Distribution of Prime numbers

Having discovered prime numbers as basic building blocks, for multiplication, a natural question is, whether there are finite numbers of primes or there are infinitely many primes. Euclid in his elements [Elements (Book IX, proposition 20) [Aigner and Ziegler2010], in 300 BC, that is more than 2300 years ago, proved that there are infinitely many primes. The next question which arises is about their distribution. Given x, how many primes are there less than or equal to x ? Let us explore the answer to this second question. Let the number of primes $\leq x$ be denoted as $\pi(x)$.

For example, $\pi(10) = 4$ as there are 4 primes ≤ 10, namely 2, 3, 5, and 7. Similiarly, $\pi(3) = 2$ and $\pi(25) = 9$. (Primes ≤ 25 are 2, 3, 5, 7, 11, 13, 17, 19, and 23). If we look at the sequence of prime numbers ≤ x, as x increases, they get sparser and rarer on the average. Just, have a look at the table below of prime numbers upto 300, we find:

Range	Prime numbers
2 - 10	2, 3, 5, 7
11 - 20	11, 13, 17, 19
21 -30	23, 29
31 - 40	31, 37
41 - 50	41, 43, 47
51 - 60	53, 59
61 - 70	61, 67
71 - 80	71, 73, 79
81 - 90	83, 89
91 - 100	97
101 - 110	101, 103, 107, 109
111 - 120	113
121 - 130	127
131 - 140	131, 137, 139
141 - 150	149
151 - 160	151, 157
161 - 170	163 ,167
171 - 180	173, 179
181 - 190	181
191 - 200	191, 193, 197, 199
201 - 210	--
211 - 220	211
221 - 230	223, 227, 229
231 - 240	233, 239
241 - 250	241
251 - 260	251, 257
261 - 270	263, 269
271 - 280	271, 277
281 - 290	281, 283
291 - 300	293

Table 1: Table of prime numbers upto 300

$$\pi(100) = 25$$

$$\pi(200) = 21 + 25$$

$$\pi(300) = 16 + 25 + 21$$

See, number of primes up to 100 is 25, whereas between 100 and 200 is 21 and between 200 and 300 is 16 & so on. Following are the graphs for primes up to 100, 1000 and 10,000 respectively.

1. $\pi(x)$: No. of primes for $x \leq 100$

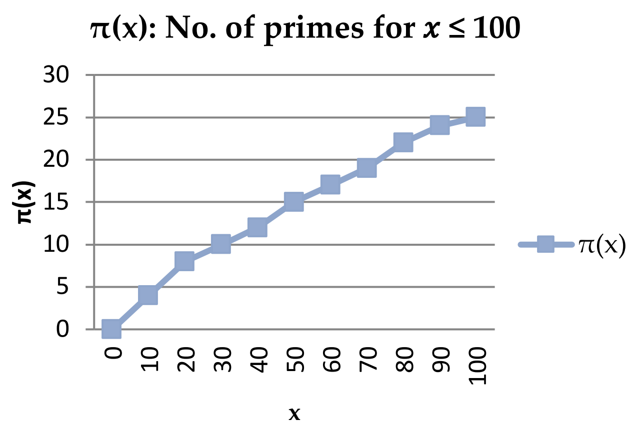

Figure 1

2. $\pi(x)$: No. of primes for $x \leq 1000$

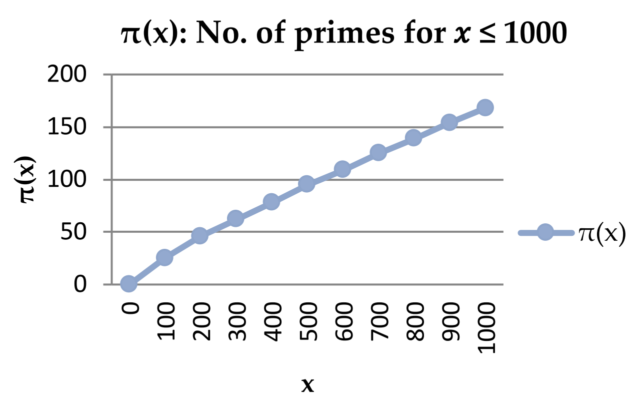

Figure 2

3. $\pi(x)$: No. of primes for $x \leq 10000$

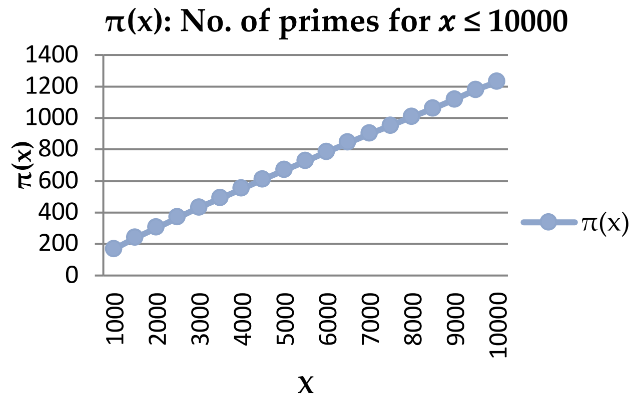

Figure 3

Distribution of $\pi(x)$ is not uniform; determining the behavior of $\pi(x)$ has been quite difficult and inspiring problem for researchers. On one side, for any $n \geq 2, n! + 2, n! + 3, \ldots, n! + n$ are all composite, giving a stretch of $n - 1$ consecutive numbers containing no prime number. On the other side, it is conjectured that there are twin prime numbers, that is, prime numbers that differ by 2, (like 3 and 5, 11 and 13), no matter how up on integer number line, one goes. In spite of this random behavior of primes, after examining tables of prime numbers, Legendre the French mathematician conjectured in 1798 that

$$\pi(x) \approx \frac{x}{\log x - 1.08366}$$

Gauss conjectured that rate of increase of $\pi(x)$ is same as that of the functions:

$$x/\log x \text{ and } \quad Li(x) = \int_2^x \frac{dt}{\log t} \qquad \text{(Li : Logarithmic integral)}$$

But, neither Legendre nor Gauss could prove their conjectures. The prime number theorem which describes the asymptotic distribution of prime numbers was finally proved in 1896, by two French and Belgian mathematicians independently. The most interesting part of the proof is that it is based on complex analysis (uses Riemann zeta function in complex plane), though there are no complex variables in the statement of this theorem. Later in 1949, two mathematicians , this time Norwegian and Hungarian independently again found proofs of the prime number theorem without using complex variables , but the proofs are not easy to understand.

Prime Number Theorem: The ratio of $\pi(x)$ to $x/\log x$ approaches 1 as x increases without bound. That is , $\pi(x) \approx x/\ln(x)$; which is same as,

$$\text{limit } x \to \infty \ \frac{\pi(x)}{\left(\frac{x}{\ln(x)}\right)} = 1$$

$\pi(x)$ is called prime counting function and $\ln(x) = log_e x$, that is natural logarithm of x. The following figures show relationship between x, $\pi(x)$ and its approximations for

(i) $x \leq 100$

(ii) $x \leq 1000$
(iii) $x \leq 10000$

1. Comparison for $x \leq 100$

$\pi(x)$, $x/\log x$ and $x/\log(x - 1.08366)$ for $x \leq 100$

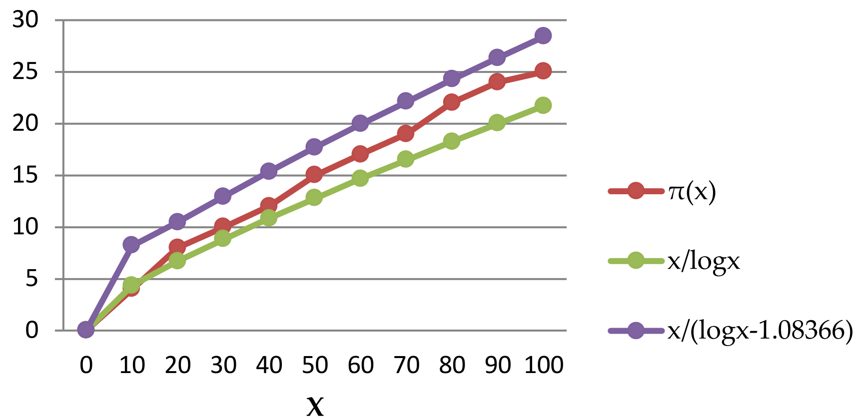

Figure 4

2. Comparison for $x \leq 1000$

$\pi(x)$, $x/\log(x)$ and $x/\log(x-1.08366)$ for $x \leq 1000$

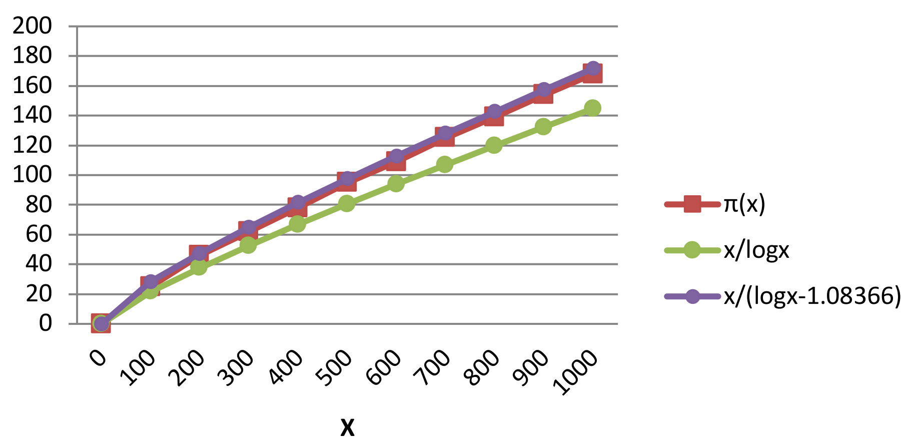

Figure 5

3. Comparison for $x \leq 10000$

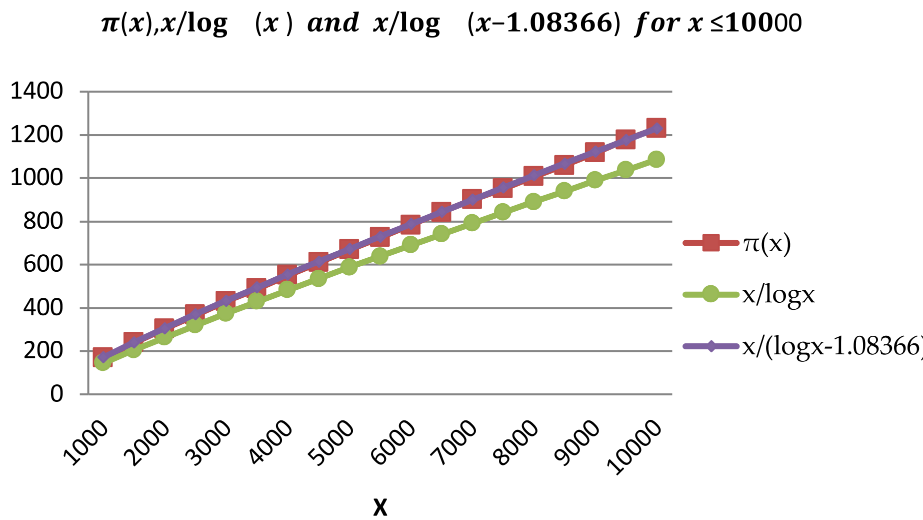

Figure 6

Though, **prime number theorem** approximates $\pi(x)$ with $x/\ln(x)$, it has been established (also evident from preceding graphs in *Figure 4, 5 and 6*) that $\frac{x}{\log x - 1.08366}$, is a very good approximation to $\pi(x)$. After a long gap, with the availability of computers, different researchers have worked on $\pi(x)$ taking help of computers. Distributed computer project was done to derive $\pi(10^{23})$ in 2007 - 2008. $\pi(10^{25})$ has been calculated in May, 2013 using analytical method.

One of the applications of prime numbers is in cryptography. RSA cryptography makes use of prime numbers. Large prime numbers are used to make the process more secure, almost unbreakable for all practical purposes. Suppose, we want to use a 50 digit prime, wish to estimate how many prime numbers of 50 digits exist, then, number

$$= \pi(10^{50}) - \pi(10^{49})$$

$$\approx \frac{10^{50}}{\ln 10^{50}} - \frac{10^{49}}{\ln 10^{49}}$$

$$= \frac{1}{\ln 10}\left(\frac{10^{50}}{50} - \frac{10^{49}}{49}\right)$$

$$\cong \frac{1}{50 \ln 10} 10^{49}(10 - 1)$$

$$\cong 10^{47} \times 18 / 2.3026$$

$$\cong 10^{47} \times 18 / 2.3026$$

$$\cong 10^{47} \times 7.82$$

Thus, there is quite large number of primes of 50 digits of the order of 10^{47}.

Now, let us apply prime number theorem to determine the probability of a randomly chosen number in the range $[1, x]$ to turn out to be a prime number. From prime number theorem, the number of prime numbers $\leq x : \pi(x)$ is approximately $x/\ln(x)$, and there are x numbers, it can be interpreted that, if we pick at random, a sufficiently large number, say x chance of it being prime is approximately $(x / \ln x) / x$, that is, $1 / \ln x$. So, for numbers 50 digit long, chance is $\frac{1}{\ln 10^{50}} = \frac{1}{50 \ln 10} = \frac{1}{115.13}$ means that, 1 in 115 is prime. That is, roughly, in the worst scenario, one may have to check 115 numbers of 50 digits to get a prime number of this size.

The next question is, how to check whether a number is prime? The interested readers can refer for some of the efficient methods, like Fermat's method and Miller-Rabin primality test, illustrated with examples in the book [link --- Cryptography and Network Security]. Proofs of the two base properties used in Miller-Rabin primality test (*section 4.6.2: Miller-Rabin primality test)* is available at https://www.savitagandhi.com/articles/miller-rabin-primality-test. Another common primality test, namely, Solovay-Strassen primality test is discussed, explained, and illustrated with examples at https://www.savitagandhi.com/articles/solovay-primality-test

****