Notes on Number Theory and Discrete Mathematics
ISSN 1310-5132
Vol. 21, 2015, No. 1, 18–30

Distribution of prime numbers by the modified chi-square function

Daniele Lattanzi

Former ENEA-FUS, Frascati, Roma, Italy
Private address: Via La Spezia 81 – 00182 Roma, Italy
e-mail: lattanzio.lattanzi@alice.it

Abstract: The statistical distribution of prime numbers represents an open problem in number theory still nowadays. The methodology of experimental mathematics has not yet been attempted in this field, thus the present report treats prime numbers as raw experimental data and as elements of larger and larger finite sequences $\{P_m\}$. The *modified chi-square function* $\mathcal{X}^2{}_k(A, x/\mu)$ with the ad-hoc A, k and $\mu = \mu(k)$ parameters is the best-fit function of the differential distribution functions of both prime finite sequences $\{P_m\}$ and truncated progressions $\{n^\alpha\}$ with $\alpha \in (1, 2)$ so that an injective map can be set between them through the parameter k of their common fit function $\mathcal{X}^2{}_k(A, x/\mu)$ showing that the property of scale invariance does not hold for prime distribution. The histograms of prime gaps, which are best fitted by standard statistical distribution functions, show unexpected clustering effects.

Keywords: Prime distribution, Prime sequences, Modified chi-square function, Progressions.
AMS Classification: 11N13, 11N05, 11B25.

1 Introduction

The statistical distribution of prime numbers has always been a challenge in number theory and still nowadays it remains one of the major open problems notwithstanding the many theoretical successes achieved both historically [1–4] and in recent years [5–11].

Thus, in the attempt to answer to the questions on what is the actual distribution of prime numbers as experimentally assessed and whether and how it can be approximated, in the present report an innovative approach is suggested starting from the experimental viewpoint [12–16] and using the *modified chi-square function* $\mathcal{X}^2{}_k[A, x/\mu] = \mathcal{X}^2{}_k\ [A,\ x/\mu(k)]$ with $k \in (1, 2)$:

$$\mathcal{X}^2{}_k\left[A, \frac{x}{\mu}\right] = \left(\frac{A}{2 \cdot \Gamma(\frac{k}{2})}\right) \cdot \left[\frac{x}{2 \cdot \mu(k)}\right]^{\frac{k}{2}-1} \cdot e^{-\left[\frac{x}{2 \cdot \mu(k)}\right]} \tag{1}$$

as the best-fit function along the whole study to match the statistical distributions of prime numbers and progressions $\{n^{\alpha}\}$ having domain $\mathbb{N}$ and co-domain $\mathbb{R}^+$ with $\alpha \in (1, 2) \subset \mathbb{R}^+$. The rationale underlying the entire issue is to use this function taking advantage of the adjustment of its three parameters k, A and $\mu = \mu(k)$, which allow to optimize the fits as much as possible, up to 99% and even more whenever possible. In other words a *plot & fit* algorithm is set up. The modified chi-square function with k degrees of freedom is a new general form of the standard chi-square function [17, 18] used also in statistics [19, 20] where the two new parameters A and μ have been used and $\Gamma = \Gamma(k/2) = \Gamma_{k/2}$ is the standard gamma function necessary to the normalization. It is easy to identify $\mu(k)$ as a *decay parameter*, for which the limit holds:

$$lim_{k\to 2}\,\mu(k) = +\infty$$

so that the further limit holds:

$$lim_{k\to 2}\mathcal{X}^2_k[A, x/\mu(k)] = \text{ const.} = \langle \mathcal{X}^2_k \rangle$$

hence marking a basic difference with the standard chi-square function for which:

$$lim_{k\to 2}\mathcal{X}^2_k(x) = \mathcal{X}^2_2(x) = \tfrac{1}{2}\cdot e^{-x/2}$$

Apart from the usual improvements of the statistical values, for any truncated progression $\{n^{\alpha}\}$ the results in no way depend on the number of its terms, what is a consequence of the scale invariance of the progressions themselves, whereas for prime numbers just larger and larger finite sequences, subsets of their whole infinite sequence, have been examined, that is sequences of the kind: $\{2 \;\; 3 \;\; 5 \;\; 7 \;\; 11 \;\ldots\; P_{h-1} \; P_h\} \equiv \{P_h\} \subset \{P_i\} \subset \{P_j\} \subset \ldots \subset \{P_n\} \subset \ldots$ being of course $h < i < j < \ldots < n < \ldots$. The reason for such an unconventional choice is a strict consequence of the scale non-invariance of prime distribution and of scaling laws holding for them, as shown later on.

2 Statistical treatments

The statistical distribution of the terms of truncated numerical progressions $\{n^{\alpha}\}$ with $\alpha \in (1, 2)$ and of finite sequences of prime numbers $\{P_m\}$ are examined and reported as well as of prime finite differences $\{\Delta^h P_m\}$ and $\{\delta^j P_m\}$. The accuracies, error sources, error propagations and reliability of the results have been investigated too, being these issues crucial to the whole algorithm. After all, what has been done is just what is usually done in treating experimental raw data, a procedure that is common to all fields of experimental physics. The only difference is to treat numbers just as experimental data in a broad sense, to which all these concepts and criteria can be applied, with the further undisputable advantage of having zero inaccuracy (i.e. no systematic errors) and zero imprecision (no random errors) on the base data and zero inaccuracy though not zero imprecision (owing to the approximations) on the final results.

In the fit performed between the two differential distribution functions (DDFs) - i.e. the experimental or *sample* one given by the actual counts of $\{n^{\alpha}\}$ or $\{P_m\}$ and the theoretical or *parent* one given by $\mathcal{X}^2_k[A, x/\mu]$ - the first issue has been to get the correct value of $k \in (1, 2)$ (usually up to the third decimal digit that is with a precision $\delta k/k \approx 0.5$ ‰ and in some cases even up to the fourth decimal digit) by means of a *trial and error* procedure, letting k vary and, at any value of it, balancing the average value $\langle C\rangle$ of the counts to the average value $\langle F\rangle$ of the fit function, up to the 12th decimal digit, by acting on the value of the decay parameter μ thus sweeping the plane (k, μ) within the ranges $k \in (1, 2^{-})$ and $\mu < 1E308 = 1\cdot10^{308}$ (the upper numerical limit of the software). At the same time both the *Bravais–Pearson* correlation coefficient $R = R(C, F) < 1$ and the non-linear index of correlation $I = I(C, F) < 1$ between the two DDFs of the counts C and of the fit F have been calculated and maximized during the variations of k and μ. In addition, even the two standard deviations of the means σ_C and σ_F have been examined in order to ascertain that each of them would be much lower than its respective mean $\langle C\rangle$ and $\langle F\rangle$. Finally, two further gauges of the fit have been minimizing the value of the *least square sum* (LSS) according to the principle of maximum likelihood and the value of the *chi-square test* ($\mathcal{X}^2_{test}$), in that both of these variables measure the goodness of the fit.

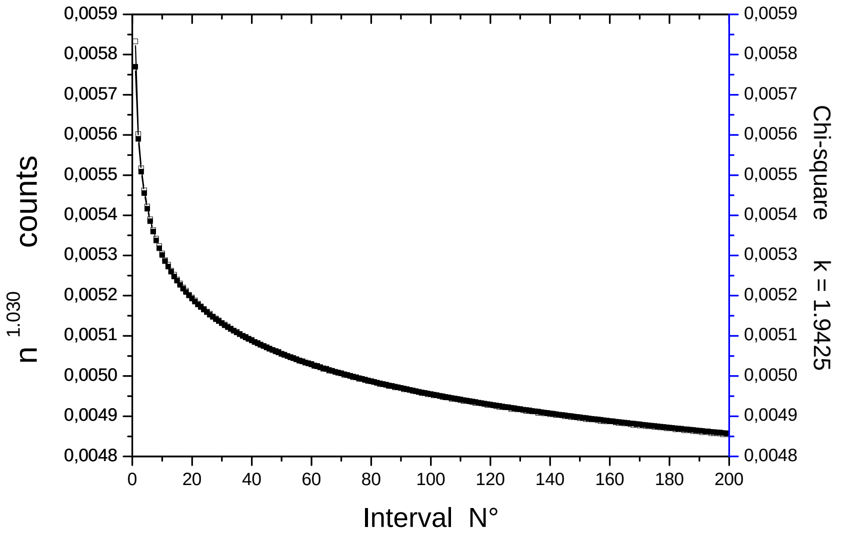
In addition, for any progression and prime sequence even the pertaining cumulative distribution function (CDF = $\sum$ DDF) has been examined and compared to the connected CDF of the $\mathcal{X}^2_k[A, x/\mu(k)]$ fit function with the ad hoc parameters, finding that the matching of the two cumulative distributions is much better than the two DDF pairs increasing from values R, $I \approx 0.99\ldots$ up to values $R, I \approx 0.999999\ldots$ and beyond.

2.1 Truncated numerical progressions

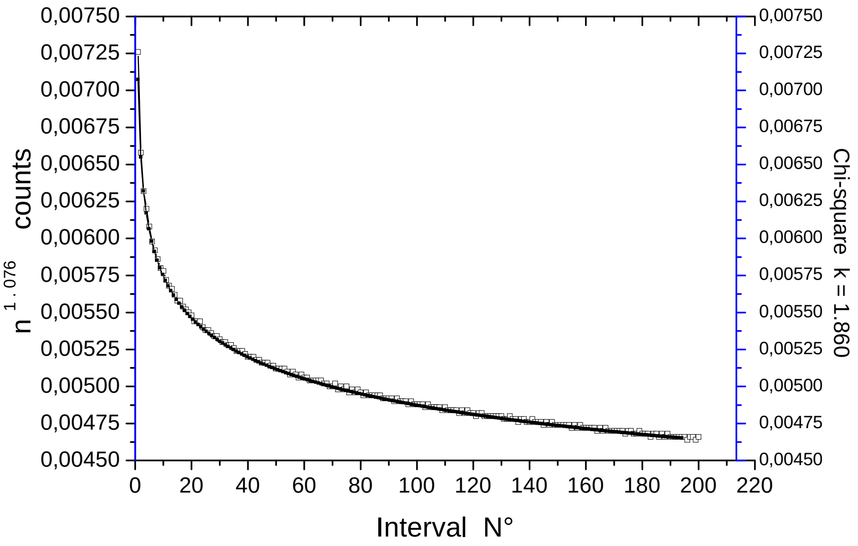
The first problem to cope with is the statistical treatment of the terms of truncated positive ascending numerical progressions $\{n^\alpha\}$ with the index $n \in \mathbb{N}$ running from 1 thru $m\approx10^5$ and the exponent $\alpha \in [1, 2] \subset \mathbb{R}^+$. Thus for any $\{n^\alpha\}$ progression the statistical distribution of a finite number of its terms has been examined getting the histogram that is the plot of its differential distribution function for n_Δ intervals (typically 200) that has been fitted, at n_Δ data points, by the modified chi-square function with the appropriate values of the parameters k, A and $\mu = \mu(k)$. Speaking in a strict and formal way any $\{n^\alpha\}$ progression can be analytically continued to the function $f(x) = x^\alpha$ and also to the function $g(x) = \mathcal{X}^2_k[A, x/\mu(k)]$; both functions are analytic on the whole $\mathbb{R}^+$ plane. In other words the $\mathcal{X}^2_k[A, x/\mu(k)]$ function is an interpolating function of the $\{n^\alpha\}$ progressions for non-integer $n \in \mathbb{R}^+$, within the chosen accuracy, just like the x^α function.

The Figure 1(a) shows the example of the DDF of the progression $\{n^{1.030}\}$ with 800K terms and the relative fit of the function $\mathcal{X}^2_k[A, x/\mu]$ with $k = 1.9425$ ($\Gamma_{k/2} = 1.01743472$) $n_\Delta = 200$ the fit parameters being: $R = 0.999715$ $I = 0.998889$ LSS = 0.11646 $\mathcal{X}^2_{\text{test}} = 8.76697\text{E}{-7}$ while $\langle C\rangle = \langle F\rangle = 0.500000000000$ (i.e. balanced up to the 12th decimal digit) $\sigma_C = 0.000150$ $\sigma_F = 0.000146$ $\mu = 2.1101944\text{E}{+61}$ with $A = 200/800{,}000 = 0.00025$.

The Figure 1(b) shows the example of the DDF of the progression $\{n^{1.076}\}$ with 50K terms and the relative fit of the function $\mathcal{X}^2_k[A, x/\mu]$ with $k = 1.860$ ($\Gamma_{k/2} = 1.04558808$) $n_\Delta = 200$ the fit parameters being: $R = 0.999346$ $I = 0.998199$ $LSS = 0.18856$ $\mathcal{X}^2_{\text{test}} = 8.26567\text{E}{-6}$ whilst $\langle C\rangle = \langle F\rangle = 0.500000000000$ $\sigma_C = 0.000379$ $\sigma_F = 0.000370$ $\mu = 1.85117918\text{E}{+10}$ with $A = 0.004$. The very good superposition of the data points and the fit curve is evident in both plots. Noteworthy, the value of k does not depend on the number of terms of the progression, unlike the value of μ.



(a)



(b)

Figure 1. Histograms of the progressions $\{n^{1.030}\}$ (a) and $\{n^{1.076}\}$ (b) and their fits

In such a way, the values of k and α have been found for all the 60 $\{n^{\alpha}\}$ progressions examined, thus setting up the existence of a one-to-one correspondence or bi-injection map between them, inside the closed ranges $\alpha, k \in [1, 2]$, that is:

$$\alpha \Leftrightarrow k_\alpha \qquad \text{or likewise} \qquad k_\alpha = k(\alpha) \;\text{ and }\; \alpha = \alpha(k).$$

The related plot of the function $k_\alpha = k(\alpha)$ is shown in Figure 2(a). The best-fit function inside the whole closed ranges $\alpha, k \in [1, 2]$ with the values $\mathcal{X}^2_{\text{test}} = 1.2697\text{E}{-5}$ and $R^2 = 0.9998$ is:

$$k_\alpha = k(\alpha) \approx (0.737 \pm 0.007) + (6.24 \pm 0.07)\cdot \mathrm{e}^{-\alpha/(0.627 \pm 0.007)}.$$

Of course in the inner range $\alpha \in [1, 1.125]$ and $k \in [2, 1.78]$ the exponential decay law can be approximated by the linear fit:

$$k_\alpha = k(\alpha) \approx (3.87 \pm 0.02) - (1.87 \pm 0.02)\cdot\alpha \tag{2}$$

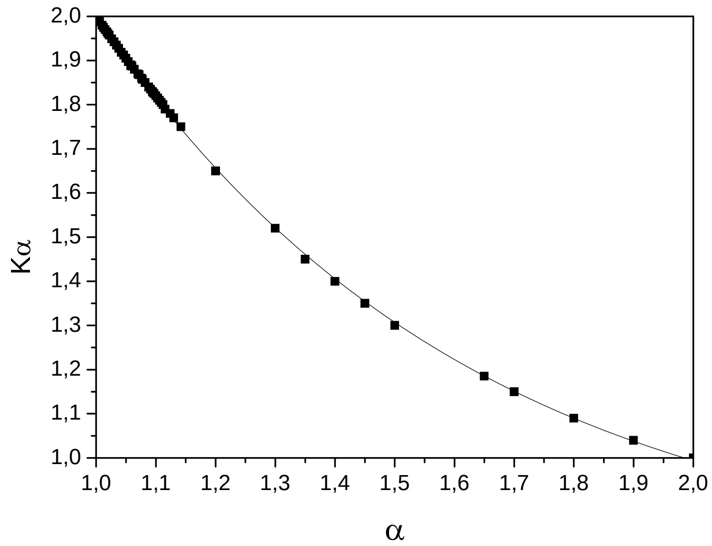
with a correlation coefficient value $R = 0.9995$ a standard deviation $\sigma = 0.0022$ and a probability $p < 0.0001$ that a data point may fall off the average within a distance $> \sigma$.

The Figure 2(b) shows the data points inside this range with their linear fit. In addition, near the point $(\alpha, k) \equiv (1^+, 2^-)$ i.e. $\alpha \in [1.04, 1]$ and $k \in [1.92, 2]$ the plain linear relationship $k_\alpha \approx 2\cdot(2-\alpha)$ holds, too.

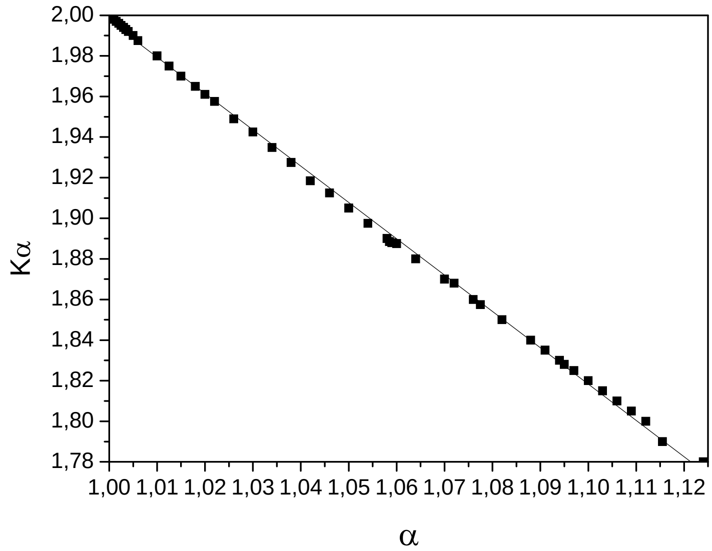
Just at this point the modified chi-square function reduces to a simple constant function owing to the asymptotic trend of the decay parameter

$$lim_{k\to 2^-}\mu(k) = lim_{\alpha\to 1^+}\mu(\alpha) = +\infty$$

as already said and as the author has thoroughly checked, at least for very high values of μ up to ~1E308.



(a)



(b)

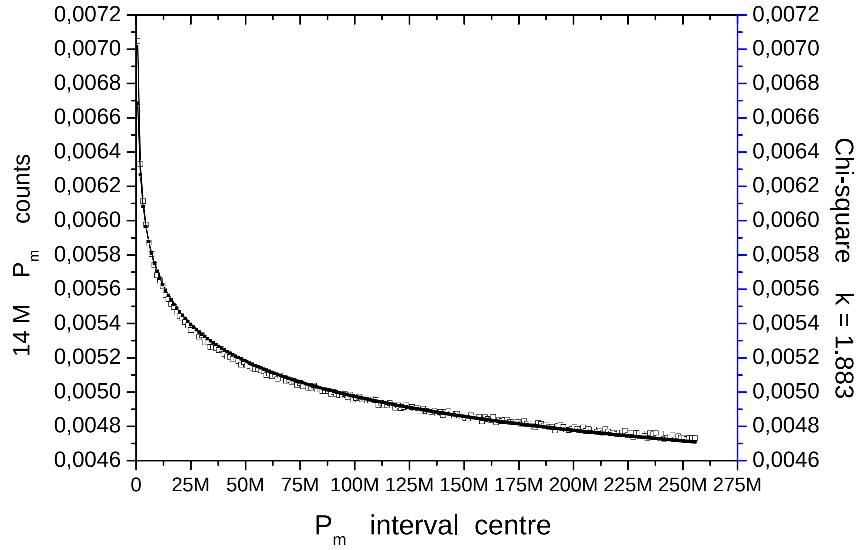
Figure 2. The relationships $k_\alpha = k(\alpha)$ with $\alpha, k \in [1, 2]$ (a) and $\alpha \in (1, 1.125)$ $k \in (1.78, 2)$ (b)

2.2 Finite sequences of prime numbers

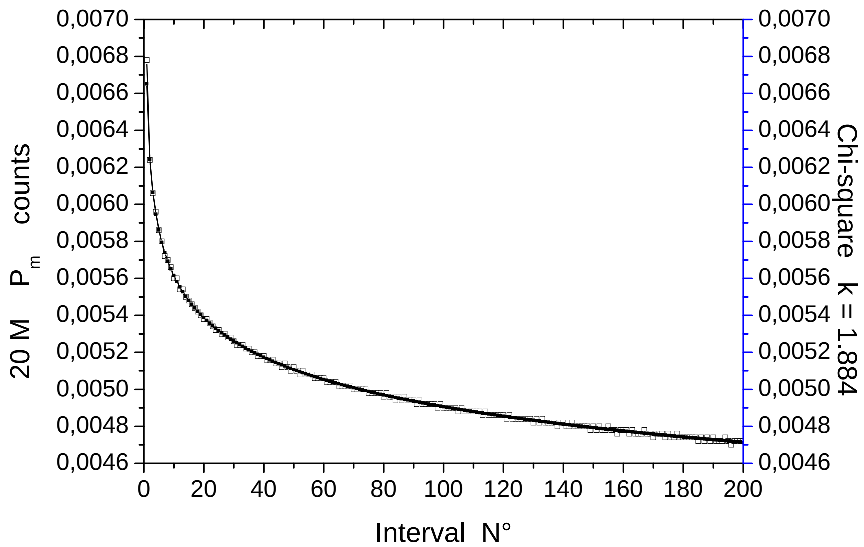
Passing to the statistics of prime numbers that is of the terms of larger and larger finite sequences of primes $\{P_m\}$ (up to $m_p = 5{\cdot}10^7 = 5E7 = 50M \rightarrow P_{50M} = 982{,}451{,}653$ [21]) it is easy to verify the scale non-invariance of their distribution.

The Figure 3(a) shows the example of the first $m_p = 1.4E7 = 14M$ prime numbers $\{P_{14M}\} \equiv \{256, 203, 161\}$ with the values of the fit by $\mathcal{X}^2_k[A, x/\mu]$: $A = n_\Delta/m_p = 200/14M$ $k = 1.883$ $\Gamma_{k/2} = 1.0373458$ $\mu = 3.693939E{+}56$ $\langle C\rangle = \langle F\rangle = 0.005000000000$ $\sigma_c = 0.000311$ $\sigma_F = 0.000306$ $R = 0.995286$ $I = 0.990477$.

The Figure 3(b) shows the sequence of the first 20M primes $\{P_{20M}\} \equiv \{373, 587, 883\}$ fitted by $\mathcal{X}^2_k[A, x/\mu]$: $A = n_\Delta/m_p = 200/20M$ $k = 1.8835$ $\Gamma_{k/2} = 1.037170175$ $\langle C\rangle = \langle F\rangle = 0.00500...$ $\sigma_c = 0.000303$ $\sigma_F = 0.000302$ $R = 0.99572$ $I = 0.99140$ $\mu = 3.97421857E{+}59$.



(a)

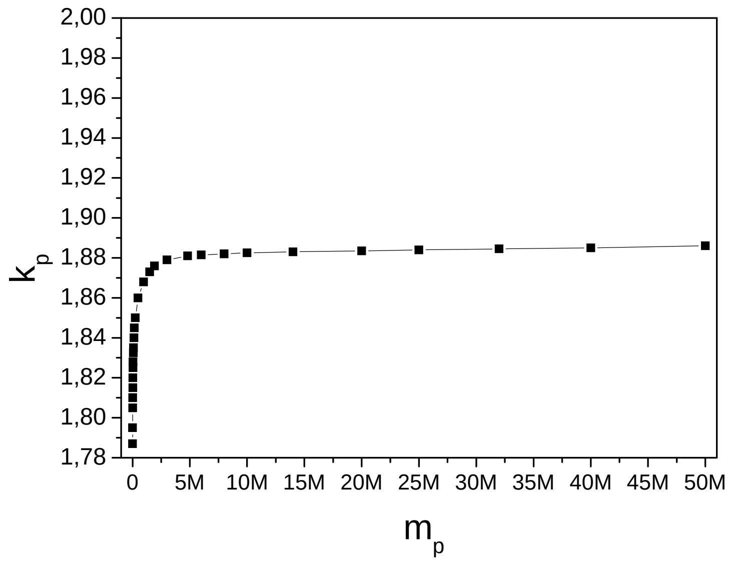


(b)

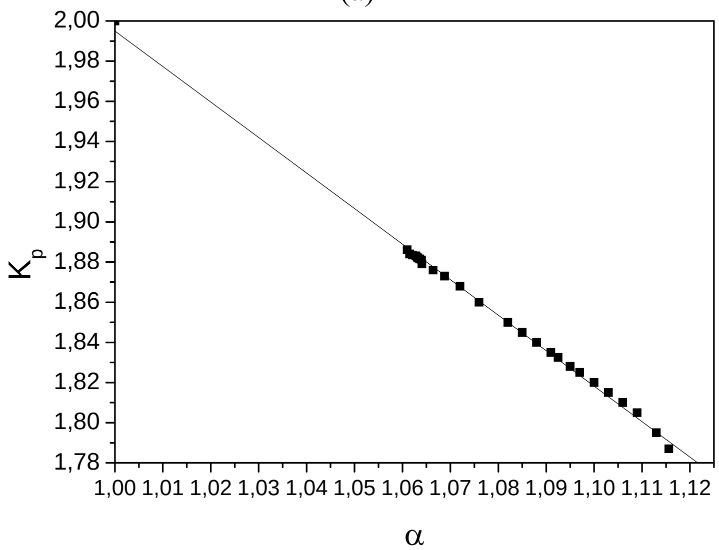
Figure 3. Histograms of the prime sequences $\{P_{14M}\}$ (a) and $\{P_{20M}\}$ (b) with their fits

Thus for any sequence $\{P_m\}$ the related value of $k_p = k(P_m) = k(m_p)$ has been calculated.

The Figure 4(a) illustrates the trend of $k_p = k(m_p)$ for the 30 cases examined of $m_p \in$ [3K– 50M] that is $P_m \in$ [27, 449–982, 451, 653] and $k \in$ [1.787–1.886] clearly showing that, from the statistical viewpoint, prime numbers have not the property of scale invariance.



(a)



(b)

Figure 4. The relationships $k_p = k(m_p)$ (a)
and $k_p = k(\alpha)$ (b) for the finite sequences $\{P_m\}$

It is plain to correlate the two relationships $k_\alpha = k(\alpha)$ and $k_p = k(m_p)$ one each other via their common parameter k. As a matter of fact the author has actually verified that any histogram of $\{P_m\}$ can be fitted by the corresponding (via k) histogram of $\{n^\alpha\}$ with the same k value at a confidence level R and $I > 99\%$. Thus there is a correspondence or injective mapping between the distribution of the prime finite sequences $\{P_m\}$ and that of the finite progressions $\{n^\alpha\}$ $\alpha \in (1, 2)$ by means of their common parameter $k = k_\alpha = k_p$ namely:

$$\{P_m\} === [\mathcal{X}^2_k] === \rightarrow k_p = k = k_\alpha \leftarrow === [\mathcal{X}^2_k] === \rightarrow \{n^\alpha\}$$

as well as a one-to-one correspondence or bi-injective map between the two parameters k_p and α as shown in Figure 4b so that:

$$k_p = k_p(\alpha) \approx (3.81 \pm 0.03) - (1.81 \pm 0.03)\cdot\alpha \tag{3}$$

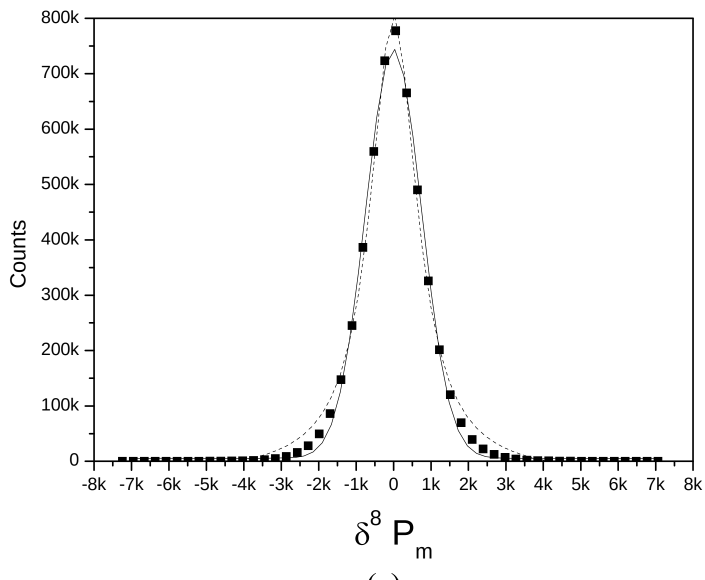
with $R = 0.99905$ $\sigma = 0.0019$ and $p < 0.0001$. The good matching between (2) and (3) is clear, though it is worthy and necessary to investigate the range $1 < \alpha < 1.061$ corresponding to $1.886 < k_p < 2$ that is $m_p > 50M$ and $m_p >> 50M$. The asymptotic trend of $k_p = k(m_p)$ can be immediately inferred from the Figures 4(a) and 4(b) as $lim_{m_p \to +\infty} k_p(m_p) = lim_{P_m \to +\infty} k_p(P_m) = = lim_{\alpha \to +1^+} k_p(\alpha) = 2^-$ and $lim_{m_p \to +\infty}[\Delta k(m_p)/\Delta m_p] = 0^+$. As for the decay parameter $\mu = \mu(k)$ it is easy to check that the limits hold: $lim_{k \to 2^-} \mu(k) = +\infty$ and $lim_{m_p \to \infty} (m_p/\mu) = 0^+$.

The relationship $\alpha = \alpha(m_p)$ has been examined too, the analogue $\alpha = \alpha(P_m)$ being similar at all though shifted ahead, finding that it is represented by a *logistic* curve with all the parameters depending on m_p, owing to the usual scale non-invariance, and that the limits hold: $\lim_{m_p \to +\infty} \alpha(m_p) = \lim_{P_m \to +\infty} \alpha(P_m) = 1^+$.

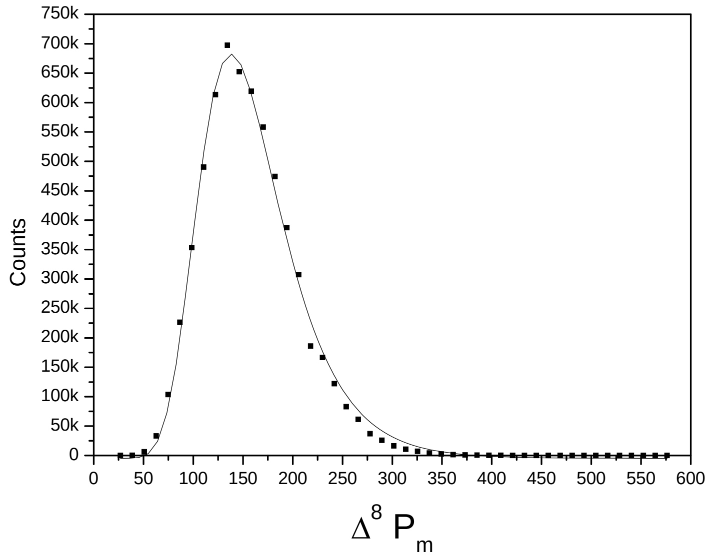
2.3 Distribution of prime gaps

In treating the distribution of prime numbers from the experimental standpoint one cannot miss to study the statistics of their differences or *gaps*.

The differences in the form $\delta^j P_m = \sum_{i\,=\,1 thru j} (-1)^i \cdot (j \;\; i) \cdot P_{m-i}$ with $j \geq 2$ and $(j \;\; i)$ the binomial coefficients follow the *pseudo-Voigt* DDF ${}_pV(x) = f \cdot G(x) + (1 - f) \cdot L(x)$ $(0 \leq f \leq 1)$ that is a simple linear combination of *Gauss* $G(x)$ and *Lorentz* $L(x)$ distribution functions with the parameter f giving the weight of each one as in Figure 5(a) in which the example of the histogram of the $\delta^8 P_m$ for 5M and $n_\Delta = 50$ with Gauss fit at $R^2 = 0.99625$ (full line ——) and Lorentz fit at $R^2 = 0.99289$ (dashed line ------) is reported showing that the *peak* of the DDF is best fitted by the Lorentzian distribution function (------), the *wings* (or *tails*) are approximately half-way between the $G(\delta^j)$ and the $L(\delta^j)$ functions. The linear gaps $\Delta^h P_m = P_m - P_{m-h}$ $(h \geq 2)$ follow a statistical distribution of the type *ExpExp* also called *extreme* function [22] having the form $E(x) = E_o + A \cdot e^{-\exp(-y) - y + 1}$ with $y = (x - x_o)/w$ being A the amplitude, E_o the baseline value, x_o the center of the DDF, $E(x_o) = E_o + A$ the top value and w the width as in Figure 5(b) in which the histogram of $\Delta^8 P_{50M}$ for $n_\Delta = 50$ with the extreme fit at $R^2 = 0.99588$ is shown.



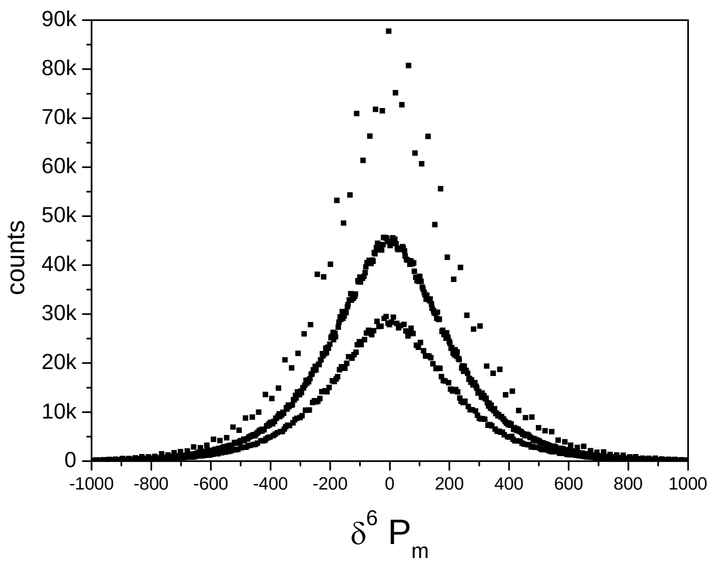
(a)



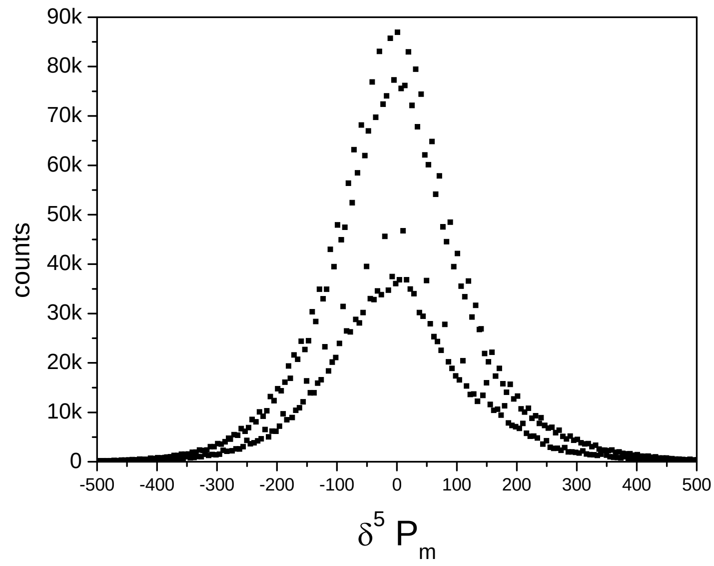
(b)

Figure 5. Distributions of the prime gaps $\delta^8 P_{5M}$ (a) and $\Delta^8 P_{50M}$ (b), both with $n_\Delta = 50$

The mixing of the two DDFs (Gauss and Lorentz) for the $\delta^j P_m$ into a pseudo-Voigt DDF suggests [23] and references therein, [24] the arising of inner structures of $\delta^j P_m$ in increasing the value of n_Δ, the number of the histogram intervals, that is of innermost groups of $\delta^j P_m$. This has been thoroughly checked and shown in the two examples of Figure 6a holding for m_p = 10M→P_{10M} = 179, 424, 673 n_Δ = 2,000 $\delta^6 P_m$ and of Figure 6(b) for m_p = 5M→P_{5M} = 86, 028, 121 n_Δ = 750 $\delta^5 P_m$ which highlight the existence of five or six concentric distributions with the same mean value $\langle\delta^6 P_m\rangle = 0$ and $\langle\delta^5 P_m\rangle = 0$ though with different maximum values and different standard deviations σ.



(a)



(b)

Figure 6. Distributions of the prime gaps $\delta^6 P_{10M}$ $n_\Delta = 2{,}000$ (a) and $\delta^5 P_{5M}$ $n_\Delta = 750$ (b)

In Figure 6(a), the symmetry of the resulting DDFs for $\delta^6 P_m$ is manifest, unlike the *skewness* of the case of $\delta^5 P_m$, just a consequence of the binomial coefficients for the even and odd cases of j respectively.

The same effect of data point clustering is visible in the gaps of the kind $\Delta^h P_m$, though not shown, and even the first order gaps of prime numbers (that is $\delta P_m = \Delta P_m = P_m – P_{m-1}$, i.e. $j = h = 1$) display, despite their different plots, the same effect of data clustering with increasing values of n_Δ. In these cases too the spreads observed in the data points are not to be ascribed to statistical reasons, though just to the aforesaid data point gathering into innermost structures.

At the present stage of the research the distribution of prime numbers themselves does not show this effect of data point bunching into a multitude of similar curves with increasing n_Δ; such an effect is just a feature of the distribution of prime gaps.

3 Conclusions

The new algorithm presented in this report, that treats the statistical distributions of truncated progressions $\{n^\alpha\}$ with $\alpha \in (1, 2)$ and of finite sequences of prime numbers $\{P_m\}$ from the experimental viewpoint by the modified chi-square function, represents an innovative approach leading to remarkable findings which, at this early stage, are:

- The statistical distributions of both truncated progressions $\{n^\alpha\}$ with $\alpha \in (1, 2) \subset \mathbb{R}^+$ and prime number finite sequences $\{P_m\}$ can be fitted by the modified chi-square function $\mathcal{X}^2_k[A, x/\mu]$ with $k \in (1, 2) \subset \mathbb{R}^+$;
- There is a correspondence between the two above said distributions of $\{P_m\}$ and of $\{n^\alpha\}$;

- The prime number distribution has not the property of scale invariance, holding the scale non-invariance relationship $k_p = k(m_p)$ of the modified chi-square function. Thus it has no meaning, in examining the distribution of primes, to study just one single sequence of them, however great or huge it might be, in that just many greater and greater prime sequences must be examined to reach significant conclusions;
- Owing to the above told correspondence, it can be inferred that the infinite sequence of primes $\{P_\infty\}$ has a uniform statistical distribution in the same way as $\{n^1\} \equiv \{n\}$, i.e. the succession of natural numbers;
- The distributions of prime gaps of any kind and order show unexpected clustering effects.

However much more is still to be done in the matter, first of all as for the enlargement of the prime data base (m_p >> 50M) and the improvement of the accuracy of the calculations: to reach these goals an ad-hoc computer code running on a mainframe can give a valuable aid.

Thanks and acknowledgements

The author wishes to thank Prof. D. Levi of the University *Roma-Tre*, Roma, I., and Prof. G. Dattoli, Director of the Unit *Theoretical Physics and Applied Mathematics* of ENEA, Frascati, Roma, I., for having read the manuscript and given valuable comments.

References

[1] Smith, D. E. (1929, 1959) *A Source Book in Mathematics*. Dover Publications, Inc. New York, 1959, unabridged republication of the 1rst edition, originally published in 1929 by McGraw-Hill Book Co., Inc.

[2] Du Sautoy, M. (2003, 2004) *L'enigma dei numeri primi*. 2004 RCS Libri S.p.A. Milano, ISBN 978-8817-05022-7. *The Music of the Primes* 2003, Marcus du Sautoy.

[3] Derbyshire, J. (2003, 2006) *L'ossessione dei numeri primi – Bernhard Riemann e il principale problema irrisolto della matematica*. 2006 Bollati Boringhieri Editore s.r.l. Torino. *Prime Obsession. Bernhard Riemann and the Greatest Unsolved Problem in Mathematics* Joseph Henry Press (National Academic Press), Washington D.C., 2003, John Derbyshire.

[4] Languasco, A., & Zaccagnini, A. (2005) *Alcune proprietà dei numeri primi, I* e *II* Sito web Bocconi-Pristem, `http://matematica.uni-bocconi.it/LangZac/home.htm`

[5] Goldstone, D. A., Pintz, J., & Yildirim, C.Y. (2007) The path to recent progress on small gaps between primes, *Clay Mathematics Proceedings*, 7, 125–135.

[6] Granville, A. (2008) Prime Number Patterns, *The American Mathematical Monthly*, 115(4), 279–296.

[7] Granville, A., & Martin, G. (2006) Prime Number Races, *The American Mathematical Monthly*, 113(1), 1–33.

[8] Ball, P. (2003) Prime numbers not so random?, *Nature* 24 March 2003 doi: 10.1038/news030317-13.

[9] Granville, A., Harald Cramér and the Distribution of Prime Numbers, based on a lecture presented on 24th September 1993 at the Cramér Symposium in Stockholm. https://www.dartmouth.edu/~chance/chance_news/for_chance_news/Riemann/cramer.pdf.

[10] Holdom, H. (2009) Scale-invariant correlations and the distribution of prime numbers, *J. Phys. A, Math. Theor*. 42, Art. ID 345102 (10 pp.).

[11] Jeong, S., Lee, G. & Kim, G. (2013) Statistical and structural analysis of the appearance of prime Numbers, *J. Appl. Math. Comp*. 41, 283–299.

[12] Bailey, D. H., & Borwein, J.M. (2011) Exploratory Experimentation and Computation. *Notices of the AMS*, 58(10), 1410–1419. http://escholarsship.org/uc/item/26m6noob

[13] Andeberhan, T., Medina, L.A. & Moll, V.H. (Eds) (2009) *Contemporary Mathematics – 517 – Gems in Experimental Mathematics*. AMS Special Session, Experimental Mathematics, January 5, 2009, Washington D.C.

[14] E Silva, T. O. *Goldbach Conjecture verification*, Dep. De Electronica, Telecomuniçaões e Informatica, Universidade de Aveiro, Portugal. http://www.ieeta.pt/~tos/primes.html

[15] Sørensen, H. K. *Exploratory experimentation in experimental mathematics: A glimpse at the PSLQ algorithm.* Institut for Videnskabsstudier, Aarhus Universitet, 8000 Ärhus C, Denmark.

[16] Borwein, J. M. (2009) *Exploratory Experimentation: Digitally-Assisted Discovery and Proof*, University of Newcastle, Australia.

[17] Mathews, J. & Walker, R. L. (1964) Caltech *Mathematical Methods of Physics*. W.A. Benjamin Inc. New York N.Y.

[18] Babusci, D., Dattoli, G., Del Franco, M. (2010) *Lectures on Mathematical Methods for Physics* RT/2010/58/ENEA, ENEA-Roma-I.

[19] Morice, E. (1967, 1971) *Dizionario di statistica*. ISEDI. Milano, 1971. *Dictionnaire de Statistique* Dunod, Paris, 1967.

[20] Young, H. D. (1962) *Statistical Treatment of Experimental Data.* Carnegie Institute of Technology. McGraw-Hill Book Company, Inc.

[21] Caldwell, C., K., The first fifty million primes. http://primes.utm.edu/lists/small/millions

[22] http://originlab.com/www/helponline/origin/en/UserGuide/Extreme.html

[23] Fubini, A., Alberini, R., & Lattanzi, D. (1973) $^{203}Tl(n, \gamma)$ Reaction and Level Structure of ^{204}Tl , *Il Nuovo Cimento (S.I.F.)* 11(18), 711–725. and references therein. http://link.springer.com/article/10.1007%2FBF02727587#

[24] Colao, F., Fantoni, R. & Lattanzi, D. (2005) *LIPS analysis of samples of tree trunks* ALT 04 (Rome and Frascati, Sept. 10–15, 2004) Advanced Laser Technologies 2004, Shcherbakov, I.A., Giardini, A. Vitali Konov, I., & Pustovoy, I.V. (Eds) Proceedings of the S.P.I.E., 5850, 166–173.