+2 \in \mathbb{P})} = 2 \prod_{p>2} \frac{(p-2)/(p-1)}{(p-1)/p} = 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right).$$

Теперь, используя результат, полученный в предположении о независимости событий $x \in \mathbb{P}$ и $x+2 \in \mathbb{P}$, получаем

$$\pi_2(x) \sim \frac{x}{\ln^2 x} \cdot 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right). \tag{15}$$

Значение константы

$$C_2 \stackrel{\text{def}}{=} 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \tag{16}$$

можно вычислить приближенно: $C_2 \approx 0.66$.

Формула (15) означает, что от 1 до x чисел-близнецов примерно в $(\ln x)/C_2$ раз меньше, чем простых чисел.

Проверим точность аппроксимаций $\pi(x)$ и $\pi_2(x)$ правыми частями соотношениями (12) и (15), например, при $x = 100$. Согласно (12) и (15) имеем $\pi(x) \approx 22$ и $\pi_2(x) \approx 9$, в то время как на самом деле $\pi(x) = 25$ и $\pi_2(x) = 8$.

Наилучшим из известных достижений в поисках асимптотики чисел-близнецов является оценка

$$\pi_2(x) \le \kappa \cdot C_2 \cdot \frac{x}{\ln^2(x)}, \qquad \kappa = 7. \tag{17}$$