The distribution of prime numbers ≥ 5 within sequences of natural numbers

Adrian M. Stokes

17th April 2024

Abstract

The prime numbers ≥ 5 within a finite sequence of natural numbers can be found arithmetically by calculating all of the values of $6n-1$ and $6n+1$ that fall within the sequence and subtracting the composites given by $(6n_1 \pm 1)(6n_2 \pm 1)$, where n is a natural number. For a given value of n_1, successive $(6n_1-1)(6n_2+1)$, $(6n_1-1)(6n_2-1)$ and $(6n_1+1)(6n_2+1)$ composites occur at a regular interval, which increases by 36 from one value of n_1 to the next. When combined, these regular but different intervals create disorder in the sequences of $6n-1$ and $6n+1$ composites, which in turn creates the apparent randomness of the primes in the sequence of natural numbers. Furthermore, $\{(6n_1-1)(6n_2-1)\}$ and $\{(6n_1+1)(6n_2+1)\}$ numbers are subsets of $6n+1$ composites whereas the only subset of the $6n-1$ composites is $\{(6n_1-1)(6n_2+1)\}$. This creates a slight inequality in the proportions of composites and primes between the sets $\{6n-1\}$ and $\{6n+1\}$, which otherwise have an equal number of members overall.

1. The $6n-1$ and $6n+1$ sets

Prime numbers ≥ 5 are either members of the sets $\{6n-1\}$ or $\{6n+1\}$, where n is any natural number. To see this, Table 1 shows a sequence of natural numbers from 2 to 127 divided into rows of six consecutive numbers. All of the prime numbers within the sequence are shown in **bold**. By arranging the natural numbers in this way, six distinct columns of related numbers are formed, which would not be the case if the first number in the table was 1. Instead, 1 is shown on its own since it is the only common factor for all of the natural numbers and therefore cannot be attributed to any particular column.

Note that the numbers of the first and third columns are all multiples of 2 and are therefore composites with the exception of 2. The second column comprises of multiples of 3 and all are composites except for 3. All of the numbers in the fifth column are composites that are multiples of 2 and 3 and therefore of 6. The fourth and sixth columns are respectively made up of numbers that are 1 less or 1 more than a multiple of 6 and these cannot be divided by 2 and/or 3. Since prime numbers can only be divided by 1 and themselves, all of the primes ≥ 5 are thus located in these latter columns given the remaining columns consist almost entirely of composites with the only exceptions being 2 and 3. Furthermore, since the rows of the table can be continued indefinitely, it must be the case that primes ≥ 5 are always located in these columns and thus belong to the sequences $6n-1$ or $6n+1$.

Table 1

1

x2	**x3**	**x2**	**6*n* - 1**	**x6**	**6*n* + 1**
2	**3**	4	**5**	6	**7**
8	9	10	**11**	12	**13**
14	15	16	**17**	18	**19**
20	21	22	**23**	24	25
26	27	28	**29**	30	**31**
32	33	34	35	36	**37**
38	39	40	**41**	42	**43**
44	45	46	**47**	48	49
50	51	52	**53**	54	55
56	57	58	**59**	60	**61**
62	63	64	65	66	**67**
68	69	70	**71**	72	**73**
74	75	76	77	78	**79**
80	81	82	**83**	84	85
86	87	88	**89**	90	91
92	93	94	95	96	**97**
98	99	100	**101**	102	**103**
104	105	106	**107**	108	**109**
110	111	112	**113**	114	115
116	117	118	119	120	121
122	123	124	125	126	**127**

It also follows that the composites in the $6n - 1$ and $6n + 1$ columns cannot be multiples of 2 and/or 3, which in turn means that the factors of these composites can only belong to either $\{6n - 1\}$ or $\{6n + 1\}$. There are three permutations to consider for multiplying two factors that produce these composites:

$$(6n_1 - 1)(6n_2 + 1) = 6n - 1$$

$$(6n_1 - 1)(6n_2 - 1) = 6n + 1$$

$$(6n_1 + 1)(6n_2 + 1) = 6n + 1$$

where n_1 and n_2 are any natural numbers and can be the same number, and n is a natural number dependent on the values of n_1 and n_2.

Again, this can be seen by considering the $6n - 1$ and $6n + 1$ composites in Table 1. Applying the formula for $6n - 1$ composites results in the following:

(6 x 1 – 1)(6 x 1 + 1) = 5 x 7 = 35

(6 x 1 – 1)(6 x 2 + 1) = 5 x 13 = 65

(6 x 1 – 1)(6 x 3 + 1) = 5 x 19 = 95

(6 x 1 – 1)(6 x 4 + 1) = 5 x 25 = 125

(6 x 2 – 1)(6 x 1 + 1) = 11 x 7 = 77

(6 x 3 – 1)(6 x 1 + 1) = 17 x 7 = 119

This system of calculations keeps the value of n_1 constant whilst increasing the value of n_2 by 1 for successive calculations until reaching the composite after which the next composite will exceed the upper limit of the number sequence. For the above, this composite is reached when $n_1 = 1$ and $n_2 = 4$. Now n_1 is increased by 1 and n_2 is reset to its lowest value and the method is repeated. The process continues until it is no longer possible to perform calculations that produce composites lower than the upper limit of the chosen number sequence, in this case when $n_1 = 3$ and $n_2 = 1$.

Comparing these results to the values of the composites in the $6n - 1$ column in Table 1 reveals a perfect match.

This process can also be followed for the $6n + 1$ column but bearing in mind there are two possible equations as shown above. Note that when $(6n_1 - 1)(6n_2 - 1) = (6n_2 - 1)(6n_1 - 1)$ and $(6n_1 + 1)(6n_2 + 1) = (6n_2 + 1)(6n_1 + 1)$ only one of these permutations is used in each case to avoid the unnecessary duplication of results:

(6 x 1 – 1)(6 x 1 – 1) = 5 x 5 = 25

(6 x 1 – 1)(6 x 2 – 1) = 5 x 11 = 55

(6 x 1 – 1)(6 x 3 – 1) = 5 x 17 = 85

(6 x 1 – 1)(6 x 4 – 1) = 5 x 23 = 115

(6 x 2 – 1)(6 x 2 – 1) = 11 x 11 = 121

(6 x 1 + 1)(6 x 1 + 1) = 7 x 7 = 49

(6 x 1 + 1)(6 x 2 + 1) = 7 x 13 = 91

Once again comparing the above results to the values of the composites in the $6n+1$ in Table 1 reveals a perfect match.

If Table 1 was extended indefinitely the above results would also be extended indefinitely. This can be summarised generally by:

$$(6n_1 \pm 1)(6n_2 \pm 1) = 6n \pm 1$$

2. Finding primes ≥ 5 arithmetically

The conclusions above lead to an arithmetical method of finding the complete subsets of primes within a sequence of natural numbers with an upper limit of x. This is because there are two ways to calculate the $6n-1$ and $6n+1$ composites but only one way to calculate the corresponding primes. The summary equation above shows that composites can be calculated by using $6n \pm 1$ or via the factors $(6n_1 \pm 1)(6n_2 \pm 1)$ whereas the factors method cannot be used to calculate the primes. This is because the smallest value of $6n \pm 1$ is 5 and therefore both factors must be greater than 1 when calculating a number using $(6n_1 \pm 1)(6n_2 \pm 1)$. Since the factors for primes can only be 1 and the prime itself, that means that prime numbers ≥ 5 can only be written in the $6n \pm 1$ form.

Finding all of the prime numbers in a sequence of natural numbers can therefore be achieved by first calculating all of the $6n-1$ and $6n+1$ numbers within the sequence. This results in two sets each comprising a mix of primes and composites (essentially the $6n-1$ and $6n+1$ columns in Table 1 but with a revised upper limit, x). Next the three possible permutations of $(6n_1 \pm 1)(6n_2 \pm 1)$ are used to generate subsets of the composites. Subtracting the subsets from the corresponding sets results in two further subsets, which when combined represent all of the primes for the number sequence [1]. This can be summarised as follows:

$$\{6n-1\} - \{(6n_1 - 1)(6n_2 + 1)\} = \{6n-1\}_p$$

And;

$$\{6n+1\} - \{(6n_1 - 1)(6n_2 - 1)\} - \{(6n_1 + 1)(6n_2 + 1)\} = \{6n+1\}_p$$

where $_p$ is prime.

3. An explanation of the distribution of primes ≥ 5

Calculating primes in this way is quite revealing and offers an explanation for the apparent randomness in the distribution of primes. Tables 2.1 – 2.3 show why.

These tables represent the calculation method for finding primes in action. The number sequence chosen is 5 – 205 so that the tables each fit on a page whilst remaining readable. The best way to describe how the tables work is by way of example. Table 2.1 represents $\{6n-1\}$ – $\{(6n_1-1)(6n_2+1)\}$ = $\{6n-1\}_p$. The first column gives the values for n in $6n-1$, the results of which appear in the second column. Note that the entries for the third column, $6n+1$, are greyed out for now. The next two columns calculate the values of $6n_2+1$ and these are followed by the columns calculating the values of $6n_1-1$.

Turning to the example, when $n=6$, $6n-1=35$. Reading across the table, when written as the multiplication of two factors, $(6n_1-1)(6n_2+1)$, $n_2=1$ and $n_1=1$. The result is shown in the $n_1=1$ column and is again 35. This is matched against the $6n-1$ value and the cell shaded. This is repeated for all of the possible $(6n_1-1)(6n_2+1)$ calculations for which the result is ≤ 205. Note that these calculations are colour coded according to the value of n_1. The reason for this will become clear soon. What are left is all of the $6n-1$ primes, which are shown in red.

Table 2.1

					n_1	1	2	3	4	5
n	$6n - 1$	$6n + 1$	n_2	$6n_2 + 1$	$6n_1 - 1$	5	11	17	23	29
1	5	7								
2	11	13								
3	17	19								
4	23	25								
5	29	31								
6	35	37	1	7		35				
7	41	43								
8	47	49								
9	53	55								
10	59	61								
11	65	67	2	13		65				
12	71	73								
13	77	79	1	7			77			
14	83	85								
15	89	91								
16	95	97	3	19		95				
17	101	103								
18	107	109								
19	113	115								
20	119	121	1	7				119		
21	125	127	4	25		125				
22	131	133								
23	137	139								
24	143	145	2	13			143			
25	149	151								
26	155	157	5	31		155				
27	161	163	1	7					161	
28	167	169								
29	173	175								
30	179	181								
31	185	187	6	37		185				
32	191	193								
33	197	199								
34	203	205	1	7						203

Tables 2.2 and 2.3 build on the results of Table 2.1 and together represent the equation for $\{6n+1\}_p$:

$$\{6n+1\} - \{(6n_1-1)(6n_2-1)\} - \{(6n_1+1)(6n_2+1)\} = \{6n+1\}_p$$

Table 2.2 covers the first part of the equation, $\{6n+1\} - \{(6n_1-1)(6n_2-1)\}$, which is then completed in Table 2.3 by subtracting $\{(6n_1+1)(6n_2+1)\}$.

In Table 2.2, the $6n+1$ primes are shown in red so that the two-step process for eliminating the composites can be more clearly seen. Whilst many of the composites are matched and shaded, some remain unmatched. The latter are matched in Table 2.3 and the two subsets of primes for the number sequence are revealed.

Note that in Table 2.2 and 2.3 the duplication of composite results mentioned earlier is noted and this is captured by showing the n_2 value in the relevant n_1 column rather than repeat the composite again. For the purpose of the exercise, finding replicated composites is not problematic since the goal is to identify composites for elimination in the quest to find primes. How many times they are identified is immaterial. They do however represent a practical challenge in terms of the amount of excessive data they create, a topic that is discussed later.

Finally the issue of the distribution of primes can now be addressed. Going back to Table 2.1, all of the composites in the $n_1 = 1$ column occur at regular intervals of 30 i.e. 35, 65, 95…, 185. For $n_1 = 2$, the interval is 66 i.e. 77, 143. Although this table is not large enough to show it, the difference between successive n_1 intervals is always 36 e.g. 66 – 30 = 36 for the two n_1 columns mentioned here.

The exact same pattern of intervals is shown in Table 2.2 except that the starting composite is 25 rather than 35 as in Table 2.1. In Table 2.3, the first composite is 49 and the interval between composites in the $n_1 = 1$ column is 42. Nevertheless, the interval between successive columns increases by 36 in the same way as in the other two tables.

In summary, the composites are generated in a completely predictable and regular way, unsurprisingly given the formulae used to generate them. However, when these different intervals for the composites are combined the pattern becomes disordered. This is reflected in the n_2 and $6n_2+1$ columns where each interval's colour shows the mixing of values generated by the different intervals. Table 2.1 is the best example.

The final effect is unveiled by the grey shaded cells of Table 2.3 where the disorder of the composites is plain to see. It is also clear that it is this disorder that in turn gives rise to the apparent randomness in the distribution of the primes.

Table 2.2

					n **$_1$**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
n	6*n* - 1	**6*n* + 1**	***n*** **$_2$**	**6*n*** **$_2$** **- 1**	**6*n*** **$_1$** **- 1**	**5**	**11**	**17**	**23**	**29**	**35**	**41**
1	5	7										
2	11	13										
3	17	19										
4	23	25	1	5		25						
5	29	31										
6		37										
7	41	43										
8	47	49										
9	53	55	2	11		55	($n_2 = 1$)					
10	59	61										
11		67										
12	71	73										
13		79										
14	83	85	3	17		85		($n_2 = 1$)				
15	89	91										
16		97										
17	101	103										
18	107	109										
19	113	115	4	23		115			($n_2 = 1$)			
20		121	2	11			121					
21		127										
22	131	133										
23	137	139										
24		145	5	29		145				($n_2 = 1$)		
25	149	151										
26		157										
27		163										
28	167	169										
29	173	175	6	35		175					($n_2 = 1$)	
30	179	181										
31		187	3	17			187	($n_2 = 2$)				
32	191	193										
33	197	199										
34		205	7	41		205						($n_2 = 1$)

Table 2.3

					n_1	1	2	3	4
n	$6n-1$	$6n+1$	n_2	$6n_2+1$	$6n_1+1$	7	13	19	25
1	5	7							
2	11	13							
3	17	19							
4	23	25							
5	29	31							
6	35	37							
7	41	43							
8	47	49	1	7		49			
9	53	55							
10	59	61							
11	65	67							
12	71	73							
13	77	79							
14	83	85							
15	89	91	2	13		91	($n_2 = 1$)		
16	95	97							
17	101	103							
18	107	109							
19	113	115							
20	119	121							
21	125	127							
22	131	133	3	19		133		($n_2 = 1$)	
23	137	139							
24	143	145							
25	149	151							
26	155	157							
27	161	163							
28	167	169	2	13			169		
29	173	175	4	25		175			($n_2 = 1$)
30	179	181							
31	185	187							
32	191	193							
33	197	199							
34	203	205							

4. Inequalities in the ratios of primes and composites between $6n - 1$ and $6n + 1$ sequences

The fact that the $\{6n + 1\}$ composites are formed from two subsets versus one for the $\{6n - 1)$ composites raises a question about the composition of primes and composites for the $\{6n - 1\}$ and $\{6n + 1\}$ sets relative to each other.

Using the approach and methodology described above, a model was created in Excel to find all of the primes in the number sequence 1 – 102,001. The selection was chosen partly because the equipment used was an old, budget laptop and this selection was at the limit of what it could handle, and partly to ensure an equal number of members between the $\{6n - 1\}$ and the $\{6n + 1\}$ sets. i.e. 17,000 apiece.

Interestingly there is very little difference in composition between the sets. For $\{6n - 1\}$, there are 12,104 composites and 4,896 primes. For $\{6n + 1\}$, there are 12,131 composites and 4,869 primes. So there are indeed more $6n + 1$ composites but only 27 more. Dividing the number sequence into a hundred subsequences, the first being 1 to 1,021 with subsequent divisions of 1,020 numbers, revealed that the numerical difference in composites in favour of $\{6n + 1\}$ grows cumulatively with each added subsequence but the rate of this growth is slow and erratic. Table 3 below shows how this difference evolves using divisions of 10,200 for convenience (the first division is actually for the subsequence 1 – 10,201).

Table 3

Natural Number Sequence (1 - 102001)	10201	20401	30601	40801	51001	61201	71401	81601	91801	102001
6*n* -1 Numbers	1700	3400	5100	6800	8500	10200	11900	13600	15300	17000
(6*n* 1-1)(6*n* 2+1) Composites	1070	2243	3440	4657	5878	7107	8348	9590	10848	12104
6*n* -1 Primes	630	1157	1660	2143	2622	3093	3552	4010	4452	4896
6*n* +1 Numbers	1700	3400	5100	6800	8500	10200	11900	13600	15300	17000
(6*n* 1-1)(6*n* 2-1) Composites	700	1469	2259	3063	3874	4692	5521	6349	7186	8022
(6*n* 1+1)(6*n* 2+1) Composites	638	1378	2153	2946	3754	4571	5399	6235	7082	7924
Duplicated 6*n* +1 Composites	258	593	952	1335	1727	2128	2537	2955	3383	3815
Total 6*n* +1 Composites	1080	2254	3460	4674	5901	7135	8383	9629	10885	12131
6*n* +1 Primes	620	1146	1640	2126	2599	3065	3517	3971	4415	4869
(6*n* +1)-(6*n* -1) Composites	10	11	20	17	23	28	35	39	37	27

So why are the differences so small? The main reason is that nearly half of the results given by $(6n_1 - 1)(6n_2 - 1)$ and $(6n_1 + 1)(6n_2 + 1)$ are duplicates. When tabulated, the duplicated results are mirrored either side of a diagonal of square numbers. Note that these types of duplicates are not shown in Table 3 because the model was built to exclude them. In contrast, $(6n_1 - 1)(6n_2 + 1)$ does not produce this kind of duplication. Lesser reasons include a duplication of results across the $(6n_1 - 1)(6n_2 - 1)$ and $(6n_1 + 1)(6n_2 + 1)$ results (these are shown in Table 3) due to different combinations of factors yielding the same composite. This also happens within the results for each subset but the added duplication across two sets of results can only happen for the two $\{6n + 1\}$ subsets. A third consideration is the larger intervals between composites for the $\{(6n_1 + 1)(6n_2 + 1)\}$ subset producing fewer composites for a given number sequence relative to the $\{(6n_1 - 1)(6n_2 - 1)\}$ subset.

In summary, the effect of two $\{6n+1\}$ subsets is largely countered by the effects of duplicated results and differences in intervals between composites. However, the proportions of primes and composites between the $\{6n+1\}$ and $\{6n-1\}$ sets do not balance completely, at least for the number sequences explored here, and result in more composites and fewer primes for the set $\{6n+1\}$ compared to the set $\{6n-1\}$.

5. Thoughts on the practicalities of the modelling method

The main advantage of this method for finding primes is that it identifies every prime ≥ 5 in a finite sequence of natural numbers without the need for primality testing.

The model created to underpin these findings comprised five tables, one for each the sets, $\{6n+1\}$ and $\{6n-1\}$, and one each for the three composite subsets. The composites in the subset tables were matched against the composites in the tables of the sets leaving the primes as unmatched numbers.

The main problem is the number of calculations needed to produce the composite subsets. Removing duplicate values is key to minimising the size of the datasets. For the model used here, many repeated composites were avoided or deleted in groups (multiples of the intervals discussed earlier produce duplicate values too) but individual duplicates produced by different combinations of factors were not removed so there is scope for further improvement. Ultimately, the goal should be to calculate each composite only once, which would be 24,235 calculations for number sequence in this model (i.e. one calculation per composite number) in order to find every prime ≥ 5 in the number sequence 1 to 102,001. Although this is many, it should be remembered that there are, in fact, 92,234 composites for this sequence in total allowing for multiples of 2, 3 and 6 in an extended version of Table 1. In this sense, 24,235 calculations would seem to represent “good value”.

Other improvements would be to use a more powerful computer or computers, more appropriate software and perhaps an algorithm that completely removes the need to produce the composite subsets.

References

[1] A. M. Stokes. Isolating the Prime Numbers. *http://viXra.org/abs/2402.0163*. 2024.

Contact

adestokes71@live.com