The Density of Primes Less or Equal to a Positive Integer up to 20,000: Fractal Approximation

[1]Rodel B. Azura, [1]Dennis A. Tarepe, [2]Mark S. Borres and [3]Joy Panduyos

Abstract

The highly irregular and rough fluctuations of the number of primes less or equal to a positive integer x for smaller values of x (x≤20,000) renders the approximations through the Prime Number Theorem quite unreliable. A fractal probability distribution more specifically, a multifractal fit to the density of primes less or equal to x for small values of x, is tried in this study. Results reveal that the multifractal fit to the density of primes in this situation outperforms the Prime Number Theorem approximation by almost 200% viz. the prediction error incurred by using the PNT approximation is double that of the multifractal fit to the density of primes. The study strongly suggests that a better multifractal distribution exists, even for large x, than the Prime Number approximation to the density of primes.

Keywords: density of primes, prime number theorem, multifractal distribution
AMS classification: number theory, applied mathematics

1.0 Introduction

This study explores the possibility of fitting a fractal density to the distribution of prime numbers less than or equal to a positive integer X (X ≤ 20,000). This is the first in a series of papers which ultimately culminates in providing for a fractal approximation to the density of primes less than or equal to any positive integer. For larger values of X, the prime number theorem (PNT) provides an asymptotic approximation namely:

1. $$\frac{\pi(x)}{x} \approx \frac{1}{\log x}$$, for large X, (Legendre and Gauss (1760))

where $\pi(x) =$ number of prime less than or equal to X.

$\log(x) =$ natural logarithm of X.

For small values of X, e.g $x = 10^n, n = 1,2,3,4,5,6,$ the difference between the actual $\pi(x)$ and that estimated by PNT (say, $\pi_{PNT}(x)$)is quite large (Padua, 2012).

The PNT is a smooth approximation of the density of primes. Consequently, the fluctuation in the values of $\frac{\pi(x)}{x}$ cannot be captured by this approximation especially for small values of X. On the other hand, fractal analysis provides a convenient platform for modelling repeated large and small scale fluctuations in the values of $\frac{\pi(x)}{x}$. Figure 1 shows the graph of the actual approximation for small values of X, $x = 1,2,3,.20$:

[1]Mindanao University of Science and Technology
[2]University of San Jose-Recoletos
[3]Surigao del Sur State University

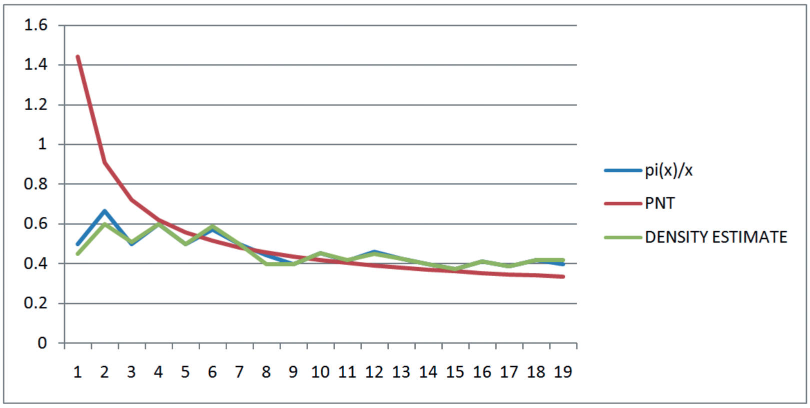

Figure 1: Actual Density of Primes, PNT and Fractal Approximation X=2,3,...20

After the proof of M. De La Valee Poussin and J. Hadamard (1854) of the prime number theorem (PNT), there had been several attempts to improve on its smoothing approximation particularly for small values of X. Notable among these attempts were those provided by Leibnitz and Bernoulli:

2. $$\frac{\pi(x)}{x} \approx \frac{1}{A\log x + B}$$

where A and B are suitably chosen constants. Riemann later showed that $A = 1,\ B = 0$ remain the best choices which brings us back to the PNT. In fact, Riemann demonstrated the truth of the remarkable statement that "among all possible smooth approximation to $\frac{\pi(x)}{x}$, the logarithmic smoothing is the best." (Erdos and Dudley, 1983)

Thus, any attempt to improve on the approximation of $\frac{\pi(x)}{x}$ must necessarily stay out of the context of smooth function. This study is one such attempt. We organize the paper as follows: Section 2 discusses the importance of the density of primes via the Prime number theorem (PNT) in number theory; Section 3 introduces the concept of Fractal statistical analysis as developed by Padua et al. (2012) with the end-in-view of using the statistical methodology therein, to address the shortcoming of PNT for small x (up $x = 10^6$); section 4 gives the research methodology; section 5 gives the results and conclusion of the study.

2.0 Density of Primes

Prime numbers are positive integers greater than 1 whose divisors are one (1) and itself only. The first and only even prime is 2. Euclid (n.d.) proved the famous fundamental theorem of arithmetic: "There are infinitely many prime numbers" by *reductio ad absurdum*. Moreover, he also demonstrated that every positive integer can be decomposed into the product of primes less than that integer:

3. $$x = \prod_{i=1}^{n} P_i^{r_i} \text{ , where } P_i / x \text{ for all } i.$$

It is this property of primes that makes them important in mathematics: they are the fundamental building blocks of positive integers.

To date, there is no readily available formula for generating prime numbers. However, a famous conjecture by Berndhardt Riemann (1870) states that the location of the zeroes of the zeta function:

4. $$\xi(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \text{ , } (s > 1)$$

considered as an analytic continuation to the entire complex plane is on the line $x = \frac{1}{2}$ (viz, $z = \frac{1}{2} + i$). If true, the location of the zeroes of the zeta function relates to the distribution of primes. He asserted then that degree to which the density of primes less or equal to X differs from $\frac{1}{\ln(x)}$ varies in a "systematically random fashion" implying that although it is not possible to predict when the next prime will occur, the general pattern of primes appears " regular "(Dunham, (1990)).

The connection between primes and the Riemann zeta function was demonstrated by Euler. Note that we can write:

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\cdots$$
$$=\left(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots\right)\left(1+\frac{1}{3}+\frac{1}{9}+\cdots\right)\left(1+\frac{1}{5}+\cdots\right)\cdots$$

In the region of absolute convergence of the zeta function, we have:

$$\zeta(s)=\frac{1}{1^s}+\frac{1}{2^s}+\frac{1}{3^s}+\frac{1}{4^s}+\frac{1}{5^s}+\frac{1}{6^s}+\frac{1}{7^s}+\cdots$$
$$=\left(1+\frac{1}{2^s}+\frac{1}{4^s}+\frac{1}{8^s}+\cdots\right)\left(1+\frac{1}{3^s}+\frac{1}{9^s}+\cdots\right)\left(1+\frac{1}{5^s}+\cdots\right)\cdot$$
$$=\prod_p\sum_{k=0}^{\infty}p^{-ks}$$
$$=\prod_p\left(1-\frac{1}{p^s}\right)^{-1},$$

This age-old problem has fundamental applications in computer security today. Computer security codes are expressed in term of huge numbers which the entails factoring these numbers into their prime factorization. The larger the number (say about 250 digits), the more difficult it is to decrypt the code. Cryptographers and security experts are worried that if powerful enough techniques are developed (in relation to the Riemann hypothesis) which will lead to better factoring algorithms, then the current cryptosystems will become vulnerable.

3.0 Fractal Statistics

Rather than deal directly with a difficult Riemann hypothesis in approximating the number of primes less or equal to x, we prefer to approach the problem of estimating $\frac{\pi(x)}{x}$ outside the realm of continuous and smooth techniques. This approach is anchored on fractal Geometry which thrives on the analysis of roughness, fluctuation and irregularities,

Fractal Geometry was firs placed in formal mathematical setting by Benoit Mandelbrot (1967) in his book "Fractal: the Geometry of nature". He argued, essentially, that nature and natural processes are highly irregular, fluctuating and rough for which current mathematical methods(accumulated over more than two thousand years) of employing smooth and regular curves are inadequate. He then proceeded to describe "roughness" through the fractal dimension (λ) of a natural geometric object. A geometric object is said to be "fractal" if (a) it exhibits ruggedness or irregularities repeated at various scales, and (b) it possesses a fractional (non-integer) dimension. The box-counting fractal dimension of an object is:

5. $$\lambda = \frac{\log(m)}{\log(n)}$$

where m = Number of copies of the observed fluctuations
n = Scaling factor

For instance, the fractal dimension of the famous Cantor set is λ =0.63. The Cantor set is obtained by successively removing the middle third of a unit interval. After doing the process over and over again, what remains are called " fractal dusts" which constitute the Cantor set.

Mandelbrot passed away in October, 2010 without extending his work from Geometry to the other fields of Mathematics: Statistics, Analysis and Algebra. In December, 2012 Padua et al. (2012)

published an initial attempt to extend Fractal Geometry to Fractal Statistics in the paper: "Data Roughness and Fractal Statistics." Essentially, the paper exploited the connection between fractal dimension (λ) and space filling property of a fractal object with information theory.

Let X be a random variable whose probability density function obeys the power law:

6. $f(x) = \left(\frac{\lambda - 1}{\theta}\right)\left(\frac{x}{\theta}\right), x \geq \theta, \theta \geq 0, \lambda > 0$

The random variable X is then called a fractal random variable and $f(x)$ is its fractal probability distribution. The first moment of X (its mean) will not exist for $\lambda < 2$. Consequently, the second moment (its variance) will also not exist for $\lambda < 2$. The parameter λ of (6) is called the fractal dimension of X.

For $\lambda \leq 2$, the non-existence of the second moment or variance of X implies that observation from fractal distribution are highly erratic, fluctuating and rough. In fact, the Central Limit Theorem fails to apply in cases where the observation come from fractal distribution.

For $\lambda > 2$, the variance σ^2 exist and is related to λ by:

7. $\lambda = 1 + \theta\sigma$ (Padua et al. (2012))

In other words, when the variance exists, the fractal dimension λ describes the variability of the data around the mean just as the standard deviation (σ) does. Further, the fractal dimension, λ, of X is a more general description of data variability than σ.

From (6), the maximum likelihood estimator of λ is easily obtained as

8. $\hat{\lambda} = 1 + n\left(\sum_{i=1}^{n} \log\left(\frac{x_i}{\theta}\right)\right)^{-1}$,

for $x_1, x_2, \ldots, x_n$, iid $f(x)$, Similarly, the cumulative distribution function (cdf), $F(x)$, is:

9. $F(x) = P(X \leq x) = 1 - \left(\frac{x}{\theta}\right)^{1-\lambda}$

Equation (9) gives the probability that an observation X is less or equal to x.

4.0 Study Objectives

The study aims to make use of the ability of fractal statistics to describe irregular and highly fluctuating series of observation to deduce the distribution of primes less or equal to a number x, $\frac{\pi(x)}{x}$. Specifically:

a. Obtain the graph of $\frac{\pi(x)}{x}$ versus x for $x = 2,3,\ldots,10^6$.
b. Obtain the numerical approximation due to the prime number theorem (PNT) of $\frac{\pi(x)}{x} \approx \frac{1}{\log(x)}$ for $x = 2,3,\ldots,10^6$ and compare the results with the actual $\frac{\pi(x)}{x}$.
c. Fit a fractal distribution to $\frac{\pi(x)}{x}$ for $x = 2,3,\ldots,10^6$ both for monofractal and multifractal cases.
d. Obtain the numerical approximation due to the fractal distribution fit made in (3) for $x = 2,3,\ldots,10^6$
e. Compare the numerical approximation of PNT and the fractal fit (FF).

5.0 Research Design and Methodology

Data for the prime, P_i, are obtained from the freely accessible WOLFRAM.MATHWORLD website. The website contains all prime known up to $x = 10^{25}$. The listing of primes was used to construct the actual density of primes less or equal to x:

10. $\frac{\pi(x)}{x} = \frac{no. of\ primes\ less\ or\ equal\ to\ x}{x}$, $x = 2,3,\ldots,10^6$

An EXCEL program was developed to automate the process of computing (10). The pairs $\left(x_i, \frac{\pi(x_i)}{x}\right)$ were saved to a file. Next, we computed the PNT approximation:

11. $PNT = \frac{1}{\ln(x)}$, $x = 2,3,\ldots,10^6$

and saved the pairs $\left(x_i, \frac{1}{\ln(x_i)}\right)$ to a separate file.

Finally we computed the fractal fit approximation (FF) in a series of steps. From the data, we first computed for the maximum-likelihood estimator of λ:

12. $\hat{\lambda} = 1 + n\left(\sum_{i=1}^{n} \log\left(\frac{x_i}{\theta}\right)\right)^{-1}$, where $\theta = 2$ (minimum of x's)

Next, we plugged in the value $\hat{\lambda}$ obtained in (12) to the cumulative distribution function $F(x)$:

13. $\hat{F}_n(x) = 1 - \left(\frac{x}{\theta}\right)^{1-\hat{\lambda}}$, $\theta = 2$, $x = 2,3,\ldots,10^6$, $n = 10^6$

The pairs $\left(x_i, \hat{F}_n(x_i)\right)$ are saved in a separate file.

Using MATLAB we then proceeded to graph the following curves in a single figure:

$S_1 : \left(x_i, \frac{\pi(x_i)}{x}\right)$ $S_2 : \left(x_i, \frac{1}{\ln(x_i)}\right)$ and $S_3 : \left(x_i, \hat{F}_n(x_i)\right)$

In order to compare the performance of PNT versus FF, we computed the absolute error:

14. $MAE = \frac{\sum_{i=1}^{n} |(actual\ density - proposed\ density|}{n}$, where $n = 10^6$

Multi-Fractal Distribution Fit (MFF)

The fit provided by (9) assume that there is a single exponent (fractal dimension) λ that would explain the global behaviour of $\frac{\pi(x)}{x}$. In the event that (10) proves to be large for the FF approximation using only one $\hat{\lambda}$, we modify (9) and assume several fractal dimensions (or multi fractal system). In this case, we assume that:

15. $\frac{\pi(x)}{x} \propto \frac{1}{x^{\lambda}}$ $\theta = 2$

We solve for the value of λ as follows:

16. $\lambda = 1 - \frac{\ln(\pi(x))}{\ln(x)}$

and then obtain several approximation $\hat{F}_n(x)$:

17. $\hat{F}_n(x) = \frac{1}{x^{\lambda}}$, , $x = 1,2,\ldots,n$, $n = 10^6$

The pairs $\left(x_i, \hat{F}_i\right)$ are saved and (14) is computed. We will refer to this as the multifractal fit (MFF).

Prediction Model

For prediction purposes, we regress the values of λ obtained in (16) to the values of x:

. $\lambda_i = a + bh(x_i)$, $i = 1,2,\ldots,n$

For our prediction of the value of λ_{n+k}, for k =1,2,..., we use:

. $\lambda_{n+k} = a + bh(x_{n+k})$, $k = 1,2,\ldots$

where h(.) is a function of x.

6.0 Results and Discussions:

6.1 Initial Experiment for X ≤ 1,000

Table 1 shows the mean absolute error for the PNT, FF (using the maximum likelihood estimator of the fractal dimension) and the MFF (multifractal estimator of the density of primes) for X = 2,3,...,1000:

Table 1: Mean Absolute Error of the Three Estimators of the Density of Primes for X ≤ 1000

ESTIMATOR	MAE	SD
PNT	0.03113	0.03112
FF	0.10801	0.05967
MFF	0.00000	0.00000

Note that the PNT outperforms the monofractal fit (FF) using a single fractal dimension $\lambda = 1.1321$ but is inferior to the multifractal fit (MFF). The latter has zero mean absolute error which means that the multifractal fit perfectly predicted the actual density of primes for all X less than 1000. Figure 2 shows the graphs of the three estimators with the actual density of primes:

Since the MFF was found to be the most superior estimator of the actual density of primes, we fitted a prediction line on it to forecast future values of the fractal dimensions beyond X = 1000. The fitted regression line is given below:

The regression equation is
lambda1 = 0.204 + 0.395/LN(X)
R-SQUARED : 82.10%

Table 2: Actual vs. MFF Prediction Values

X	new lambda	new est.	Actual
1001	0.261174	0.164575	0.168000
1002	0.261166	0.164542	0.167665
1003	0.261157	0.164508	0.164980
1004	0.261149	0.164475	0.167331
1005	0.261141	0.164442	0.167164
1006	0.261133	0.164408	0.166998
1007	0.261124	0.164375	0.166832
1008	0.261116	0.164342	0.166667
1009	0.261108	0.164308	0.167493
1010	0.261111	0.165109	0.177438

MEAN ABSOLUTE PREDICTION ERROR: 0.00292 = 0.292%

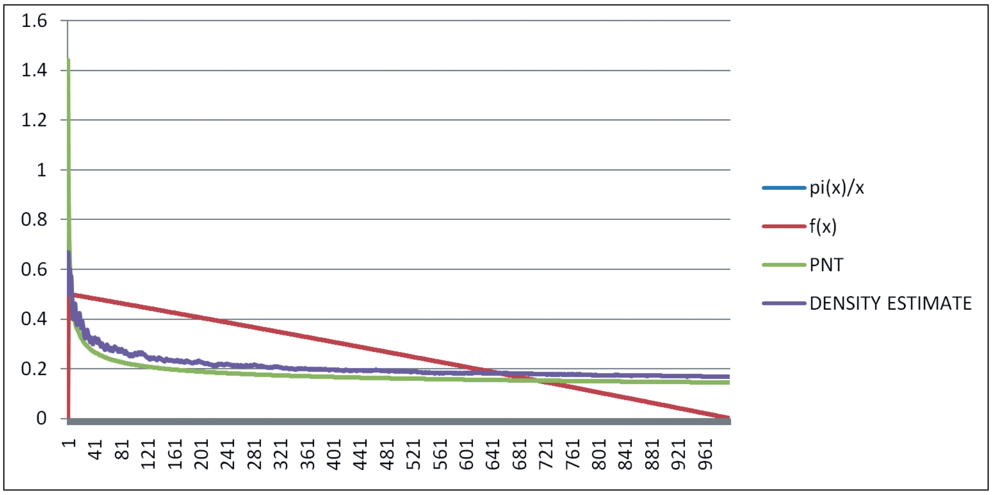

Figure 2: Comparison of PNT, FF, MFF with the Actual Density of Primes less or equal to 1000

Figure 3 shows the actual density of primes for the next 10 values of X and the MFF fit.

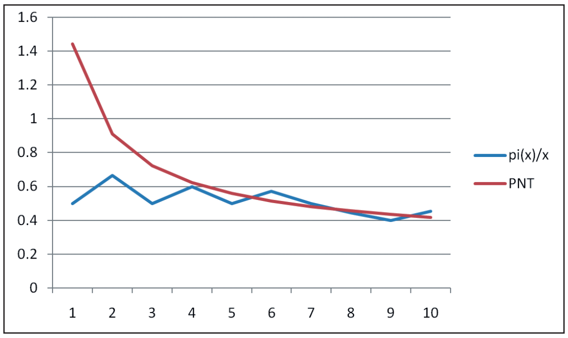

Figure 3: Predicted density of primes by MFF versus actual density of primes

The best fitting regression curve gives an r-squared value of 82.10%. Moreover, we note that, as expected, the absolute error of prediction gradually increases as X gets farther and farther away from the last given value at X = 1000. On the average, however, the mean absolute prediction error is less than 1%.

6.2 Second Experiment for X ≤ 10,000

Next, we continued exploring the data set on primes by performing the same experiment on a larger sample size (ten(10) times larger than the initial experiment). We have eliminated FF from consideration since the initial experiment showed that it does not perform well. We concentrate on the performance of the PNT and the MFF fit. Table 3 shows the Mean Absolute Error for PNT and MFF :

Table 3: Mean Absolute Error of PNT and MFF for X ≤ 10,000

ESTIMATOR	MAE	SD
PNT	0.01820	0.01098
MFF	0.00000	0.00000

Again, the MFF perfectly predicted the actual density of primes while the PNT incurred a mean absolute error of 1.820%. Figure 4 shows the plot of PNT, MFF and the actual density of primes:

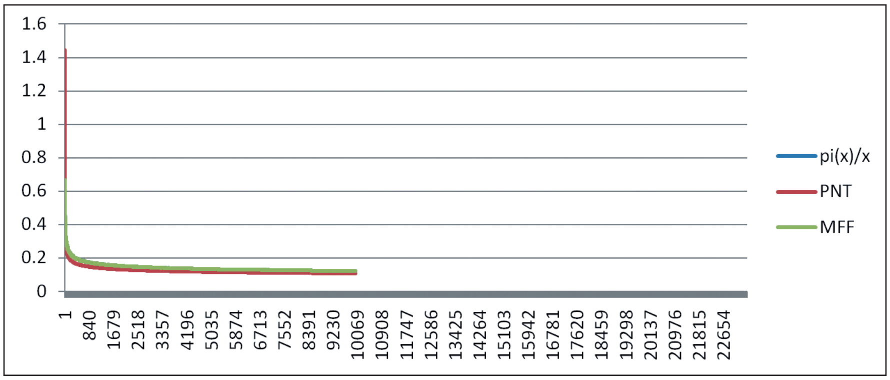

Figure 4: MFF, PNT versus Actual Density of Primes

Next, we performed the usual regression on the values of lambda versus the logarithm of X: The regression equation is:lambda = 0.175 + 0.524/ln(X)with: R-squared : 86.9%

Table 4: Mean Absolute Prediction Errors for MFF and PNT for X=10,001 TO 10,010

X	actual density	PNT	lambda	MFF	MFF error	PNT error
10001	0.122900	0.108572	0.231892	0.118147	0.0047532	0.0143276
10002	0.122888	0.108571	0.231891	0.118145	0.0047429	0.0143164
10003	0.122875	0.108570	0.231891	0.118143	0.0047327	0.0143053
10004	0.122863	0.108569	0.231890	0.118141	0.0047225	0.0142942
10005	0.122851	0.108568	0.231889	0.118139	0.0047123	0.0142831
10006	0.122839	0.108567	0.231889	0.118137	0.0047021	0.0142720
10007	0.122826	0.108565	0.231888	0.118134	0.0046919	0.0142609
10008	0.122814	0.108564	0.231888	0.118132	0.0046817	0.0142498
10009	0.122802	0.108563	0.231887	0.118130	0.0046715	0.0142387
10010	0.122789	0.108562	0.231886	0.118128	0.0046613	0.0142276
MEAN ABSOLUTE PREDICTION ERROR:					**0.00471**	**0.01428**

Figure 4 shows the graph of MFF,PNT and the actual density of primes within the prediction region from X = 10001 to X = 10010:

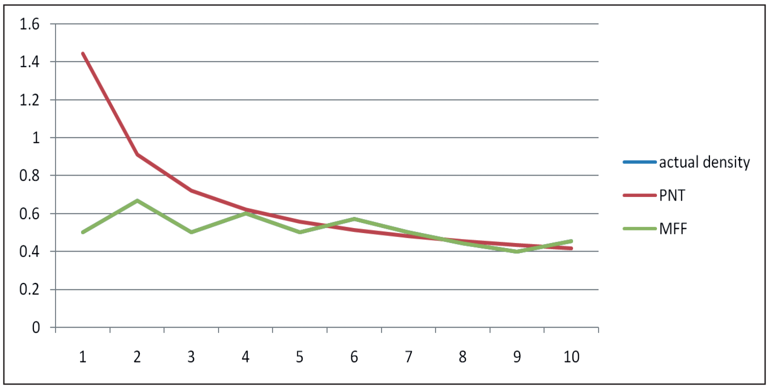

Figure 4: Values of MFF(predicted), PNT (predicted) and Actual Density of Primes

6.3 Third Experiment with X ≤20,000

Our third experiment increased the sample size further by doubling the previous number of primes. Table 4 shows the mean absolute errors of the PNT and MFF for this experiment.

Note that while the MAE of the PNT increased to 0.02317, the MFF remained in perfect synchronization with the actual density of primes for X = 2 to X = 20,000. Figure 5 shows the graph of PNT, MFF and the actual density of primes:

Table 4: Mean Absolute Error of PNT and MFF for X ≤ 20,000

ESTIMATOR	MAE	SD
PNT	0.02317	0.01367
MFF	0.00000	0.00000

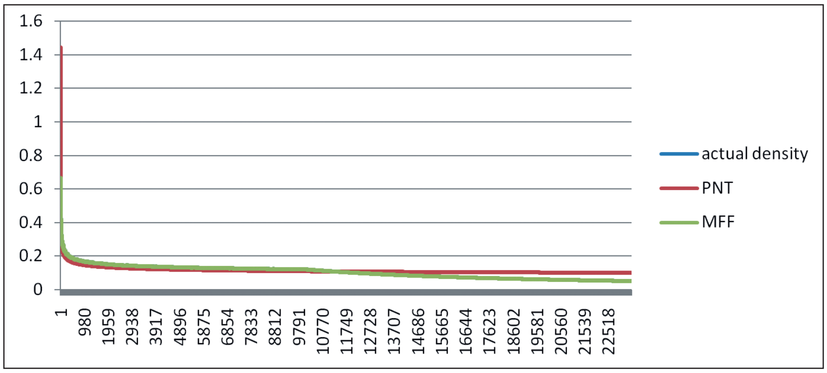

Figure 5: MFF, PNT and the Actual Density of Primes for X ≤ 20,000.

For prediction purposes, we regressed the values of the fractional dimensions, lambda, with the logarithm of X as before to obtain:

log(lambda) = 0.413 - 0.483 lnX + 0.0312 (lnX) 2 , R-squared: 60.70%,

Table 5: Mean Absolute Prediction Error for PNT and MFF X = 20001 to X = 20010

density	new PNT	lambda1	NEW MFF	PNT(ERROR)	MFF(ERROR)
0.061447	0.100974	0.269736	0.0691593	0.039527	0.0077123
0.061444	0.100974	0.269738	0.0691593	0.039530	0.0077153
0.061441	0.100973	0.26974	0.0691593	0.039532	0.0077183
0.061438	0.100972	0.269742	0.0691593	0.039535	0.0077213
0.061435	0.100972	0.269743	0.0691593	0.039537	0.0077243
0.061432	0.100971	0.269745	0.0691593	0.039540	0.0077273
0.061428	0.100971	0.269747	0.0691593	0.039542	0.0077313
0.061425	0.100970	0.269749	0.0691593	0.039545	0.0077343
0.061422	0.100970	0.269751	0.0691593	0.039548	0.0077373
0.061419	0.100969	0.269753	0.0691593	0.039550	0.0077403
MEAN ABSOLUTE PREDICTION ERROR:				**0.03954**	**0.0077262**

The ratio of the prediction error PNT to the prediction error MFF is: 5.1176 which means that the MFF is about 5 times more efficient than the PNT approximation. Figure 6 shows the graph of the prediction error curves MFF and PNT:

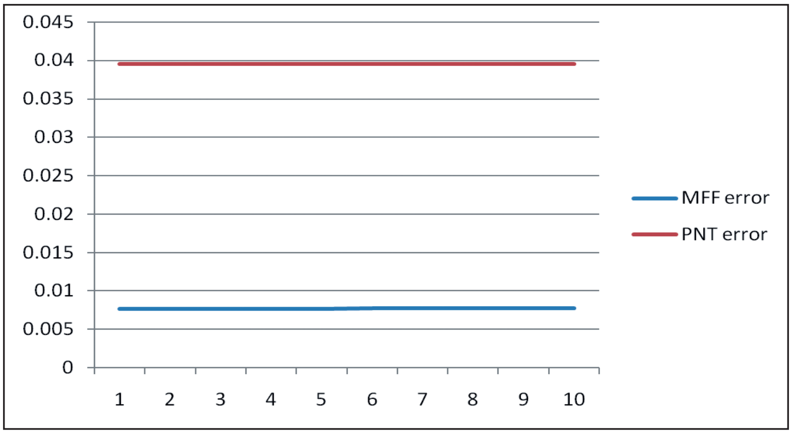

Figure 6: Prediction Error Curves for MFF and PNT

The graph of log(lambda) versus log(x) is quite revealing of the behaviour of the fractal dimension. We show this below:

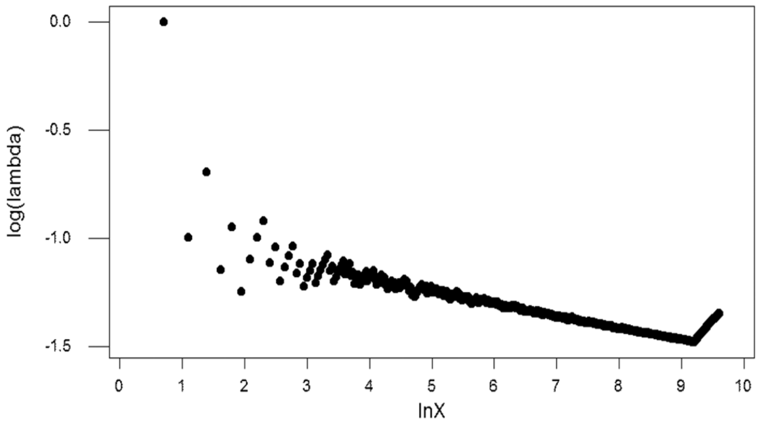

Figure 7: Graph of log(lambda) versus log(X)

Figure 7 shows that the fractal dimension decreases monotonically up to ln(X) ~9.25, and then it increases thereafter. The implication is that towards the tail of the series of values, the density of primes becomes more rugged.

References

Chris K. Caldwell. (2010)"Conditional Calculation Chris K. Caldwell. (2010)"Conditional Calculation of pi(1024)"..: "Computing π(x) Analytically)". Retrieved Jul. 25, 2012.

D. J. Newman (1980). "Simple analytic proof of the prime number theorem". American Mathematical Monthly 87 (9): 693–696.

Glaisher, J. W. L.(1891) "On the Sums of Inverse Powers of the Prime Numbers." Quart. J. Math. 25, 347-362.

Graybill, W. (1980). General Linear Models. (Wiley Series, NY)

Hardy, G. H. & Littlewood, J. E. (1916). "Contributions to the Theory of the Riemann Zeta-Function and the Theory of the Distribution of Primes". Acta Mathematica 41: 119–196.

Ingham, A.E. (1990). The Distribution of Prime Numbers. Cambridge University Press. pp. 2–5.

Merrifield, C. W. (1881)"The Sums of the Series of Reciprocals of the Prime Numbers and of Their Powers." Proc. Roy. Soc. London 33, 4-10.

Nils (2009) "Formalizing an analytic proof of the Prime Number Theorem". Journal of Automated Reasoning. volume 43, pp. 243—261.

O. Shanker (2006). "Random matrices, generalized zeta functions and self-similarity of zero distributions". J. Phys. A: Math. Gen. 39 (45): 13983–13997.

Padua, R, Palompon, D and Ontoy, D. (2012). "Data Roughness and Fractal Statistics" (CNU Journal of Higher Education, CHED-JAS – Category A, Vol. 6, No. 2).

Pierre Dusart (1999). "The kth prime is greater than k(ln k + ln ln k−1) for k ≥ 2". Mathematics of Computation68: 411–415.

Tao, Terence (2007) "Video and slides of Tao's lecture on primes", UCLA January 2007.

Similarity Index: 5%
Paper ID: 374701795