Hustota prvočísel

Nechť $\pi(n)$ značí počet prvočísel menších nebo rovných n. Platí

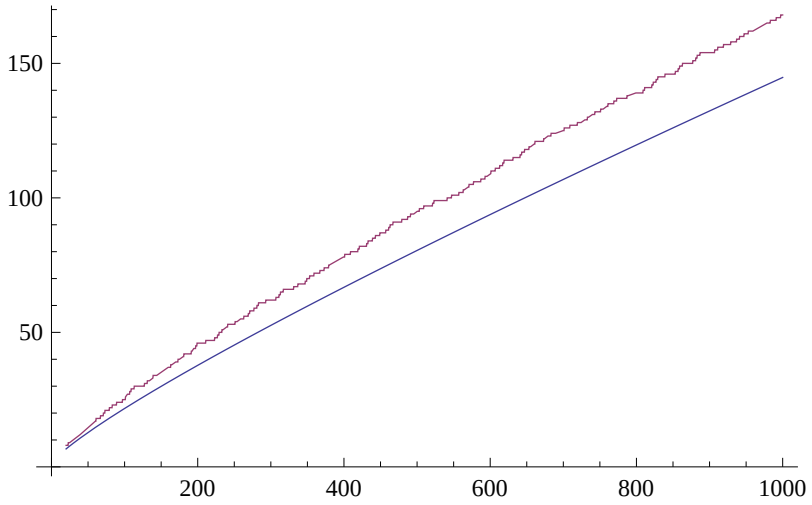
$$\lim_{n\to\infty} \frac{\pi(n)\cdot \log n}{n} = 1,$$

kde log značí přirozený logaritmus.

Chování funkce $\pi(n)$ nás bude zajímat pro velká n. Pro malá n je možné spočítat $\pi(n)$ přesně. Pro $n < 20$ máme:

n	$\frac{n}{\log n}$	$\pi(n)$	n	$\frac{n}{\log n}$	$\pi(n)$
2	2.9	1	11	4.6	5
3	2.7	2	12	4.8	5
4	2.9	2	13	5.1	6
5	3.1	3	14	5.3	6
6	3.3	3	15	5.5	6
7	3.6	4	16	5.8	6
8	3.8	4	17	6.0	7
9	4.1	4	18	6.2	7
10	4.3	4	19	6.5	8

a pro $20 \le n \le 1000$ ukazuje vztah $\frac{n}{\log n}$ a $\pi(n)$ následující graf



podle nějž je tedy $\frac{n}{\log n}$ spodním odhadem počtu prvočísel.

Prvočísel není příliš mnoho. Pro velká n ukážeme horní odhad počtu prvočísel

$$\pi(n) \le \frac{3n}{\log n}.$$

Základem je Čebyševův odhad (dokázaný níže), podle nějž platí

$$\prod_{p\le n} p < 4^n.$$

Otázka zní, jak z Čebyševova odhadu dostat informaci o $\pi(n)$. Přirozené je výraz zlogaritmovat:

$$\sum_{p\le n} \log p < n \log 4.$$

Vzhledem k tomu, že se zabýváme vztahem

$$\pi(n) \log n \sim n,$$

potřebujeme odhadnout vztah mezi $\log n$ a aritmetickým průměrem $\log p$.

Součet logaritmů prvočísel menších než n můžeme zdola odhadnout součtem logaritmů prvočísel mezi $\sqrt{n}$ a n, jejichž logaritmus odhadneme logaritmem $\sqrt{n}$. Tedy

$$\begin{aligned} n \log 4 > \sum_{p\le n} \log p > \sum_{\sqrt{n}\le p\le n} \log p > (\pi(n) - \pi(\sqrt{n})) \log \sqrt{n} > \\ > (\pi(n) - \frac{3\sqrt{n}}{\log \sqrt{n}}) \log \sqrt{n}. \end{aligned}$$

Úpravou získáváme

$$\pi(n) \log n < n \cdot 2 \log 4 + 6\sqrt{n}.$$

Tím jsme hotovi, protože pro velká n platí

$$n \cdot 2 \log 4 + 6\sqrt{n} < 3n.$$

Konkrétně to platí pro všechna $n > 696$. Pro menší n je možné (a kvůli indukci nutné) vztah ověřit přímo (viz graf a tabulka na první straně).

Důkaz Čebyševova odhadu. Vyjdeme ze vztahu

$$2^n = \sum_{i=0}^{n} \binom{n}{i},$$

který plyne např. z binomické věty pro $(1+1)^n$ nebo z toho, že $\binom{n}{j}$ je počet binárních čísel délky n s j jedničkami a počet všech čísel délky n je 2^n. Pro $n = 2k+1$ z toho dostáváme

$$2^{2k+1} \ge \binom{2k+1}{k} + \binom{2k+1}{k+1} = 2 \cdot \binom{2k+1}{k},$$

a tedy

$$2^{2k} \ge \binom{2k+1}{k}.$$

Nyní už můžeme dokázat Čebyševův odhad indukcí. Součin prvočísel menších než $2k+1$ rozdělíme na dvě části:

$$\prod_{p\le 2k+1} p = \prod_{2}^{k+1} p \cdot \prod_{k+2}^{2k+1} p\,.$$

První sčítanec odhadneme pomocí indukčního předpokladu. Pro druhý platí

$$\prod_{k+2}^{2k+1} p < \frac{(k+2)\cdot(k+3)\cdots 2k\cdot(2k+1)}{1\cdot 2\cdots(k-1)\cdot k} = \binom{2k+1}{k} \le 2^{2k}.$$

Celkem tedy

$$\prod_{p\le 2k+1} p = \prod_{2}^{k+1} p \cdot \prod_{k+2}^{2k+1} p < 4^{k+1}\cdot 4^k = 4^{2k+1}.$$

Pro sudá složená čísla máme

$$\prod_{p\le 2k+2} p = \prod_{p\le 2k+1} p < 4^{2k+1} < 4^{2k+2}.$$

□

Prvočísel není příliš málo. Zamysleme se nad kombinačním číslem

$$\binom{2k}{k} = \frac{(k+1)\cdot(k+3)\cdots(2k-1)\cdot 2k}{1\cdot 2\cdots(k-1)\cdot k}.$$

Je vidět, že pro zvolené m obsahují čitatel a jmenovatel zhruba stejné množství násobků m, totiž

$$\left\lfloor \frac{k}{m} \right\rfloor.$$

Pro jmenovatel je tento počet přesný, čitatel může obsahovat o jeden násobek víc. Např. pro $k = 17$ obsahuje čitatel tři násobky šesti $(18, 24, 30)$, zatímco jmenovatel jen dva $(6, 12)$.

Platí, že čitatel obsahuje

$$\left\lfloor \frac{2k}{m} \right\rfloor - \left\lfloor \frac{k}{m} \right\rfloor$$

násobků m. Nechť $k = hm + z$, kde $z < m$. Pak

$$\left\lfloor \frac{2k}{m} \right\rfloor = 2h + \left\lfloor \frac{2z}{m} \right\rfloor,$$

takže počet násobků m je v čitateli o jednu větší než ve jmenovateli, právě když $z \ge m/2$; jinak je stejný.

Předchozí úvahy jsou inspirací pro následující tvrzení.

Lemma. Označme

$$v = \mathrm{val}_p\binom{2k}{k}.$$

Pak

$$p^v \le 2k.$$

Důkaz. Na základě předchozích úvah lze důkaz neformálně vyjádřit tak, že každá mocnina p^j přispěje do p-valuace nejvýše jedničkou. a to právě když $k \mod p^j \ge p^j/2$.

Podrobněji: Platí

$$\mathrm{val}_p\binom{2k}{k} = \mathrm{val}_p((2k)!) - 2\mathrm{val}_p(k!),$$

přičemž pro každé n máme

$$\mathrm{val}_p(n!) = \sum_{j=1}^{\log_p n} \left\lfloor \frac{n}{p^j} \right\rfloor.$$

Máme tedy

$$\mathrm{val}_p\binom{2k}{k} = \sum_{j=1}^{\log_p 2k} \left\lfloor \frac{2k}{p^j} \right\rfloor - 2\left\lfloor \frac{k}{p^j} \right\rfloor,$$

přičemž

$$\left\lfloor \frac{2k}{p^j} \right\rfloor - 2\left\lfloor \frac{k}{p^j} \right\rfloor$$

je buď 0, pokud $k \mod p^j < p^j/2$, nebo 1, pokud $k \mod p^j \geq p^j/2$. □

Důsledkem předchozího lematu je odhad

$$\binom{2k}{k} \leq (2k)^{\pi(2k)}.$$

Ze vztahu

$$2^{2k} = \sum_{i=0}^{2k} \binom{2k}{i},$$

plyne

$$\frac{2^{2k}}{2k+1} \leq \binom{2k}{k},$$

protože $\binom{2k}{k}$ je ze všech kombinačních čísel v sumě největší. Dostáváme tedy základ pro dolní odhad pro $\pi(n)$

$$\frac{2^{2k}}{2k+1} \leq (2k)^{\pi(2k)}.$$

Po zlogaritmování vychází

$$\pi(2k) \geq \frac{2k}{\log 2k} \cdot \left(\log 2 - \frac{\log(2k+1)}{2k}\right),$$

přičemž pro $k = 20$ už je $\log 2 - \frac{\log(2k+1)}{2k} > 0.6$.

Prvočísla jsou rozložena rovnoměrně. Ukážeme pouze tzv. Bertrandův postulát, podle nějž mezi n a $2n$ vždy leží nějaké prvočíslo, a jeho zobecnění, podle nějž je mezi n a $2n$ alespoň r prvočísel pro každé $n \geq n_r$.

Uvažujme opět číslo

$$m := \binom{2k}{k}.$$

Platí zřejmě

$$m = \prod p^{\mathrm{val}_p(m)}.$$

Rozdělme prvočísla menší než $2k$ do čtyř intervalů:

- $I_1 :\ p \leq \sqrt{2k}$, kde použijeme odhad
$$p^{\mathrm{val}_p(m)} < 2k.$$
- $I_2 :\ \sqrt{2k} < p \leq \frac{2}{3}k$, kde už platí $p^2 > 2k$, takže $\mathrm{val}_p(m)$ je nejvýše jedna (viz úvahy o počtu násobků p ve jmenovateli a čitateli).
- $I_3 :\ \frac{2}{3}k < p \leq k$; v tomto intervalu je $k \mod p < p/2$, takže $\mathrm{val}_p(m) = 0$.
- $I_4 :\ k < p \leq 2k$; počet prvočísel v tomto intervalu, který nás zajímá, označme r.

Dostáváme

$$\binom{2k}{k} < (2k)^{\sqrt{2k}} \cdot \prod_{p \in I_2} p \cdot (2k)^r.$$

Za použití odhadu

$$\frac{2^{2k}}{2k+1} < \binom{2k}{k}$$

a Čebyševova odhadu

$$\prod_{p \in I_2} p < \prod_{p \leq \frac{2}{3}k} p < 4^{\frac{2}{3}k}$$

dostáváme

$$\frac{2^{2k}}{2k+1} < (2k)^{\sqrt{2k}} \cdot 4^{\frac{2}{3}k} \cdot (2k)^r,$$

tedy

$$4^{\frac{k}{3}} < (2k+1)(2k)^{\sqrt{2k}+r-1} < (2k)^{\sqrt{2k}+r+2}.$$

Odtud po zlogaritmování dostáváme

$$r > \frac{\log 4}{3} \cdot \frac{k}{\log(2k)} - \sqrt{2k} - 2,$$

což je rostoucí funkce, která garantuje $r > 0$ pro $k \geq 507$. Bertrandův postulát pro menší čísla lze snadno ověřit posloupností prvočísel 3, 5, 7, 13, 23, 43, 83, 163, 317 a 631.