About the distribution of the prime numbers:
elementary approaches.

Introduction:

The prime numbers are those integers that can only be divided by themselves and 1.

While prime numbers seem to be distributed irregularly along the number line, they exhibit fascinating patterns and properties upon closer study. The Prime Number Theorem offers a mathematical description of this distribution, indicating that primes become less frequent as numbers grow larger, yet they follow a predictable pattern, revealing a deeper order within their seemingly random arrangement.

The prime number theorem tells us something about how the prime numbers are distributed among the other integers.

Definition 1: A positive integer number is called a *prime number* if it is greater than 1 and can only be divided by itself and 1. The integer numbers greater than 1 that are not prime are called *composite* numbers which can be written as a product of prime numbers in a unique way, for example $30 = 2 \times 3 \times 5$.

Definition 2: We define $\pi(x)$ the prime-counting function that gives the number of primes less than or equal to a real number x,

$$\pi(x) = \sum_{p \leq x} 1$$

The Prime Number Theorem.

Theorem 1: (Euclid, 300 BC). The sequence of the prime numbers $p_1 = 2 < p_2 = 3 < p_3 = 5 < \cdots < p_n < \cdots$ is infinite and for any integer $n \geq 1$, we have this inequality:

$$p_n \leq 2^{2^{n-1}}$$

No matter how far you go along the number line, you will always find more primes.

Proof: By induction $p_1 = 2 \leq 2^{2^0}$ and $p_2 = 3 \leq 2^2$ satisfy the inequality.

Assuming for any integer $n \geq 3$, and for any $k = 1, 2 \ldots, n-1$ we have $p_k \leq 2^{2^{k-1}}$.

The integer number $p_1.p_2.p_3 \ldots p_{n-1} - 1$ is not divisible by any prime number $p_1, p_2, p_3, \ldots, p_{n-1}$.

Therefore $\quad p_n \leq p_1.p_2.p_3 \ldots p_{n-1} \leq 2^{2^0+2^1+\cdots+2^{n-2}} = 2^{2^{n-1}-1} \leq 2^{2^{n-1}}$

This theorem shows that the sequence of the prime numbers is infinite, and $\pi(x)$ has a lower bound. $\square$

Corollary 2: For any number $x \geq 2$, we have:

$$\pi(x) \geq \ln(\ln(x)).$$

Proof: From Euclid's theorem, we have $\pi\left(2^{2^{n-1}}\right) \geq n.$

Let $x \geq 2$, consider the number integer n defined by $n = 1 + \left\lfloor \frac{\ln\left(\frac{lnx}{ln2}\right)}{ln2} \right\rfloor$, we take the integer part of the fraction.

The integer n satisfies the following double inequality: $2^{2^{n-1}} \leq x \leq 2^{2^n}$.

As the function $x \to \pi(x)$ is increasing, we deduce the following: $\pi(x) \geq \pi\left(2^{2^{n-1}}\right) \geq n \geq \frac{\ln\left(\frac{lnx}{ln2}\right)}{ln2}$.

Now, we need to show that for any number $x \geq 2$, we have $\frac{\ln\left(\frac{lnx}{ln2}\right)}{ln2} \geq \ln(\ln(x))$.

This inequality is clearly equivalent to $\quad (1-\ln(2))\ln(\ln(x)) \geq \ln(\ln(2))$.

However, the function $x \to (1-\ln(2)\ln(\ln(x))$ is increasing, then for any $x \geq 2$, we have

$$(1-\ln(2)\ln(\ln(x)) \geq \left(1-\ln(2)\right).\ln(\ln(2)).$$

As $\ln(\ln(2)) \leq 0$, this implies $\quad \left(1-\ln(2)\right).\ln(\ln(2)) \geq \ln(\ln(2))$. $\square$

The approximation provided by Euclid's Theorem has highly been improved by **C. Gauss** conjecture 1792.

$$\pi(x) \sim \frac{x}{ln(x)}, \quad x \to +\infty$$

In using complex analysis methods, notably by examining the analytical properties of the Riemann zeta function ς, (in particular $\varsigma(0) \neq 0$, for any $s \in \mathcal{R}e(s) \geq 1$). In 1896 that conjecture has been independently proven by J. Hadamard and C.J de la Vallee Poussin and since it has been known as ***the Prime Number Theorem*.**

Theorem 3: (The prime number theorem)

We have: $$\pi(x) \sim \frac{x}{ln(x)}, \quad x \to +\infty.$$

The **Prime Number Theorem** describes the asymptotic distribution of prime numbers, stating that as n grows large, the number $\pi(n)$ of primes less than or equal to n approximates $\frac{n}{ln(xn)}$.

Consequently, the density of prime numbers around a large number n is approximately $\frac{1}{\ln(n)}$, indicating that that primes become less frequent as numbers increase.

This distribution is intimately connected to the **Riemann** zeta function, a complex function central to number theory. The **Riemann Hypothesis** posits that all non-trivial zeros of this function have a real part of 1/2. Proving this hypothesis would have profound implications for understanding the distribution of prime numbers, providing a precise description of their irregularities and regularities.

Inequalities of Chebyshev type.

Next, we are going to show that

$$a.\frac{x}{ln(x)} \leq \pi(x) \leq b.\frac{x}{ln(x)}$$

Where a and b are two constants.

Theorem 4: For any $x \geq 2$, we have:

$$\frac{\ln 2}{6}.\frac{x}{ln(x)} \leq \pi(x) \leq 4.\frac{x}{ln(x)}$$

Notes these constants are not the best, the goal is to show how this elementary approximation can lead to interesting results. The proof of this theorem relies on four lemmas.

Lemma 5: For any integer $m \geq 1$, we have:

$$\prod_{m+1<p\leq 2m+1} p \leq \binom{2m+1}{m} \leq 4^m$$

Proof: Given

$$\binom{2m+1}{m} = \frac{(2m+1)(2m)(2m-1)\dots(2m+1-m+1)}{m!} = \frac{1}{m!}\prod_{k=m+2}^{2m+1} k.$$

Let p be a prime number, $m+1 < p \leq 2m+1$. From the above equality p is a factor of $\binom{2m+1}{m} m!$. As p does not divides $m!$, it must necessarily divide $\binom{2m+1}{m}$. In other terms, $\binom{2m+1}{m}$ is multiple of every prime number p, $m+1 < p \leq 2m+1$, therefore multiple of their product. Consequently

$$\prod_{m+1<p\leq 2m+1} p \leq \binom{2m+1}{m}$$

The second inequality, one notices that

$$2^{2m+1} = (1+1)^{2m+1} = \sum_{k=0}^{2m+1} \binom{2m+1}{k} \geq \binom{2m+1}{m} + \binom{2m+1}{m+1} = 2\binom{2m+1}{m}$$

$$2^{2m+1} = 2.(2)^{2m} = 2.4^m \geq 2\binom{2m+1}{m}$$

Which gives the inequality wanted. □

Lemma 6: For any integer $n \geq 1$, we have:

$$\prod_{p\leq n} p < 4^n$$

Proof: By induction, for $n = 2,$ or 3 the inequality is obvious.

Assuming the result remains true for a certain integer n $n \geq 1$ and any $k = 1, 2, 3, \dots, n-1$.

If n is even, we have

$$\prod_{p\leq n} p = \prod_{p\leq n-1} p \leq 4^{n-1} < 4^n$$

If n is odd, $n = 2m+1,$ from **Lemma 5,** we have

$$\prod_{m+1<p\leq 2m+1} p \leq \binom{2m+1}{m} \leq 4^m$$

$$\prod_{p\leq n} p = \left(\prod_{p\leq m+1} p\right)\left(\prod_{m+1<p\leq 2m+1} p\right) \leq 4^m\left(\prod_{p\leq m+1} p\right)$$

According to the induction assumption we have

$$\prod_{p\le m+1} p < 4^{m+1}$$

Therefore

$$\prod_{p\le 2m+1} p < 4^{2m+1} = 4^n$$

Corollary 7: For any real number $x \ge 1$, we have:

$$\prod_{p\le x} p < 4^x$$

Proof: Let $x \ge 1$, and $n = \lfloor x \rfloor$. By using **Lemma 6,** we have

$$\prod_{p\le x} p = \prod_{p\le n} p < 4^n = 4^x$$

□

Lemma 8: For any integer $n \ge 1$, we have:

$$\frac{4^n}{2n} \le \binom{2n}{n} \le 4^n$$

Proof: For the first inequality, it suffices to notice that:

$$\binom{2n}{n} < (1+1)^{2n} = 2^{2n} \le 4^n$$

For the second inequality, we set

$$2^{2n-1} = \sum_{k=0}^{2n-1} \binom{2n-1}{k}$$

Notice that pour tout $k = 1, 2, 3, \dots, 2n-1$

$$\binom{2n-1}{k} \le \binom{2n-1}{n-1}$$

Then

$$2^{2n-1} \le 2n \binom{2n-1}{n-1}$$

Given

$$\frac{4^n}{2n} = \frac{2^{2n-1}}{n} \le 2\binom{2n-1}{n-1} = \binom{2n-1}{n-1} + \binom{2n-1}{n} = \binom{2n}{n}$$

□

Lemma 9: Assuming that p^{α} divide $\binom{2n}{n}$, for a certain prime number p a certain integer number $\alpha \geq 1$. Then

$$p^{\alpha} \leq 2n.$$

Proof: From the Legendre's Formula

$$v_p(m!) = \sum_{k=1}^{+\infty} \left\lfloor \frac{m}{p^k} \right\rfloor$$

We have

$$v_p\left(\binom{2n}{n}\right) = v_p\big((2n)!\big) - 2v_p(n!) = \sum_{k=1}^{+\infty} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2\left\lfloor \frac{n}{p^k} \right\rfloor \right)$$

Note if $k > \frac{\ln(2n)}{\ln(p)}$ then

$$\left\lfloor \frac{2n}{p^k} \right\rfloor = 2\left\lfloor \frac{n}{p^k} \right\rfloor$$

Moreover, the function $x \to \varphi(x) = \lfloor 2x \rfloor - \lfloor x \rfloor$ is 1-periodic and we have:

$$\varphi(x) = \begin{cases} 0 \ \ if \ 0 \leq x \leq \frac{1}{2} \\ 1 \ \ if \ \frac{1}{2} \leq x < 1 \end{cases}$$

Therefore

$$\alpha \leq v_p\left(\binom{2n}{n}\right) = \sum_{k=1}^{\left\lfloor \frac{\ln(2n)}{\ln(p)} \right\rfloor} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2\left\lfloor \frac{n}{p^k} \right\rfloor \right) \leq \left\lfloor \frac{\ln(2n)}{\ln(p)} \right\rfloor \leq \frac{\ln(2n)}{\ln(p)}$$

This implies

$$p^{\alpha} \leq 2n. \square$$

Now, let us prove **Theorem 4**

First let us show that

$$\pi(x) \leq \frac{4x}{ln(x)}$$

Note

$$\prod_{p \leq n} p > \prod_{\sqrt{n} < p \leq n} p \geq \left(\sqrt{n}\right)^{\pi(n) - \pi(\sqrt{n})}$$

Using **Lemma 6**, we obtain

$$n^{\frac{1}{2}\left(\pi(n)-\pi(\sqrt{n})\right)} < 4^n$$

Thus

$$\pi(n) - \pi\left(\sqrt{n}\right) < \frac{2nln(4)}{\ln(n)} < \frac{3n}{\ln(n)}$$

Because $4ln(2) < 3$.

Since, for $n \geq 2$, we have:

$$\pi\left(\sqrt{n}\right) \leq \sqrt{n} < \frac{n}{\ln(n)}$$

Finally, we obtain

$$\pi(n) < \frac{4n}{ln(n)}$$

Note, the function $x \to \frac{t}{\ln(t)}$ is increasing over interval $[e, +\infty[$, so we deduce for any $x \geq 2$,

$$\pi(x) = \pi(\lfloor x \rfloor) < \frac{4\lfloor x \rfloor}{ln\lfloor x \rfloor} < \frac{4x}{ln(x)}$$

Now let us show that

$$\pi(x) > \frac{\ln(2)}{6} \cdot \frac{x}{ln(x)}$$

In one hand, **Lemma 9** implies that

$$\binom{2n}{n} = \prod_{p \leq 2n} pv_p\binom{2n}{n} \leq \prod_{p \leq 2n} 2n = (2n)^{\pi(2n)}$$

The other hand, according to **Lemma 8,** we have

$$\frac{2^{2n}}{2n} \leq \binom{2n}{n}$$

Therefore

$$\frac{2^{2n}}{2n} \leq (2n)^{\pi(2n)}$$

Apply logarithm to both sides of the above inequality, we obtain, for $n \geq 1$:

$$\frac{2nln(2)}{ln(2n)} - 1 \leq \pi(2n)$$

We easily verify

$$\frac{nln(2)}{ln(2n)} \le \frac{2nln(2)}{ln(2n)} - 1 \le \ \pi(2n)$$

Thus

$$\pi(2n) \ge \frac{ln(2^n)}{ln(2n)}$$

To conclude, note that the inequality

$$\pi(x) \ge \frac{xln(2)}{6ln(x)}$$

Is true for $2 \le x < 3$. For $x \ge 3$, we set $n = \left\lfloor \frac{x}{2} \right\rfloor$, we have then

$$2n \le x < 2(n+1)$$

This implies that

$$\pi(x) \ge \pi(2n) \ge \frac{ln(2^n)}{ln(2n)} \ge \frac{nln(2)}{ln(2n)}$$

$$\pi(x) \ge \left(\frac{x}{2} - 1\right) . \frac{ln(2)}{ln(x)}$$

$$\pi(x) \ge \left(\frac{x}{2} - \frac{x}{x}\right) . \frac{ln(2)}{ln(x)}$$

$$\pi(x) \ge \left(\frac{1}{2} - \frac{1}{x}\right) . \frac{xln(2)}{ln(x)}$$

$$\pi(x) \ge \left(\frac{1}{2} - \frac{1}{x}\right) ln(2) . \frac{x}{ln(x)}$$

Since $x \ge 3$, we deduce

$$\pi(x) \ge \left(\frac{1}{2} - \frac{1}{3}\right) ln(2) . \frac{x}{\ln(x)} = \frac{ln(2)}{6} . \frac{x}{ln(x)}$$

$$\pi(x) \ge \frac{x}{\ln(x)} . \frac{ln(2)}{6}$$

□.

Chebyshev functions.

Apart from $\pi(x)$, two other sums appear in the study of the distribution of prime numbers. There are two Chebyshev functions.

$$\Theta(x) = \sum_{p \le x} \ln p$$

And

$$\Psi(x) = \sum_{p^m \leq x} \ln p$$

In this section we will show how the three functions are closely related together. Let us give a first elementary estimate arising from **Corollary 7**.

Lemma 10: We have

$$\Theta(x) \leq (2ln2)x$$

Proof: Let apply the exponential function

$$e^{\Theta(x)} = \prod_{p \leq x} p \leq 4^x$$

Thus

$$\Theta(x) = xln(4) \leq (2ln2)x$$

□.

The following result shows the link between the two functions $\Theta(x)$ and $\Psi(x)$.

Lemma 11: For $x \geq 2$, we have:

$$\Theta(x) \leq \Psi(x) \leq \Theta(x) + 2\sqrt{x}.ln(x)$$

Proof: Let k be the greatest integer such that $\sqrt[k]{x} \geq 2$, equivalently $k \geq \left\lfloor \frac{ln(x)}{ln(2)} \right\rfloor$. We have then

$$\Psi(x) = \sum_{m=1}^{k} \sum_{p \leq x^{1/m}} \ln p = \Theta\left(x^{\frac{1}{m}}\right)$$

In particular,

$$\Psi(x) = \Theta(x) + \sum_{m=2}^{k} \Theta\left(x^{\frac{1}{m}}\right) \geq \Theta(x)$$

However, by using **Lemma 10**, we obtain

$$\Psi(x) = \Theta(x) + (2ln2) \sum_{m=2}^{k} x^{\frac{1}{m}} \leq \Theta(x) + (2ln2)k\sqrt{x}$$

As $k \leq \frac{ln(x)}{ln(2)}$ we deduce that

$$\Psi(x) \leq \Theta(x) + 2\sqrt{x}.ln(x)$$

□

Below, we present the link between the two functions $\Theta(x)$ and $\pi(x)$.

Lemma 12: For $x \geq 2$, we have:

$$\frac{\Theta(x)}{ln(x)} \leq \pi(x) \leq \frac{\Theta(x)}{ln(x) - 2lnln(x)} + \frac{x}{(ln(x))^2}$$

Proof: The first inequality occurs from the following estimation

$$\Theta(x) = \sum_{p \leq x} \ln p \leq \ ln(x) \sum_{p \leq x} 1 = \ \pi(x).ln(x)$$

For the second inequality, we note that for $2 \leq y < x$, we have

$$\pi(x) - \pi(y) = \sum_{y<p\leq x} 1 \leq \frac{1}{ln(y)} \sum_{y<p\leq x} ln(p)$$

$$\pi(x) - \pi(y) \leq \frac{1}{ln(y)} \left(\Theta(x) - \Theta(y)\right)$$

This implies

$$\pi(x) \leq \frac{\Theta(x)}{ln(y)} + \pi(y) \ \leq \frac{\Theta(y)}{ln(y)} + y$$

If we substitute $y = \frac{x}{(ln\, x)^2}$, we deduce that

$$\pi(x) \leq \frac{\Theta(x)}{ln(x) - 2\ln{(lnx)}} + \frac{x}{\left(ln(x)\right)^2}$$

□.

Theorem 12: The following assertions are equivalent.

i. $\pi(x) \sim \frac{x}{ln(x)}$, when $x \to +\infty$
ii. $\Theta(x) \sim x$, when $x \to +\infty$
iii. $\Psi(x) \sim x$, when $x \to +\infty$

By using **Lemma 11** and **Lemma 12,** we have

$$\frac{\Theta(x)}{x} \leq \frac{\Psi(x)}{x} \leq \frac{\Theta(x)}{x} + \frac{2ln(x)}{\sqrt{x}}$$

And

$$\frac{\Theta(x)}{x} \leq \frac{\pi(x)\ln(x)}{x} \leq \frac{\Theta(x)}{x} . \frac{ln(x)}{ln(x) - 2\ln{(ln(x)}} + \frac{1}{ln(x)}$$

Thus, the result has been proved. □