Prime Gaps

The Structure and Properties of the Differences between Successive Prime Numbers

Jerrad Neff

Table of Contents

1. Defining the Prime Gap Function

In the field of mathematics, few sequences have been as studied and closely examined as that of the prime numbers; positive integers only wholly divisible by one and themselves. A simple observation of even the first several prime numbers (whose primality can be checked by hand;) 2, 3, 5, 7, 11, 13, 17, 19, 23, 29… reveals a few apparent possible properties of these famously unpredictable numbers.

We can easily see that most are odd (obvious, since every even integer greater than two is divisible by two,) and a surprising proportion of the first part of the sequence are two numbers apart. (These are known as Twin Primes.) Four seems to be another relatively popular difference in successive primes, but this raises a question. Even though we see no concrete patterns in the sequence of prime numbers, will there be any properties of the *difference* between these numbers that may shed more light on their structure and attributes? To do so, let us consider the *prime gap function g_n*, defined by:

$$g_n = p_{n+1} - p_n$$

Where g_n is the n-th prime gap, and p_n are the n-th prime numbers. For example, $g_{30} = p_{31} - p_{30} = 131 - 127 = 4$. Like the prime numbers themselves, if we take a look at the first several terms in this sequence, we can also make some interesting observations. The first terms of g_n are as follows: 1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6, 4, 2, 4, 6, 6, 2, 6, 4, 2, 6, 4, 6, 8… We can clearly see that near the beginning of the sequence, two seems to be the most popular number, but six begins occurring often towards the end of the visible terms.

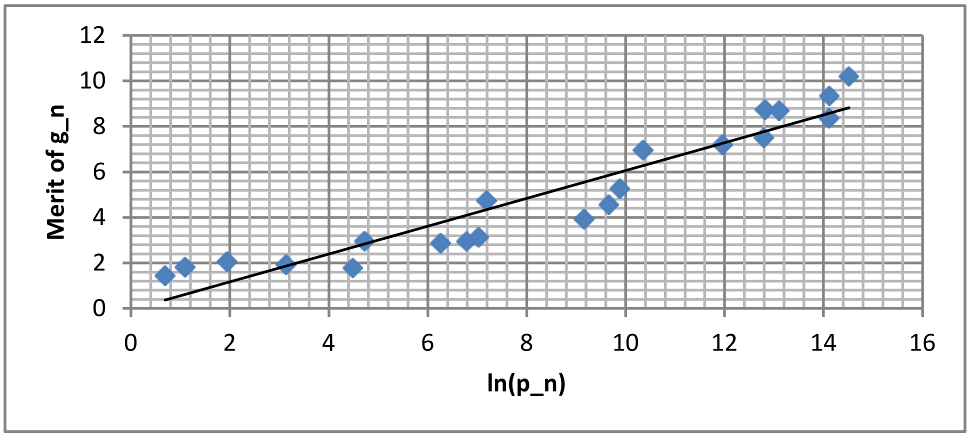
Also, we may wonder if, even though the largest number appearing in the sequence (also known as the **maximal gap**) appears to increase the farther and farther we go out in the terms of the sequence, can the behavior of these terms be predicted with any kind of accuracy? Or is there any way we can put an upper or lower bound on these maximums?

2. Discoveries and Data

So now that we have a function defined as the differences between these concurrent prime numbers, we should take a look at the data it contains when extrapolated out over a wide range of primes. As the sequence carries on, larger and larger gaps appear as the prime numbers in question increase. These largest such gaps in the sequence are called **maximal gaps**, and the first several observed are the following:

G_n	After #	Merit	G_n	After #	Merit	G_n	After #	Merit
1	2	1.4426950408889634	20	887	2.9464432455549328	86	155921	7.1923765672871003
2	3	1.8204784532536746	22	1129	3.1298514635393411	96	360653	7.5025370554622377
4	7	2.055593369479003	34	1327	4.7283454069777955	112	370261	8.7350116471909871
6	23	1.9135739334228061	36	9551	3.9282435862321434	114	492113	8.6979984149315523
8	89	1.7822784785922416	44	15683	4.5547086018837817	118	1349533	8.3597413994470227
14	113	2.9614663608904754	52	19609	5.2611642311336286	132	1357201	9.3478228987480492
18	523	2.875591619557091	72	31397	6.9535202195900787	148	2010733	10.197044174552893

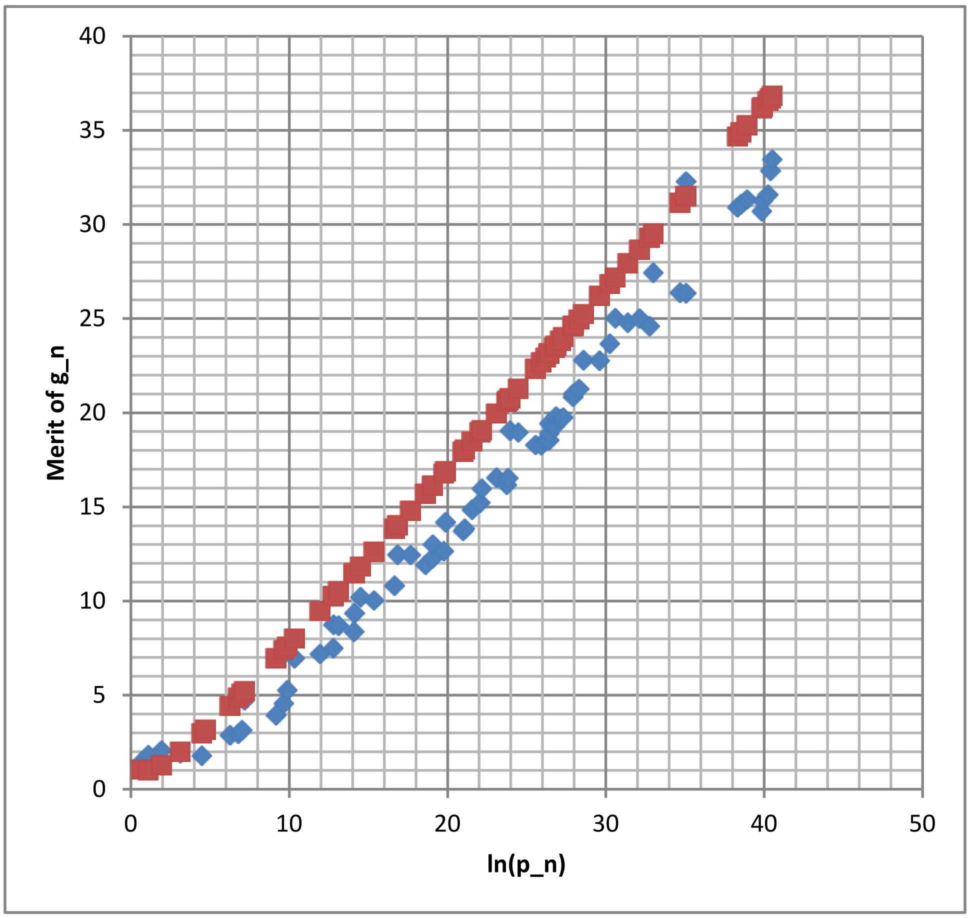
The **merit** of the prime gaps in the table above is a normalized number representing how "soon" in the sequence a prime gap appears, relative to the logarithm of the larger prime. This sequence is important in the context of prime gaps because if we graph them out, they rise approximately as the logarithm of the prime gap scaled by a constant, and are represented by the equation $(p_{n+1} – p_n)/(\log p_n)$. A graph of the merits of the first twenty-one prime gaps against the logarithm of their leading prime is below:



3. A Short Conjecture on Merit

By observing the table at the end of section 2 in which the merit of the first 21 prime gaps appears to rise almost linearly, we can ask the question: does this linear rise continue out past these first terms? To investigate, we will take a look at a complete list of all known (last updated January 2010) maximal prime gaps, and plot it out against their merit. (Below, in blue.)

Let us then take a look at the logarithm of the prime counting function estimate, ln(x/ln(x)), and plot it against the data given above, (in red,) and a correlation arises:



Neff's Merit Conjecture:

$$\textit{Merit of Maximal Gaps } g_n \sim \ln\left(\frac{p_n}{\ln(p_n)}\right)$$

4. Twin Primes and Beyond

One of the questions that arises from observing the pattern of the prime gap series is if, in this infinite sequence of numbers, does $g_n = 2$ occur infinitely often? These primes spaced exactly two apart are more often referred to as twin primes, the first few pairs of which being (3, 5); (5, 7); (11, 13); etc. Although it seems obvious that this is true by just looking at a large list of prime numbers and observing that even as the prime numbers become large, these twin primes still occur relatively often, the Twin Prime Conjecture, which states exactly this, is still a fully open problem in number theory.

But the same problem also remains and is yet to be proven for larger gaps; if there are an infinite number of gaps of size 2, will this mean that there are infinite gaps of size 2n for any n? It is suspected so, but like the Twin Prime Conjecture, it is yet to be proven. However, a good exercise to visualize the possibility of this is to ask the question: Does g_n ever get large enough to cover all such 2n?

The answer is yes; there is no absolute upper bound on g_n. To find any gap of size N or larger, first consider the sequence of numbers (N! + 2), (N! + 3), (N! + 4), …, (N!+N). We can clearly see that 2 | (N! + 2) since it divides each of the terms separately, therefore it is composite. 3 | (N! + 3), 4 | (N! + 4), and so on all the way up to N | (N! + N) The terms surrounding all of the guaranteed composite terms of the sequence, N! + 1 and N! + N + 1, are the closest possible prime numbers. (Though they are still not guaranteed to be prime!) Using this, we can say that any gap of size N or larger must occur before N! + N + 1.

Even though we have this infinitely high bound for the size of g_n, we are still not guaranteed that every even value for this function will be hit by some prime gap. We know that any such gap size will be exceeded, but the conditions for which N! + 1 and N! + N + 1 are guaranteed to be prime is still not known, and according to my sources, no theorems or conjectures exist for such a condition.

5. Jumping Champions

Taking a look back at the table and graphs presented in the beginning of section 2, we could clearly see that the farther and farther we go out in our list of primes, the larger the maximal gap seems to appear, seemingly asymptotic to the logarithm of the prime number in question. However, we have yet to address if these maximums are the only thing growing as the list goes on. Let us take a look at instead how often certain gap sizes occur for the first primes p_n, and a few select gap sizes:

	$g_n = 2$	$g_n = 4$	$g_n = 6$	$g_n = 8$	$g_n = 10$	$g_n = 12$
n < 10	5	3	1	0	0	0
n < 20	8	6	5	0	0	0
n < 40	12	12	12	1	1	0
n < 80	21	21	22	4	6	2
n < 160	35	39	40	14	16	7
n < 320	66	65	81	27	33	21
n < 640	122	115	155	54	63	47

We can see that $g_n = 2$ starts out as the most popular gap size until around $p_{40} = 179$, at which point $g_n = 6$ begins to clearly dominate the gap size distribution. Mathematicians have dubbed these "most popular" prime gaps as **jumping champions**, and upon observation far out into the list of known prime numbers, the value of the jumping champion seems to increase in a very peculiar manner.

It is conjectured that beginning with 2 and 6 as seen in the table above, the next gap size to become the jumping champion is 30, followed by 210, and so on; these numbers are in the sequence known as **primorial numbers**, simply the product of the first however-many prime numbers. (6 = 2*3, 30 = 2*3*5, 210 = 2*3*5*7, etc.) Currently, the computing power available to even compute the prime gaps past where $g_n = 30$ is the jumping champion is not available, which occurs around 10^{425}. Even using simplified formulas to estimate the number of occurrences of a given gap under any given prime (such as conjectured by Brent, 1974), we can't even come close to estimating when 2310 (the next conjectured jumping champion after 210) takes over as the jumping champion.

However, there is yet another conjecture still standing, proposed by Marek Wolf, that the next jumping champion, $(p_5)\# = 2310$ will become the jumping champion at approximately $5^{(5\wedge5)} \sim 1.9*10^{2184}$. It may very well be a long time until we can even use another conjecture to verify that this result may be anywhere in the neighborhood of the correct value, let alone before we will actually have the CPU time to calculate all the prime gaps to this extraordinarily large number.

6. References

Caldwell, Chris K.. "The Prime Pages". The University of Tennessee at Martin. 2/27/2010 http://primes.utm.edu/.

Herzog, Siegfried. “Frequency of Occurrence of Prime Gaps”. 2/27/2010. http://mac6.ma.psu.edu/primes/

Nicely, Thomas R.. “Some Results of Computational Research in Prime Numbers”. 2/27/2010 http://www.trnicely.net/.

Weisstein, Eric W. "Jumping Champion." From *MathWorld*--A Wolfram Web Resource. 2/27/2010. http://mathworld.wolfram.com/JumpingChampion.html

Weisstein, Eric W. "Prime Difference Function." From *MathWorld*--A Wolfram Web Resource. 2/27/2010. http://mathworld.wolfram.com/PrimeDifferenceFunction.html

Wolf, Marek. “First Occurrence of a given gap between consecutive primes.” 2/27/2010. http://www.ift.uni.wroc.pl/~mwolf/