CS 2051: Patterns in Primes

Anthony Hong
Georgia Institute of Technology

Yongyu Qiang
Georgia Institute of Technology

Aravinth Venkatesh Natarajan
Georgia Institute of Technology

March 12th, 2023

1 Background

- Definition of $\pi(x)$, the number of primes $\leq x$.
- Euclid's proof that $\lim_{x\to\infty} \pi(x) = \infty$
- Introduce the Riemann zeta function, $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$
- Notation for prime gaps: $g_n = p_{n+1} - p_n$
- A prime gap can be arbitrarily large: Consider $n! + 2, n! + 3, \ldots, n! + n$
- Are there infinite small prime gaps? Introduce twin prime conjecture
- Move on to some other main topic?

(hook to be added here)

A common result many students learn early on in number theory is about the infinitude of prime numbers. This fact is also known as Euclid's theorem, and we include a short summary of Euclid's original proof.

Euclid's Theorem. *The set of all prime numbers is larger in cardinality than any finite collection of prime numbers.*

Proof. Consider $p_1, p_2, \ldots, p_n$, some arbitrary finite collection of prime numbers. Let $N = p_1 p_2 \ldots p_n$, and consider $P = N + 1$. P is either prime or not prime.

First, let P be prime. Then, we have constructed a new prime number and we are done.

Now, let P not be prime. Let g be a prime factor of P. We propose that $g \notin \{p_1, p_2, \ldots, p_n\}$. To show this, suppose that $g \in \{p_1, p_2, \ldots, p_n\}$. Then, since $p_1, p_2, \ldots, p_n$ are all factors of N, we have $g|N$. $g|P$ and $g|N$, so we must also have $g|(P - N)$, i.e. $g|1$. But $g > 1$ (g is prime), so g cannot possibly divide 1. Therefore, $g \notin \{p_1, p_2, \ldots, p_n\}$, and we have found a new prime, as required. $\square$

2 Main result

A interesting problem is the rate of growth of prime gaps. $\pi(n)$ denotes the prime counting function, which counts the number of primes less than or equal to n. The Prime Number Theorem (PMT) states that $x/\ln(x)$ approximates $\pi(n)$. To be more precise, the PMT states:

$$\lim_{x\to\infty} \frac{\pi(x)}{\left\lfloor \frac{x}{\ln(x)} \right\rfloor} = 1$$

This means, as x gets larger, $x/\ln(x)$ will get better as an approximation for $\pi(x)$. This also implies that the average size of the gaps between consecutive primes up until x asymptotically approaches $\ln(x)$. So for a random number n in the interval $[x, x+kx]$ for large x and fixed k, the probability that n is prime is approximately $1/x \approx 1/n$.

Cramér's random model uses this idea as a naive approach to emulate the distribution of prime numbers. Consider a random subset of the natural numbers, where the independent probability that a number n is chosen is $1/\ln(n)$. Let's call this random set P', where P is the set of actual prime numbers. Cramér conjectured that P', which consists of our "fake primes", accurately models the distribution of P.

According to this heuristic, we have the resulting claim, which is known as Cramér's conjecture:

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{(\ln p_n)^2} = 1$$

where p_n denotes the n-th prime.

(Some directions to take these ideas):

- Problems with Cramér's naive model and ways we can improve it (with modern results)
- How Cramér's model fares depending on the size and location of the interval, calculating asymptomatic statistics

Another problem in number theory is finding the bounds in which you would find a prime number. Of these, one of the paramount significance is **Bertrand's Postulate**:
Bertrand's postulate states that for an integer $i > 1$, there is at least one prime number p

$$i \leq n \leq 2i$$

The proof for it is as follows:
We'll start by proving Lemma 1:

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

$$4^n = (1+1)^{2n} = \sum k = 0^{2n} \binom{2n}{k}$$

Since, $\binom{2n}{0}$ is 1 and $\binom{2n}{2n}$ is 1, this is the same as

$$\equiv 2 + \sum k = 1^{2n-1} \binom{2n}{k}$$

Since the largest term in this summation is $\binom{2n}{n}$ (since for $\binom{n}{k}$, $k = n/2$ will give the largest term) and there are $2n$ terms,

$$(2 + \sum k = 1^{2n-1} \binom{2n}{k}) \leq (2n * \binom{2n}{n})$$

Therefore,

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

Let's now prove Lemma 2:
For a given prime p, let's define r as the greatest number for which $p^r | \binom{2n}{n}$. Lemma 2 is as follows, for such an r,

$$p^r \leq 2n$$

Firstly, we have to introduce Legendre's Formula:
Legendre's Formula states that for any prime number p, and any integer n, let's define the function $v_p(n)$ as the exponent of the largest power of p that divides n. Let L = $\lfloor \log_p n \rfloor$
Legendre's Formula is:

$$v_p(n!) = \sum_{i=1}^{L} \left\lfloor \frac{n}{p^i} \right\rfloor$$

$\binom{2n}{n}$ can also be written as $\frac{(2n)!}{n!*n!})$
Finding the largest exponent of p, r, that divides $\frac{(2n)!}{n!*n!}$ is the same as finding the largest exponent of p, r, that divides each component of $\frac{(2n)!}{n!*n!}$, i.e. $(2n)!$, $n!$

In this case, $L = \lfloor \log_p 2n \rfloor$
Writing this in terms Legendre's Formula, we get that:

$$v_p(\binom{2n}{n}) = \sum_{i=1}^{L} \left\lfloor \frac{2n}{p^i} \right\rfloor - 2\sum_{i=1}^{L} \left\lfloor \frac{n}{p^i} \right\rfloor$$

This is equivalent to:

$$v_p(\binom{2n}{n}) = \sum_{i=1}^{L} \left\lfloor \frac{2n}{p^i} \right\rfloor - 2 \left\lfloor \frac{n}{p^i} \right\rfloor$$

Thinking intuitively, every term in $\sum_{i=1}^{L} \left\lfloor \frac{2n}{p^i} \right\rfloor - 2 \left\lfloor \frac{n}{p^i} \right\rfloor$ must either be 0 or 1.
If $(\frac{2n}{p^i} \bmod 1) \geq 0.5$ then the term would be 1, otherwise the term is 0. Therefore, the maximum value of this function would be if all the terms were equal to 1. Since we are only dealing with positive numbers, and the exponent is a monotonic function if both numbers are positive numbers,

$$v_p(\binom{2n}{n}) = r \leq L$$

$$\equiv r \leq \log_p 2n$$

Therefore, it follows that

$$p^r \leq p^{\log_p 2n} = 2n$$

We are able to prove our initial Lemma 2,

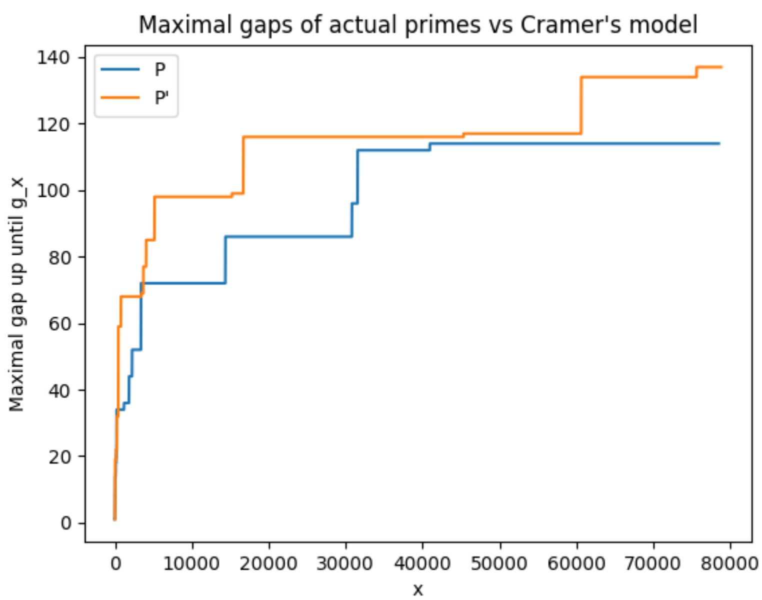
$$p^r \leq 2n$$

3 Extension/application/generalisation

- Connections from Cramér's conjecture to the Riemann hypothesis
- Other ways to use Cramér's technique of random modeling

4 Preliminary Code and Illustrations

Cramér's random model allows us to heuristically test properties of primes. In this example, we graphically compare the maximal prime gap of the model and the actual primes.




```
def sieve(n):
    prime = [True for i in range(n + 1)]
    i = 2
    while i * i <= n:
        if prime[i]:
            for j in range(i * i, n + 1, i):
                prime[j] = False
        i += 1

    last_prime = -1
    gaps = []
    for i in range(n + 1):
        if prime[i]:
```

```
            if last_prime != -1:
                if len(gaps) == 0:
                    gaps.append(i - last_prime)
                else:
                    gaps.append(max(gaps[-1], i - last_prime))
            last_prime = i

    return gaps

def cramer_model(n):
    # Assume 2 is in the model because 1 / ln(2) > 1
    gaps = []
    last_prime = 2
    for i in range(3, n + 1):
        if random.random() <= 1 / np.log(i):
            if len(gaps) == 0:
                gaps.append(i - last_prime)
            else:
                gaps.append(max(gaps[-1], i - last_prime))
            last_prime = i

    return gaps
```

5 Reflection/Conclusion

Under heuristic testing with variations of Cramér's model, we should be able to support strong statements such as Bertrand's postulate and Legendre's conjecture. However, comparing with the actual primes, it should be clear that Cramér's model is inaccurate. Further work would involve the creation and tuning of other random models, to more closely emulate prime distributions.

6 References

- Terence Tao's Blog
- Wikipedia