On the Infinitude of Twin Primes

Benny Lim

December 2023

1 Abstract

This is by no means a rigorous proof of the Twin Prime Conjecture. However, if certain hand-waving assumptions made in this paper can be proven, I believe the methods presented here can certainly contribute to the Conjecture's eventual proof.

2 Introduction

We introduce the idea of the number of Candidates C_n being the upper limit of twin prime pairs within the range $(p_n^2,\, p_{n+1}^2)$ for $n > 2$.

Using C_n as the upper limit, we derive a crude estimate E_n as an approximation of the true number of twin prime pairs.

Applying some restrictions on E_n, we assert a lower limit L_n of twin prime pairs in the range $(p_n^2,\, p_{n+1}^2)$, which turns out to be non-zero for all $n > 2$. Therefore if we assume L_n to be true, then the Twin Prime Conjecture must also be true.

3 Counting Twin Primes in the range (p_n^2, p_{n+1}^2)

3.1 The Sixes

Given that all p_n where $n > 2$ have the form $6m-1$ or $6m+1$, the boundaries p_n^2 and p_{n+1}^2 must both be 1 greater than a multiple of 6.

$$\begin{aligned}(6m-1)^2 &= 36m^2 - 12m + 1 \\ &= 6 \cdot (6m^2 - 2m) + 1\end{aligned} \tag{1}$$

$$\begin{aligned}(6m+1)^2 &= 36m^2 + 12m + 1 \\ &= 6 \cdot (6m^2 + 2m) + 1\end{aligned} \tag{2}$$

Since the boundaries are both 1 greater than a multiple of 6, their difference must be a multiple of 6.

$$p_{n+1}^2 - p_n^2 = 6m \tag{3}$$

In the range $p_n^2 < x < p_{n+1}^2$, the count of x where x is a multiple of 6 is simply $\frac{p_{n+1}^2 - p_n^2}{6}$.

In a low-magnitude example where $p_3 = 5$, the boundaries would be $p_3^2 = 5^2 = 25$ and $p_4^2 = 7^2 = 49$, and all x which are multiples of 6 are in the set $\{30, 36, 42, 48\}$. The size of this set being:

$$\begin{aligned}\frac{p_{n+1}^2 - p_n^2}{6} &= \frac{49 - 25}{6} \\ &= \frac{24}{6} \\ &= 4\end{aligned} \tag{4}$$

3.2 The Candidates

A twin prime Candidate is defined as a tuple of the form $(6m-1,\ 6m+1)$ which exists in the range $(p_n^2,\ p_{n+1}^2)$. To generate a set of all possible Candidates, we first derive the subset of x which are multiples of 6. To this subset, we transform each element into its corresponding tuple of neighbours (± 1).

The set from the example above $\{30,\ 36,\ 42,\ 48\}$ would thus be transformed into $\{(29,\ 31),\ (35,\ 37),\ (41,\ 43),\ (47,\ 49)\}$

We note that since the upper bound p_{n+1}^2 is always of the form $6m+1$, then the largest multiple of 6 in the range $(p_n^2,\ p_{n+1}^2)$ must be $p_{n+1}^2 - 1$, hence the largest Candidate $(p_{n+1}^2 - 2,\ p_{n+1}^2)$ always contains the upper bound p_{n+1}^2 which is composite.

Therefore, we must always disqualify the largest Candidate in the set of Candidates generated for any $(p_n^2,\ p_{n+1}^2)$. Revisiting the example of $p_3 = 5$, the updated set of Candidates would be $\{(29,\ 31),\ (35,\ 37),\ (41,\ 43)\}$, having removed $(47,\ 49)$.

We are not as concerned with the values of the Candidates as we are with the *number* of Candidates, which we denote C_n.

$$C_n = \frac{p_{n+1}^2 - p_n^2}{6} - 1 \tag{5}$$

For $n > 2$, C_n is the upper limit of twin prime pairs in the range $(p_n^2,\ p_{n+1}^2)$.

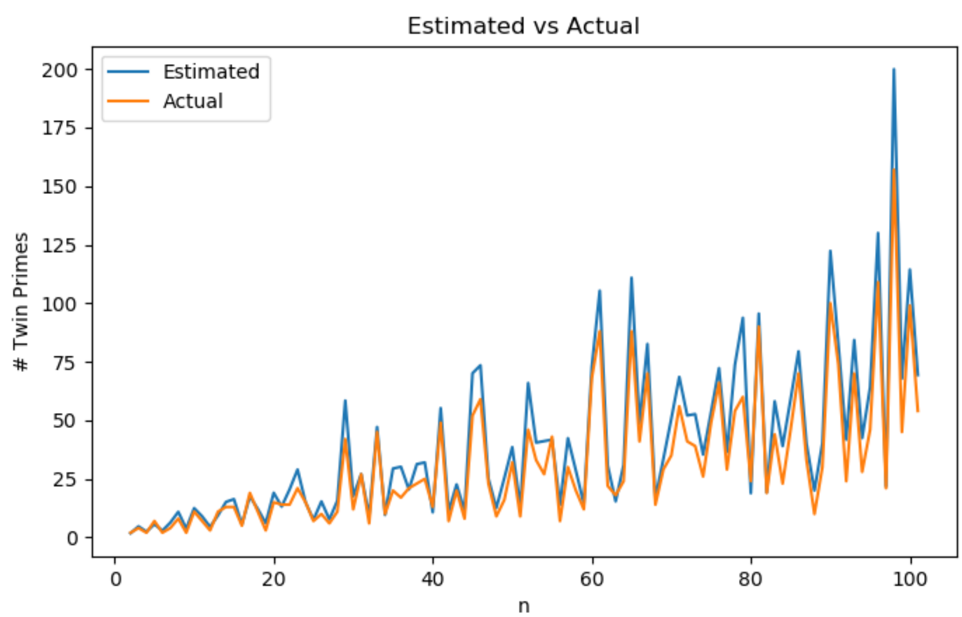
3.3 The Estimation

For large n, a reasonable estimate E_n for the number of twin prime pairs in the range $(p_n^2,\, p_{n+1}^2)$ would be:

$$\begin{aligned} E_n &= C_n \left(\frac{p_3 - 2}{p_3}\right)\left(\frac{p_4 - 2}{p_4}\right)\cdots\left(\frac{p_{n-1} - 2}{p_{n-1}}\right)\left(\frac{p_n - 2}{p_n}\right) \\ &= C_n \left(\frac{3}{5}\right)\left(\frac{5}{7}\right)\left(\frac{9}{11}\right)\left(\frac{11}{13}\right)\cdots\left(\frac{p_{n-1} - 2}{p_{n-1}}\right)\left(\frac{p_n - 2}{p_n}\right) \end{aligned} \tag{6}$$

Since each prime p in the range $[p_3,\, p_n]$ has a $\frac{2}{p}$ chance of being a factor of any Candidate, we calculate E_n by taking C_n multiplied by the finite product series of $\frac{p-2}{p}$. This process stochastically eliminates the Candidates which are rendered composite by each prime in the range $[p_3,\, p_n]$.

Though E_n is prone to slightly overestimating, its correlation with the actual count is strong.



Graph 1 – E_n vs Actual for $2 \leq n \leq 102$

3.4 The Hand-waving Lower Limit

In equation (6) above, the estimation E_n is not particularly easy to reduce. To simplify things, we make an extremely hand-waving assumption: by substituting the finite product series with **all the odd numbers up to** p_n instead of sequential prime numbers, we will instead derive the lower limit L_n of twin prime pairs in the range $(p_n^2,\, p_{n+1}^2)$.

Essentially pretending that all odd numbers are prime, we are eliminating far more Candidates than necessary. Doing this also allows us to efficiently reduce the product series:

$$
\begin{aligned}
L_n &= C_n \left(\frac{3}{5}\right)\left(\frac{5}{7}\right)\left(\frac{7}{9}\right)\cdots\left(\frac{p_n-4}{p_n-2}\right)\left(\frac{p_n-2}{p_n}\right) \\
&= C_n \left(\frac{3}{\cancel{5}}\right)\left(\frac{\cancel{5}}{\cancel{7}}\right)\left(\frac{\cancel{7}}{\cancel{9}}\right)\cdots\left(\frac{\cancel{p_n-4}}{\cancel{p_n-2}}\right)\left(\frac{\cancel{p_n-2}}{p_n}\right) \\
&= C_n \left(\frac{3}{p_n}\right)
\end{aligned}
\tag{7}
$$

Since the set of primes is a minuscule subset of the set of odd numbers (with the exception of 2), then $L_n \ll E_n$ for sufficiently large n.

3.5 Infinitely many Twin Primes

Given $n > 2$, we can denote p_{n+1} as $p_n + k$ for some even k.

Expanding C_n,

$$\begin{aligned} C_n &= \frac{p_{n+1}^2 - p_n^2}{6} - 1 \\ &= \frac{(p_n + k)^2 - p_n^2}{6} - 1 \\ &= \frac{p_n^2 + 2kp_n + k^2 - p_n^2}{6} - 1 \\ &= \frac{2kp_n + k^2}{6} - 1 \end{aligned} \tag{8}$$

Assuming n is large,

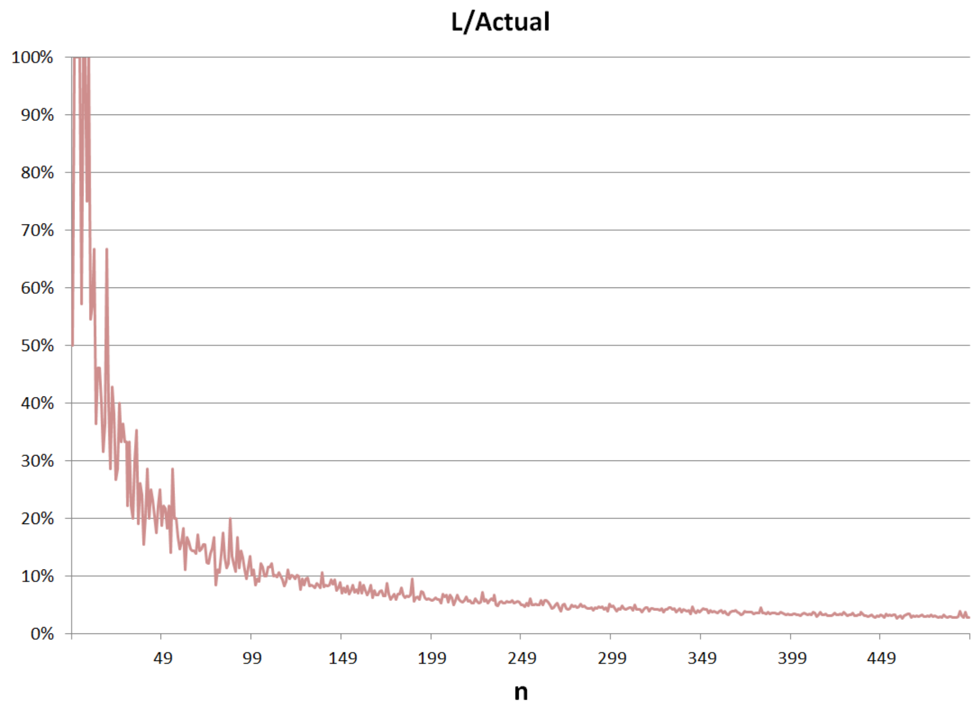
$$C_n \approx \frac{2kp_n}{6} \tag{9}$$

Now taking the lower limit L_n defined in equation (7),

$$\begin{aligned} L_n &= C_n \left(\frac{3}{p_n} \right) \\ &= \frac{2kp_n}{6} \left(\frac{3}{p_n} \right) \\ &= k \end{aligned} \tag{10}$$

An astonishing result! By showing $L_n = k$ for large n, we are asserting that the number of twin prime pairs in the range (p_n^2, p_{n+1}^2) can be no fewer than k, where $k = p_{n+1} - p_n$.

If the assumption of L_n being a valid lower limit of twin prime pairs is true for **all** $n > 2$, then there must be infinitely many twin primes.



Graph 2 – L_n/Actual for $3 \leq n \leq 499$

Graph 2 above shows L_n as a percentage of the actual count of twin prime pairs in the range $(p_n^2,\ p_{n+1}^2)$. Note that L_n never exceeds the actual count and maintains a steadily decreasing percentage as n grows. These empirical observations further support the assertion that L_n serves as a valid lower limit.