An Observation on Distribution of Prime Numbers

Dariush Bahrami

Faculty of New Sciences & Technologies, University of Tehran, Tehran, Iran

Email: dariush.bahrami@ut.ac.ir; Tel.: +98-910-192-8706

June 7, 2021

Abstract

In this research first, a sequence of properties called delta is assigned to each prime number and then examined. Deltas are only dependent on the distribution of prime numbers, so the results obtained for the delta distribution can be considered as a proxy for the distribution of prime numbers. The first observation was that these properties are not unique and different prime numbers may have the same value of delta of a given order. It was found that a small number of deltas cover a large portion of prime numbers, so by recognizing repetitive deltas, the next prime numbers can be predicted with a certain probability, but the most important observation of this study is the normal distribution of deltas. This research has not tried to justify the obtained observations and instead of answering the questions, it seeks to ask the right question.

1. Introduction

The study of prime numbers, apart from satisfying human curiosity, can be very important because, regardless of the choice of counting basis and independent of mathematical notation, prime numbers will have a clear and unambiguous meaning among intelligent beings. Despite the efforts of many curious and smart people, the mystery of prime numbers, specifically their distribution, remains an unsolved problem in mathematics. This study is another attempt to investigate the distribution of prime numbers. For this purpose, in section 2 a series of properties called delta is defined, which is only dependent on the distribution of prime numbers. In Section 3, it was found that the different orders of the delta are not of equal importance, and in particular, the first-order delta is rarer than the higher-order ones, so in the next sections, the first-order delta is examined. Interesting results appeared after observing the distribution of different first-order deltas. In particular, in section 4 it was found that unique deltas have a normal distribution; In addition, it was observed that a small fraction of first-order deltas covers a large portion of prime numbers. By using the observations made in this research, the prediction of prime numbers can be done in a probabilistic approach; This potential application was discussed in section 5.

The nature of this research is statistical and the obtained statistical results can be used as probabilities to predict unstudied prime numbers. But it should be noted that this research seeks to ask the right question rather than trying to answer difficult ones.

2. Definition of Delta

The sequence of prime numbers is as follow:

$$P = (p_1, p_2, p_3, \ldots, p_n, \ldots) = (2,3,5,\ldots, p_n, \ldots)$$

For each prime number, p_n, a sequence of properties, called Δ_n, is defined as follows:

$$\Delta_n = \left(\delta_n^k\right)_{k=1}^{n-1} = \left(\frac{p_{n+k} - p_n}{p_n - p_{n-k}}\right)_{k=1}^{n-1} = (\delta_n^1, \delta_n^2, \ldots, \delta_n^{n-1}) \qquad (1)$$

All results reported in this study are based on the sequence defined in equation 1 which describes how prime numbers are positioned relative to each other. The parameter k in equation 1 is called order. for each prime number p_n, $n-1$ order of delta, δ, can be defined. According to the position of a prime number relative to the prime numbers around it, a sequence of properties called delta is defined in equation 1. The first-order delta of a prime examines the position of primes that are exactly adjacent to the prime number itself. As the order of delta increases, the position relative to the farther away primes is involved in the delta value. According to this definition, the first prime number for which delta can be defined is the number 3. For number 3, only one delta can be defined which is of order 1. Based on equation 1, the delta of order 1 for prime number 3 is defined as follows:

$$\delta_2^1 = \frac{5-3}{3-2} = \frac{2}{1} = 2$$

The first prime number for which the delta of order m is defined is p_{m+1}. Therefore, the first prime number for which delta of order 2 is defined is p_3 which is 5. The Δ sequence for prime number 5 is defined as follows:

$$\Delta_3 = \left\{\frac{p_4 - p_3}{p_3 - p_2}, \frac{p_5 - p_3}{p_3 - p_1}\right\} = \left\{\frac{7-5}{5-3}, \frac{11-5}{5-2}\right\} = \left\{\frac{2}{2}, \frac{6}{3}\right\} = \{1,2\}$$

To clarify the definition given in equation 1 several other Δ sequences are shown in table 1:

Table 1 - Examples of different orders of delta for first 9 prime numbers

n	p_n	δ^1	δ^2	δ^3	δ^4	δ^5	δ^6	δ^7	δ^8
2	3	2							
3	5	1	2						
4	7	2	$3/2$	2					
5	11	$1/2$	1	1	$4/3$				
6	13	2	1	$5/4$	$8/5$	$18/11$			
7	17	$1/2$	1	$6/5$	$7/6$	$10/7$	$8/5$		
8	19	2	$5/3$	$3/2$	$3/2$	$11/7$	$3/2$	$28/17$	
9	23	$3/2$	$4/3$	$7/5$	$3/2$	$5/4$	$4/3$	$3/2$	$12/7$

Assuming n first-order delta is known, the prime number p_{n+1} can be calculated according to equation 2. It should be noted that equation 2 does not provide any additional information because the prime number p_{n+1} is needed to calculate the δ_n^1 which is part of this equation.

$$p_{n+1} = \left(\sum_{i=3}^{n} \left(\prod_{j=2}^{i} \delta_j^1 \right) \right) + \delta_2^1 + 3 \tag{2}$$

3. Rareness of Delta

By examining Table 1, it can be seen that the delta is not unique and, for example, a few specific numbers are repeated for the first order of delta. Judging by 8 deltas does not seem wise. Therefore first, all prime numbers below one billion are collected, and then the number of unique deltas is examined. Figure 1 shows the quantity of unique deltas by increasing the quantity of prime numbers.
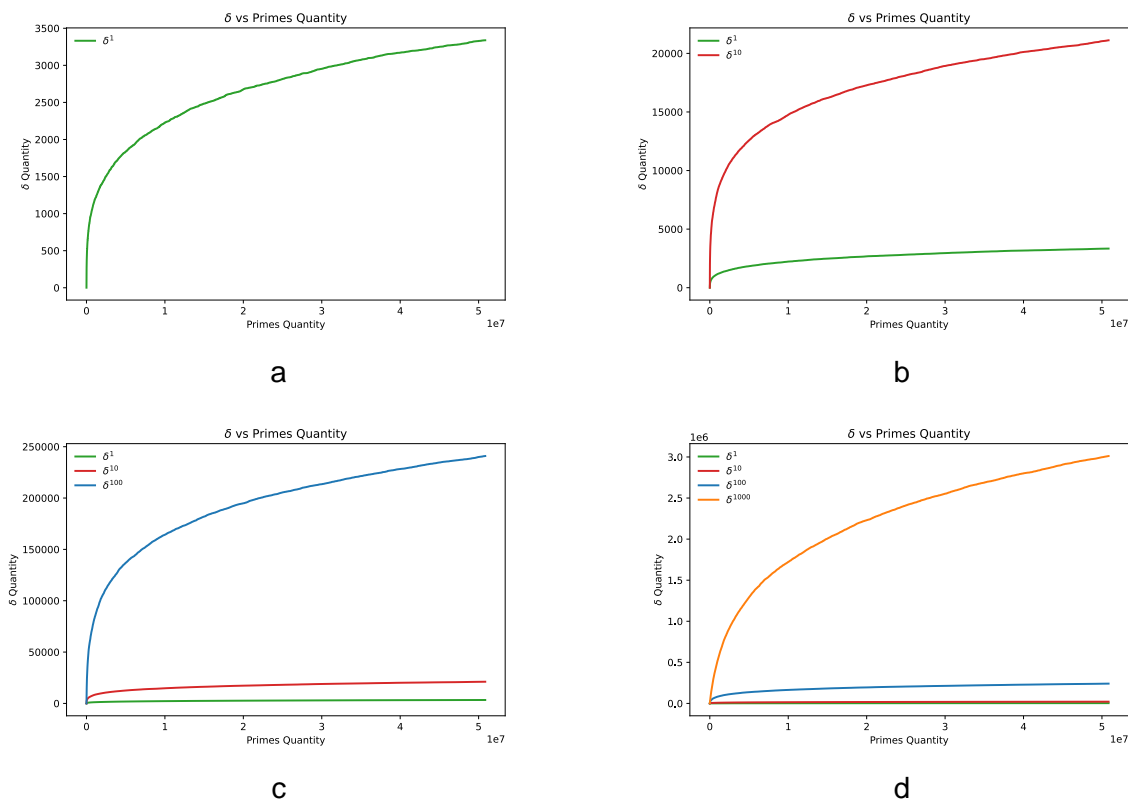

Figure 1 - Quantity of unique deltas at different orders versus quantity of prime numbers

By examining the plots drawn in figure 1, it can be concluded that the first-order delta is much rarer than the higher-order ones.

Table 2 - Quantity of unique delta at different orders for prime numbers below 1 billion

δ Order	Quantity	Quantity Relative to Order 1
1	3338	1.000
10	21120	6.327
100	240930	72.178
1000	3011700	902.247

According to the data shown in table 2, the number of unique deltas, for prime numbers below one billion, increases with the order. It is observed that by increasing the order of delta, the number of unique deltas of order *k* relative to the number of unique deltas of order 1 tends to k. This relation described in equation 3:

$$\lim_{k \to \infty} \frac{q_{\delta^k}}{q_{\delta^1}} \approx k \tag{3}$$

In equation 3, q_{δ^1} and q_{δ^k} are quantity of unique deltas of order 1 and k respectively. The error of equation 3 which is quantity ratio estimation error, QREE, is defined in equation 4:

$$QREE = \frac{\left| \frac{q_{\delta^k}}{q_{\delta^1}} - k \right|}{k} \tag{4}$$

Table 3 shows the accuracy of equation 3 by examining the QREE reduction trend.

Table 3 - Decrease of QREE by increasing delta order

δ Order	QREE
10	0.367
100	0.278
1000	0.098

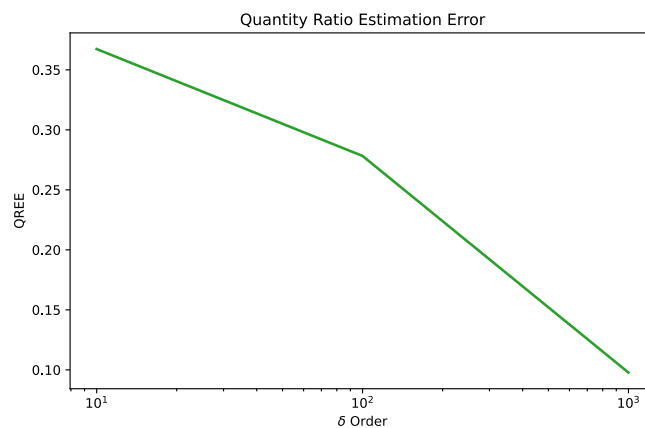The data in table 3 are also shown in figure 2.

Figure 2 - Reduction of QREE by the increase of delta order

The relative quantity of unique deltas for different orders is shown in figure 3. Relative delta quantity is the quantity ratio of unique deltas to the primes.
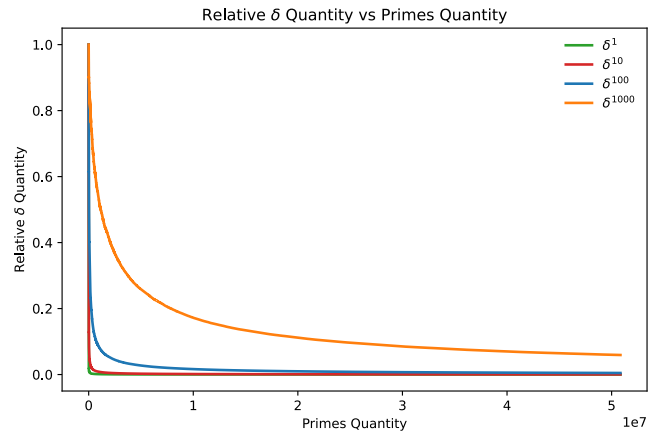

Figure 3 – Relative quantity of unique deltas versus quantity of prime numbers

As shown in figure 3, as well as in table 4, the quantity of first-order unique deltas is very small compared to the number of prime numbers.

Table 4 - Quantity of unique delta relative to quantity of primes at different orders for prime numbers below 1 billion

δ Order	δ Quantity Relative to Primes Quantity
1	6.565E-05
10	4.154E-04
100	4.738E-03
1000	5.923E-02

In short, first-order deltas are much rarer than higher-order ones. This observation can have very interesting consequences for practical applications of prime numbers which will be discussed in the following sections. There is an infinite number of prime numbers [1], and there may be an infinite number of first-order deltas, But the size of the two "infinities" are different. In addition, not all deltas are equally important for predicting the next prime numbers. As will be explored in the next section, a small percentage of first-order deltas cover up a large portion of prime numbers. In simpler terms, some deltas are very repetitive, and by finding them one can make a better guess about the next prime numbers.

4. First-Order Delta Analysis

In the previous section, it was investigated that the first-order delta is much rarer than the higher-order ones. Also, according to the data in Table 1, it was found that the first-order delta, and the delta in general, is not unique. This section will investigate the distribution of the first-order deltas. First of all, repetitive deltas are examined. Figure 4 shows the most frequent deltas.
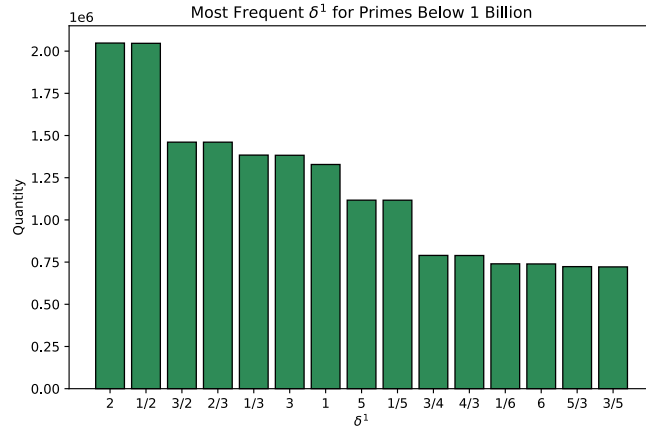

Figure 4 - Most frequent deltas of order 1

A very interesting observation of figure 4 is the equal frequency of the multiplicative inverse, also knowns as reciprocal, deltas. The product of each pair of reciprocal deltas is by definition equal to 1. Another observation is the high frequency of 2 and 1/2 in deltas of order 1. To understand the share of each of these delta values in the coverage of prime numbers, figure 5 shows the coverage percentage of each delta.
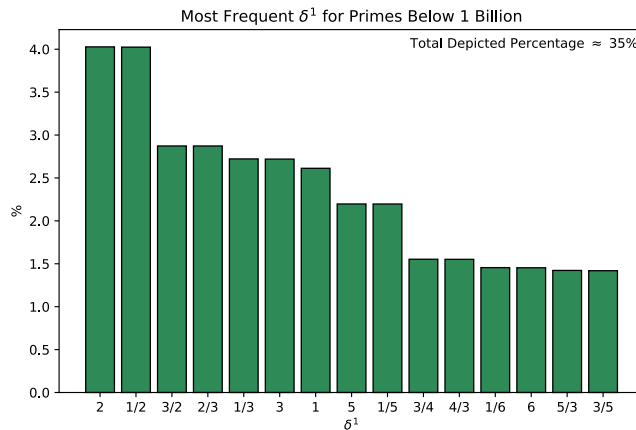

Figure 5 – Coverage percentage of most frequent deltas

As shown in figure 5, the first 15 most frequent deltas cover about 35% of the total prime numbers. 2 and 1/2 alone account for 8% of all primes. Table 5 shows the total share of most frequent deltas in prime numbers. From this table, it can be seen that less than 1% of deltas, i.e., 30 deltas, cover 50% of prime numbers. It can be seen that about 6.2% of deltas cover about 90% of prime

numbers. However, when we want to cover more prime numbers, we will need more deltas exponentially.

Table 5 - Total Share of most frequent deltas in primes

Primes Coverage (%)	Required Unique Deltas	Required Unique Deltas (%)
25	8	0.23966447
50	30	0.898741762
75	89	2.666267226
90	206	6.171360096
95	317	9.496704614
99	649	19.44278011
99.9	1269	38.01677651
99.99	2053	61.50389455
99.999	2861	85.71000599
99.9999	3287	98.47213901
99.99999	3332	99.82025165

The data in table 5 are shown in Figure 6. As can be seen, more unique deltas are needed to cover more prime numbers exponentially.
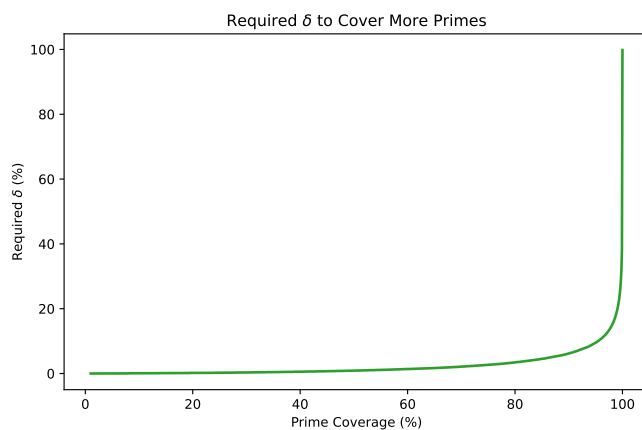

Figure 6 - Required deltas to cover more prime numbers

The logarithm function can be used to convert a multiplicative inverse relationship between two numbers to an additive inverse.

$$\log_b \frac{1}{x} = -\log_b x$$

To plot a delta histogram, the logarithm of deltas is used so that the pairs of reciprocal deltas are positioned symmetrically on the horizontal axis. The base of the logarithm, b, is set to 2. The reason for choosing number 2 is the high frequency of deltas with values of 2 and 1/2. Figure 7 shows the histogram of the base 2 logarithms of the first-order deltas.
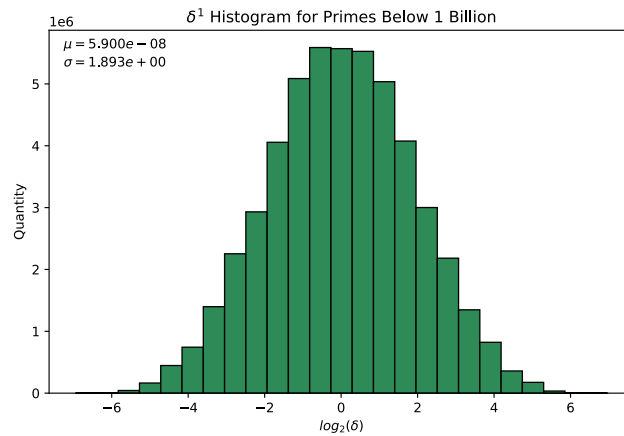
Figure 7 - Histogram of the first-order deltas

The number of histogram bins is obtained using the Sturges formula [2] which depicted in equation 5:

$$k = \lceil \log_2 n \rceil + 1 \tag{5}$$

In equation 5, k is the number of histogram bins and n is the data number which in this case is the number of first-order deltas for primes below 1 billion which is equal to 50847532. As can be seen, the distribution of the logarithm of delta appears to be normal. By calculating the probability density of the delta's logarithm frequency, the histogram is plotted again in figure 8. In this diagram by using the mean and standard deviation of the delta's logarithms, the probability density function, PDF, of the normal distribution is plotted. Equation 6, which is the famous PDF equation of Gaussian distribution [3], is used to plot normal distribution diagram:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{6}$$
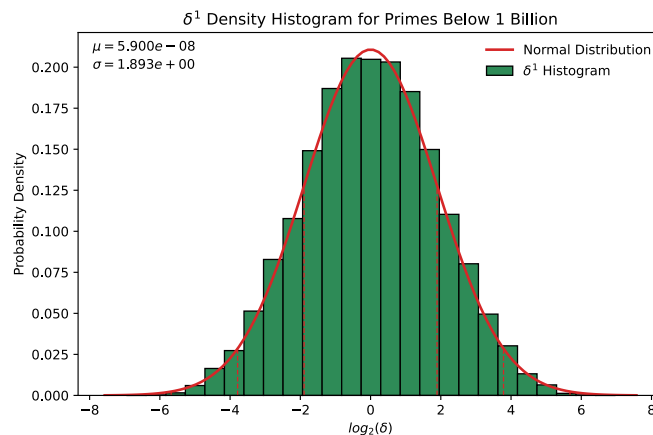

Figure 8 - Histogram of the probability density of the first-order deltas

From the diagram in Figure 8, Surprisingly, it can be seen that the delta's logarithm distribution is Gaussian.

5. Practical Applications of First-Order Delta

The first-order delta equation for each prime number, p_n, is as follows:

$$\delta_n = \frac{p_{n+1} - p_n}{p_n - p_{n-1}} \tag{7}$$

By rearranging equation 7 to calculate the first prime number after p_n, one can write:

$$p_{n+1} = \delta_n(p_n - p_{n-1}) + p_n \tag{8}$$

Assuming p_n and p_{n-1} are known and the appropriate delta is selected, the next prime number can be guessed. The criterion for selecting an appropriate delta is to check its frequency for prime numbers smaller than p_{n+1}. Figure 5 shows the fifteen deltas that had the most abundance. These 15 deltas represent 35% of the total first-order deltas for prime numbers below one billion. According to Table 5, by selecting 206 deltas, in order of frequency, one can 90% likely find the next prime number. By increasing the number of selected deltas to 649, the probability of correctly guessing the next prime number can be increased to 99%. Because deltas are rational numbers in the form of a fraction like $\frac{a}{b}$ the computations for guessing the next prime number can be reduced. By substituting the $\frac{a}{b}$ fraction for delta in equation 8, one can write:

$$p_{n+1} = \frac{a}{b}(p_n - p_{n-1}) + p_n \tag{9}$$

Since p_{n+1} and p_n are integers, the expression $\frac{a}{b}(p_n - p_{n-1})$ must also be an integer. For this purpose, $(p_n - p_{n-1})$ must be divisible by b. Common methods in number theory can be used for divisibility testing. For example, if b is equal to 3, one can recursively test the divisibility of the sum of digits by 3. These methods can be useful for very large prime numbers. Because instead of performing subtraction, multiplication, and division operations and then testing for the correctness of the guess, some deltas can be crossed out by much less computation.

6. Conclusion

A sequence of properties called delta was defined to examine the position of prime numbers relative to each other. Examining few prime numbers, it became clear that these properties are not unique. The number of deltas is much smaller than the number of prime numbers, and among the various orders, it was found that the first-order delta, which examines adjacent prime numbers, is rarer than the higher-order ones. In addition, the first-order delta was found to have a Gaussian distribution. It was also found that a small number of first-order deltas cover a large portion of prime numbers, a result that could potentially be used as a way to generate new prime numbers.

7. Future Works

As mentioned before, this study has a statistical approach and does not try to justify the observations. Therefore, the following questions have not been answered in this study and need future work:

- Why is first-order delta rarer than higher-order ones?
- As described in equation 3, the number of unique deltas of order n relative to the number of unique deltas of order 1 tends to n. Why is that?
- Why do first-order reciprocal deltas have the almost same frequency?
- Do prime numbers that have a common first-order delta have more in common? If so, what are those common features?
- We need more first-order deltas to cover more prime numbers exponentially; This fact implies that some prime numbers have unusual first-order deltas. Are these prime numbers more suitable for security applications than other prime numbers that have a usual first-order delta?
- Why do first-order delta's logarithms have a Gaussian distribution?
- How efficient is it to predict prime numbers with the help of repetitive first-order deltas compared to existing methods?

Acknowledgments

References

[1]   Euclid and J. Williamson, "The Elements of Euclid. With dissertations," 1781.

[2]   H. A. Sturges, "The Choice of a Class Interval," *J. Am. Stat. Assoc.*, vol. 21, no. 153, pp. 65–66, 1926, doi: 10.1080/01621459.1926.10502161.

[3]   C. F. Gauss, "Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium," *Werke*, pp. 1–2, 2012, doi: 10.1017/cbo9781139058285.001.