How frequently is a prime the smallest divisor

Daniel Eid

December 2024

Abstract

In this paper I will cover how to calculate how frequently a prime is the smallest divisor of all positive integers . I will use that information to approximate how likely specific integers are to be prime . And how to find positive large integers which are more likely to be prime. Finally, I will construct a proof utilizing these approximations.

Key Definitions

1. **Coprime (Relatively Prime) Integers:** Two positive integers a and b are coprime if their greatest common divisor is 1:

$$\gcd(a, b) = 1$$

$$\gcd(7, 3) = 1.$$

2. **Primorial :** The primorial of p_n, denoted P_n, is the product of the first n primes :

$$P_n = \prod_{p \leq p_n} p.$$

$$P_4 = \prod_{p \leq 7} p = 7 \cdot 5 \cdot 3 \cdot 2 = 210.$$

3. **Totient Function :** The Euler totient function $\phi(a)$ is calculated using the prime factorization of a. If a has the prime factorization:

$$a = p_1^{e_1} p_2^{e_2} \cdots p_b^{e_b},$$

then the totient function is given by:

$$\phi(a) = a \prod_{i=1}^{b} \left(1 - \frac{1}{p_i}\right),$$

where $p_1, p_2, \ldots, p_b$ are the distinct prime factors of a. This formula represents the count of integers $x < a$ that do not share any factors with a .

$$\phi(6) = \phi(2 \cdot 3) = 6 \cdot (1 - \frac{1}{2}) \cdot (1 - \frac{1}{3}) = 6 \cdot \frac{1}{2} \cdot \frac{2}{3} = 2$$

1 Prime Frequency Theorem

We will calculate the number of times in which the nth prime is the smallest factor of any positive integer. We will express this count in terms of n using an approximation from Mertens third theorem. We will compare its count to the interval that it is measured in , to determine a relative frequency. This provides a heuristic to determine how frequently any positive integer is uniquely factorable by some prime .

Objective

The objective is to leverage unique factorization to account for all positive integers . Given the fundamental theorem of arithmetic we know that all positive integers can be written as a unique product of primes . Each positive integer will have only one prime as its smallest divisor . This will be the premise for unique factorization . We will calculate how frequently some prime divides all positive integers as the smallest factor . This is done so each positive integer is counted only once , when some unique factorization condition is met .

Using the Totient Function to Determine the number of times a Prime is the Smallest Factor Over a Fixed Interval

$P_n = a \cdot p_n$

Consider the product $a \cdot p_n$, where $a = p_1 \cdot p_2 \cdot \cdots \cdot p_{n-1}$ is the product of the first $n-1$ primes , and p_n is the n-th prime . We wish to determine the number of times p_n is the smallest factor of integers in the interval $[1, a \cdot p_n]$.
Let $n = 3$ then $p_3 = 5$. a is the product of all primes less than $p_3 = 5$.

$$a = 2 \cdot 3 = 6$$

$$P_n = a \cdot p_n = 6 \cdot 5 = 30$$

Integers Factorable by p_n

All integers in $[1, a \cdot p_n]$ that are divisible by p_n take the form:

$$p_n \cdot 1,\ p_n \cdot 2,\ p_n \cdot 3,\ \ldots,\ p_n \cdot a.$$

This set contains a elements, as a represents the range of possible values for the multiplier m.

Integers divisible by no smaller primes than p_n

For p_n to be the smallest factor of an integer $p_n \cdot m$, m must not share any factors with a, if m shares a factor with a , then m is divisible by a prime less than p_n , since a is composed of all the primes less than p_n. The number of integers j in

$[1, a]$ that are coprime to a is given by $\phi(a)$. Thus, there are $j = \phi(a)$ integers in $[1, a \cdot p_n]$ which take the form $p_n \cdot m$, where p_n is the smallest factor . The totient function $\phi(a)$ allows us to determine the exact number of times p_n is the smallest factor of integers in the interval $[1, a \cdot p_n]$. This is precisely $\phi(a)$, as it counts the values of m coprime to a .
Use the totient function to exclude any m divisible by a factor of a , when $a = 6$ and $p_n = 5$.

$$5 \cdot 1, \cancel{5 \cdot 2}, \cancel{5 \cdot 3}, \cancel{5 \cdot 4}, 5 \cdot 5, \cancel{5 \cdot 6},$$

$$\phi(6) = 6 \cdot (1 - \frac{1}{2}) \cdot (1 - \frac{1}{3}) = 6 \cdot \frac{1}{2} \cdot \frac{2}{3} = 2$$

Frequency of p_n as the Smallest Factor

To compute the relative frequency in $[1, a \cdot p_n]$ where p_n is the smallest factor , we will use the ratio of the number of times that p_n is the smallest factor , in the interval of its primorial P_n:

$$\text{Frequency} = F_n = \frac{\phi(a)}{a \cdot p_n} = \frac{\phi(P_{n-1})}{P_n}$$

For $p_3 = 5$ and $a = 6$

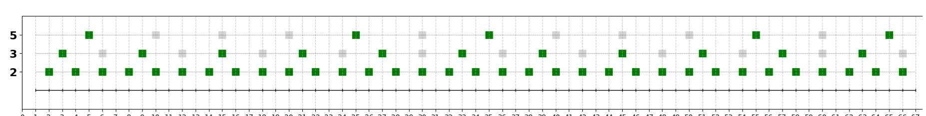
$$F_3 = \frac{\phi(a)}{a \cdot p_n} = \frac{\phi(6)}{6 \cdot 5} = \frac{\phi(6)}{30} = \frac{2}{30}$$

We can say that $p_3 = 5$ is the smallest divisor of $\frac{2}{30}$ of all integers .

Case for F_1 , $p_1 = 2$

F_1 is a special case since there are no primes smaller than 2 . So the frequency for 2 will be calulated heuristically . Since 2 divides $\frac{1}{2}$ of all integers and there are no primes smaller than 2 we will say that F_1 is $\frac{1}{2}$. 2 is the smallest divisor of $\frac{1}{2}$ of all integers .

Visual Proof that the frequency repeats



Green ticks are when p_n is the smallest factor of an integer . Grey ticks are for when p_n divides an integer but there are smaller primes which also divide that integer . We can see that the frequency of p_n to divide an integer as the smallest factor, repeats over the interval of its primorial . This is because frequency is measured relative to smaller primes and its primorial is the least common factor $a \cdot p_n$.

Approximating the behavior of F_n for large n

We start from

$$F_n = \frac{\phi(P_{n-1})}{P_n}.$$

$$F_n = \frac{\phi(P_{n-1})}{P_n} = \frac{P_{n-1}\prod_{i=1}^{n-1}\left(1-\frac{1}{p_i}\right)}{P_{n-1}\cdot p_n}.$$

$$F_n = \frac{\cancel{P_{n-1}}\prod_{i=1}^{n-1}\left(1-\frac{1}{p_i}\right)}{\cancel{P_{n-1}}\cdot p_n} = \frac{\prod_{i=1}^{n-1}\left(1-\frac{1}{p_i}\right)}{p_n}.$$

We can use Mertens' third theorem aka Mertens' product formula , to approximate the numerator asymptotically , we have

$$\prod_{i=1}^{n-1}\left(1-\frac{1}{p_i}\right) \sim \frac{e^{-\gamma}}{\ln p_{n-1}}.$$

where $\gamma \approx .577$ is the Euler–Mascheroni constant. This provides us with an asymptotic approximation for the numerator .
Since $p_{n-1} \approx p_n$ for large n, this gives

$$\prod_{i=1}^{n-1}\left(1-\frac{1}{p_i}\right) \approx \frac{e^{-\gamma}}{\ln p_n}.$$

To complete the original definition we must factor in the denominator .

$$F_n = \frac{1}{p_n}\prod_{i=1}^{n-1}\left(1-\frac{1}{p_i}\right) \approx \frac{1}{p_n}\cdot\frac{e^{-\gamma}}{\ln p_n}.$$

Given

$$F_n \approx \frac{e^{-\gamma}}{p_n \ln p_n},$$

and $p_n \sim n\ln n$, we have

$$F_n \approx \frac{e^{-\gamma}}{n\ln n\cdot\ln(n\ln n)}$$

$$F_n \approx \frac{e^{-\gamma}}{n\ln n\cdot(\ln n + \cancel{\ln\ln n}))}$$

For large n , F_n will grow as $\ln n$ as opposed to $\ln\ln n$, since $\ln n \gg \ln\ln n$ for large n .

$$F_n \approx \frac{e^{-\gamma}}{n(\ln n)^2}$$

$$
\begin{aligned}
F_{11} &\approx \frac{e^{-\gamma}}{11\,(\ln 11)^2} \\
&\approx \frac{0.561459}{63.265169} \\
&\approx 0.008878
\end{aligned}
$$

Conclusion

Above I have provided a means for calculating and approximating unique prime frequency in terms of n as $F_n = \frac{\phi(P_{n-1})}{P_n}$. This is useful in problem contexts where divisibilty with respect to certain primes are necessary . Existing methods only utilize simple heuristics to determine the frequency that a prime is a factor of all integers ie: $\frac{1}{p_n}$, which is limited in its usefulness depending on the problem setup .

2 Divisibilty Decay Theorem

I will show that as n increases, the likelihood that any primes larger than p_n are the smallest divisor of some positive integer q , diverges to 0.

Sum of frequencies

To determine the likelihood that an integer q is composite , we must consider the likelihood that q is divisible by any smaller prime. This can be done using a sum with respect to the frequencies of all smaller primes . The total sum of frequencies S_n of integers in $[1, P_n]$ that are divisible by some prime p_i up to p_n, for which p_i is the smallest factor, is given by:

$$S_n = \sum_{i=1}^{n} \frac{\phi(P_{i-1})}{P_{i-1} \cdot p_i} = \sum_{i=1}^{n} \frac{\phi(P_{n-1})}{P_n} = \sum_{i=1}^{n} F_i \approx \sum_{i=1}^{n} \frac{e^{-\gamma}}{i(\ln i)^2}.$$

S_n is the measure of the likelihood of an integer q to share a factor with a prime less than or equal to p_n . Let $G = \{p_i \leq p_n | p_i\}$, then the likelihood that an integer is divisble by G is S_n . S_n strictly increases for subsequent n , because all terms are positive .

Divisibility by primes not in G

Let $K = \{p_i > p_n | p_i\}$. If q is not divisble by an element in G, then it must be divisble by an element in K , since G and K inlcude all primes, and via the fundamental theorem of arithmetic at least one of these primes must factor q. The likelihood that an integer is not factorable by G is equal to the tail sum of S_n , we will call this tail/end sum E_n . E_n is the likelihood that an element of K divides some q.

Quantifying the Tail Sum of S_n as E_n

We have:

$$E_n = \sum_{m>n} F_m,$$

To estimate E_n, consider the sum:

$$E_n = \sum_{m>n} \frac{e^{-\gamma}}{m(\ln m)^2}.$$

For large n, we can approximate sums by integrals. Specifically:

$$\sum_{m>n} \frac{1}{m(\ln m)^2} \approx \int_n^\infty \frac{dx}{x(\ln x)^2}.$$

Simplify and Evaluate the Integral

$$u = \ln x$$

$$\frac{du}{dx} = \frac{d}{dx}(\ln x) = \frac{1}{x} \quad \Rightarrow \quad du = \frac{1}{x}dx \quad \Rightarrow \quad dx = x\,du$$

$$\frac{1}{x(\ln x)^2} = \frac{1}{xu^2}$$

$$I = \int_n^\infty \frac{dx}{xu^2} = \int_{\ln n}^\infty \frac{x\,du}{xu^2} = \int_{\log n}^\infty \frac{du}{u^2}$$

$$\int \frac{du}{u^2} = \frac{1}{u^2} = u^{-2} = \frac{u^{-2+1}}{-2+1} = \frac{u^{-1}}{-1} = -\frac{1}{u} + C$$

$$I = \left[-\frac{1}{u}\right]_{\ln n}^\infty = \left(-\frac{1}{\infty}\right) - \left(-\frac{1}{\ln n}\right) = 0 - \left(-\frac{1}{\ln n}\right) = \frac{1}{\ln n}$$

Thus:

$$\int_n^\infty \frac{dx}{x(\ln x)^2} = \frac{1}{\ln n}.$$

Incorporating the constant $e^{-\gamma}$:

$$E_n \approx e^{-\gamma} \cdot \frac{1}{\ln n} \quad \text{as } n \to \infty.$$

E_n is on the order of $\frac{1}{\ln n}$. If the sum defining E_n is not from n to ∞ but to some finite upper limit M we can write:

$$\int_n^M \frac{dx}{x(\ln x)^2} = \left(-\frac{1}{\ln M}\right) - \left(-\frac{1}{\ln n}\right).$$

If M is extremely large the dominant term is $1/\ln n$, ensuring it remains on the order of $1/\ln n$ for large n.

$$E_n \approx \frac{e^{-\gamma}}{\ln n}.$$

$$\begin{aligned} E_{15} &\approx \frac{e^{-\gamma}}{\ln 15} \\ &\approx \frac{0.561459}{2.708050} \\ &\approx 0.2073 \end{aligned}$$

$$\begin{aligned} E_{16} &\approx \frac{e^{-\gamma}}{\ln 16} \\ &\approx \frac{0.561459}{2.772589} \\ &\approx 0.2026 \end{aligned}$$

$$E_{15} > E_{16}$$

Conclusion

E_n is defined as the likelihood that some element of K divides some q as the smallest factor. We can see that $E_n \to 0$ as $n \to \infty$. We have proved that the likelihood of any element in K being the smallest divisor decreases as n increases. Since S_n strictly increases , and E_n goes to 0 for large n , we can deduce that as n increases if q is coprime to G then it has a decreasing likelihood to be factorable by K . If q is coprime to G then it has an increasing likelihood of being prime for large n . Existing approaches to divisibility of q do not permit for such fine grained control of unique relative divisibility.

3 Adjacent Coprime Theorem

Finding prime integers

We will identify a point C_n on the number line, which will be factorable by all elements in G . I will then show that $C_n \pm 1$ are not factorable by G . Lastly, I will show that the likelihood that $C_n \pm 1$ is composite decreases for larger n.

Proving the likelihood that either $C_n + 1$ or $C_n - 1$ are prime increases for large n

Let C_n be the product of the first n primes $G = \{2, 3, 5, ..., p_n\}$. C_n may also be defined as the product of the first z positive integers where $G = \{p_i \leq z | p_i\}$. This means C_n is factorable by all elements in G . It can be shown that $C_n \pm 2$ is factorable by 2 , $C_n \pm 3$ is factorable by 3 , $C_n \pm 5$ is factorable by 5 ... $C_n \pm p_n$

is factorable by p_n . It can be deduced that $C_n \pm 1$ is not factorable by any of the elements in G . $C_n \pm 1$ is coprime to G . There are primes $p_{n+1}, p_{n+2}, ...$ in the interval $[p_n, C_n + 1]$, which may be a factor of $C_n \pm 1$. This is defined as $E_n = \sum_n^M F_m$, where $M = C_n + 1$. Let K be the set of primes accounted for in E_n. If $C_n \pm 1$ is coprime to G and coprime to K then $C_n \pm 1$ is prime. As n increases , the likelihood of $C_n \pm 1$ being divisible by K decreases via the Divisibility Decay Theorem . The likelihood that either $C_n + 1$ or $C_n - 1$ is composite is approximately E_n for each , so the combined likelihood that either is composite is at most $\approx E_n + E_n = 2 \cdot E_n$. Therefore, as n increases, the likelihood that either $C_n \pm 1$ are composite decreases as $2 \cdot E_n$. It follows that as $n \to \infty$, the likelihood that both $C_n - 1$ and $C_n + 1$ are both prime is at least $(1 - 2 \cdot E_n)$, which increases for large n . Thus, the likelihood that $C_n \pm 1$ is prime increases for larger n.

Conclusion

I have proven that the likelihood of $C_n \pm 1$ being prime increases for large n. This is useful because it assists in finding primes at specific increasingly large integers, with an increasing likelihood for large n. This is opposed to current exhaustive methods which find primes with decreasing likelihoods for larger integers .

Infinite Twin Primes Proof

Twin primes are two primes seperated by a distance of 2 .
Let A denote the infinite ordered set of all prime numbers. We construct a subset $B \subseteq A$ consisting of P_n where $P_n \pm 1$ are both prime. The construction proceeds iteratively as follows:

1. For each positive integer n, define the primorial P_n as the product of the first n primes:
$$P_n = \prod_{i=1}^{n} p_i,$$
2. Define B as the set of primorials P_n for which both $P_n - 1$ and $P_n + 1$ are prime. Formally,
$$B = \left\{ P_n \in A \,\middle|\, P_n - 1 \text{ and } P_n + 1 \text{ are both prime} \right\}.$$
3. The likelihood that $P_n \pm 1$ are both prime increases as n increases, by the Adjacent Coprime Theorem.
4. The set B is expected to contain more elements as n grows.
5. Since B is constructed by iteratively including primorials where $P_n \pm 1$ are both prime, and the likelihood of inclusion increases with n, we infer that B is an infinite subset of A.
6. There are an infinite number of twin primes.