Estimating the Probability of Primes with in a Finite Integral Domain in Higher Order Range

Mehroz Mir[1*]
Crescent Public School, Naseem Bagh,
Hazratbal, Srinagar, J&K-190006, India

Jyotiraditya Jadhav[2]
[2]Presidency School, Bondanthila, Kelarai Post,
Neermarga, Mangalore, Karnataka-575029, India

Corresponding Authors:- Mehroz Mir[1*]

Abstract:- Prime number distribution is a fundamental concept in multiple areas such as cryptography and statistical analysis. In this paper, we have developed a highly close approximation of the probability of primes in a finite integral domain for higher order range of [0,10000]. Through rigorous mathematical derivations, we have established up- per bound and lower bound probability limits. Linear regressions observed through graphical representations showcase prime prob- ability distributions between the respective modular differences of upper and lower limits of probability with the actual probabilistic values. The observed convergence between the actual probability of primes and the developed relation represented by graphs val- idates the propositions. This convergence of the proposed limits contributes a deeper understanding of prime number distribution in finite integer domains which is being reported for the first time in this study.

Keywords:- *Prime Distribution, Probability Limits, Linear Re- Gression, Approximation, Order Ranges.*

I. INTRODUCTION

An integer ≥ 2 which can be only divided by 1 and the number itself is known as a prime number [1]. The fundamental theorem of arithmetic which factor- izes each integer into a set of unique primes amplifies the definition of a prime number [2]. There are two foregoing interpretations of probability: Frequentist probability and Bayesian probability [3], the former interpretation categorizes the probability of an event as the culmination of the results of repeated exper- iments; which means the probability of an event is obtained by the repetition of experiments for that particular event. The latter interpretation categorizes probability of an event as subjective or as an individual's belief that this is the likelihood of the event happening.

The Prime Number Theorem formalizes the idea that the existence of prime numbers within large positive integers is not very common. The theorem was first proved by Jacques Hadamard and Charles Jean de la Vall´ee Poussin in 1896 [4, 5]. The interest in prime distribution within integers had its inception as a result of Euclid's theorem [6] which was followed through a more sophisticated proof by Euler [7] which visualizes the addition of reciprocals of prime numbers [8]. Furthermore, the speculation regarding the distribution of primes by the 14-year-old Carl Frierdrich Gauss when he examined a table that listed all prime numbers below 102,000 led to his approximation for the density of primes [8]

$$\frac{1}{\log n}$$

As When more tables were available, Gauss further developed his hypothesis and formalized the density of primes in the expression as follows:

$$\int_2^x \frac{dn}{\log n}$$

In the first number theory textbook, Adrien Marie Legendre gave a similar approximation for the density of primes and used a constant in the denominator to determine the closest approximation for the density of primes.

$$\frac{x}{\log x - 1.08366}$$

[8] which is stated as Legendre in the 1808 edition of his book [9], gave another conjecture on the arithmetic progression of prime numbers which stated that the difference be- tween any natural number and its predecessor is constant. The progression was suitable for infinitely many numbers except the number 2 and the odd primes can be represented in the form of 4n+1 and 4n+3. Neither Gauss nor Leg- endre gave any proof or showcased how they reached these approximations or conjectures. Although, in 1837 Peter GuStav Lejeune Dirichlet proved Legen- dre's conjecture; which is now known as Dirichlet's theorem of the infinitude of primes in arithmetic progressions [8, 10]. It wasn't till 1848 that a profound step towards the proof of the prime number theorem was given by Pafnuty Lvovich Chebyshev. The Russian mathematician postulated that the equation has a limit *x* which tends to infinity [11].

ISSN No:-2456-2165

$$\lim_{x \to \infty} \frac{\pi[x] \times \log x}{x} = 1$$

Chebyshev was unable to prove that the ratio tends to 1 for the given limit, although he further theorized it into a finite integral limit. Furthermore, Chebyshev's functions [12] became the base for the proof of the prime number theorem. In a series of final attempts by various mathematicians, Georg Friedrich Bernhard Riemann made the next significant step towards proving the theorem. He speculated that Euler's product identity would help in proving the theorem since the RHS side of the identity involves primes [13]. Furthermore, Riemann replaced the exponent s which was greater than the number 1, with a complex exponent with the same symbol. Riemann developed the equation $s = \sigma + ti$ where σ and t are natural numbers and $i = \sqrt{-1}$ [14] and used this notation. Riemann connected the distribution of primes with the properties of $\zeta[s]$, where ζ determines the function of the variable s, and represented it by the infinite series [15]:

$$\zeta[s] = \sum_{n=1}^{\infty} 1/n^s$$

Further, Riemann showcased the relation between the distribution of primes and the location of the zeroes of the zeta function i.e., the points in the complex plane at which $\zeta[s] = 0$. The prime number theorem could now be proved by showing that there were no zeroes of the zeta function on the line where $\sigma = 1$ On the strategic basis of Reimann, Jacques Hadamard and Charles Jean de la Vall´ee Poussin in 1896 showed that there were no zeroes of the zeta function on the line $\sigma = 1$ [16].

It is to note that the prime number theorem describes the asymptotic relation for all primes $\leq x$ as x [17]. It states that the average density of primes for $\frac{x}{\log x}$ large positive integral values is approximated as the ratio $\log x$. Inducing from this, the density of primes for very large numbers gradually decreases, which inversely results in the approximated value for the probability that a number is prime also decreases [18].

The prime number theorem's assertion [19] is presented below; one of its first elementary proofs was given by Atle Selberg [20]:

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\log x} = 1 \qquad (1)$$

By using the Prime Number Theorem (1), we can find the probability of primes in an interval set; however, the values deviate with large errors for higher order range.

In this study we use analytical data to develop a relation which gives the probabilistic value of primes within a finite integral domain for higher order range, with minimum errors. The results are proved by mathematical deriva- tion and graphical representations respectively. This work provides a valuable probabilistic framework for prime number distribution analysis, enhancing pre- dictive accuracy in various mathematical and computational contexts. The approximations showcased in this study have been observed for the first time and gives the most accurate estimates for the probability of primes within a finite integral domain.

II. METHODOLOGY

The research topic was undertaken when we noticed that the probabilistic value of all ratios in the interval sets like [0,1], [0,2],. [0,100] can never exceed the ratio $\frac{1}{2}$; this set up the upper limit for our proposition which was later proved by rigorous mathematical derivation. Following the similar technique used for finding the upper-limit, we discovered the lower limit for the probability of primes within an interval set. Later proved, by using Bertrand Chebyshev theorem [21]. The modular difference between the lower limit and actual value of the probability of primes within an interval set was calculated via a python code and a graph of the expressions was plotted by taking the numerator as the *x*-axis and the denominator as the *y*-axis for 10,000 numbers. Similarly, the modular difference between the upper limit and the actual value of probability of primes was calculated and a graph was plotted by taking the numerator as the *x*- axis and the denominator as the *y*-axis for 10,000 numbers. The main result was obtained by equalising and approximating the upper-value modular difference and the lower-limit modular difference to the actual probabilistic value of primes within an interval set. The proposition has been backed up by analytical data for 10,000 numbers.

III. RESULTS AND DISCUSSIONS

➢ *Theorem 3.1:*

Probability of prime distribution in a closed set of natural num- bers [0, *n*] is always less than $\frac{1}{2}$

- *Proof:*

Suppose n is an even number then total number of odd numbers in closed set is $\frac{n}{2}$ and as 1 is not prime and 2 is a prime, assuming those odd numbers are primes then maximum probability of primes in set [0 *n*] is $\frac{n}{2(n+1)}$ and as $\frac{n}{n+1} < 1$ then $\frac{n}{2(n+1)} < \frac{1}{2}$ for all even numbers n.

If n is odd then the total number of the odds in the set1 [0, *n*] is $\frac{n-1}{2} + 1 = \frac{n+1}{2}$ Assuming that all those odds are primes we get $\frac{n+1}{2(n+1)} = \frac{1}{2}$ But we know after 7 no such odd number exists where all the odds before it are primes. Hence the number of primes $n + 1$ for $n > 9$, thus again the probability of primes when n is odd is

ISSN No:-2456-2165

$< \frac{1}{2}$, Therefore, only three such values of n exist when probability is $\frac{1}{2}$ and all other values are strictly $\leq \frac{1}{2}$.

- *Corollary:*

The following theorem is based on the foundation of the Bert- nard Chebyshev theorem which is proved by Shiva Kintali [22].

Bertnard Chebyshev theorem: Bertnard's postulate states that for every natural number n, where $n \geq 1$ there exists atleast one prime number for $n \leq p \leq 2n$ [21]

➢ *Theorem 3.2:*

Given an integer n, let 2^k be the largest power of 2 less than or equal to n. Then, the lower limit is:

$$\frac{k}{2^k+1}.$$

- *Proof* :

By Bertrand-Chebyshev theorem we know there is atleast one prime number in the set $(\frac{x}{2}, x]$ where $x \in \mathbf{N}$ so the total number of primes in the set [0, n] can be found by the number of such intervals by first finding the closest 2power below n and then the least number of primes possible in that range would be the power of 2 which we have found. This results in finding the probability for [0, 2^k] which is $\frac{k}{2^k+1}$.

The following conjecture relates the modular difference of the lower-limit value and the actual value of the probability of primes within a number set.

➢ *Conjecture 3.3:*

When we plot the probability values on a scatter plot taking the numerator as the x-axis and the denominator as the y-axis we find that graphed points fall under a linear region, as shown in Figure 1.

$$\left|\frac{\pi(x)}{x+1} - \frac{k}{2^k+1}\right| = R_1 \tag{2}$$

The numbers in Figure (1) to Figure (7) were obtained via a python code. Figure (1) and Figure (2) have been plotted for 10,000 numbers which were ob- tained from equation (2) and equation (3), whereas Figure (7) has been plotted for 5,000 numbers which were obtained from relation (4).

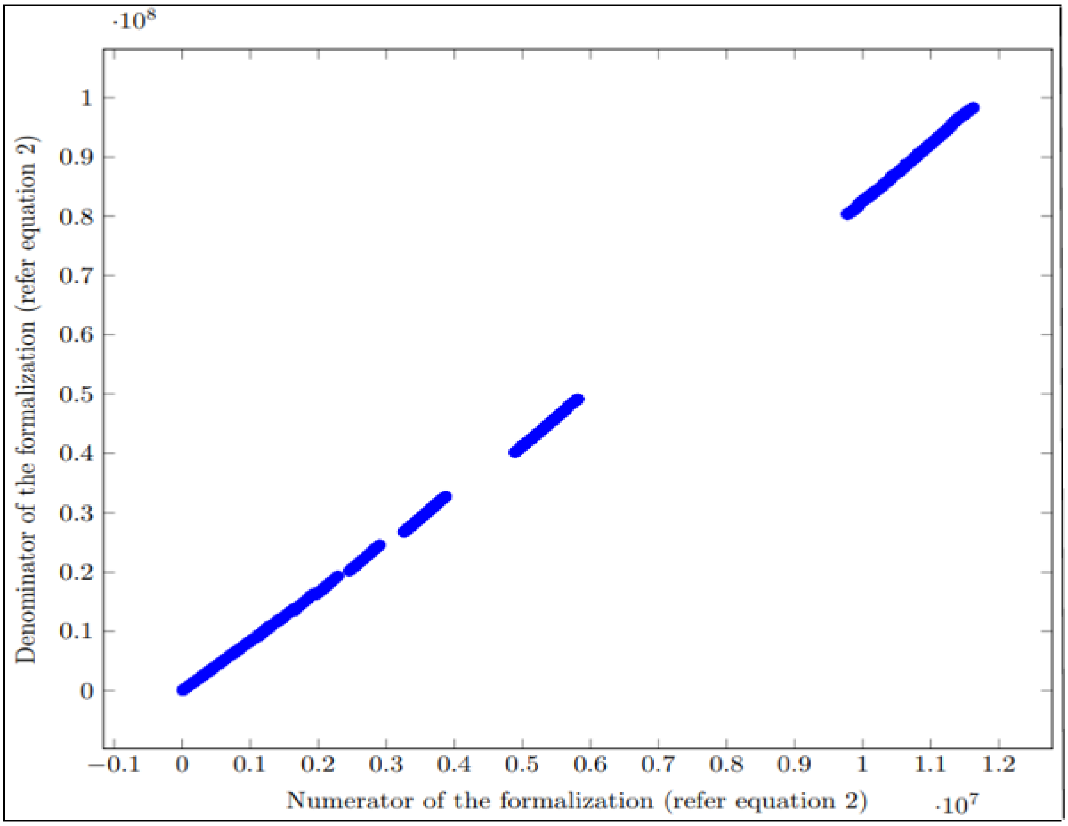

Fig 1 Variation of denominator of the formalization with variation of nu- merator of the formalization (refer equation 2).

In accordance with Proposition 3.3, Figure 1 prominently illustrates a lin- ear ascending segment, wherein all data points are, essentially, constituents of both the numerator and denominator. This phenomenon is observed for n over the initial 10000 numerical values, signifying a fundamental mathematical re- lationship of numerator to denominator which can be approximated to a fixed ratio.

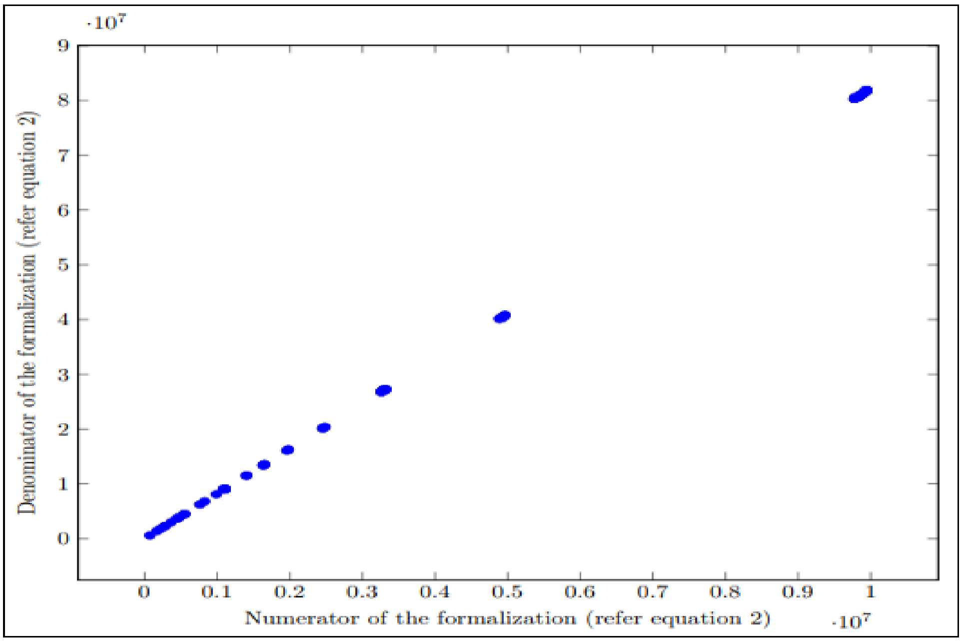

Fig 2 A Detailed examination of higher magnitudes *n* in the preceding graph.

Figure 2 serves as a magnified representation, focusing on a narrower range of *n* values spanning from 9800 to 10000. Within this constrained interval, the graph consistently reveals a linear ascending pattern. Notably, as the numerical values increase, a convergence occurs between the points constituting both the numerator and denominator, resulting in the emergence of a distinct, sharply defined linear region.

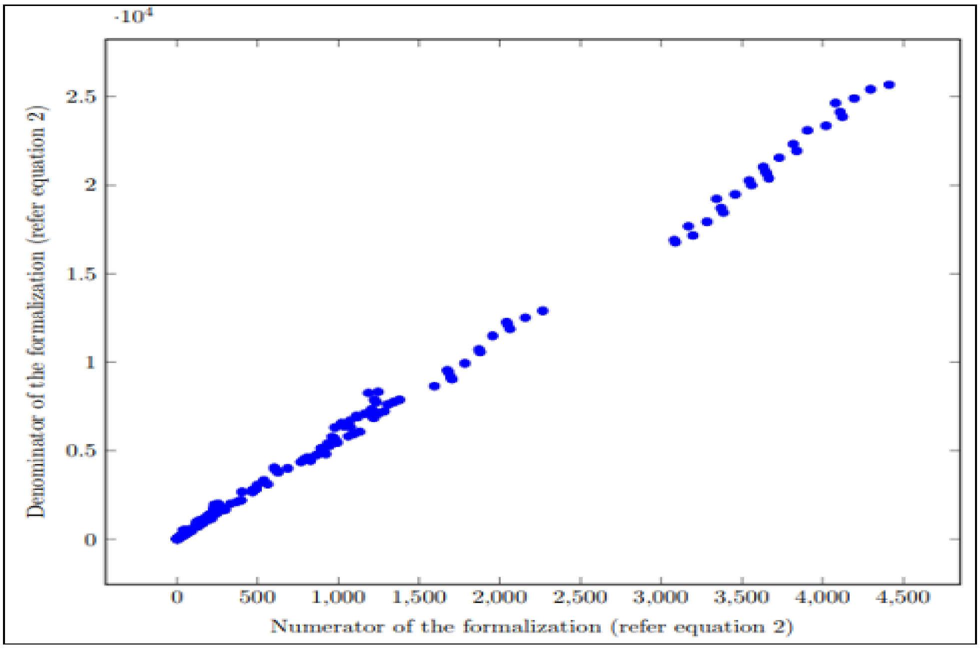

Fig 3A Detailed examination of lower magnitudes *n* in the preceding Graph.

Figure 3, constructed for the range of *n* values from 0 to 200, initially por- trays a discernible yet not highly defined linear region. However, it is imperative to note that as we extend our analysis to encompass larger magnitude numbers, the observed linear region progressively sharpens and becomes more pronounced. This trend underscores the dynamic nature of the relationship under investiga- tion, with the degree of linearity intensifying as we approach higher numerical magnitudes.

The following conjecture relates the modular difference of the upper-limit value and the actual value of probability of primes within a number set.

➢ *Conjecture 3.4:*

When we plot the probability values on a scatter plot taking the numerator as the x-axis and the denominator as the y-axis we find

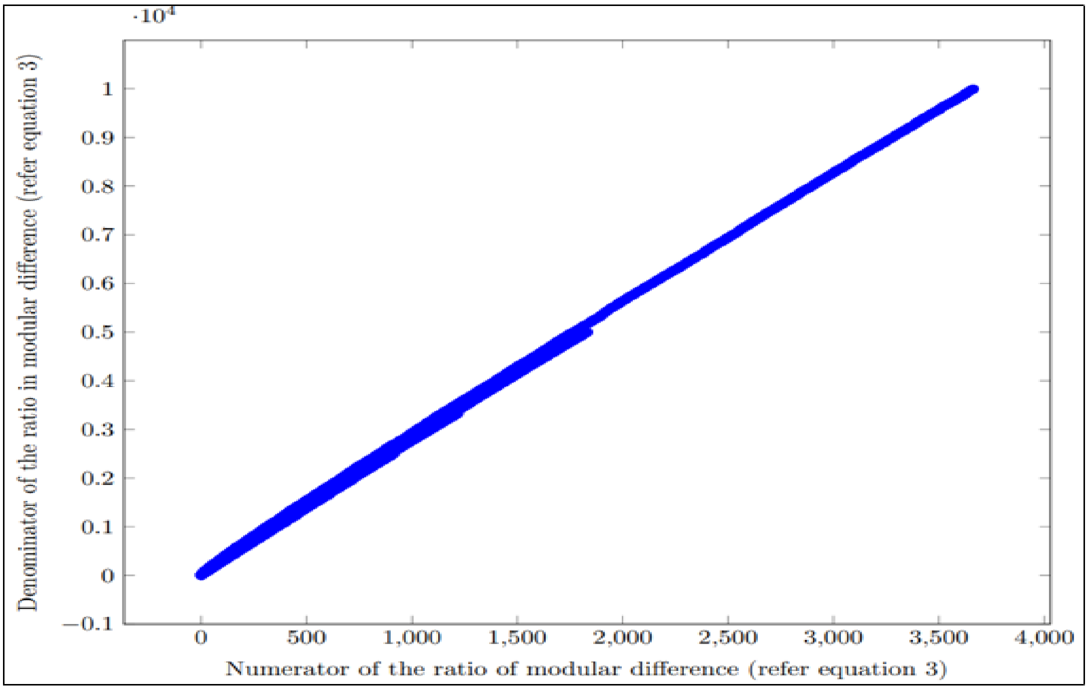

Fig 4 Variation of denominator of the formalization with variation of nu- merator of the formalization (refer equation 3). that graphed poin$_4$ts fall under a linear region, as shown in Fig 4.

Figure 4 represents the plot of both the numerator and denominator of the difference, as per Conjecture 3.4. Much like the preceding graphs encompassing *n* values over the initial 10,000 numbers, this graph likewise exhibits a discernible linear region, consistent with the principles elucidated in Conjecture 3.3. This observation underscores a recurring pattern, emphasizing the establishment of a clear relationship between the numerator and denominator.

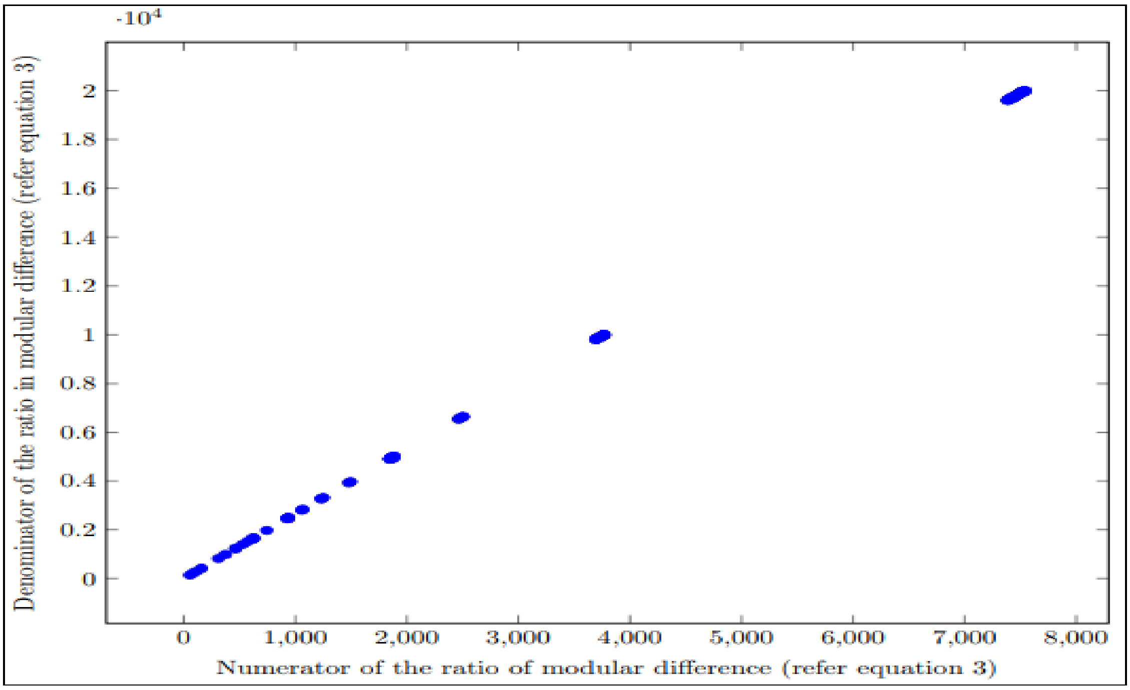

Fig 5 A Detailed examination of higher magnitudes *n* in the preceding Graph.

Constructed for the interval of *n* values ranging from 9800 to 10000, Figure 5 consistently reveals a linear ascending region. Notably, the data points comprising this linear segment are closely clustered, forming a remarkably sharp and well-defined linear region. This pronounced alignment of coordinates un- derscores the precision and significance of the observed mathematical pattern within the confines of this particular range, providing valuable insights for our analytical considerations.

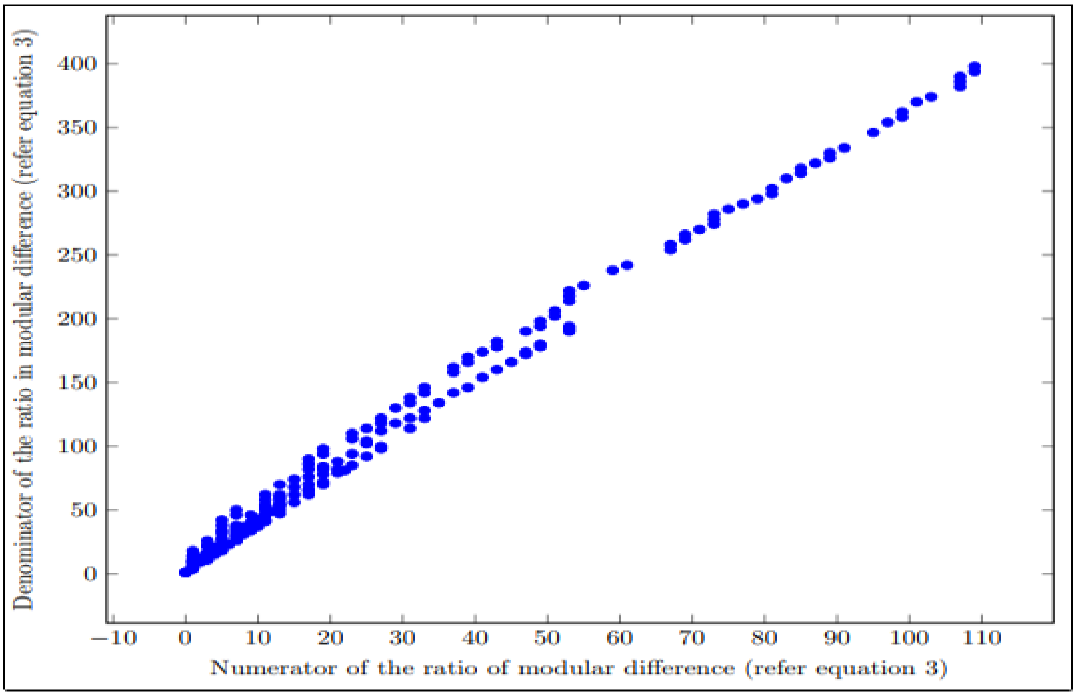

Fig 6 A Detailed examination of lower magnitudes *n* in the preceding Graph.

Figure 6 serves as a magnified rendition of the graphical representation eluci- dated in Conjecture 3.4, focused specifically on a range characterized by smaller magnitude numbers. Comprising *n* values within the first 200 numbers, this graph, akin to its counterparts, initially exhibits a discernible linear ascending region.

As we extend our examination to encompass a wider magnitude range, a salient trend emerges: the graph progressively evolves to form a more conspic- uously defined and sharp linear region.

The relation between the upper-limit and the actual value of the probability of primes can be stated as follows:

$$\left|\frac{1}{2} - \frac{\pi(x)}{x+1}\right| = R_2 \qquad (3)$$

From Conjecture 3.3 and Conjecture 3.4 we can approximate a ratio which is inverse of the slope of the line around which all the points lie. From this we can make a relation between the actual probability of primes in the set $[0, n]$ to the lower limit $\frac{k}{2^k+1}$ and the upper limit $\left(\frac{1}{2}\right)$.

From Equation (2) and (3):

$$\left|\frac{\pi(x)}{x+1} - \frac{k}{2^k+1}\right| = R_1$$

$$\left|\frac{1}{2} - \frac{\pi(x)}{x+1}\right| = R_2$$

Based on Equations (2) and (3), we are able to derive the mean value of the probability associated with the prime numbers as follows:

$$\frac{\pi(x)}{x+1} = \frac{1}{2} \times \left(R_1 - R_2 + \frac{k}{2^k+1} + \frac{1}{2}\right)$$

Since the relation gives a very close approximate, we have:

ISSN No:-2456-2165

$$\frac{\pi(x)}{x+1} \sim \frac{1}{2} \times \left(R_1 - R_2 + \frac{k}{2^k+1} + \frac{1}{2}\right) \quad (4)$$

Relation (4) stated above gives an approximate value for the probability of primes within a finite integral set [0, n].

Figure 7 below showcases the interplay between the actual value of the proba- bility of primes and the approximated value of the probability of primes derived through relation (4). Multiple value-disparities are observed for lower order ranges. These discrepancies show the prime distribution between the actual probabilistic values and our approximations for smaller numbers. However, as the graph progressively moves onto higher order ranger; the disparities between the actual value of probability and the approximated values shrinks and the two graphs seem to converge. This showcases the prime distribution between the actual probabilistic values and the approximated values in higher order ranges.

Notably, the two graphs start converging. This convergence marks the cul- mination of our investigation, signifying the attainment of exceptionally close approximations through the utilization of relation (4).

The given constraints of computational resources, results the analysis being conducted over a finite integral set of 5000 numbers. Due to the limitations

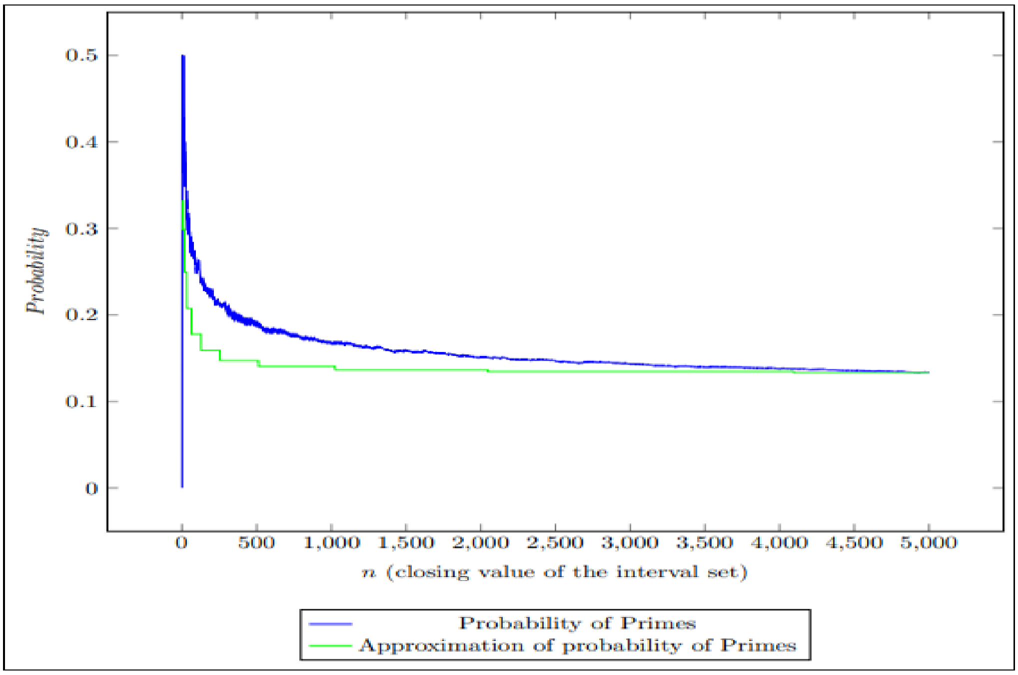

Fig 7 Comparison between actual probability of primes and approximation of probability obtained by relation (4).

Imposed by available computational power, we are unable to showcase the con- vergence behavior for significantly larger values within this paper. Nonetheless, this limitation in presentation should not overshadow the significance of the ob- served trends and their implications within the pattern of primes distribution in natural number sets.

IV. CONCLUSION

This study delves into the probabilistic nature of prime numbers within a finite integral domain of [0, 10000]. Here, we derived two theorems through rigorous mathematical derivations which laid a foundational basis for our propositions and in asserting the domain of our relation. Furthermore, the observed linear regression in our propositions was showcased through graphical representations and a relation which approximated the probability of primes within a finite integral domain was developed. An observed converge of the actual probability and approximated probability of primes through our relation justified our claims and proved the relation. Through elementary proofs, the relation can be further extended to an integral domain [a, n] where $a, n \in \mathbf{W}$.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Dr. Khurshed A. Shah, Assistant Professor, Post-Graduate Department of Physics, Sri Pratap College, Cluster University Srinagar, J&K, India and Dr. Arun Kumar, Associate Pro- fessor, Department of Mathematics, Indian Institute of Technology, Ropar, India for thorough discussions, peer-reviewing and providing literature which helped them to complete this research work.

ISSN No:-2456-2165

REFERENCES

- [1]. David G Wells. Prime numbers. Wiley Hoboken, 2005.
- [2]. Artur Korni-lowicz and Piotr Rudnicki. Fundamental theorem of arithmetic. Formalized Mathematics, 12(2):179–186, 2004.
- [3]. Sheldon M Ross. Introduction to probability and statistics for engineers and scientists. Academic press, 2020.
- [4]. Jacques Hadamard. Sur la distribution des z´eros de la fonction\ zeta (s) et ses cons´equences arithm´etiques. Bulletin de la Societ´e mathematique de France, 24:199–220, 1896.
- [5]. Charles De La Valee-Poussin. Recherches analytiques sur la th´eorie des nombres premiers. Ann. Soc. Sc. Bruxelles, 1896.
- [6]. Hendra Gunawan. On the fundamental theorem of arithmetic and euclid's theorem. reason, 1:1, 2016.
- [7]. Keith Conrad. The infinitude of the primes.
- [8]. Tom M Apostol. A centennial history of the prime number theorem. En- gineering and Science, 59(4):18–28, 1996
- [9]. Adrien Marie Legendre. Essai sur la theorie des nombres; par AM Legendre, membre de l'Institut et de la Legion d'Honneur.. chez Courcier, imprimeur-libraire pour les mathematiques, quai des Augustins, 1808.
- [10]. Atle Selberg. An elementary proof of dirichlet's theorem about primes in an arithmetic progression. Annals of Mathematics, 50(2):297–304, 1949.
- [11]. Benjamin Fine and Gerhard Rosenberger. An epic drama: the development of the prime number theorem. Scientia, Series A: Mathematical Sciences, 20:1–26, 2010.
- [12]. Albert Edward Ingham. The distribution of prime numbers. Number 30. Cambridge University Press, 1990.
- [13]. Jonathan Sondow and Eric W Weisstein. Euler prod- uct. From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/EulerProduct.html.
- [14]. Sy M Blinder. Guide to essential math: a review for physics, chemistry and engineering students. 2013.
- [15]. Aleksandar Ivic. The Riemann zeta-function: theory and applications. Courier Corporation, 2012.
- [16]. Edward Charles Titchmarsh and David Rodney Heath-Brown. The theory of the Riemann zeta-function. Oxford University Press, 1986.
- [17]. Don Zagier. Newman's short proof of the prime number theorem. The American Mathematical Monthly, 104(8):705–708, 1997.
- [18]. Graham James Oscar Jameson. The prime number theorem. Number 53. Cambridge University Press, 2003.
- [19]. Dorian Goldfeld. The elementary proof of the prime number theorem: A historical perspective. In Number Theory: New York Seminar 2003, pages 179–192. Springer, 2004.
- [20]. Atle Selberg. An elementary proof of the prime-number theorem. Annals of Mathematics, pages 305–313, 1949.
- [21]. Andrea Asperti and Wilmer Ricciotti. A proof of bertrand's postulate. Journal of Formalized Reasoning, 5(1):37–57, 2012.
- [22]. Shiva Kintali. A generalization of erd¨os's proof of bertrand-chebyshev the- orem. Georgia Institute of Technology, 2008.