

Hobby Number Theory: $\pi(n)$

Geoffrey Dixon
gdixon@7stones.com



A really really very good smooth approximation to $\pi(n)$.

Caveats

- Although I have toyed with number theory since my early teens (my high school Science Fair project was a big hit), and although I have advanced degrees in mathematics and physics, I've never had a course in number theory.
- I've read a book or three, doing research to fill in knowledge gaps when my number theory play needed augmenting, but I will offer no references. You're on your own.
- I have utterly no idea if there are precedents for some of the formulae I shall develop, and, quite frankly, I don't care. Number theory is a very easy branch of mathematics to toy with. The curious amateur can pass many an odd hour rediscovering old truths, and that's what this paper will cover, viz., my odd hours and what they wrought, old or not.
- If you're reading this you doubtless know enough prime number theory to follow what I have to say. Some of my thinking I will explicate, but mostly the content is not deep and I shan't bother.

Throughout the rest, \mathbb{N} is the set of positive integers, the variables $i, j, k, m, n \in \mathbb{N}$, $p, q \in \mathbb{N}$ are prime numbers, and x, y, z will be real number variables. Some of these may never get used.

Prime Number Theorem and its Shortcomings

The deep end of prime number theory frequently involves the prime number theorem and things like the Riemann zeta function. On the shallow end are the following ideas.

Let

$$\pi(x) = \text{number of primes } < x, x \geq 1.$$

The principle supposition guiding the rest of this paper is that there exists an optimal smooth function that provides a best approximation to $\pi(x)$ (which, of course, is not at all smooth, being a step function).

Let

$$\rho(x) = \text{density of primes at } x,$$

where $\rho(x)$ is a hypothetical optimal smooth function. (For n a positive integer, $\rho(n)$ is the probability that n is prime.) So

$$\pi(n) \approx \int_2^n \rho(x) dx. \quad (1)$$

Two smooth approximations feature prominently in the prime number theorem:

$$\pi(x) \approx \frac{x}{\ln x}, \quad (2)$$

and

$$\rho(x) \approx \frac{1}{\ln x}. \quad (3)$$

This gives rise to another approximation for $\pi(n)$:

$$\pi(n) \approx Li(n) \equiv \int_2^n \frac{dx}{\ln(x)}. \quad (4)$$

The approximation for $\pi(n)$ in (4) is much better than (2), but neither is really very stellar, and as a consequence neither has interested me particularly. As a prime number hobbyist, none of this offered me anything I could sink my teeth into. It has been proven that

$$\lim_{n \rightarrow \infty} \pi(n)/(n/\ln(n)) = \lim_{n \rightarrow \infty} \pi(n)/Li(n) = 1, \quad (5)$$

and that's most of what I know about the prime number theorem.

The problem for me is that the limits in (5) are not all that impressive. Consider that

$$\lim_{x \rightarrow \infty} \frac{x + \sqrt{x}}{x} = 1,$$

and yet

$$\lim_{x \rightarrow \infty} ((x + \sqrt{x}) - x) = \infty.$$

So there is no way that $x + \sqrt{x}$ is an optimal (or even good) approximation to x . And in fact, consider that

$$\pi(10^7) = 664,579,$$

and

$$\pi(10^7) - \frac{10^7}{\ln 10^7} = 44,158,$$

and

$$Li(10^7) - \pi(10^7) = 339,$$

and as one carries on beyond 10^7 these differences grow impressively more positive, both seemingly without bound. Neither of these approximations to $\pi(n)$ behave even remotely how I'd expect a (the) true optimal smooth approximation to behave.

Looking Elsewhere: $lcm(n)$

Much of my time as a number theory hobbyist was spent plotting step functions stemming from prime numbers. A couple of these involved

$$lcm(n), \tag{6}$$

the least common multiple of the integers from 1 to n . This is the smallest positive integer such that for all $1 \leq k \leq n$,

$$\frac{lcm(n)}{k} = m \in \mathbb{N}. \tag{7}$$

Here is a small table of values:

n	$lcm(n)$	added factor
1	1	1
2	2	2
3	6	3
4	12	2
5	60	5
6	60	1
7	420	7
8	840	2
9	2520	3
10	2520	1

Three things should be clear from this:

- $lcm(n)$ is in fact a step function;
- $lcm(n)$ steps up iff $n = p^k$ for some $k \in \mathbb{N}$, and p a prime;
- If $n = p^k$, then $lcm(n) = p lcm(n-1)$.

(Just a quick aside: finite mathematical fields exist of order n iff $n = p^k$, p prime.)

Clearly lcm grows rather quickly, which makes it difficult to plot. So instead I plotted

$$LL(n) \equiv \ln(lcm(n)). \tag{8}$$

The result surprised the hell out of me: it looked linear. After a lot of plotting, looking things up in tables (there was no viable internet when I began this effort), I decided that $LL(n)$ had a simple smooth approximation, optimal in exactly the way I expect optimal to look:

$$LL(n) \approx n - 1. \tag{9}$$

Why had I never seen this in any of my readings in number theory? Surely it must be known. Was it considered unimportant? To this day I've still not stumbled across this notion in any of my admittedly not very assiduous perambulations in published number theory. Well, known or not known, in the absence of any cogent argument against looking further, I did.

Looking Further

Years (decades) ago I used (9) to derive a formula for $\pi(n)$ that I subsequently determined was something that Riemann played with over a century earlier. I'm too lazy to dig into my piles of mathematical notes to find it, and in any case that would be less fun than just digging into (9) anew and see where it led. I was not disappointed.

Discussion: Contributing to $lcm(n)$

Our underlying assumption is $LL(n) \approx n - 1$, which leads to

$$lcm(n) \approx e lcm(n-1). \quad (10)$$

In reality, however,

$$lcm(n) = lcm(n-1), \text{ if } n^{1/k} \text{ is not prime for all } k \in \mathbb{N},$$

$$lcm(n) = n^{1/k} lcm(n-1), \text{ if } n^{1/k} = p, \text{ a prime.}$$

Each of these latter possibilities occurs with a certain probability, $\rho_k(n)$, the density of prime powers, p^k , at n .

Discussion: $\rho_k(n)$

If there are m primes between $n^{1/k}$ and $n^{1/k} + j$, then these same m primes account for all (primes) ^{k} between $(n^{1/k})^k = n$ and $(n^{1/k} + j)^k$. Likewise, at the infinitesimal level, if there is an amount of "primeness" dm between $n^{1/k}$ and $n^{1/k} + dj$, then dm is also the amount of (primeness) ^{k} between n and $(n^{1/k} + dj)^k$.

Note: let $Pi(x)$ be our hypothetical optimal smooth approximation to $\pi(x)$.

Then

$$dm = Pi(n^{1/k} + dj) - Pi(n^{1/k}).$$

The density of primes at $n^{1/k}$ is

$$\rho(n^{1/k}) = \frac{dm}{(n^{1/k} + dj) - n^{1/k}} = \frac{dm}{dj}. \quad (11)$$

Therefore, the density of p^k at n is

$$\rho_k(n) = \frac{dm}{(n^{1/k} + dj)^k - n} = \frac{dm}{kn^{\frac{k-1}{k}} dj} = \frac{\rho(n^{1/k})}{kn^{\frac{k-1}{k}}}, \quad (12)$$

where in the denominator of the third term we can safely ignore all higher powers of the infinitesimal, dj .

Discussion: Average contribution to $lcm(n)$

Suppose we have a function $s(n)$, $n \in \mathbb{N}$, and that

$$s(n) = j s(n-1),$$

where $j = 2$ or $j = 3$, and that the probability that $j = 2$ is $P_2 = \frac{2}{3}$, and that $j = 3$ is $P_3 = \frac{1}{3}$. So, for example, the expected (average) value of

$$s(61) = 2^{40} 3^{20} s(1).$$

So the average contribution at each step is

$$(2^{40} 3^{20})^{\frac{1}{60}} = 2^{P_2} 3^{P_3}.$$

We'll apply this same reasoning to $lcm(n)$. In incrementing from $lcm(n-1)$ to $lcm(n)$, the value either stays the same, or it changes by some factor in the set

$$\{n, n^{1/2}, n^{1/3}, \dots, n^{1/k}, \dots\}.$$

The probability that $lcm(n) = n^{1/k} lcm(n-1)$ is $\rho_k(n)$. Therefore, the expected, or average, factor is

$$\prod_{k=1}^{\infty} (n^{1/k})^{\rho_k(n)} = \prod_{k=1}^{\infty} n^{\rho(n^{1/k}) / (k^2 n^{(k-1)/k})} = e,$$

where that last equality is the point of all this. Taking the \ln of both sides and dividing by $\ln(n)$ gives us:

$$\frac{1}{\ln(n)} = \rho(n) + \frac{\rho(n^{1/2})}{2^2 n^{1/2}} + \frac{\rho(n^{1/3})}{3^2 n^{2/3}} + \dots + \frac{\rho(n^{1/k})}{k^2 n^{(k-1)/k}} + \dots \quad (13)$$

Discussion: Möbius function, $\mu(n)$

If $n > 1$ has a prime factorization $p_1^{i_1} p_2^{i_2} \dots p_k^{i_k}$, $i_j \geq 1$ for all $1 \leq j \leq k$, then the Möbius function of n is defined as follows:

$$\mu(n) = 0, \text{ if any } i_j > 1,$$

$$\mu(n) = (-1)^k \text{ otherwise.}$$

And $\mu(1) = 1$.

As an example of how this gets used, let

$$s = \sum_{k=1}^{\infty} \frac{1}{k}.$$

Ignore the fact that this sum diverges. Using $\mu(n)$ we can sum multiples of s to extract just that first term, $\frac{1}{1} = 1$. If you're unfamiliar with $\mu(n)$ it is worth your while to prove to yourself that

$$\sum_{m=1}^{\infty} \mu(m) \frac{s}{m} = \frac{1}{1} = 1. \quad (14)$$

Discussion: Use $\mu(n)$ to solve for $\rho(n)$

I'm going to apply this same method to express $\rho(n)$ as an infinite sum involving $\ln(n)$ and $\mu(n)$ starting from (13). If I want to get rid of the k^{th} term in the sum, I need to modify the first term, $\rho(n)$, to make it look like the k^{th} term,

$$\frac{\rho(n^{1/k})}{k^2 n^{(k-1)/k}}.$$

To start, then, I need to replace n by $n^{1/k}$:

$$\frac{1}{\ln(n^{1/k})} = \frac{k}{\ln(n)} = \rho(n^{1/k}) + \dots$$

so

$$\frac{1}{\ln(n)} = \frac{\rho(n^{1/k})}{k} + \dots$$

leading to

$$\frac{1}{k n^{(k-1)/k} \ln(n)} = \frac{n^{1/k}}{k n \ln(n)} = \frac{\rho(n^{1/k})}{k^2 n^{(k-1)/k}} + \dots \quad (15)$$

Now, just as I did in the last section in equation (14), I'm going to apply the function μ to (15) to get

$$\begin{aligned} \rho(n) &= \sum_{k=1}^{\infty} \frac{\mu(k)}{k n^{(k-1)/k} \ln(n)} = \frac{1}{\ln(n)} \sum_{k=1}^{\infty} \frac{\mu(k)}{k n^{(k-1)/k}} \\ &= \frac{1}{\ln(n)} \left(1 - \frac{1}{2n^{1/2}} - \frac{1}{3n^{2/3}} - \frac{1}{5n^{4/5}} + \frac{1}{6n^{5/6}} - \dots \right). \end{aligned} \quad (16)$$

So, on the assumption I'm right (and I most certainly am), we now see why (3) fails to be any better an approximation to $\rho(n)$ than it is, and it isn't really very good: it's because it's missing all those terms after the 1 inside the parentheses in the second line of (16). Admittedly their contribution is not large, but I hope to convince you that it's exactly what's needed.

Discussion: Let a computer take over from here

A lot of work has been published on how much $\rho(n)$ differs from $\frac{1}{\ln(n)}$. Glancing at discussions on the internet I see little to excite. So I'm going to ignore it, for the most part. And I'm also not going to do any heavy lifting when it comes to integrating my $\rho(n)$ in (16). I also assume that

$$\pi(n) \approx \int_2^n \rho(x) dx,$$

but the thought of trying to work out that integral makes me tired. Instead I decided to approximate the integral by replacing dx with $\Delta x = 1$, and change the integral for a sum:

$$\pi(n) \approx \sum_{k=2}^n \rho(k), \quad (17)$$

using (16) for ρ .

Originally I tested this idea with javascript code, but higher values of n required something more robust, so I used Fortran. I incorporated the first 80 values of $\mu(n)$.

```

program pi
DOUBLE PRECISION, DIMENSION(80) :: mu
DOUBLE PRECISION nr,kr,psu,su

mu = (/1,-1,-1,0,-1, 1,-1,0,0,1, -1,0,-1,1,1, 0,-1,0,-1,0, 1,1,-1,0,0, &
& 1,0,0,-1,-1, -1,0,1,1,1, 0,-1,1,1,0, -1,-1,-1,0,0, 1,-1,0,0,0, &
& 1,0,-1,0,1, 0,1,1,-1,0, -1,1,0,0,1, -1,-1,0,1,-1, -1,0,-1,1,0, &
& 0,1,-1,-1,0/)

N = 1000000000
su = 0.0

do n = 2,N+1
  psu = 1.0
  nr = 1.0*n
  do k = 2,size(mu)
    kr = 1.0*k
    psu = psu + mu(k)/(kr*(nr**((kr-1)/kr)))
    !print *, psu, nr**((kr-1)/kr)
  enddo
  !print *,psu, log(nr)
  psu = psu/log(nr)
  su = su + psu
enddo
print *,'dog 80 = ', su
end program pi

```

So, is the result any good? You be the judge.

n	$\pi(n)$	$Li(n) - \pi(n)$	$Pi(n) - \pi(n)$
10	4	1	-1
10^2	25	4	-1
10^3	168	9	-1
10^4	1229	16	-3
10^5	9592	37	-6
10^6	78498	129	28
10^7	664579	338	87
10^8	5761455	753	96
10^9	50847534	1701	-80

Certainly this is better than $Li(n)$. Indeed, from 10^4 onward the difference $Li(n) - \pi(n)$ more than doubles in size between n and $10n$. At the highest value for which I can find data,

$$Li(10^{25}) - \pi(10^{25}) = 55,160,980,938.$$

As I do not own a computer that can test $Pi(n)$ at an n that big (I just barely managed $n = 10^9$), I'll just assume my result would be significantly better. It's also worth noting that in the table above, the $Pi(n) - \pi(n)$ uses two approximations: it replaces the integral by a sum; and it uses only 80 values of the Möbius function, $\mu(k)$. Still, unlike $Li(n)$, $Pi(n)$ does not stray ever farther from $\pi(n)$ even in this pretty good approximation, there being a positive difference at 10^8 , and negative at 10^9 . I could not wish for anything better.

Discussion: You can bank on it

This all started from the supposed optimal smooth approximation:

$$LL(n) = \ln(\text{lcm}(n)) \approx n - 1.$$

Why $n - 1$? Well, the plot looks a little less than n , and it's an equality at $n = 1$ if I set $LL(n) = n - 1$. But that doesn't change the implication that on average

$$\text{lcm}(n) = e \text{lcm}(n - 1),$$

and it was from this that my smooth "optimal" approximation, $PI(n)$, to $\pi(n)$ arose.

In closing, I find it difficult to believe that all of this doesn't exist in the literature somewhere. Even so, and even admitting I was too lazy to dig too deeply for it, it still remains true, in my view, that it should not require deep digging. This is a cool result and should be right up there with the prime number theorem.