A Taste of Analytic Number Theory, Part I

Ayan Nath

HS Student, Kaliabor College

Abstract. These series of articles (three in total) are aimed at olympiad contestants, focuses on solving olympiad Number Theory problems using analytic techniques and making contestants familiar with common techniques and results in this topic. We start with the Prime Number Theorem, give an elementary proof of the weak version and establish a few well known estimates for the two Chebyshev functions. We also show Mertens' first theorem on the fly and discuss Mertens' second theorem. Asymptotic density, Equidistribution theorem are also added.

1. The Prime Number Theorem

1.1. Introduction

Primes are the building blocks of the integers, just as molecules and atoms are the building blocks of nature, hence it makes great sense to study the primes, and in particular, the distribution of primes. I think you know that there are infinitely many primes and most likely you already know Euclid's proof of it, but it doesn't tell us anything significant about the distribution of primes. This is exactly what our objective is, i.e. try to understand the distribution of primes. It is natural to define the function

$$\pi(x) = \text{No. of primes at most } x.$$

We would like to find a "formula" for $\pi(x)$ in terms of x, it turns out finding an exact formula is not really possible due to the raggedy nature of primes. Instead we try to estimate it. The well known estimate which we call the Prime Number Theorem (PNT) asserts that:

Theorem 1.1 (Prime Number Theorem (PNT))**.**

$$\pi(x) \sim \frac{x}{\log x}$$

here we write $f(x) \sim g(x)$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$,. The way to think about this is, $f(x)$ is approximated by $g(x)$, the larger the x, the better the approximation. We do not prove this here, but instead we establish a weaker estimate that there exist positive real numbers a and b such that

$$\frac{ax}{\log x} < \pi(x) < \frac{bx}{\log x}.$$

Before we move on to the proof, it will help to get comfortable with the big-$\mathcal{O}$ and little-o notation:

Digression 1.2. If f and g are two functions then we say that $f(x)$ is $\mathcal{O}(g(x))$ if and only if there exist some constant C such that $|f(x)| < Cg(x)$ for all large x. And we say that $f(x)$ is $o(g(x))$ if and only if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$ where $g(x)$ should be non-zero for all large enough x. For example, $\sin x + e^{91}$ is $\mathcal{O}(1)$, $x + (\log x)^{10}$ is $\mathcal{O}(x)$, $\sin x + \log x$ is $o(x^{0.001})$. If $f(x)$ is $\mathcal{O}(g(x))$ then this is often expressed as $f(x) = \mathcal{O}(g(x))$ and similarly $f(x) = o(g(x))$ if $f(x)$ is $o(g(x))$, but you should remember that this is an abuse of notation. For example, we write $\lfloor x \rfloor = x + \mathcal{O}(1), \log x = o(x^{0.001}), n! = \mathcal{O}(n^n)$ in this article. It will help to get comfortable with these notations, it lets us ignore stuff we don't care about. I suggest reading this : https://en.wikipedia.org/wiki/Big_O_notation

1.2. Proving weak PNT

Let me define some functions, bear with me for now, I will explain the motivation in a moment. Define the **von Mangoldt function** $\Lambda : \mathbb{N} \to \mathbb{R}$ as

$$\Lambda(n) = \begin{cases} \log p, & n = p^k \text{ for some prime } p \text{ and positive integer } k \\ 0, & \text{otherwise.} \end{cases}$$

One can also think of this as weighting all the prime powers p^k with $\log p$. You can easily see that

$$\sum_{d|n} \Lambda(d) = \log n.$$

Whenever we try to find bounds it is a common theme to look at the "big picture" at once, also called "global" methods, or in simple terms, double counting. Since we are looking for bounds related to primes, it is somewhat motivated to "sum" everything up and try to double count:

$$\begin{aligned}
\sum_{n\leq x} \log n &= \sum_{n\leq x} \sum_{d|n} \Lambda(d) \\
&= \sum_{d\leq x} \Lambda(d) \sum_{n\leq x, d|n} 1 \\
&= \sum_{d\leq x} \Lambda(d) \left\lfloor \frac{x}{d} \right\rfloor \\
&= \sum_{d\leq x} \Lambda(d) \left(\frac{x}{d} + \mathcal{O}(1) \right) \\
&= x \sum_{n\leq x} \frac{\Lambda(n)}{n} + \mathcal{O}(1) \sum_{n\leq x} \Lambda(n)
\end{aligned}$$

Now where did we double count? We double counted when we swapped the summations. The left hand side is very easy to estimate accurately. For now let us focus on the RHS, notice how the RHS is related to primes while the LHS is not, clearly the above equation has some information about primes encoded via the von Mangoldt function. We would want to estimate $\sum_{n\le x}\Lambda(n)$ to get rid of the awkward $\mathcal{O}(1)$ multiple. It now makes sense to define

$$\psi(x)=\sum_{n\le x}\Lambda(n).$$

This is called the **Second Chebyshev Function**, we take the domain as $\mathbb{R}$ instead of $\mathbb{N}$ to avoid writing floors whenever we have a non-integral input, we do this with almost all the discussed functions in this article. In what follows, p always denotes a prime number. Observe the following *rough* calculation

$$\psi(x)=\sum_{n\le x}\Lambda(n)=\sum_{p\le x}\log p\left\lfloor \log_p x\right\rfloor=\sum_{p\le x}\log p\left\lfloor\frac{\log x}{\log p}\right\rfloor\approx\sum_{p\le x}\log x=\pi(x)\log x.$$

This suggests that $\psi(x)\sim x$ (which is indeed true). We are deliberately ambiguous about what $\approx$ means. We just need a rough estimate for $\psi(x)$, even $\psi(x)=\mathcal{O}(x)$ should do. For now assume this is true, we will get that

$$\sum_{n\le x}\log n=x\sum_{n\le x}\frac{\Lambda(n)}{n}+\mathcal{O}(x). \tag{1}$$

We now want an estimate for the LHS.

Lemma 1.3 (Weak Stirling's Approximation)**.** $\sum_{n\le x}\log n>x\log x-x$ for all $x\in\mathbb{N}$ and in particular, $\sum_{n\le x}\log n=x\log x+\mathcal{O}(x)$ for all $x\in\mathbb{R}$.

Proof. The most common and natural way to prove this would be direct integration but we won't do that here. Look at the expansion of e^x where $x>0$ is an integer:

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots.$$

Clearly $\frac{x^x}{x!}$ is a term in the expansion, therefore

$$e^x>\frac{x^x}{x!}\implies x>x\log x-\log x!\implies\sum_{n\le x}\log n>x\log x-x.$$

The lemma is proved now because $\sum_{n\le x}\log n<x\log x$ is trivial. □

Using the above lemma and equation (1) we have

$$x\log x+\mathcal{O}(x)=x\sum_{n\le x}\frac{\Lambda(n)}{n}+\mathcal{O}(x)\implies\sum_{n\le x}\frac{\Lambda(n)}{n}=\log x+\mathcal{O}(1).$$

The above relation is clearly useful since the LHS encodes primes in it, and gives us an estimate of a "global" sum involving primes. What is left is to prove that $\psi(x)=\mathcal{O}(x)$. The main idea is to write

$$\begin{aligned}\psi(x)&=\sum_{p\le x}\log p+\sum_{p^2\le x}\log p+\sum_{p^3\le x}\log p+\cdots\\&=\sum_{p\le x}\log p+\sum_{p\le\sqrt{x}}\log p+\sum_{p\le\sqrt[3]{x}}\log p+\cdots.\end{aligned}\tag{$\star$}$$

For brevity let us define

$$\theta(x) = \sum_{p \le x} \log p.$$

This is called the **First Chebyshev Function**, again the domain here is $\mathbb{R}$. Another way to motivate the first Chebyshev function is :

$$\begin{aligned}
\psi(x) = \sum_{n \le x} \Lambda(n) &= \sum_{p \le x} \log p \left\lfloor \log_p x \right\rfloor \\
&= \sum_{p \le x} \log p \left(\frac{\log x}{\log p} + \mathcal{O}(1) \right) \\
&= \pi(x) \log x + \mathcal{O}(1) \underbrace{\sum_{p \le x} \log p}_{\theta(x)} .
\end{aligned}$$

The above relation also suggests that $\theta(x) \sim x$ (which is indeed true, but we don't need that here). We can write $(\star)$ concisely as

$$\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \cdots .$$

This sum is not an infinite sum, the terms become zero eventually. Note that $\theta(x)$ is the "biggest" term in the RHS and the rest of them are "small", let us try to prove $\theta(x) = \mathcal{O}(x)$, the fact that $\psi(x) - \theta(x)$ is "small" (compared to $\mathcal{O}(x)$ of course) will be automatically implied.

Lemma 1.4. $\theta(n) < 4n \log 2$, in particular $\theta(x) = \mathcal{O}(x)$.

Proof. The main idea is to consider a number which is divisible by many consecutive primes but the size of the number is not too large. A crude example is $x!$. Clearly the primes less than x divide $x!$, therefore $\prod_{p \le x} p \le x! \implies \theta(x) \le \log x! = x \log x + \mathcal{O}(x)$, yes this is a trivial bound, but the point is that we want to do something similar.

We consider $\binom{2n}{n}$. Note that this number is divisible by all primes in the interval $[n+1, 2n]$. Therefore, $\prod_{n < p \le 2n} p \le \binom{2n}{n}$. By the binomial theorem, $\binom{2n}{n}$ is trivially bounded above by $(1+1)^{2n} = 2^{2n}$. Therefore, taking logarithms we obtain

$$\theta(2n) - \theta(n) \le 2n \log 2.$$

So we have that

$$\begin{aligned}
\theta(2^k) - \theta(2^{k-1}) &\le 2^k \log 2 \\
\theta(2^{k-1}) - \theta(2^{k-2}) &\le 2^{k-1} \log 2 \\
&\vdots \\
\theta(2) - \theta(1) &< 2 \log 2
\end{aligned}$$

Summing up, $\theta(2^k) \le 2^{k+1} \log 2$. Therefore for general n, it holds that

$$\theta(n) \le \theta(2^{\lceil \log_2 n \rceil}) \le 2^{\lceil \log_2 n \rceil + 1} \log 2 < 2^{\log_2 n + 2} \log 2 < 4n \log 2.$$

$\square$

The following implies that $\psi(x) - \theta(x)$ is "small", I suggest you to try to prove this on your own.

Lemma 1.5. $\psi(x) = \theta(x) + \mathcal{O}(\sqrt{x})$, and in particular $\psi(x) = \mathcal{O}(x)$.

Proof. We have that

$$\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \cdots .$$

This summation is not infinite, we can write it as

$$\begin{aligned}\psi(x) = \theta(x) + \theta(x^{1/2}) + \sum_{k=3}^{\lceil \log_2 x \rceil} \theta(x^{1/k}) &= \theta(x) + \mathcal{O}(x^{1/2}) + \frac{\log x}{\log 2}\mathcal{O}(x^{1/3}) \\ &= \theta(x) + \mathcal{O}(\sqrt{x}).\end{aligned}$$

This finally implies that $\psi(x) = \mathcal{O}(x)$ since $\theta(x) = \mathcal{O}(x)$. $\square$

And we can now state the result we obtained initially as the proof is now complete:

Theorem 1.6.

$$\sum_{n \le x} \frac{\Lambda(n)}{n} = \log x + \mathcal{O}(1).$$

Recall the rough calculation we did for $\psi(x)$, it implied that $\psi(x)$ is roughly $\pi(x) \log x$. But we already got $\psi(x) = \mathcal{O}(x)$, so this should somehow imply estimates on $\pi(x)$ right? Yes, but with some work. Doing the calculations properly:

$$\begin{aligned}\psi(x) = \sum_{n \le x} \Lambda(n) &= \sum_{p \le x} \log p \left\lfloor \log_p x \right\rfloor \\ &= \sum_{p \le x} \log p \left(\frac{\log x}{\log p} + \mathcal{O}(1) \right) \\ &= \pi(x) \log x + \mathcal{O}(1)\theta(x) \\ &= \pi(x) \log x + \mathcal{O}(x).\end{aligned}$$

Therefore we conclude that $\pi(x) = \mathcal{O}\left(\frac{x}{\log x}\right)$, this proves that there exist a positive constant b such that $\pi(x) < \frac{bx}{\log x}$, if you really want to you can get an explicit constant b easily, but this is not terribly important.

Remark 1.7. By elementary means you can show that $b = 1.6$ works.

We are left to prove a lower bound. There is a completely elementary proof[1] of the lower bound but here I will discuss a proof which uses the theory developed till now and also proves Mertens' First Theorem on the fly.

Consider Theorem 1.6, for majority of the terms $n \le x$, $\Lambda(n)$ is zero. And if n is not a prime but a power of a prime p, then the denominator of $\frac{\Lambda(n)}{n}$ becomes very large compared to the numerator,

[1] https://math.stackexchange.com/a/1890792

for those who know about $p-$series convergence it is probably immediate that the sum of terms when n is not a prime is bounded above by a constant, or in other words, $\mathcal{O}(1)$.

Digression 1.8. Consider $S_p = \sum_{n=1}^{\infty} \frac{1}{n^p}$, called the $p-$series, it is well known that S_p converges for $p > 1$ and diverges for $p \leq 1$. Proving this is not hard, see here: https://math.stackexchange.com/a/29466

Theorem 1.9 (Mertens' First Theorem)**.**

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + \mathcal{O}(1).$$

Proof.

$$\begin{aligned}
\sum_{n \leq x} \frac{\Lambda(n)}{n} &= \sum_{p \leq x} \frac{\log p}{p} + \sum_{p^k \leq x, k \geq 2} \frac{\log p}{p^k} \\
&= \sum_{p \leq x} \frac{\log p}{p} + \sum_{p \leq x} \sum_{2 \leq k \leq \log_p x} \frac{\log p}{p^k} \\
&< \sum_{p \leq x} \frac{\log p}{p} + \sum_{p \leq x} \log p \sum_{k=2}^{\infty} \frac{1}{p^k} \\
&= \sum_{p \leq x} \frac{\log p}{p} + \sum_{p \leq x} \frac{\log p}{p^2 - p}.
\end{aligned}$$

Note that $\sum_{n=2}^{\infty} \frac{\log n}{n^2 - n}$ is convergent because

$$\frac{\log n}{n^2 - n} < \frac{n^{0.1}}{n^2 - n} = \frac{1}{n^{1.9} - n^{0.9}} < \frac{1}{n^{1.8}}$$

holds for all sufficiently large n. We are now done by Theorem 1.6. □

Fix a large constant c. The above result implies that

$$\sum_{\frac{x}{c} < p \leq x} \frac{\log p}{p} = \sum_{p \leq x} \frac{\log p}{p} - \sum_{p \leq \frac{x}{c}} \frac{\log p}{p} = \log c + \mathcal{O}(1).$$

Here we pick c very large so that the RHS is positive, say the RHS is bounded below by $\delta > 0$. We do some trivial bounding,

$$(\pi(x) - \pi(x/c)) \frac{\log \frac{x}{c}}{x/c} \geq \sum_{\frac{x}{c} < p \leq x} \frac{\log p}{p} > \delta.$$

Now just neglect $\pi(x/c)$ we get

$$\pi(x) > \delta \cdot \frac{x}{c(\log x - \log c)}.$$

Thus we have established that there exists some a such that $\pi(x) > \frac{ax}{\log x}$. Again, with some hard work you can get an explicit constant but that is not very important.

Question 1.10. Put together a logical write-up of the proof of the weak PNT.

Example 1.11 (Generalisation of Bertrand's Postulate)**.** Let $\varepsilon > 0$. Prove that there exist a prime between n and $(1+\varepsilon)n$ for all large n, in particular there always exist a prime between n and $2n$ for $n > 1$.

Demonstration. Just use PNT for the first part. Proving that there always exist a prime between n and $2n$ for $n > 1$ is doable[2] without the full power of PNT though. Hint: Consider $\binom{2n}{n}$.

Example 1.12. Fix $1 > \varepsilon > 0$. For some natural n, let $g(n)$ be the number of divisors of n in $(\sqrt{n}, (1+\varepsilon)\sqrt{n})$. Prove that $g : \mathbb{N} \to \mathbb{Z}_{\geq 0}$ is surjective.

Demonstration. 1. Check that $n = p^k$ for some prime p won't work.

2. Take $n = p^{2a}q^{2b}$ for two primes p and q.
3. Set some arbitrary k. You want to ensure that $p^a q^b < p^x q^y < (1+\epsilon)p^a q^b$ has k solutions in $0 \leq x \leq 2a, 0 \leq y \leq 2b$.
4. Fix x. At most how many possibilities for y are there?
5. It would be nice to have something like this: all such divisors are given by $p^a q^b (\frac{p}{q^t})^i$ for $i = 1, 2, \ldots, k$. Why do we expect this? When can this happen?
6. Obviously we want p/q^t to be very close to 1 and $tk = b$.
7. Finish using Generalised Bertrand's Postulate.

1.3. Asymptotics for primes

We define p_n to be the nth prime number. It will be quite nice to find a smooth function $f(n)$ such that $p_n \sim f(n)$. It turns out this is quite easy using PNT, the reader may try this on their own.

Theorem 1.13. $p_n \sim n \log n$

Proof. Obviously $\pi(p_n) = n$. Therefore

$$\frac{p_n}{\log p_n} \sim n \implies \frac{p_n}{n \log n} \sim \frac{\log p_n}{\log n}.$$

But see that

$$n \sim \frac{p_n}{\log p_n} \implies \log n \sim \log p_n - \log\log p_n \implies \frac{\log p_n}{\log n} \sim 1 - \frac{\log\log p_n}{\log p_n}.$$

Thus it follows that $p_n \sim n \log n$. □

Remark 1.14 (Rosser's Theorem)**.** $p_n > n \log n$.

[2] https://www.cut-the-knot.org/arithmetic/algebra/BertrandPostulate.shtml

This result is useful for ad-hoc calculations to get a feel about whether a statement or a conjecture should be true. Let me state a result without proof just to summarise:

Theorem 1.15. The following are equivalent :

- $\pi(x) \sim x/\log x$,
- $\theta(x) \sim x$,
- $\psi(x) \sim x$,
- $p_n \sim n \log n$.

If you are interested then you may trying proving it. Finally, here's a real olympiad problem:

Example 1.16 (EMMO 2016 Sr, Anant Mudgal)**.** We call a sequence of positive integers $\{a_n\}_{n\in\mathbb{N}}$ as a scouter if it is strictly increasing and $a_n < 9000n$. We call an integer $i \geq 1$ as divisor friendly if a_i divides the least common multiple of all previous terms of the sequence and call divisor-unfriendly otherwise. Is it necessarily true that a scouter has infinitely many

(a) divisor friendly

(b) divisor unfriendly

indices?

Demonstration. 1. Do part (b).

2. Assume that a_n is divisor unfriendly for all $n \geq N$.
3. Look at the prime factorisation of a_n, precisely, look at the exponents.
4. Conclude that there is a sequence of prime powers d_n such that $d_n \mid a_n$ and $d_i = d_j$ if and only if $i = j$.
5. Intuitively, d_n should grow faster than $9000n$.
6. Prove it by estimating the proportion of prime powers less than some large fixed number M. (We will discuss this idea in detail in the following section)
7. You may need to split the summation into the intervals $[1, \sqrt{9000n}]$ and $(\sqrt{9000n}, 9000n]$ and bound them separately, there are a lot of other ways to do this though.
8. Conclude.
9. Bonus: Strengthen the bound. You can relax the upper bound for a_n to $\delta n \log n$ for some sufficiently small $\delta > 0$.

In the next part of the series we will discuss the concept of 'density' and prove some results. The concluding part will contain some problems for practise.