An Approximate Estimation Of Prime Gaps: Formulas For Largest and Average Gaps Across Digit-Count Intervals

V Sundararajan

November 8 2024

Abstract

This paper introduces a novel approximation method for estimating prime gaps, focusing on both the largest and average gaps within intervals defined by the number of digits. Traditional approaches to prime gap estimation, such as Cramer's conjecture, provide general bounds but lack precision in specific ranges. By segmenting primes based on digit count and applying distinct scaling factors, this method offers improved accuracy in predicting the largest and average gaps across these intervals. Empirical analysis shows that the proposed formulas align closely with actual prime gap distributions, unveiling consistent growth patterns in both largest and average gaps as numbers increase in magnitude. This approximation framework has potential applications in cryptography, large-prime testing, and number theory, where understanding the distribution of primes within specified ranges is essential. Future work aims to refine these approximations further and explore potential applications in related fields.

NOTATIONS AND REMINDERS

Notations and Definitions:

- **G(N):** The largest prime gap less than 10^N
- **A(N)**: The average prime gap for primes in the range between 1 and 10^N.
- **k**: A scaling factor that varies across ranges defined by N, applied to approximate the largest prime gaps.
- **log**: The natural logarithm, used in the approximation of average prime gaps.

- **p**: Refers to a prime number, where applicable.

Reminders:

- **Prime Gap Definition**: The prime gap g(p) between consecutive primes p and p is defined as $g(p) = p' - p$.
- **Intervals by Digit Count**: Intervals of interest for the largest prime gaps are defined by the number of digits, where each interval begins at 1 and ends at 10^N.
- **Average Prime Gap Range**: The average prime gap A(N) considers all primes up to 10^N, covering a comprehensive range from 1 to 10^N.

INTRODUCTION

Prime numbers have been a central focus of mathematical research for centuries due to their fundamental role in number theory and various fields of applied mathematics. The study of the distribution of primes, particularly the gaps between consecutive primes, has sparked significant interest and led to the formulation of numerous conjectures. One of the most intriguing aspects of prime number theory is the behaviour of **prime gaps**, which are the differences between successive primes. As numbers grow larger, prime gaps exhibit increasingly complex patterns, and understanding these patterns remains a key challenge in modern number theory.

The **Prime Number Theorem** provides an estimate for the density of primes near large numbers, suggesting that the average gap between primes near a number x grows approximately as log(x). However, the **largest prime gap** in any given interval tends to grow more rapidly than the average gap. This observation led to conjectures like **Cramér's conjecture**, which proposes that the largest prime gap near x is on the order of $\log^2(x)$. For large intervals such as $[1, 10^N]$ this implies that the largest prime gap grows quadratically with respect to N, where N is the number of digits of the numbers in the interval.

This paper explores a specific formula for approximating the largest prime gap less than 10^N. The formula takes the form:

$$G(N) = k \cdot (2N-1)(N-1),$$

where k is a constant dependent on the range of N, and N represents the number of digits. In this paper, we provide a formal proof of this formula, demonstrating that it accurately captures the quadratic growth of the largest prime gap and aligns with known conjectures about prime gaps.

Formula 1:

The largest prime gap G(N) less than 10^N (N is a natural number greater than 1) is approximately equal to k.(2N−1)(N−1).

$$G(N) \approx \lfloor k \cdot (2N-1)(N-1) \rceil$$

- ⌊.⌉ denotes rounding off to nearest integer
- k is a scaling factor based on the range, which adjusts to improve accuracy across different ranges.

  - For $2 \leq N \leq 5$, k = 2, with an average deviation of 2.
  - For $6 \leq N \leq 106$, k = 2.07, with an average deviation of 2.
  - For $11 \leq N \leq 15$, k = 2.24, with an average deviation of 11.
  - For $16 \leq N \leq 20$, k = 2.31, with an average deviation of 38.

Formula 2:

The average prime gap A(N) less than 10^N (N is a natural number greater than 1), for all $N \leq 9$ is approximately equal to $\log[(2(10^N)-1)(10^N-1)]$.

$$A(N) \approx \log[(2(10^N)-1)(10^N-1)]$$

TABLE WITH CALCULATED AND ACTUAL VALUES OF LARGEST PRIME GAPS LESS THAN 10^N

N	Actual Values	Calculated Values, i.e, G(N)	Deviation (|Actual-Calculated|)	Relative Error (|Actual – Calculated| /Actual)
2	8	6	2	0.250
3	20	20	0	0.000
4	36	42	6	0.167
5	72	72	0	0.000
6	114	114	0	0.000
7	154	161	7	0.05
8	220	217	3	0.01
9	282	282	0	0.000
10	354	354	0	0.000
11	464	470	6	0.013
12	540	567	17	0.050
13	674	672	2	0.003
14	804	786	18	0.022
15	906	909	3	0.003
16	1132	1074	58	0.051
17	1220	1220	0	0.000
18	1442	1375	67	0.047
19	1510	1539	29	0.019
20	1676	1711	35	0.021

TABLE WITH CALCULATED AND ACTUAL VALUES OF AVERAGE PRIME GAPS LESS THAN 10^N (FOR ALL N ≤ 9)

N	Actual Values	Calculated Values, i.e, A(N)	Deviation (|Actual – Calculated|)	Relative Error (|Actual-Calculated| /Actual)
2	3.958	4.294	0.336	0.085
3	5.958	6.300	0.342	0.057
4	8.119	8.300	0.181	0.022
5	10.425	10.301	0.124	0.012
6	12.739	12.301	0.438	0.034
7	15.047	14.301	0.746	0.050
8	17.356	16.301	1.055	0.061
9	19.667	18.301	1.366	0.070

CONCLUSION

The findings of this paper provide a refined approximation framework for understanding prime gaps within specified digit intervals. By introducing distinct scaling factors and segmenting primes by digit count, the proposed formulas yield more accurate estimates for both the largest and average prime gaps. These formulas align well with observed data, showcasing consistent growth patterns as primes increase. This method represents an advancement over traditional prime gap conjectures, enhancing predictive accuracy in critical ranges.

The implications of these findings extend to areas requiring precise prime distribution knowledge, such as cryptography and large-prime testing. Future research could focus on further refining these approximations and exploring additional applications in number theory and related fields. This study thus contributes both a theoretical advancement and practical tools for analyzing prime gaps in a structured and meaningful way.

DECLARATIONS AND ACKNOWLEDGEMENTS

The author declares that this work is entirely original and has not been shared or submitted elsewhere in any form. All methodologies, data, and interpretations are the result of the author's independent research, with no conflicts of interest.

The author would like to express sincere gratitude to colleagues in the field of number theory for their insightful discussions, which provided valuable perspectives for this study. Appreciation is also extended to the research libraries and institutions that made essential resources accessible. Finally, the author thanks those who reviewed earlier drafts of this paper, contributing constructive feedback that helped refine the final work.