2016

# A Power Law Governing Prime Gaps

Raul Matsushita, *University of Brasilia*
Sergio Da Silva, *Federal University of Santa Catarina*

Open Access Library Journal
2016, Volume 3, e2989
ISSN Online: 2333-9721
ISSN Print: 2333-9705

# A Power Law Governing Prime Gaps

**Raul Matsushita[1], Sergio Da Silva[2*]**

[1]Department of Statistics, University of Brasilia, Brasilia, Brazil
[2]Graduate Program in Economics, Federal University of Santa Catarina, Florianopolis, Brazil
Email: *professorsergiodasilva@gmail.com

**How to cite this paper:** Matsushita, R. and Da Silva, S. (2016) A Power Law Governing Prime Gaps. *Open Access Library Journal*, **3**: e2989.
http://dx.doi.org/10.4236/oalib.1102989

**Received:** August 18, 2016
**Accepted:** August 29, 2016
**Published:** September 2, 2016



## Abstract

A prime gap is the difference between two successive prime numbers. Prime gaps are casually thought to occur randomly. However, the "*k*-tuple conjecture" suggests that prime gaps are non-random by estimating how often pairs, triples and larger groupings of primes will appear. The *k*-tuple conjecture is yet to be proven, but a very recent work presents a result that contributes to a confirmation of the *k*-tuple conjecture by finding unexpected biases in the distribution of consecutive primes. Here, we present another contribution to confirmation of the *k*-tuple conjecture based on statistical physics. The pattern we find comes in the form of a power law in the distribution of prime gaps. We find that prime gaps are proportional to the inverse of the chance of a number to be prime.

## Subject Areas

Applied Statistical Mathematics, Number Theory

## Keywords

Prime Numbers, Power Laws, Prime Gaps, *k*-Tuple Conjecture, Number Theory

## 1. Introduction

Prime numbers are divisible only by themselves and 1. Primes are the building blocks of the entire number line because all the other numbers are created by multiplying primes together. Thus, primes are the core of arithmetic.

Whether a number is prime or not is pre-determined, as evidenced by innumerous laws already proven. For instance, the prime number theorem states that the average length of the gap between a prime $p$ and the next prime is $\ln p$. Such prime gaps have been extensively studied; however many questions and conjectures remain unanswered. In particular, there is no way to predict with certainty which numbers are

prime and so, for practical purposes, prime gaps are considered to occur randomly. Actually, the "*k*-tuple conjecture" [1] allows for a non-random explanation by estimating how often pairs, triples and larger groupings of primes will appear. The *k*-tuple conjecture is yet to be proven, but a very recent work [2], however, presents a result that contributes to a confirmation of the *k*-tuple conjecture by finding unexpected biases in the distribution of consecutive primes.

Apart from 2 and 5, all prime numbers end in 1, 3, 7 or 9, and each of the four endings is supposedly equally likely. But the authors in Ref. [2] find that primes ending in 1 were less likely to be followed by another prime ending in 1. That is unexpected if the primes are truly random. The authors [2] find that in the first hundred million primes, a prime ending in 1 is followed by another ending in 1 just 18.5 percent of the time. (Incidentally, we confirmed this pattern using a sample of 50,000 prime numbers, but found 15.8 percent of the time instead. Perhaps the difference can be explained by our smaller sample, or perhaps we are right after all and 18.5 is just a typo.) If the primes were distributed randomly, the authors argue that one would expect to see two 1 s next to each other 25 percent of the time. Primes ending in 3 and 7 follow a 1 about 30 percent of the time, while a 9 follows a 1 around 22 percent of instances.

The authors then show that the last-digit pattern can be explained by the groupings given by the *k*-tuple conjecture. However, as the primes tend to infinity, the pattern vanishes and the primes become genuinely random. Here, we contribute to the literature by presenting further evidence that the *k*-tuple conjecture can be true. In line with the authors in Ref. [2] we also find unexpected biases in the distribution of consecutive primes. However, we adopt a statistical physics perspective. The patterns we show to occur come in the form of power laws in the distribution of prime gaps.

There is already substantial literature on primes adopting the statistical physics perspective. In line with our finding, the histograms in the distribution of gaps between primes divided into "congruence families" are shown to be scale invariant [3]. A theory to explain the origin of the unexpected periodic behavior of gaps between primes has been linked to the *k*-tuple conjecture, the statistical mechanics of spin systems and the Sierpinski fractal [4]. Higher-order gaps between the primes have been analyzed by Fourier decomposition to show patterns in the resulting power spectra [5]. The gaps between primes have been shown to exhibit a period six oscillation [6]. The histogram of the increments (the difference between two consecutive gaps between primes) has been shown to follow an exponential distribution with superposed periodic behavior of period three [7]. Rather than prime gaps, for the distribution of primes themselves there have been claims of self-similarity [8], small world networks [9] and even chaos [10].

## 2. Power Law Pattern

As the authors in Ref. [2], we investigate whether all the prime numbers ending in 1, 3, 7 or 9 are equally likely. First, we transform the sequence of 50,000 numbers of such four endings into a binary sequence, as in **Table 1**.

Then we let $P_n$ be the relative frequency of a number $n$ to be prime. After, we

**Table 1.** Binary transform of the sequence of numbers ending in 1, 3, 7 or 9.

| Number $n$ | Prime? 0 = no, 1 = yes |
|---|---|
| 1 | 0 |
| 3 | 1 |
| 7 | 1 |
| 9 | 0 |
| 11 | 1 |
| 13 | 1 |
| 17 | 1 |
| 19 | 1 |
| 21 | 0 |
| 23 | 1 |
| 27 | 0 |
| 29 | 1 |
| 31 | 1 |
| 33 | 0 |
| 37 | 1 |
| ... | ... |

consider a Bernoulli random variable $X_n$ such that $X_n = 1$ if $n$ is prime and $X_n = 0$ otherwise. Our aim is to find the functional form of the relative frequency $P_n = P(X_n = 1)$. We could observe the occurrences of $X_n = 1$ tend to decrease as $n$ grows. This pattern cannot be easily visualized considering a plot of $X_n$ against $n$ due to the very fact that our series is a sequence of 0 s and 1 s.

To overcome this difficulty, we devised the following: Assume $X_n$ is a nonstationary Bernoulli process, where the probability of success $P_n = P(X_n = 1)$ is not constant. Such a probability can be estimated using a nonlinear Kalman filter [11], which linearizes an estimate of the current mean and covariance. **Figure 1** (top) shows the estimates of $\ln P_n$ versus $\ln \ln n$, and the bottom shows the adjusted straight line for $n > 500$ in red font. It is implied that

$$\ln P_n \approx 1.25 - \ln \ln n. \tag{1}$$

Thus, for $n > 500$ the probability $P_n$ behaves according to a power law that takes the form:

$$P_n \approx \frac{3.5}{\ln n}. \tag{2}$$

This power law can be translated in terms of prime gaps, as in **Table 2**. First note that the prime gap distribution is approximately geometric.

Then, let $Y_n$ be a geometric random variable with probability of success $P_n$. Tentatively consider that "success" is the occurrence of a prime number $X_n = 1$ and that the

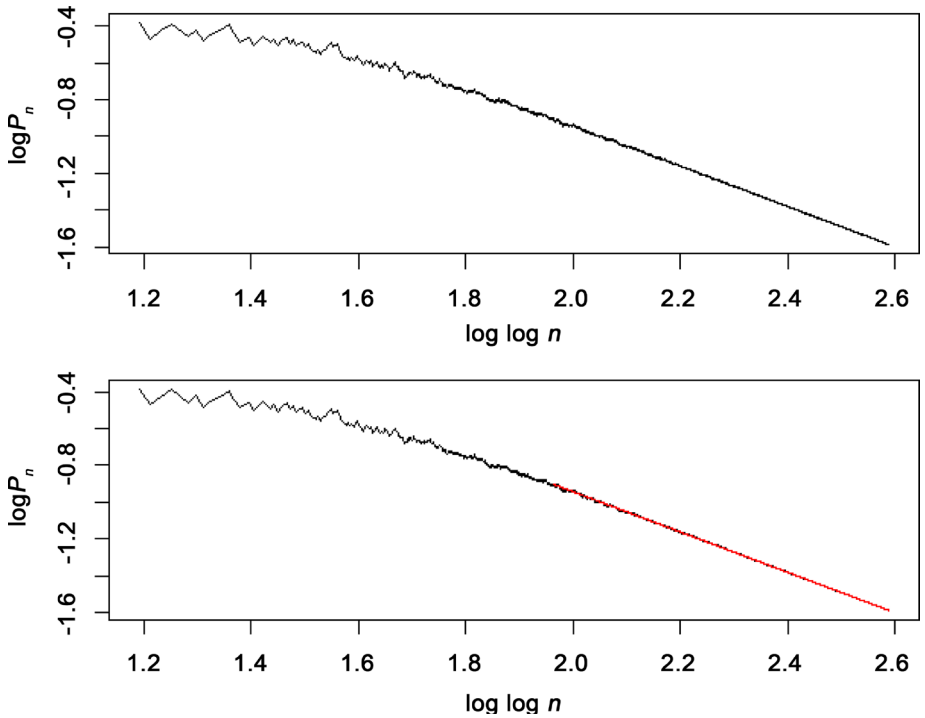


**Figure 1.** Estimations using a nonlinear Kalman filter show the chances of a number to be prime follows a power law (red font).

**Table 2.** Prime numbers and prime gaps.

| Prime number | Prime gap |
|---|---|
| 2 | 2 |
| 3 | 1 |
| 5 | 2 |
| 7 | 2 |
| 11 | 4 |
| 13 | 2 |
| 17 | 4 |
| 19 | 2 |
| 23 | 4 |
| 29 | 6 |
| ... | ... |

successive Bernoulli trials are independent. Thus, the expected number of trials until the occurrence of a subsequent prime is given by:

$$\tau_n = E\left[Y_n\right] = \frac{1}{p_n}. \tag{3}$$

However, because our Bernoulli process is nonstationary and the Bernoulli trials are not independent, one cannot expect an analytic solution such as that provided by Equation (3) to hold true. But one can empirically determine an analogous substitute for

Equation (3) by considering estimates of $\tau_n$ on the basis of the prime gaps shown in Table 2.

Using the nonlinear Kalman filter for this geometric process, Figure 2 shows the dispersion of estimates of $\tau_n$ (considering a geometric process) and estimates of $1/P_n$ (considering a Bernoulli process). As can be seen, $\tau_n \sim \frac{1}{P_n}$ as expected for a typical geometric process. However, the adjusted straight line (in red font) presents the form:

$$\tau_n \approx 0.2876\left(1+\frac{1}{P_n}\right). \tag{4}$$

Therefore, prime gaps are proportional to the inverse of the chance of a number to be prime.

Alternatively, consideration of the power law in Equation (2) yields:

$$\tau_n \approx 0.2876\left(1+\frac{\ln n}{3.5}\right). \tag{5}$$

## 3. Conclusion

Although the difference between two successive prime numbers is casually considered random, the *k*-tuple conjecture casts doubt on that. The *k*-tuple conjecture is yet to be

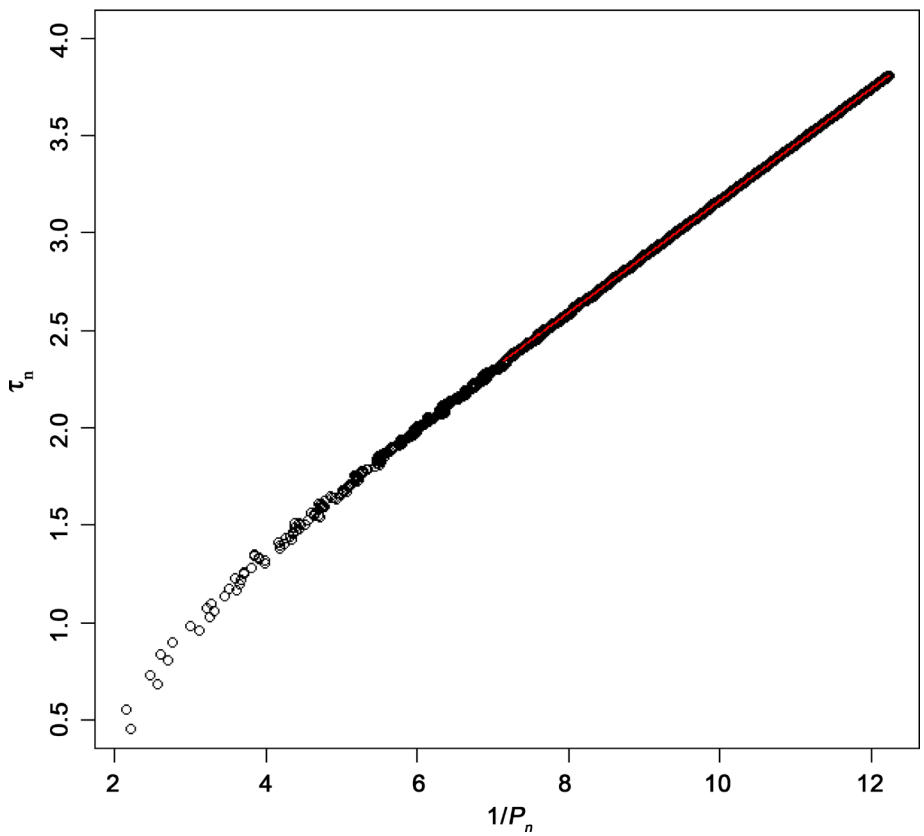


Figure 2. Estimations using a nonlinear Kalman filter show prime gaps follow a power law (red font).

proven, but finding unexpected biases in the distribution of consecutive primes provides confirmation of the *k*-tuple conjecture. Motivated by a recent mathematical study [2], we explain departures from randomness in prime gaps from a statistical physics perspective. The pattern we find that challenges the hypothesis of the genuine random behavior of prime gaps comes in the form of a power law. We find that prime gaps are proportional to the inverse of the chance of a number to be prime.

## References

[1] Hardy, G.H. and Littlewood, J.E. (1923) Some Problems of "Partitio Numerorum". III. On the Expression of a Number as a Sum of Primes. *Acta Mathematica*, **44**, 1-70. http://dx.doi.org/10.1007/BF02403921

[2] Lemke Oliver, R.J. and Soundararajan, K. (2016) Unexpected Biases in the Distribution of Consecutive Primes. http://arxiv.org/abs/1603.03720

[3] Dahmen, S.R., Prado, S.D. and Stuermer-Daitx, T. (2001) Similarity in the Statistics of Prime Numbers. *Physica A*, **296**, 523-528. http://dx.doi.org/10.1016/S0378-4371(01)00183-2

[4] Ares, S. and Castro, M. (2006) Hidden Structure in the Randomness of the Prime Number Sequence? *Physica A*, **360**, 285-296. http://dx.doi.org/10.1016/j.physa.2005.06.066

[5] Szpiro, G.G. (2007) Peaks and Gaps: Spectral Analysis of the Intervals between Prime Numbers. *Physica A*, **384**, 291-296. http://dx.doi.org/10.1016/j.physa.2007.05.038

[6] Wolf, M. (1999) Applications of Statistical Mechanics in Number Theory. *Physica A*, **274**, 149-157. http://dx.doi.org/10.1016/S0378-4371(99)00318-0

[7] Kumar, P., Ivanov, P.C. and Stanley, H.E. (2003) Information Entropy and Correlations in Prime Numbers. arXiv: cond-mat/0303110.

[8] Wolf, M. (1997) 1/f Noise in the Distribution of Prime Numbers. *Physica A*, **241**, 493-499. http://dx.doi.org/10.1016/S0378-4371(97)00251-3

[9] Chandra, A.K. and Dasgupta, S. (2005) A Small World Network of Prime Numbers. *Physica A*, **357**, 436-446. http://dx.doi.org/10.1016/j.physa.2005.02.089

[10] Gamba, Z., Hernando, J. and Romanelli, L. (1990) Are Prime Numbers Regularly Ordered? *Physics Letters A*, **145**, 106-108. http://dx.doi.org/10.1016/0375-9601(90)90200-8

[11] Harvey, A.C. (1989) Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.