Journal of Advances in Mathematics and Computer Science

Volume 39, Issue 1, Page 36-61, 2024; Article no.JAMCS.112293
ISSN: 2456-9968
(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Statistical Distributions of Prime Number Gaps

Daniele Lattanzi [a++*]

[a] *Nuclear Fusion Department, Frascati Research Centre, Frascati, Roma, Italy.*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/JAMCS/2024/v39i11861

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/112293

Received: 21/11/2023
Accepted: 25/01/2024
Published: 31/01/2024

Original Research Article

Abstract

A heuristic i.e. empirical approach to the problem of prime number gaps of many kinds and types, different degrees and orders, treated as simple raw experimental data from the statistical viewpoint is presented. The aim of the article is to show a picture of the actual situation of prime number gaps in order to describe and to try to understand the structure itself of prime gaps of various kinds and orders as well as of primes themselves. The data base comprises the finite sequences of prime number gaps up to the value P_n of the prime counter $n = 5 \cdot 10^7$ that is $P_{5E7} = P(5 \cdot 10^7) = 982{,}451{,}653$ all of them available in the net. The statistical distributions of prime gaps are best-fitted by the pseudo-Voigt fit function, a convolution of the Lorentz and the Gauss differential distribution functions, or by the so-called E-exp or exp-exp differential distribution function or by a log-linear histogram according to the kind of gaps examined, either $\delta^i P_n$ (higher order gaps) or $\Delta^k P_m = P_m - P_{m-k}$ (delta-lags) with i and $k \geq 2$ or the simple linear differences $\delta^1 P_m = \Delta^1 P_m = \Delta P_m = P_m - P_{m-1}$ respectively. One of the unexpected results of the investigation is the appearance of inner structures at high values of n_Δ, the number of the intervals of the distributions, suggesting the presence of groups or clusters strictly linked to the nature of prime numbers themselves in which the same phenomenology is present.

[++] *Former E.N.E.A.;*
**Corresponding author: Email: lattanzio.lattanzi@alice.it;*

Keywords: Prime number gaps; data fits; distribution functions; experimental mathematics; statistics.

1 Introduction

The problem of prime numbers, or primes, in mathematics has always been a challenge to face and one of the major open problems notwithstanding the many theoretical successes achieved [1-14]. In previous articles by the same author [15-19] an experimental approach has been attempted to the matter leading to valuable and remarkable findings. In the first of these articles [15] the statistical treatment of prime gaps has been just mentioned leaving the deepening of the topic to next further studies, thus the present report deals with this theme showing many cases and many different kinds, orders, degrees and ranks of prime gaps in order to try to understand the innermost nature of them and of primes themselves.

The intention of the whole article is to show a picture of the actual situation of prime number gaps in order to describe and to try to understand the structure itself of prime gaps of diverse kinds and orders as well as of primes themselves and having already treated the deterministic facet of primes it is now time to examine their stochastic or probabilistic aspect as possibly hidden in their gaps. As a matter of fact it is the author's opinion that both the former and the latter constitute the organization of prime numbers.

In this frame the aim of the present study is twofold:

1- to investigate some of the major actual features of prime number gaps of different kinds and orders describing the actual situation and to find relationships among them or at least patterns and/or configurations, if any, in order to understand their innermost nature as much as possible;
2- to trace an innovative and original investigation pathway and method that can undergo many further developments and applications in the future in the wake of experimental mathematics.

An uncommon and innovative side of the issue is that the whole topic has been examined just using experimental physics methods to analyze (and to describe too) a typical problem of mathematics, thus launching an ideal bridge between the two disciplines and setting an interdisciplinary viewpoint as already done by the same Author in his previous articles (already cited). So all along the current article an empirical language, sometimes derived from experimental physics, has been preferred to a formal and rigorous presentation, as well as the mathematical formalism and stringency have been sacrificed to the epistemological aspects of the matter in view of a pragmatic vision of the problem and the ensuing approach. As an anticipation it is to be told that the standard well-known statistical distribution functions such as Gauss, Lorentz, Voigt (or better pseudo-Voigt), E-exp or exp-exp that is $e^{\exp(-x)} = \exp[\exp(-x)]$ and log-linear, according to the type of gap, either $\delta^i P_n$ (higher order gaps) or $\Delta^k P_n$ (delta-lags) with i and k ≥ 2 or the simple linear differences $P_m – P_{m-1}$ respectively have been used like probes to investigate the behaviour of the statistical treatments of some kinds of data-points (namely of the prime gaps of different kinds and orders) that is to fit the scatter-plot of some prime gaps assumed like paradigmatic of the whole situation taking into account the induction principle. These statistical distribution functions have been chosen for the fits in that they have proven to be the best fit functions on the base of the fit parameters such as the correlation coefficient, the non-linear index of correlation and all the usual statistical methods or given by default by the PC built-in ad-hoc S.W. Other distribution function have been tested but not found to be so appropriate as those used.

As a further consideration, it is to be emphasized that the whole work, performed by the sole Author, lasted many months what means that it took approximately 1,500 person-hours in the treatment of many million data (prime numbers and their gaps). As a matter of fact an amount of about 452,000,000 i.e. 452M data has been examined and statistically treated just as reported in the next Tables 2, 4 and 5 up to the value of the prime counter n = $5 \cdot 10^7$ that is $P_{5E7}=P(5 \cdot 10^7)=982,451,653$. Of course only some i.e. few of all these prime numbers have been scrutinized among all the many primes considered owing to the limited computer memory as well as just some of the many cases examined are reported here both for space reasons and for the fact that the results shown are fully representative of the entire situation.

In this empirical and pragmatic framework and having adopted the viewpoint of computational/experimental mathematics, the computer has been a central tool in the treatment of so many data. However just a simple standard PC of commercial type has been used with 700 GB HD memory and 8 GB RAM with adequate built-in spread-sheet SW capable of supporting calculations up to 1E308 = $1 \cdot 10^{308}$ and π value with 12 decimal digits.

2 The Role of Experimental Mathematics

A straightforward way to obviate the problem that the universe of prime numbers has not yet unveiled all its deepest and most hidden secrets is to make use of experimental or computational mathematics, a matter that is now becoming more and more popular inside the scientific community [20-27] mostly owing to the large use and the wide spreading of more and more powerful and compact computers in these latest decades. Cheaper and cheaper PCs have greatly facilitated the use of such a practice among researchers worldwide. Thus no doubt that, among all the issues of mathematics, number theory is the most suitable to this experimental approach and that, inside it, prime numbers are the most apt to be treated experimentally, not only for their nature itself but even owing to the large and reliable data bases nowadays available in many websites of the net [28,29].

The central issues of how to assess the results experimentally discovered in the general frame of mathematics are reported and well explained by Bailey & Borwein from which the main goals, among all the other ones, of experimental mathematics are:

1) …….. omitted ………
2) discovering new relationships;
3) testing and falsifying conjectures;
4) exploring whether or not a plausible/possible result may deserve a formal proof;
5) suggesting approaches for further next formal proofs.

While it is evident and clear that the formal rigorous proof in its canonical meaning and formulation remains a central pillar of all the mathematical strictness and reasoning however it is the Author's opinion that in the next future mathematicians should become acquainted to manage experimental evidences or verifications.

This intellectual process is not new at all. Just as a citation: "*In arithmetic the most elegant theorems often arise experimentally as the result of a more or less unexpected stroke of luck, while their proofs lie so deeply embedded in the darkness to elude all the attempts and defeat the sharpest inquiries.*" (C.F. Gauss 1777-1855) [30]. As a matter of fact Gauss is rightly considered one of the first experimental mathematicians by many researchers.

Moreover, just to cite another great mathematician: "*If mathematics describes an objective world just like physics, there is no reason why inductive methods should not be applied in mathematics just the same as in physics.*" (Kurt Gödel, 1951) [2].

Of course this relatively new scientific discipline has been already tackled by many other authors in many scientific reports and books, as already told, nonetheless it is the Author's believe that never before now has it shown and proven all its own power, usefulness and effectiveness as in the present case of prime number gaps.
Thus the future of mathematics can rely upon both theory and experimentation like many other sciences and scientific disciplines and in such a manner this new facet of doing mathematics can display its whole power in valuably aiding classical mathematics simply asking what factually and actually happens so that, like other sciences, mathematics will show itself just like a double-sided coin where linking abstraction to computable will be its future beyond any reasonable doubt.

3 Finite Numerical Data Sets

What has been done in the present research and shown here is to consider prime number gaps just as experimental data thus doing nothing else than usually done in the statistical treatment of actual experimental data [31-34], a procedure that is common to all fields of experimental physics and many other experimental fields of science. The main feature has been to treat prime number gaps as raw experimental data in a broad sense, to which all the usual statistical 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 whilst zero inaccuracy though not zero imprecision (owing to the approximations of the fits) are present on the fits and in the final results.

A large amount of data up to the value of the prime counter $n_{Max}=5\cdot10^7$ that is $P_{5E7}=P(5\cdot10^7)= =982{,}451{,}653$ has been used. The total data points examined amount to 452 million prime number gaps (see the next tables) of many types though not all the results are shown here. As a matter of fact those here reported are just few examples of prime gaps of any kind and order though paradigmatic of the whole situation according to the induction principle.

In fact in dealing with prime numbers one cannot miss to study their finite or discrete differences or gaps what has been done by some authors [35-38] though just at the first order i.e. $P_{n+1}–P_n$ both in a classical i.e. theoretical way and experimentally, so in the present report they have been examined from the latter standpoint with a very large number of cases. Thus the present article reports the statistics of the terms of the finite differences (i.e. gaps or deltas) of primes that is $\delta^iP_n = \sum_{h=0\to i}(-1)^h\binom{i}{h}P_{n-h}$ with $\binom{i}{h} = i! / h!\cdot(i-h)!$ the binomial coefficients and $\Delta^kP_m = P_m - P_{m-k}$ where i and k are the degree or order of the gaps. Of course having to deal with an infinite number of terms in all the cases one has to choose a finite number of them in order to give a sense to the whole issue.

A brief introduction about the choice of the variables examined is necessary.

The next list reports some of the many possible variables which can be treated, both from the statistical and the analytical viewpoint, when dealing with prime finite differences.

1 - $\delta^iP_n = \sum_{h=0\to i}(-1)^h\binom{i}{h}P_{n-h} = \sum_{h=0\to i} C_hP_{n-h}$ that is the higher order gaps ($i = 2, 3, 4, \dots, n$) examined in the present study;

2 - $\Delta^kP_m = P_m - P_{m-k}$ i.e. the linear gaps or so-called delta-lags ($k = 2, 3, 4, \dots, m$) examined in the present study;

3 - $\Delta P_n = P_n - P_{n-1} = \delta P_n = \Delta^1P_n = \delta^1P_n$ examined in the present study;

4 - $dP_n = \delta P_n - \langle\delta P_n\rangle = \Delta P_n - \langle\Delta P_n\rangle$

5 - $\partial P_n = \sqrt{(P_n^2 - P_{n-1}^2)}$

6 - $\partial P_n/\sqrt{\delta P_n} = \partial P_n/\sqrt{\Delta P_n} = \sqrt{(P_n + P_{n-1})}$

7 - $\partial P_n/\delta P_n = \partial P_n/\Delta P_n = \sqrt{[(P_n + P_{n-1}) / (P_n - P_{n-1})]}$

8 - P_n / P_{n-h} (h=1, 2, 3, 4, ... n)

9 - $P_n / \Delta^hP_n = P_n / (P_n - P_{n-h})$ (h=1, 2, 3, 4, ... n)

10 - $(P_m + P_n) / (P_m - P_n)$

and so on.

Of course there are no limits to the number and kinds of the variables which can be taken into account in examining prime number gaps in order to find relationships inside all their universe. All of them can be suitable to explore the prime gaps though it is of the utmost importance to identify and select the best variables, i.e. the most appropriate to describe statistically their most significant behaviours and trends.

In the present report the first three finite differences or gaps have been taken into account when dealing with primes, having considered them the most interesting, enlightening, noteworthy and revealing that is:

1. The so-called higher-order gaps or i-order gaps or i^{th} gaps i.e. the prime differences under the form $\delta^iP_n=\delta(\delta(\delta(\delta(\dots\delta P_n))))$ = $\delta\delta\delta\delta\dots\delta P_n$ (the δ operation repeated i times) = $\delta^iP_n= =\sum_{h=0\to i}(-1)^h\binom{i}{h}P_{n-h}$ = $\sum_{h=0\to i} C_hP_{n-h}$ (n, i and h ∈ N, i ≥ 2) which is the discrete analogue of the i^{th} derivative of a standard function apart from the missing denominator;

2. The so-called linear gaps or linear differences, also called delta-lags, under the form $\Delta^k P_m = P_m - P_{m-k}$ with k, m ∈ N and k ≥ 2 that is the linear or the 1rst order difference between two primes separated by a *distance* i.e. difference k;

3. The first-order gaps $\delta P_n = \delta^1 P_n = \Delta P_n = \Delta^1 P_n = P_n - P_{n-1}$

A special attention has been devoted to the third kind of gaps that is the first order gaps in that they belong neither to the first nor to the second type and that's why they have been examined and treated apart.
The next three paragraphs give a brief description of the data base involved in the study and after having described the fitting and the statistical methodology in Ch. 4 the chapter 5 will discuss the results obtained that is the statistical distributions of the three types of gaps and their meaning in detail.

3.1 Higher order gaps $\delta^i P_n$

In the case of the higher i^{th} order gaps or gaps of order i the following equivalences hold:

$\delta^0 P_n = P_n$ conventionally

$\delta^1 P_n = \delta P_n = P_n - P_{n-1}$
$\delta^2 P_n = \delta\delta P_n = \delta(P_n - P_{n-1}) = \delta P_n - \delta P_{n-1} = P_n - P_{n-1} - (P_{n-1} - P_{n-2}) = P_n - 2P_{n-1} + P_{n-2}$
$\delta^3 P_n = \delta\delta\delta P_n = \delta\delta^2 P_n = \delta(P_n - 2P_{n-1} + P_{n-2}) = P_n - 3P_{n-1} + 3P_{n-2} - P_{n-3}$
$\delta^4 P_n = \delta\delta\delta\delta P_n = \delta(P_n - 3P_{n-1} + 3P_{n-2} - P_{n-3}) = P_n - 4P_{n-1} + 6P_{n-2} - 4P_{n-3} + P_{n-4}$
$\delta^5 P_n = \delta\delta\delta\delta\delta P_n = P_n - 5P_{n-1} + 10P_{n-2} - 10P_{n-3} + 5P_{n-4} - P_{n-5}$
....
$\delta^i P_n = \sum_{h=0\to i}(-1)^h \binom{i}{h} P_{n-h} = \sum_{h=0\to i} C_h P_{n-h}$

where, of course, n > i ≥ h ∈ N and $C_h = \binom{i}{h} = \binom{i}{i-h} = i! / [(i-h)! \cdot h!]$ are the binomial coefficients with the sum $\sum_h$ running from h = 0 thru i.

The following Table 1 reports a very short example showing the values of the first i^{th} order gaps $\delta^i P_n$ with i = 0, 1, 2, 3, … n just for the few values of i = 0 through 9 and n from 1 up to 10, i.e. from $P_1=2$ up to $P_{10}=29$. Any value of this matrix at (n, i) ≡ (P_n, $\delta^i P_n$) is the result of the difference (n, i) ≡ (n, i–1) – (n–1, i–1) ≡ (P_n, $\delta^{i-1}P_n$) – (P_{n-1}, $\delta^{i-1}P_{n-1}$) as, for example, 8 = =4 – (–4) and –12 = –4 – (8) as highlighted in the table itself. Of course the table (i.e. matrix) is infinite, and so are all the other matrices shown afterwards, just like the number of the elements (prime number gaps) so that one cannot but consider (and examine) a limited (finite) part of it. Despite that, the results are indicative i.e. paradigmatic i.e. symptomatic of the entire factual situation.

Table 1. Values of the i^{th} order gaps $\delta^i P_n = \sum_{h=0\to i}(-1)^h \binom{i}{h} P_{n-i} = \sum_{h=0\to i} C_h P_{n-i}$

n ⇨	1	2	3	4	5	6	7	8	9	10
P_n ⇨	2	3	5	7	11	13	17	19	23	29
$\delta^0 P_n = P_n$	2	3	5	7	11	13	17	19	23	29
$\delta^1 P_n = \delta P_n = P_n - P_{n-1}$		1	2	2	4	2	4	2	4	6
$\delta^2 P_n = P_n - 2P_{n-1} + P_{n-2}$			1	0	2	–2	2	–2	2	2
$\delta^3 P_n = \delta(P_n - 2P_{n-1} + P_{n-2})$				–1	2	**–4**	**4**	–4	4	0
$\delta^4 P_n = \delta\delta^3 P_n = \delta\delta\delta\delta P_n$					3	–6	**8**	–8	**8**	**–4**
$\delta^5 P_n = \delta\ \delta^4 P_n$						–9	14	–16	16	**–12**
$\delta^6 P_n$							23	–30	32	–28
$\delta^7 P_n$								–53	62	–60
$\delta^8 P_n$									115	–122
$\delta^9 P_n$										–237

As for the coefficients $C_h = \binom{i}{h} = \binom{i}{i-h} = i! / [(i-h)! \cdot h!]$ it is easy to verify that one has $\sum|C| = = \sum_{h=0\to i}|C_h| = 2^i$ (∀i) for the sum of their absolute values and $\sum C = \sum_{h=0\to i} C_h = 0$ always (i.e. ∀ i≠0) in that they are the same coefficients of those for the formulas of the binomial powers that is $(a-b)^n = \sum_{h=0\to n}(-1)^h \binom{n}{h} a^h b^{n-h}$

Table 2. Type and number of cases of $\delta^i P_n$ examined in the study (i≥2)

n	i of δiPn	1 out of …..	n° of cases
1M	i = 2, 3, 4, 5, 6, 7, 8, 9, 10	1 i.e. all	9
3M	i = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1 i.e. all	15
5M	i = 2, 3 ,4 ,5 ,6 ,7, 8	1 i.e. all	7
10M	i = 2, 3, 4, 5, 6	1 i.e. all	5
50M	i = 2	1 out of 8	1

Thus the finite differences of the prime numbers studied are the ith order gaps or $\delta^i P_n$ (i= 2, 3, 4, 5, …..) up to the primes P_{1M} P_{3M} P_{5M} P_{10M} and P_{50M} according to the previous Table 2 where an amount of 37 cases for a total of 145.25M = $145.25 \cdot 10^6$ prime gaps has been examined in the whole research as for this kind of gaps. However not all of them have been reported in this article both for space reasons and in that the same behaviour has been verified in any of them, so that just some cases paradigmatic of the entire situation are shown here according to the convenience and to their key significance. That means that the induction principle has been widely used in all the data treatments.

3.2 Linear gaps or delta-lags $\Delta^k P_m = P_m - P_{m-k}$

As for the second type of gaps, i.e. the linear ones or delta-lags, the following Table 3 reports another very short example showing the initial values of the many linear gaps or delta-lags $\Delta^k P_m = P_m - P_{m-k}$ with k = 0, 1, 2, 3, … 9 and m = 1, 2, 3, … 12 (m > k) where the table is to be read as, for instance: $\Delta^6 P_{10} = P_{10} - P_{10-6} = P_{10} - P_4 = 29 - 7 = 22$ or $\Delta^3 P_8 = P_8 - P_{8-3} = P_8 - P_5 = 19 - 11 = 8$ $\Delta^7 P_9 = =P_9 - P_{9-7} = P_9 - P_2 = 23 - 3 = 20$ etc. as also highlighted inside the table itself.

Table 3. Values of the linear gaps or Δ-lags of the kth order $\Delta^k P_m = P_m - P_{m-k}$ (k≥2)

m	1	**2**	3	***4***	**5**	6	7	**8**	**9**	***10***	11	12
P_m	2	**3**	5	***7***	**11**	13	17	**19**	**23**	***29***	31	37
$\Delta^2 P_m=P_m-P_{m-2}$			3	4	6	4	4	6	6	10	8	8
$\Delta^3 P_m=P_m-P_{m-3}$				5	8	8	10	**8**	10	12	12	14
$\Delta^4 P_m=P_m-P_{m-4}$					9	10	12	12	12	16	14	16
$\Delta^5 P_m=P_m-P_{m-5}$						11	14	14	16	18	18	20
$\Delta^6 P_m=P_m-P_{m-6}$							15	16	18	***22***	20	24
$\Delta^7 P_m=P_m-P_{m-7}$								17	**20**	24	24	26
$\Delta^8 P_m=P_m-P_{m-8}$									21	26	26	30
$\Delta^9 P_m=P_m-P_{m-9}$										27	28	32
$\Delta^{10} P_m=P_m-P_{m-10}$											29	34
$\Delta^{11} P_m=P_m-P_{m-11}$												35

Table 4. Type and number of cases of $\Delta^k P_m = P_m - P_{m-k}$ examined in the study (k≥2)

m (1out of…..)	k of $\Delta^k P_m$	n_Δ of intervals	n° cases	n°of gaps
1M (1 i.e. ALL)	2 3 4 5 6	50	5	5M
3M (1 i.e. ALL)	2 3 4 5	50	4	12M
5M (1 i.e. ALL)	2 3 4 5 6	50	5	25M
10M (1 i.e. ALL)	2 3 4 5 6	50	5	50M
50M (1 out of 8)	8	50	1	6.25M
3M (1 i.e. ALL)	2	25	1	3M
10M (1 i.e. ALL)	2	100 200	2	20M
10M (1 i.e. ALL)	4	100 200 300	3	30M
10M (1 i.e. ALL)	6	100 200 300	3	30M
50M (1 out of 8)	8	100 200 300 400 1000	5	31.25M

For these Δ-lags too the study has not been limited to the cases here reported though extended, as reported in the previous Table 4, to a sum of 20 + 14 = 34 cases for a total of 212.5E6 primes. Again the intention of these choices has been to examine and to treat the most meaningful cases that could give the most real and evident representation of the whole matter that is the statistical behaviour of prime number gaps of many orders of this kind. As usual the induction principle is a valid help to extrapolate the data and the results in generalizing from these trends of some few prime gaps to the trends of all the infinitude of prime gaps of the same kind.

The difference between the two groups is in the number of the intervals (n_Δ) into which any whole range of prime gap has been divided in order to study the variations, if any, in the results at constant n_Δ intervals and different k (the first part of the table) and at different n_Δ intervals and constant k (the second part) so to have the clearest view of the situation.

3.3 First order gaps $\delta P_n = \Delta P_n = P_n - P_{n-1}$

The third case studied, as already anticipated, is the first finite differences of prime numbers that is the first order gaps $\delta^1 P_n = \delta P_n = \Delta^1 P_n = \Delta P_n = P_n - P_{n-1}$ i.e. i=k=1 which have been left aside and examined apart owing to their special nature inside the world of prime number gaps what should have already been expected in that the first finite deltas with k = i = 1 should behave as both the $\delta^i P_n$ and the $\Delta^k P_m$ what is impossible.

Also in this situation many cases (94M) have been examined by the Author as shown in the next Table 5 and again not all the results are reported in the present article both for space constraints and because the same trend has been revealed in them all thus making use of the induction principle in these cases too.

Table 5. Type and number of cases $\delta P_n = \Delta P_n = P_n - P_{n-1}$ examined

n (1 out of …)	n_Δ = n° of intervals	n° of cases	n° of gaps
1M (all)	50	1	1 M
3M (all)	50	1	3 M
5M (all)	50	1	5 M
10M (all)	50 100 200 400 1000 2000	6	60 M
50M (1/8)	50 100 200 300	4	25 M

Thus a total of 13 cases has been examined for a sum of 94M=94E6 prime gaps just some of which (again the most paradigmatic ones) shown in the figures of the ad-hoc chapter 5.2.

4 Fitting Procedure and Statistics

In the entire study the fitting process has been a sensitive and crucial issue even due to the gauges and markers to be taken into account and to be optimized all at the same time and that's why it deserves a brief critical discussion.

The statistical treatments of the prime gaps, considered like actual experimental data, and the related fits have been performed using most of the tools available in statistics [31-34]. In the fit performed between the two frequency distribution functions (FDFs) - i.e. the experimental one given by the actual counts, i.e. the scatter-plot, and the theoretical or parent one or sample one given by the function that fits the data points - both the correlation (or Bravais-Pearson) coefficient R=R(C,F) and the non-linear index of correlation I=I(C,F) between any set of counts C (i.e. the scatter-plot) and the fit F (i.e. the parent distribution) have been calculated and maximized. The former parameter R=R(C,F) is written as

$$\mathbf{R(C,F) = \sum_i (C_i - \langle C\rangle)\cdot(F_i - \langle F\rangle) \,/\, \sqrt{\,[\sum_i (C_i - \langle C\rangle)^2 \cdot (F_i - \langle F\rangle)^2]} = 1^-}$$

while the latter I = I(C, F) is

$$I = I(C, F) = 1 - [(\sum_h (C_h - F_h)^2 \cdot \sum_i (C_i - \langle C\rangle)] = 1^-$$

In both formulas the sums extend from 1 up to the number n_Δ of the intervals used for the statistical calculations, while <C> and <F> are the average values.

Maximizing these two statistical markers means making both of them to approach the value of 1^- in order that the two FDFs could match one each other as much as possible, as well as equalizing their means <C> and <F>.

In addition the software built-in correlation coefficient

$$R' = [n_\Delta \cdot \sum_h C_h \cdot F_h - (\sum_i C_i) \cdot (\sum_j F_j)] \,/\, \sqrt{\{[n_\Delta \cdot (\sum_k C_k)^2 - (\sum_l C_l)^2] \cdot [n_\Delta \cdot (\sum_m F_m)^2 - (\sum_n F_n)^2]\}} = 1^-$$

has been considered too, which has always been found to be in agreement with the previous two. All these parameters have been used for the matching of the Cumulative Distribution Functions (CDFs) too. In addition, even the two standard deviations of the means have been examined, namely

$$\sigma^C = \sqrt{[(\sum_i (C_i - \langle C \rangle)^2)/n_\Delta]} \qquad \sigma^F = \sqrt{[(\sum_i (F_i - \langle F \rangle)^2)/n_\Delta]}$$

in order to ascertain that each of them would be much lower than its respective mean $\langle C \rangle$ and $\langle F \rangle$ and that they could be equal one each other as much as possible.

Finally, two further gauges of the fits have been minimizing the value of the Least Square Sum $LSS = \sum_i (C_i - F_i)^2 / 2\sigma^2$ according to the principle of maximum likelihood and the Chi-square Test Value $_{T\text{-}V}X^2 = \sum_h (C_h - F_h)^2 / F_h$ in that both of these variables measure the goodness of the fit. As usual the sums extend from 1 up to the n_Δ intervals used for the calculations.

Just to summarize, most of the statistical tools available in statistics to treat experimental data, as already told and cited, have been used in order to make the best possible fits with the utmost statistical reliability. However it should be taken into account that in optimizing all the gauges and markers some compromise has sometimes had to be made which, nonetheless, in no way and in no case has ever endangered the reliability of the results but just influenced occasionally though weakly their precisions.

As for the fits of the gaps of prime numbers both of type δiP_n and $\Delta^k P_n$ they have been examined making use of the well-known standard statistical differential distribution functions DDFs which have been assumed as the parents distributions such as the following ones.

Gauss or normal differential distribution function (the well-known bell-shape curve)

$$G(x) = [1/(\sigma\sqrt{2\pi})] \cdot \exp[-(x-\mu)^2/(2\sigma^2)]$$

being $\mu = \langle x \rangle$ its centre that is the mean value and σ its standard deviation specifying also the width of the distribution function FWHM = Full_Width_at_Half_Maximum = $2\sigma\sqrt{(2 \cdot \ln 2)}$ while the maximum value is $G_{MAX} = G(\mu) = G(\langle x \rangle) = 1/(\sigma\sqrt{2\pi})$

Lorentz differential distribution function

$$L(x) = (\gamma/2\pi) \,/\, [(x-\mu)^2 + (\gamma/2)^2]$$

being $\mu = \langle x \rangle$ its centre where it has its maximum value $L_{max} = L(\mu) = L(\langle x \rangle) = 2/\pi\gamma$ and FWHM = γ

The comparison between the two distributions, after their normalizations, shows that the former (G) is less peaked while the latter (L) has heavier tails (or wings) what means a larger area and higher kurtosis, the 4th statistical moment specifying the flatness of the distribution, defining kurtosis $k = \sum_i (x_i - \langle x \rangle)^4 / (N-1)\sigma^4$ what leads in the Gaussian and Lorentzian cases respectively to $k_G = 3$ and $k_L > 3$.

Both distribution functions display a null value of the skewness (a measure of their asymmetry) defined as the 3rd moment $S = \sum_i (x_i - \langle x \rangle)^3 / (N-1)\sigma^3$ so that both Gauss and Lorentz have $S = 0$ (N is the number of the data-points).

In molecular, atomic and nuclear spectroscopy (branches of experimental physics) a spectral line profile is usually described by means of the Voigt distribution function which is a convolution of the previous two DDFs i.e. $V(x) = \int G(x) L(x-y) dy$ the integral $\int$ extending from $-\infty$ up to $+\infty$ and the two DDFs being equally centred. That means that a line shape and its broadening have to be regarded as the result of two distinct phenomena: a random process around a mean value described by Gauss DF and a series of particle collective motions (of molecules, atoms or nucleons) which causes the line broadening described by Lorentz DF. It is a standard procedure for simplicity's sake to use, instead of a Voigt DF, a pseudo-Voigt DF given by $V(x) = f\cdot G(x) + (1-f)\cdot L(x) \quad (0 \leq f \leq 1)$ that is a simple linear combination of G(x) and L(x) with the parameter f (weighting or mixing ratio) giving the weight of each one.

At last, a fourth major DDF used in the present study is the ExpExp or E-exp function sometimes called extreme function [39], or Gumbel distribution, a SW built-in function, in the form $E(x) = E_o + Ae^{-\exp(-y)-(y-1)}$ with $y = (x-x_o) / w$ being A the amplitude, E_o the baseline value or offset, x_o the centre and w the width. This function is highly skewed toward the low values of x.

Going back to the finite differences of prime numbers, it is to be anticipated that the i^{th} order gaps ($i \geq 2$) or $\delta^i P_n$ follow a pseudo-Voigt DDF, the linear gaps $\Delta^k P_m$ ($k \geq 2$) follow a statistical distribution of the type E-exp (or ExpExp), while the statistical DDFs of the $\Delta P_n = \delta P_n$ (i.e. gaps with k=i=1) show very special features and it is easy to guess the reason in that they follow neither the former nor the latter distribution. That's why they are examined and treated apart and just after the previous two.

5 Results

The statistics of the finite sequences of prime gaps of many kinds and orders have been examined and reported in order to determine the presence, if any, of typical features, pathways and characteristic structures. In some cases the fits by the previously mentioned DDFs of the scatter-plots (i.e. the fits of the data-points by the parent or theoretical histograms) are shown with all their characteristics and markers while in most cases just some qualitative results are shown in that very interesting and of great importance as never attained before now. Their main feature relies on their enigmatic nature and strange configurations and at the present time just some conjectures can be made to try to solve their well identified though mysterious structure which moreover are typical of all the kinds and order of gaps thus establishing the doubt that they might be typical of prime numbers themselves. As a matter of fact that has been checked and verified as reported in the last paragraph of this chapter.

5.1 Higher order gaps or i^{th} order gaps $\delta^i P_n$ ($i \geq 2$)

The next Fig. 1 shows the case of the 2^{nd} order gaps $\delta^2 P_n = P_n - 2P_{n-1} + P_{n-2}$ as a typical example for the first 50M P_n (as well as P_{n-1} P_{n-2} and $\delta^2 P_n$ of course) 1 out of 8, a choice to make all the calculations faster and lighter therefore dealing with 25M numbers instead of 200M numbers.

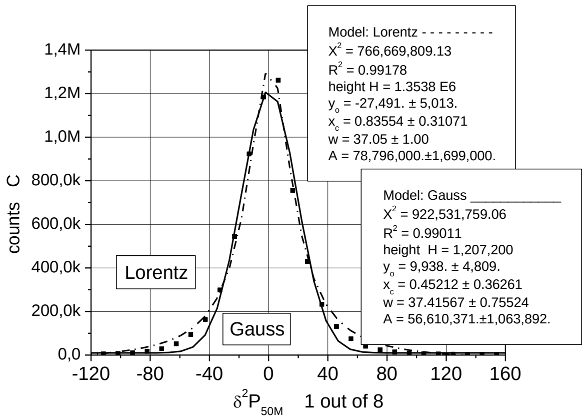
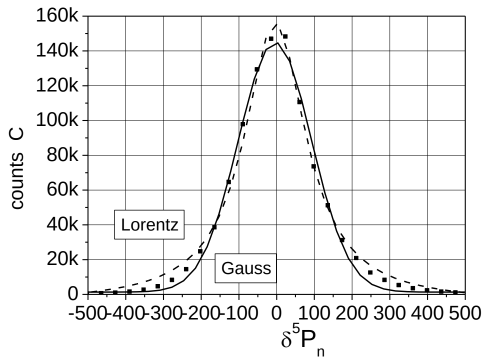


Fig. 1. Scatter-plot ■ Gauss —— & Lorentz · – · – · – fit histograms of $\delta^2 P_{50M}$ gaps 1 / 8

In this figure the whole range of $\delta^2 P_{50M}$ values from −246 up to +246 (though shown from −120 to +160) has been divided into $n_\Delta = 50$ intervals any of which having width Δ=9.84. The two fits made (the Gaussian and the Lorentzian one, each with its own parameters displayed in the insets as given by the built-in SW calculations) show clearly that this scatter-plot is well matched by a pseudo-Voigt distribution function, though at present it has not been possible to calculate the value of the parameter f ∈ (0., 1.). However it is of the utmost importance and should catch the attention the fact that in all the 37 cases examined the existence of such a situation has been carefully verified so that it is straightforward to assess that all the finite differences $\delta^i P_n = \sum_{k=0\to i} (-1)^k \binom{i}{k} P_{n-k}$ with $i \geq 2$ have statistical distributions well fitted by pseudo-Voigt DDFs that is by $V(\delta^i) = f \cdot G(\delta^i) + (1-f) \cdot L(\delta^i)$ being $f \in (0., 1.)$ and G & L the Gaussian and the Lorentzian DDFs respectively with their ad-hoc parameters shown in the insets.

The next two Figs. 2 (n = 1M and $i = 5$ that is $\delta^5 P_{1M}$) and 3 (n = 3M, $i = 16$ i.e. $\delta^{16} P_{3M}$) both of which at $n_\Delta = 50$ (zoom-in) show other two examples of what has been told just now (scatter-plots and histograms) as well as other interesting features such as**:**



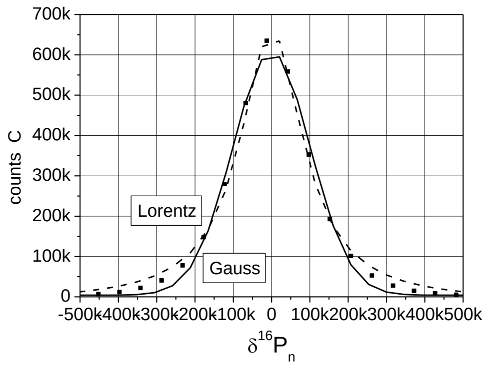


Fig. 2. Scatter-plot & fits for $\delta^5 P_{1M}$ n_Δ=50

Fig. 3. Scatter-plots & fits for $\delta^{16} P_{3M}$ n_Δ=50

the increasing values of $\delta^i P_n$ with increasing i values (clearly derivable from the binomial coefficient formula), the peaks of the DDFs which are best fitted by the Lorentzian distribution functions, the wings or tails of the DFs which are approximately half-way between the $G(\delta^i)$ and the $L(\delta^i)$ functions and finally the areas $A(i)=A(\delta^i)$ of the Gauss and Lorentz fits which increase vs. n (from 1M to 3M) and vs. the order i (from 5 to 16) of $\delta^i P_n$. As a matter of fact the features of the fits show the following results from the built-in SW calculations used.

For the case of $\delta^5 P_n$ for 1M primes i.e. $\delta^5 P_{1M}$ (Fig. 2) the fit features are**:**

Model: LORENTZ - - - - - - - - - - - - $_{t\text{-}v}X^2$ = 11,718,044.269 R^2 = 0.99307 height H = 161,890 $y_o = 5.049 \pm 0.686$ $x_c = 3.85 \pm 1.45$ $w = 198.4 \pm 4.9$ A = 50,449,164. ± 1,050,171.

Model: GAUSS ______________ $_{t\text{-}v}X^2$ = 10,011,100.434 R^2 = 0.99408 height H = 143,670 $y_o = 1.339 \pm 0.525$ $x_c = -4.41 \pm 1.41$ $w = 194. \pm 3.$ A = 35,043,836. ± 518,229.

For the case of the $\delta^{16} P_n$ gaps for 3M primes i.e. $\delta^{16} P_{3M}$ (Fig. 3) one gets**:**

Model: LORENTZ - - - - - - - - - - $_{t\text{-}v}X^2$ = 139,400,360.33 R^2 = 0.99373 $y_0 = -13.480 \pm 2.085$ $x_c = 4.91358 \pm 1.462$ w = 199,869.29 ± 4,639.53 A = 212,219,472,458. ± 3,909,146,055.

Model: GAUSS ______________ $_{t\text{-}v}X^2$ = 130,034,146.9 R^2 = 0.99415 height H = 603,320 $y_0 = 4.016 \pm 1.777$ $x_c = -52.0 \pm 1.5$ w=203,089.±3,116. A=153,566,166,178.± 2,185,919,810.

About the area A of the two curves, one can look at the next Figs. 4, 5 and 6 which show the plot of A vs. i = order of $\delta^i P_n$ for the examples of n = 1M, 3M and 5M respectively from which the exponential increases of the relationships are clearly apparent, for both Gauss and Lorentz cases, leading to a log-linear relationship in the general case (i.e. ∀ n and ∀ i).

Thus one gets from a best fit in Fig. 4**:** $LOG\_A_L 1M = (6.291 \pm 0.01791) + (0.2846 \pm 0.00274) \cdot i$
with fit values R=0.99968 SD=0.02123 N=9 p<1E−4

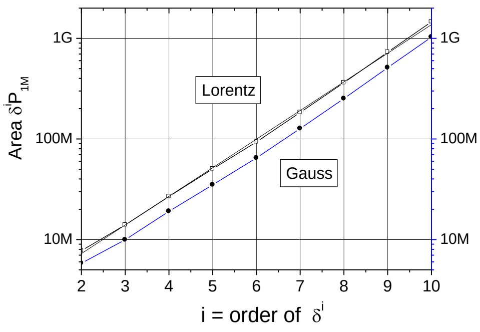


Fig. 4. Areas of the fits vs. order i for 1M

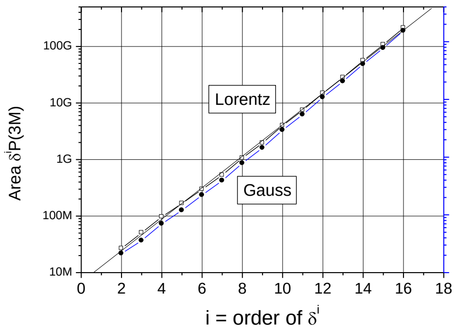


Fig. 5. Areas of the fits vs. order i for 3M

and similarly for Gauss area, while for Fig. 5 one gets the fit equation for Lorentz area

$LOG\_A_L 3M = (6.83141 \pm 0.02035) + (0.27837 \pm 0.00204) \cdot i$

with R=0.99965 SD=0.03411 N=15 p<1E−4 and the same for the Gaussian area.

The exponential increases of the relationships A = A(n) are clearly evident, for both Gauss and Lorentz cases ∀ n and ∀ i

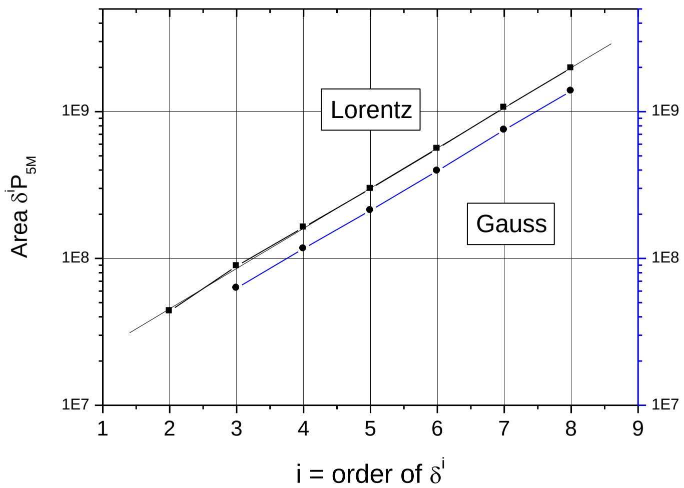


Fig. 6. Areas of Gauss and Lorentz fits vs. order i for n=5M

From the previous Fig. 6 it is easy to get for the Lorentz area of $\delta^i P_{5M}$ the fit values R=0.99985 SD=0.01114 N=7 p<1E−4 and the equation

$$LOG\_A_L = (7.11049 \pm 0.01134) + (0.273 \pm 0.002) \cdot i$$

In all the three Figs. 4 thru 6 the relationships hold both for Gauss and Lorentz case $\lg[A_L(\delta^i)] = \alpha_o(n) + \alpha(n) \cdot i$ and $\lg[A_G(\delta^i)] = \beta_o(n) + \beta(n) \cdot i$ so that it is plain to plot the coefficients α, α_o, β and β_o vs. n together with their errors as in the next two Figs. 7 and 8 that is Fig. 7

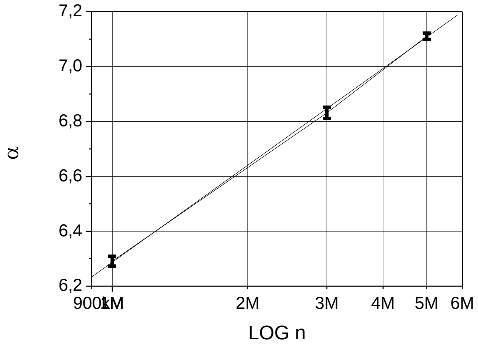


Fig. 7. Coefficient α_o for Lorentz area vs. n

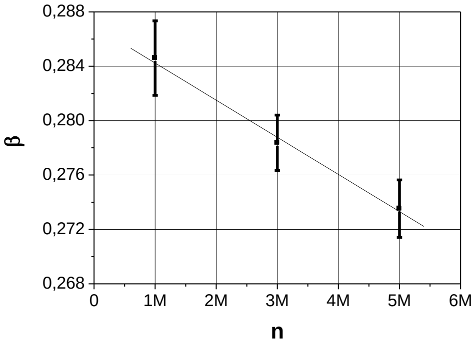


Fig. 8. Coefficient α for Lorentz area vs. n

(α_o vs. log(n)) and Fig. 8 (α vs. n) just for the Lorentzian case (the Gaussian one is similar) for the data points n = 1M, 3M and 5M with the fits shown.

As for Fig. 7 (Lorentz case) the fit gives $R=0.99976$ $SD=0.84342$ $N=3$ $p<0.01387$ with
$\alpha_o = \alpha_o(n) = -(0.75188 \pm 0.16622) + (1.17318 \pm 0.02557)\cdot\log(n)$

while for Fig. 8 (again Lorentz) the fit gives $R=0.99672$ $SD=0.25989$ $N=3$ $p<0.05156$

$\beta = \beta(n) = (0.28697 \pm \text{negligible}) - (2.73E{-}9 \pm 2.22E{-}10)\cdot n$

Of course one could proceed in the same way for what concerns the Gaussian area and, once found the value of the factor $f \in (0,1)$, go on for the pseudo-Voigt area too.

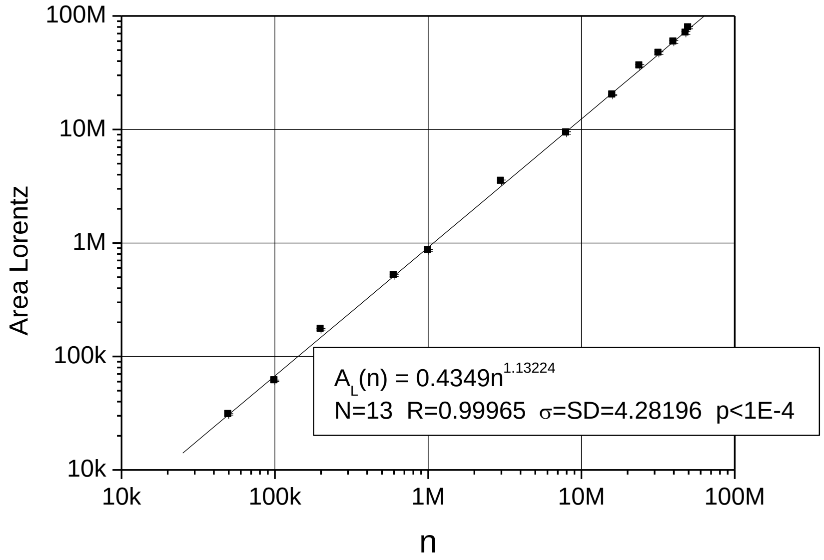


Fig. 9. Lorentz area A_L vs. n for the first 50M $\delta^2 P_{50M}$ 1 out of 8

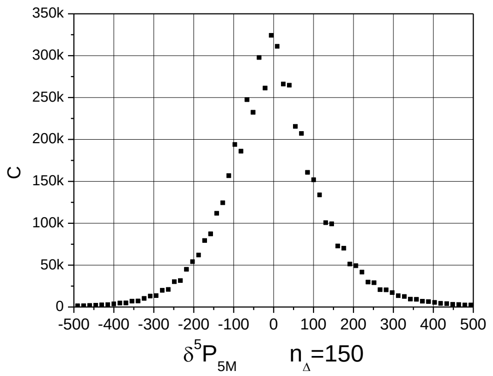
Again for the Lorentzian area the previous Fig. 9 shows the plot of $A_L(n)$ that is area vs. n for the $\delta^2 P_n$ (i.e. order i = 2 and the plotted n values), another interesting issue, among all the other ones, shown just as an example. The fit and its features are shown in the inset. Thus one can estimate that the more general relationship might hold for both the Gaussian and Lorentzian areas and so for the pseudo-Voigt too**:** $A(n) \approx \gamma_o(n)\cdot n^{\gamma(n)}$

It is now time to go back to a very important key aspect of the topic examined till now in the present paragraph wondering what is the meaning of the contemporary presence of these two distribution functions (Gauss and Lorentz) for the $\delta^i P_n$

First of all the pseudo-Voigt DF is the evidence of the mixing of the other two DFs, as detected in many molecular, atomic and nuclear spectra for the line shapes, being the sum of two effects. While the presence of the Gaussian component is the clear evidence of random processes - the random dispersion of δ^i values around a mean value (in this case null: $<\delta^i>=0$) thus perfectly reasonable and also typical of nuclear spectroscopy [40] - the presence of the Lorentzian component is less plausible. This DF tends to give a major weight to the tails thus being responsible for the line broadening that seems to suggest the occurrence of some inner structures. As a matter of fact in spectroscopy the presence of a Lorentzian shape for a spectrum line is a symptom of the existence, among the other things, of collective modes or motions governing the whole phenomenon. A typical example is the Stark effect in hydrogen as well as the line broadening of the plasma plume observed in the atomic and molecular spectroscopic technique called LIPS (Laser Induced Plasma Spectroscopy) [41].

Of course in the present circumstances (i.e. $\delta^i P_n$ and in general prime number gaps and primes themselves) certainly it is not possible to speak of collective modes or motions, but it is surely suitable to speak of gatherings or groups or clusters of numbers. Thus there is the well-founded feeling that the Lorentzian component of the pseudo-Voigt DF can be the effect of some inner structures of $\delta^i P_n$ leading to innermost groups responsible for their so-called collective behaviour i.e. their gathering.

Having ascertained the presence of the Lorentzian component in all the finite differences $\delta^i P_n$ of order i of primes P_n, though with different weight $f(i,n)\in(0,1)$, the plain next step is to look for these groups or this clustering effect the presence of which has been experimentally discovered by chance: in an effort to improve the statistics of $\delta^5 P_n$ in the case n=5M when the number $n_\Delta=50$ of the intervals of the whole histogram has been raised to $n_\Delta=150$ thus getting the Fig. 10 and to $n_\Delta=500$ as in the next Fig. 11 (zoomed scale).



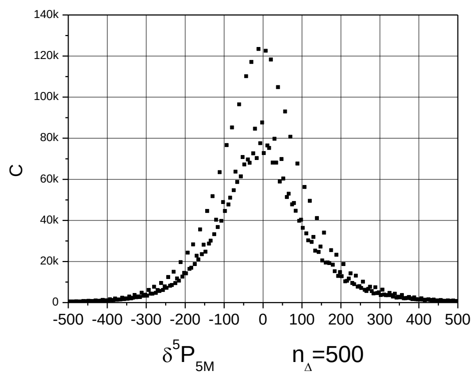
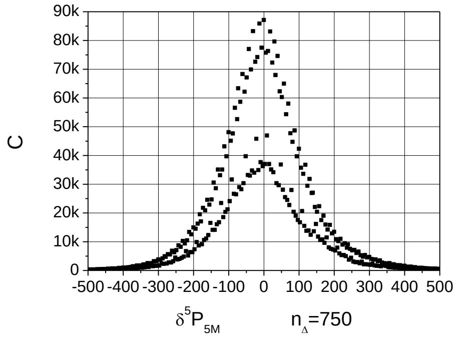


Fig. 10. Scatter-plot(s) of $\delta^5 P_{5M}$ with $n_\Delta=150$ **Fig. 11. Scatter-plots of $\delta^5 P_{5M}$ with $n_\Delta=500$**

The former suggests the existence of two scatter-plots while the latter shows at least 3 (or maybe 4) scatter-plots all of them “concentric” that is with $<\delta^5 P_n>=0$ though with different top values and FWHM. Going on in such a way the next Figs. 12 and 13 display the situation for $\delta^5 P_{5M}$ with $n_\Delta=750$ the former and $n_\Delta=1000$ the latter. It is evident for $n_\Delta=750$ the presence of at least 5 concentric DFs and for $n_\Delta=1000$ of many concentric distributions with different maximum values and different FWHM. Thus it is clear that the whole population of $\delta^5 P_n$ is subdivided into groups or clusters any of them having its own FD or fit histogram centred at $<\delta^5 P_{5M}>=0$ though with different height and different standard deviation that is different area.

In addition it can be observed the appearance of a skewness towards the negative values of $\delta^5 P_{5M}$ in all the groups. While it is not possible - at this stage of the research and with the tools available which do not seem capable of facing the situation - to understand fully all the features of these newly appeared scatter-plots, nonetheless the presence of a skewness in them is the distinctive evidence of the abandoning of the Gaussian or Lorentzian fit and so of the pseudo-Voigt curve.



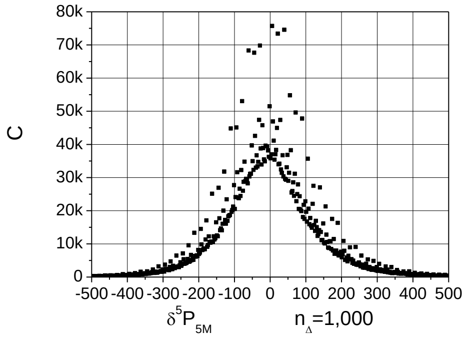
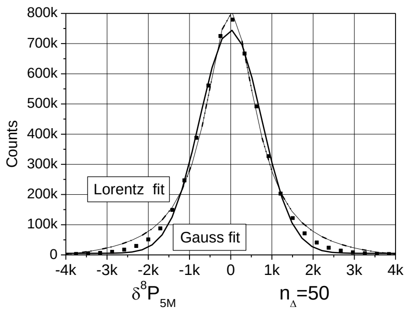


Fig. 12. Scatter-plots of $\delta^5 P_{5M}$ n_Δ=750 **Fig. 13. Scatter-plots of $\delta^5 P_{5M}$ n_Δ=1000**

It is very interesting to highlight that this effect can be seen not only in the present case of i=5 and n=5M but also in all the other cases here examined, though with odd i values i.e. i = 2j+1, what leads to infer that the whole situation holds ∀ i odd.

All the following Figs. 14 thru 19 show the case n=5M and $\delta^8 P_n$ i.e. $\delta^8 P_{5M}$ for n_Δ= 50 100 150 500 750 and 2,000 respectively.



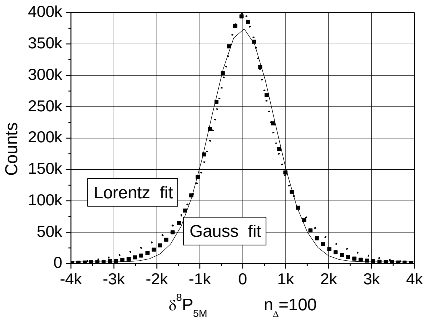


Fig. 14. Scatter-plot and fits of $\delta^8 P_{5M}$ n_Δ=50 **Fig. 15. The same for of $\delta^8 P_{5M}$ n_Δ=100**

In Fig. 14 one has: Gauss fit**:** Chi^2=1.65779E8 R^2=0.99625 A_G=1.381E9 Centre=0.17881 Width=1,491.1 Offset=5,345.0 Height=7.3938E5 while for the Lorentz fit: Chi^2=3.14327E8 R^2=0.99289 A_L=1.979E9 Centre=0.20401 Width=1,512.6 Offset=–26.658 Height=7.3293E5

It is interesting to remark that the Lorentzian fit has always a greater area (as already shown), a greater height and a greater width than the Gaussian one.

In the other Fig. 15 (n_Δ=100) the results are similar.

The splitting of the first unique scatter-plot for n_Δ = 50, 100 and 150 (Fig. 16) into two curves for n_Δ=500 (Fig. 17) and the further splitting into a multitude of further clusters for n_Δ=750 and n_Δ=2,000 as in the two Figs. 18 and 19, though gathered into two subgroups, is clearly visible. Of course also these behaviours deserve in-depth investigations in the future as well as the fact that in this case i=8 (i.e. i even) none of the curves has any asymmetry at all conserving its whole initial symmetry just like for n_Δ=50.

It is plain to recognize that at i=2h i.e. i=even the symmetry is conserved $\forall n_\Delta$ what does not happen for i=2j+1 that is for odd i as also depending from the analytical formulation of $\delta^i P_n$

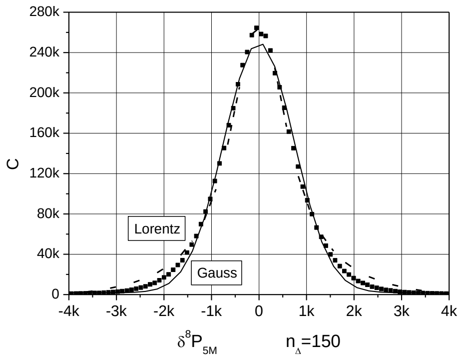


Fig. 16. Scatter-plot & fits of δ^8P_{5M} n_Δ=150

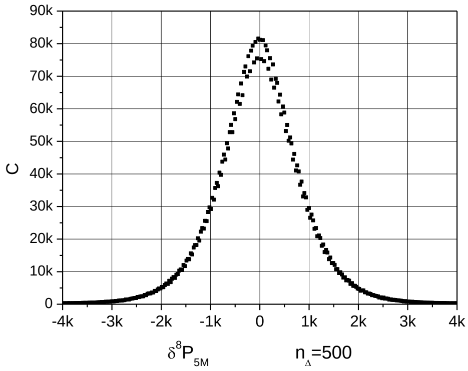


Fig. 17. Scatter-plots of δ^8P_{5M} n_Δ=500

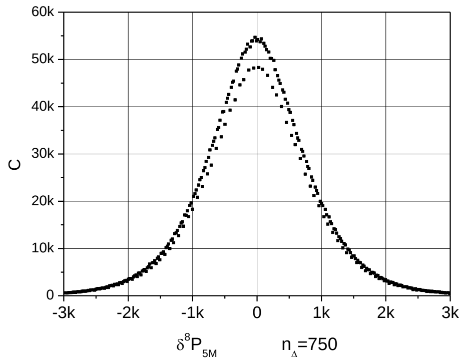


Fig. 18. Scatter-plots of δ^8P_{5M} n_Δ=750

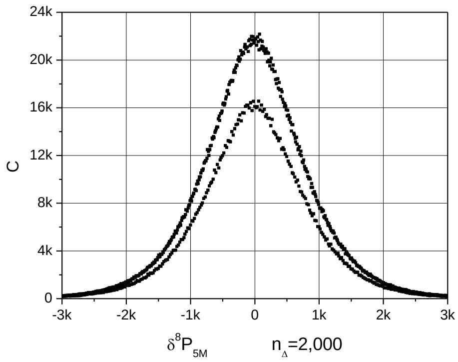


Fig. 19. Scatter-plots of δ^8P_{5M} n_Δ=2,000

All the next six Figs. 20 thru 25 illustrate the case of n=10M and δ^6P_n with n_Δ = 50 400 800 1,200 1,600 and 4,400 and they are of some interest. Again, while at n_Δ=50 there is just one scatter-plot fit by only one histogram (as usual a pseudo-Voigt DF), at n_Δ=400 there are three different concentric curves (i.e. peaked at and equally centred at $<\delta^6>=0$ and symmetric around this value), at n_Δ=800 there are 3 (maybe 4) concentric and symmetric distributions, at n_Δ=1,200 and 1,600 there are even 6 or 7, whilst finally at n_Δ=4,400 the histograms tend to gather themselves into two major groups, as shown in the Fig. 25. One can easily also see in this that there are many void intervals (that is with null counts) in that the width itself of the single intervals becomes too narrow to have counts.

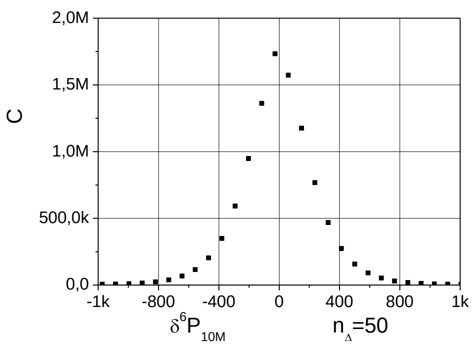


Fig. 20. Scatter-plot of δ^6P_{10M} n_Δ=50

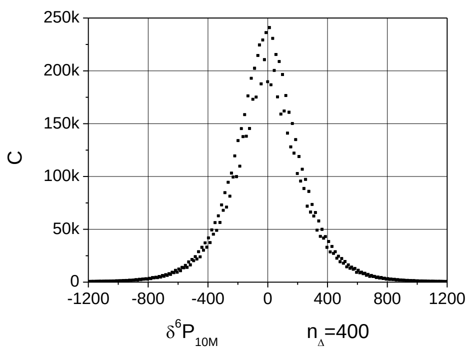


Fig.21. Scatter-plots of δ^6P_{10M} n_Δ= 400

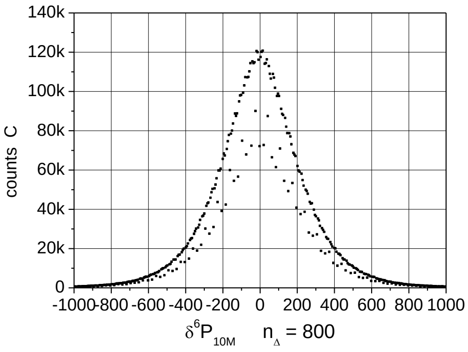


Fig. 22. Scatter-plots of δ^6P_{10M} n_Δ=800

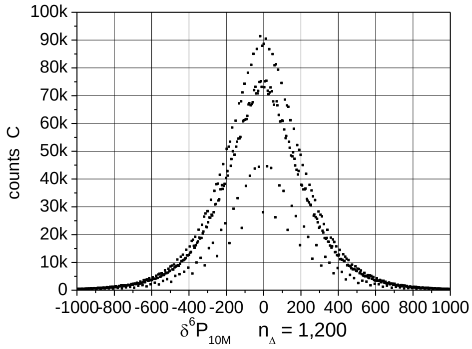


Fig. 23. Scatter-plots of δ^6P_{10M} n_Δ=1,200

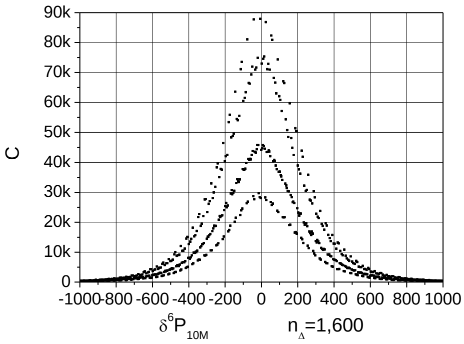


Fig. 24. Scatter-plots of δ^6P_{10M} n_Δ=1,600

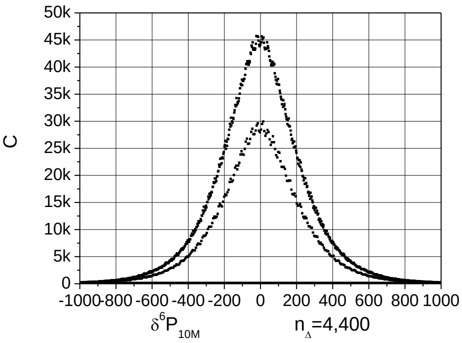


Fig. 25. Scatter-plots of δ^6P_{10M} n_Δ=4,400

It has to be remarked that most of the previous figures, as well as of the following ones, are zoomed-in as the tails of the distributions have very few or zero counts and are not important for the trend of all the data points**.**

As a final example the case of n=10M and δ^3 i.e. δ^3P_{10M} is shown in order to debate another kind of behaviour (see Figs. 26 thru 31) in which it is clear and evident the pronounced asymmetry, i.e. skewness, of the scatter-plots toward positive values.

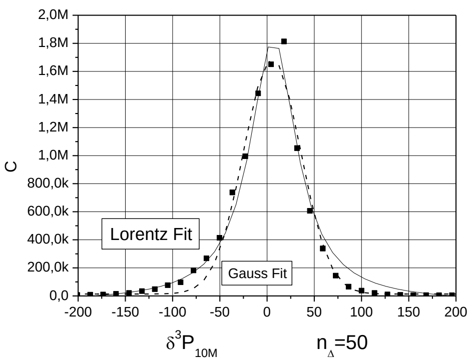


Fig. 26. Scatter-plot and fits of δ^3P_{10M} n_Δ=50

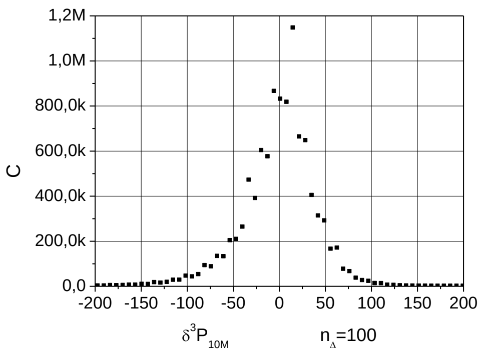


Fig. 27. Scatter-plot(s) of δ^3P_{10M} n_Δ=100

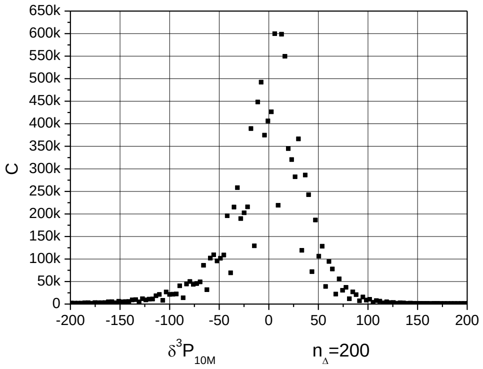


Fig. 28. Scatter-plots of δ^3P_{10M} n_Δ=200

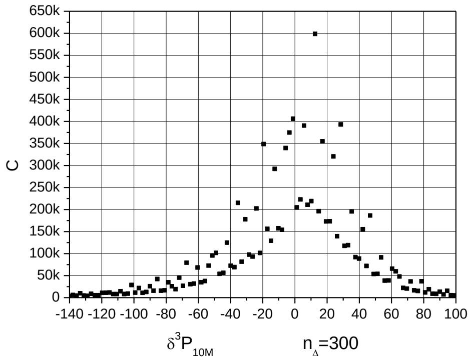


Fig. 29. Scatter-plots of δ^3P_{10M} n_Δ=300

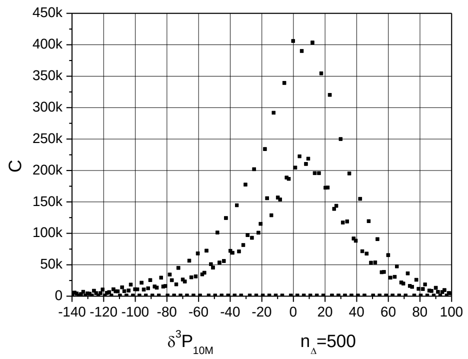


Fig. 30. Scatter-plots of δ^3P_{10M} n_Δ=500

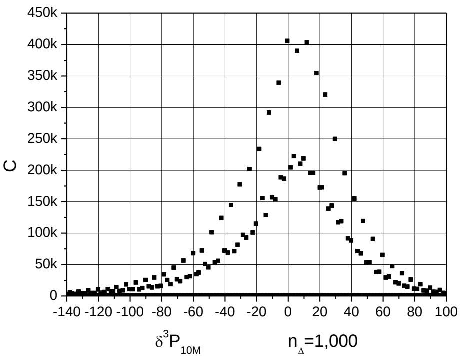


Fig. 31. Scatter-plots of δ^3P_{10M} n_Δ=1000

In this case (δ^3P_{10M} for many n_Δ) the number of scatter-plots i.e. δ^3 clusters seems to be partially limited to ~ 5 as a maximum, afterwards persisting with increasing n_Δ up to the max value examined n_Δ = 3,400 (not shown). Again, here too there are many void intervals starting from n_Δ=500 and later increasing more and more - being this the evidence of narrower and narrower interval widths with no counts at all or just one count or few counts - until all the scatter-plots become sparser and sparser and all the scatter-plots vanish at n_Δ sufficiently high. Another interesting feature of this case is its evident skewness at any value of $n_\Delta \in [50, 3400]$ as already seen, though towards positive values.

Of course one can go on, as the Author has done, at least up to an endpoint discovering an entire “zoo” of possibilities.

An important comment has to be made at this point. The presence of the clusters of prime number gaps is presumably the evidence of analogue clusters in prime numbers themselves what is nothing but the experimental evidence of Dirichlet’s theorem: *If a and b are natural numbers so that (a,b)=1 then there are infinitely many primes of the form P_n=an+b.*

In concluding this section devoted to δ^iP_n it has to be remarked that there are still a lot of aspects, issues and facets to be examined more in depth than done insofar and this will be the topic of a future work. At the present stage of the study what is important is that a methodology has been established together with some significant and remarkable findings.

5.2 Linear gaps or delta lags of the k^th order Δ^kP_m

The next theme of this research is the statistical treatment of the other kind of prime finite differences that is the linear gaps $\Delta^kP_m = (P_m - P_{m-k})$ (k=2, 3, 4, 5, ………. ∈ N and k<m) or “delta lags” the initial values of which are reported in the earlier Table 3 and the number of cases examined are shown in the previous Tab. IV.

From such matrix $|\Delta^k P_m| \equiv |P_m - P_{m-k}|$ the values of the many prime finite differences of value k to which the statistical treatment can be applied are derived soon. All the cases of $\Delta^k P_m$ with n_Δ=50 examined by the author, though not all of them reported, show that the E-exp or exp-exp or extreme function (a SW built-in fit function) can well fit the scatter-plots.

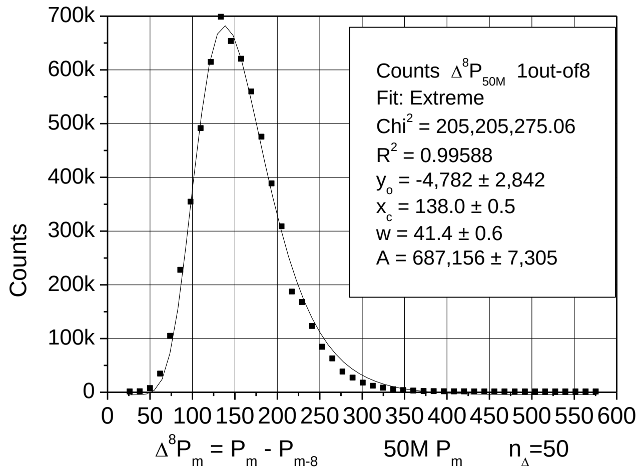


Fig. 32. Scatter-plot and fit histogram of $\Delta^8 P_{50M}$ 1 out of 8 n_Δ=50

The paradigmatic case of $\Delta^8 P_{50M}$ 1 out of 8 and n_Δ = 50 is shown in the previous Fig. 32 together with the features of the fit function i.e. the SW built-in distribution function E-exp or exp-exp is of the type $E(z)=E_o + A\cdot e^{-\exp(-z)-(z-1)}$ $z = (x-x_o) / w$ where E_o=E(0) is the baseline or offset, E_o+A=E_{Max} is the top value, w its width shown in the inset with the fit features.

As in the previous cases also here the examination has been extended to the values of n_Δ shown in the Figs. 33 (n_Δ=100), 34 (n_Δ=200), 35 (n_Δ=300), 36 (n_Δ=400), all the cases for 50M 1 value out of 8 and $\Delta^8 P_{50M}$.

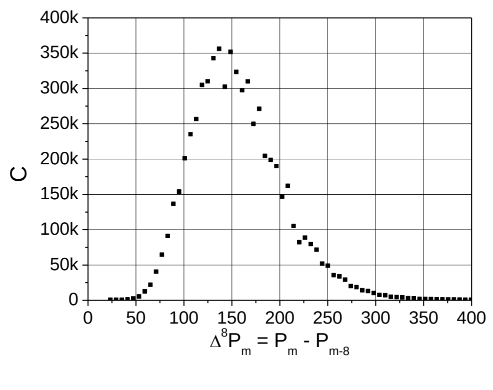


Fig. 33. Scatter-plot $\Delta^8 P_{50M}$ 1/8 n_Δ=100

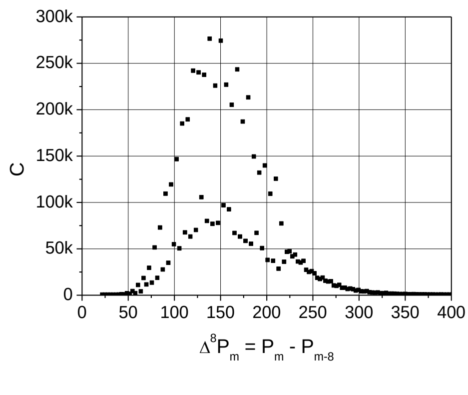
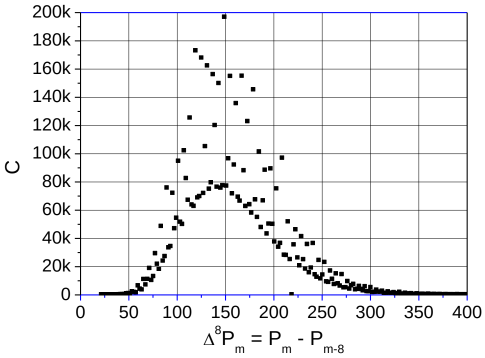


Fig. 34. Scatter-plots $\Delta^8 P_{50M}$ 1/8 n_Δ=200

The extension to many values of n_Δ has been performed in order to check, here again, the presence of $\Delta^k P_m$ clustering, an effect that seems typical of the prime finite differences both $\delta^i P_n$ and $\Delta^k P_m$ as well also of primes themselves P_m according to Dirichlet's theorem.

Also for these histograms (as in the previous δ^i cases) it can be conjectured that any inner histogram is best fitted by the same kind of function, an E-exp distribution function, having the same centre (in this case Δ^h_o=138.0±0.5) with different A and w.



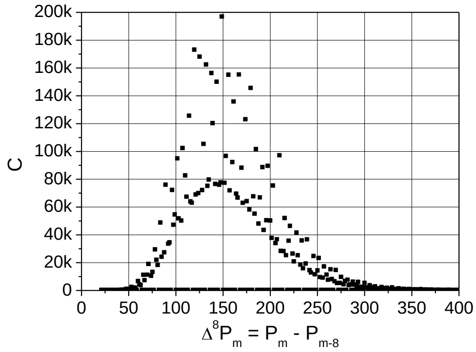
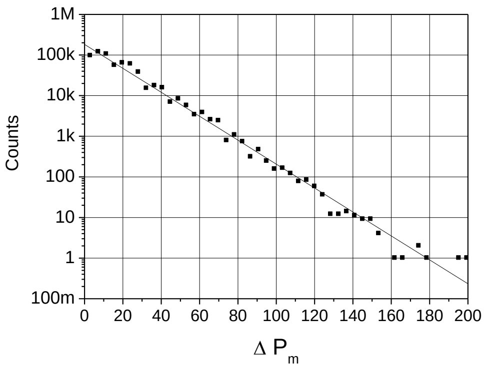


Fig.35. Scatter-plots $\Delta^8 P_{50M}$ 1/8 n_Δ=300 **Fig. 36. Scatter-plot $\Delta^8 P_{50M}$ 1/8 n_Δ=400**

The few cases described in the present study (just a part of the several cases treated) might be enough to draw some initial conclusions already reported, nonetheless much more remains to do in the field of experimental mathematics applied to prime numbers and to their gaps in order to understand and explain the whole matter. Here just some suggestions of investigations have been given to the viewpoint of both experimental and classical or theoretical mathematics.

5.3 First Order Gaps $\Delta P_m = \delta P_m$

Now it is time to face the problem of the first degree of prime number gaps, i.e. of the simple linear differences of primes $P_m - P_{m-1} = \delta P_m = \Delta P_m = \delta^1 P_m = \Delta^1 P_m$ a topic that has been left aside insofar. As a matter of fact the two Figs. 37 and 38 show the statistical behaviour of $\delta P_{5M} = \Delta P_{5M} = \delta^1 P_{5M} = \Delta^1 P_{5M}$ on a log-lin and full linear scale in the case m=5M and n_Δ=50



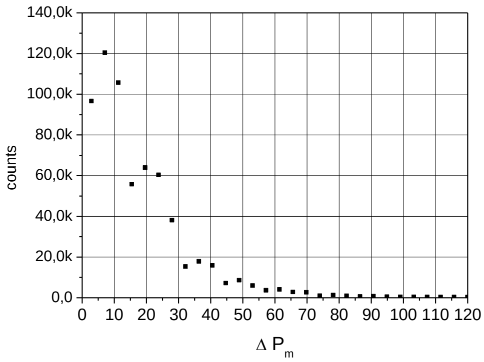


Fig. 37. Scatter-plot (log) and fit δP_{5M} n_Δ=50 **Fig 38. Scatter-plots (lin) δP_{5M} n_Δ=50**

The behaviour is completely different from that of $\delta^i P_n$ ($i \geq 2$) and $\Delta^h P_m$ ($h \geq 2$) being the best fit of the data points a seeming straight line (in the log case) that is a trend belonging neither to the first nor to the second type of gaps.

However the data plotted on a linear scale show a trend with more than one peak (maybe 2) of statistical distributions highly skewed towards the lower values of $\delta P_m = \Delta P_m$. In both figures the presence of data grouping or clustering is apparent. Despite that, a data fit can be tried in the log-lin plot leading to the equation $\log(C) = (5.26235 \pm 0.06123) - (0.02948 \pm 5.71E-4) \cdot \Delta P_{5M}$ with C=counts N=50 R=0.9924 σ=SD=0.20383 and p<1E–4

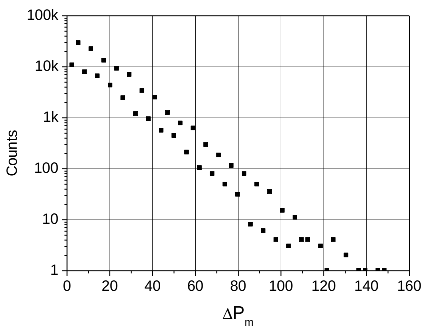


Fig. 39. Log scatter-plots of δP_{1M} n_Δ=50

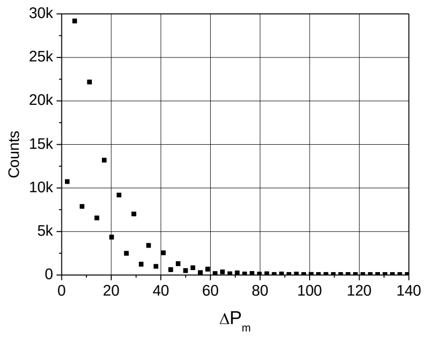


Fig. 40. Lin scatter-plots of δP_{1M} n_Δ=50

Again in all these cases it has been ascertained, beyond any reasonable doubt, that the spreads observed in the data points are to be ascribed just to the aforesaid data-point bunching or clustering i.e. to the appearance of innermost structures with increasing n_Δ that is the number of intervals of the whole range examined.

The latest two Figs. (39 and 40) show another example, even clearer, of this trend in the case of $\delta P_m=\Delta P_m$ for m=1M on a log-lin scale (Fig. 39) and on a lin-lin scale (Fig. 40) for $n_\Delta = 50$. In both these figures the presence of just two distinct statistical distributions or scatter-plots is clearly evident. These structures have been seen by other authors [35] too and this trend is clearly apparent in all the examined cases (Table 5) where there is the evident mark of their log-lin linearity.

One of the initial conclusions is that all the cases presented here surely are not exhaustive of the whole problem of prime number gaps though just symptomatic and paradigmatic of a situation that should be studied in depth and in detail having at hand many more data and many more powerful tools than at present.
Nonetheless what is astounding is the fact that the statistical distributions here reported and discussed appear new at all and never met before what contributes even more to consider prime numbers as “objects” fully atypical and uncommon in the field of number theory.

It will be up to the “classical” mathematicians the theoretical explanations of such trends and behaviours as already told. What has been done in this siege is to sketch the factual, current experimental situation of the problem giving a behaviouristic vision of prime numbers gaps and primes themselves.

5.4 Statistics of the finite sequences of prime numbers P_n

Now one should wonder whether it is possible that prime numbers themselves might show tendencies of clustering from the statistical viewpoint when changing the number of intervals n_Δ what means changing the width Δ of these intervals themselves. As a matter of fact some authors [42-46] have recognized and studied arithmetic progressions of primes within the prime themselves infinite sequence so that it would look correct to ask the aforesaid question also in view of Dirichlet’s theorem as already told. The following figures show that it is so and the effect got by changing n_Δ is the appearance of clusters again.

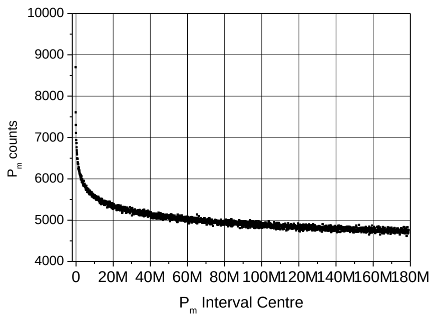


Fig. 41. Scatter-plot of P_{10M} n_Δ=2,000

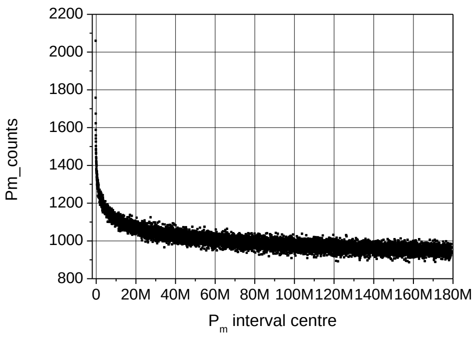
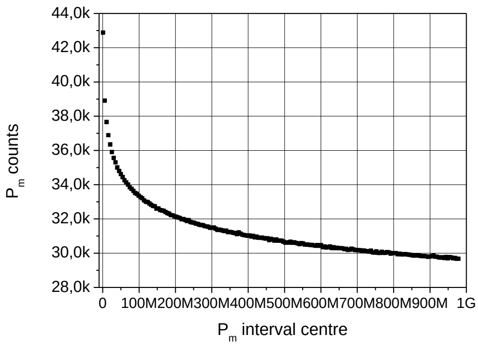


Fig. 42. Scatter-plots of P_{10M} n_Δ=10K

The previous Fig. 41 shows the scatter-plot of the first 10M primes (up to the value of P_{10M}= 179,424,673) for n_Δ=2,000 whilst the Fig. 42 illustrates the same case at n_Δ=10,000. There seems to be the clear evidence of clustering of prime numbers into a series of similar curves. of course for $n_\Delta \ll 2{,}000$ just one single well-defined curve is present [8].



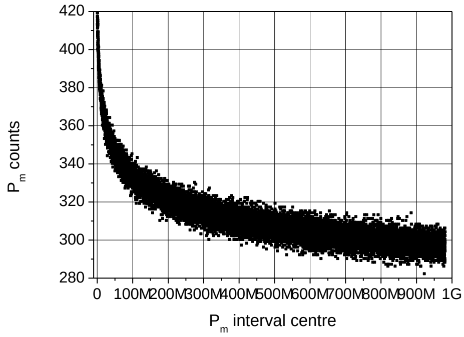


Fig. 43. Scatter-plots of P_{50M} 1/8 n_Δ=200

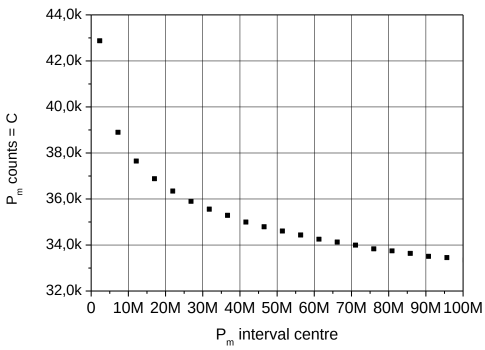
Fig. 44. Scatter-plots of P_{50M} 1/8 n_Δ=20K

An alike suggestion holds for the case m=50M (P_{50M}=982,451,653) 1 value out of 8 as reported in the previous two figures**:** the Fig. 43 for n_Δ=200 where just one single scatter-plot is apparent (fitted by a modified chi-square function with the ad-hoc value of all the parameters) and the Fig. 44 holding at n_Δ=20,000 showing many scatter-plots.

This behaviour is especially evident by zooming-in these two previous figures as shown in the next two figures 45 and 46.

The modified Chi-square function [15-18] is the best fit function for the single scatter-plot of Fig. 43 and it is to be conjectured that any scatter-plot of Fig. 44 is fitted by the same function $X^2(A,k,n/\mu)$ with the same value of A and k but different decay parameter μ. The same happens for any finite sequence of prime numbers P_n ∀n.

Though not shown, nonetheless it is interesting to get acquainted with the fact that even other variables show this effect of data point clustering or bunching into a multitude of similar curves first of all the prime frequencies $f_n=f(n)=n/P_n$ [16] so that such an effect seems typical



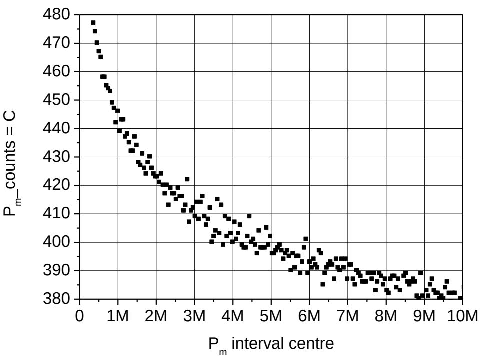


Fig. 45. Zoom-in of Fig. 43 [0, 100M]

Fig. 46. Zoom-in of Fig. 44 [0, 10M]

not only of the many types of prime number gaps but also of primes themselves. As already told that seems to be in accordance with Dirichet's theorem.

In addition some other variables reported in the list in Ch. 3 have been statistically examined (though not reported) and it has been ascertained that they show the same effect of data bunches.

6 Concluding Remarks and Future Perspectives

The use of computational mathematics in treating prime number gaps as experimental data is a powerful instrument leading to show what is the actual behaviour of prime numbers and their gaps as experimentally assessed as well as conducting to main results and findings among which the gathering of gaps (and primes themselves) into clusters or bunches.

The main conclusions of the research are:

1- The situation of prime number gaps is much more complex and complicate than ever thought in that they show unexpected features never supposed before. The spirit of this article has been to show experimentally the behaviour of prime number gaps actually and factually as never seen or even imagined before now.

2- There are many kinds of prime number gaps, different one from each other also according to their statistical behaviours, like**:**

1) the higher order gaps …….. $\delta^i P_n = \sum_{k=0\rightarrow i}(-1)^k \binom{i}{k} P_{n-k}$

2) the linear gaps or delta-lags ……… $\Delta^k P_n = P_n - P_{n-k}$

3) the first order gaps …………... $\delta P_n = \Delta P_n = P_n - P_{n-1}$

3- The scatter-plot of any of these three kinds of gaps is fitted by a different DDF i.e. respectively**:**

1) Voigt 2) E-exp 3) log-linear.

4- In increasing the number of n_Δ intervals in the statistical treatments some inner structures, i.e. clusters or bunches, appear in all the kinds of gaps as well as in prime number themselves (a probable consequence of Dirichlet’s theorem).

5- It will be up to future studies as well as to the classical, i.e. theoretical, mathematicians the deepening of this entire study and the explanation of all these behaviours.

Further comments can be made which might be useful in the future when developing and enlarging the whole research and study.

1- As already told the Gaussian function can be regarded just as a modified chi-square function with very high values of the parameter k (i.e. k>>30-40) so it can be deduced that also in the cases of $\delta^i P_n$ with i even one faces again the above-said modified X_k^2 function with the ad-hoc values of its parameters.

2- Though it has to be verified, nonetheless it can be conjectured that for the finite differences too $\Delta^h P_m$ one can face again the modified chi-square function $X_k^2(A,m/x_o)$ with k>2 as the fit function, maybe re-normalized.

3- As for the appearance of clusters in the cases of prime gaps it is obvious that the presence of Lorentz distribution function is simply a sufficient condition to the formation of such groupings though not necessary.

4- The entire trend of prime gaps with all their groupings depending on the value of n_Δ might be the clue of a chaotic behaviour (Poincare) characterized by “*strange attractors*” and “*stability islands*”. However all that must still be verified in depth especially having at one’s disposition many more data.

5- This research, though still in its first stage and to be furtherly enlarged and developed, can be of appropriate aid in future studies on prime numbers and prime gaps both from the theoretical and experimental viewpoint. Deeper studies could concern all the issues just now suggested as well as many others as derivable from the present context.

A concise though noteworthy further final remark is required.

Despite all the interesting results and conclusions, a more systematic and structured review of the whole matter is needed in that the entire subject at issue is so vast and deep to deserve many more resources at the present time not available to the Author who has been compelled to limit himself just to a bird’s-eye glance to the matter leaving any more detailed examination of all the single parts, issues and items to the next near and far researches and studies.

As for the future, one of the most important and interesting developments is the search for self-similarities, if any, in the statistical treatment of prime gaps, an issue that could contribute to shred an additional light on the matter.

Competing Interests

Author has declared that no competing interests exist.

References

[1] Hardy GH, Littlewood GE. Contributions to the theory of the riemann zeta-function and the theory of the distribution of primes. Acta Mathematic. 1916;41:119-196.
Available:http://dx.doi.org/10.1007%2FBF02422942

[2] Smith DE. 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; 1959.

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

[4] Derbyshire J. Prime obsession. bernhard riemann and the greatest unsolved problem in mathematics. Joseph Henry Press (National Academic Press), Washington D.C. 2003. ©2003 John Derbyshire. L'ossessione dei numeri primi – Bernhard Riemann e il principa- le problema irrisolto della matematica. ©2006 Bollati Boringhieri Editore s.r.l. Torino.

[5] Languasco A. and Zaccagnini A. Some properties of prime numbers, I and II" Website Bocconi-Pristem, 2005.
Available:http://matematica.uni-bocconi.it/LangZac/home.htm
https://matematica.unibocconi.eu/articoli/laltra-faccia-dei-numeri-primi-da-monolitici-monochords

[6] Goldstone DA, Pinz J, Yildirim CY. The path to recent progress on small gaps between primes. Clay Mathematics Proceedings Volume. 2007;7:125–135.

[7] Green B. © 2007 Long arithmetic progressions of primes. Clay Mathematics Proceedings. 2007;7:149–167.

[8] Oliveira e Silva T, Herzog S, Pardi S. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$.Math. Comp. 2014;83.
Available:https://doi.org/10.1090/S0025-5718-2013-02787-1

[9] Ford K, Green B, Konyagin S, Maynard J, Tao T. Long gaps between primes. J. Amer. Math. Soc. 2017;31:65–105.
Available:https://doi.org/10.1090/jams/876

[10] Carneiro E, Milinovich MB, Soundararajan K. Fourier optimization and prime gaps. Comment. Math. Helv. 2019;94:533–568.
Available:https://doi.org/10.4171/CMH/467

[11] Whittaker ET, Watson GN. A course of modern analysis, 5th edn. Cambridge University Press,Cambridge; 2021.
Available:https://doi.org/10.1017/9781009004091

[12] Brent RP, Platt DJ, Trudgian TS. The mean square of the error term in te prime number theorem. J. Number Theory. 2022;238:740–762.
Available:https://doi.org/10.1016/j.jnt.2021.09.016

[13] Stadlmann J. On the mean square gap between primes, arXiv:2212.10867, 2022;71.

[14] Erdos PL, Harcos G, Kharel SR, Maga P, Mezei TR, Toroczkai ZT. The sequence of prime gaps is graphic, Mathematische Annalen
Available:https://doi.org/10.1007/s00208-023- 02574-1

[15] Lattanzi D. Distribution of prime numbers by the modified chi-square function. Notes on Number Theory and Discrete Mathematics ISSN 1310-5132. 2015;21(1):18–30.
Available:https://nntdm.net/volume-21-2015/number-1/18-30/

[16] Lattanzi D. Scale laws of prime number frequencies by the modified chi-square function. Former British Journal of Mathematics & Computer Science, now Journal of Advances in Mathematics and Computer Science, 2016;13(6):1-21. Article n_o. BJMCS.23200, ISSN:2231-0851, Sciencedomain international.
Available:https://doi.org/10.9734/BJMCS/2016/23200

[17] Lattanzi D. An elementary proof of riemann's hypothesis by the modified chi- square function, Former British Journal of Mathematics & Computer Science, now Journal of Advances in Mathematics and Computer Science. 2016;15(5):1-14. Article n_o BJMCS.25419 ISSN: 2231-0851 Sciencedomain International.
Available:https://doi.org/10.9734/BJMCS/2016/25419

[18] Lattanzi D. Computational model of prime numbers by the modified chi-square function, Former British Journal of Mathematics & Computer Science, now Journal of Advances in Mathematics & Computer Science. 2017;20(5):1-19. Article n_o.BJMCS.31589, ISSN: 2231-851. Scencedomain International
Available:https://doi.org/10.9734/BJMCS/2017/31589

[19] Lattanzi D. Computer simulation model of prime numbers, Former British Journal of Mathematics & Computer Science, now Journal of Advances in Mathematics and Computer Science. 2023;38(8):101-121. Article n_o. JAMCS. 100586, ISSN: 2456-9968, Scencedomain International
Available:https://doi.org/10.9734/jamcs/2023/v38i81794

[20] Epstein D, Levy S. Experimentation and Proof in Mathematics, Notices of the AMS. N° 6, June 1995;42:670-674,.

[21] Brown JR. Philosophy of mathematics – A contemporary introduction to the world of proofs and pictures, 2nd edition, routledge, Taylor & Francis Group, New York and London; 2008.

[22] Porter MA, Zabusky NJ, Hu B, Campbell DK. Fermi, Pasta, Ulam and the Birth of Experimental Mathematics, American Scientist, Volume 97 © 2009 Sigma Xi, The Scientific Research Society. 2009;214-221.

[23] Andeberhan T, Medina LA, Moll VH. Editors, contemporary mathematics - 517 -Gems in Experimental Mathematics. AMS Special Session, Experimental Mathematics, Washington D.C. American Mathematical Society; January 5, 2009.

[24] Bailey DH, Borwein JM. Exploratory experimentation and computation, LBNL, Paper LBNL-3313E Lawrence Berkeley National Laboratory; Notices of the AMS Vol. 58, N° 10, November 2011;1410-1419.

[25] Borwein J, Borwein P, Girgensohn R, Parnes S. Making sense of experimental mathematics, The Mathematical Intelligencer, Springer. 2009;18(4):12-18
Available:http://dx.doi.org/10.1007/BF03026747

[26] Napoletani D, Panza M, Struppa DC. Is big data enough? A reflection on the changing role of mathematics in applications. Notices of the American Mathemat- ical Society. 2014;61(5):485-490.
Retrieved from:http://www.ams.org/notices/201405/rnoti-p485.pdf

[27] Baker A. Non-deductive methods in mathematics, The Stanford Encyclopedia of Philosophy (Summer Edition), Edward N. Zalta (ed.);2020.
Available:https://plato.stanford.edu/archives/sum2020/entries/mathematics-nondeductive

[28] Anonymous
Available:http://primes.utm.edu/lists/small/millions

[29] Anonymous
Available:http://www.bigprimes.net/cruncher/

[30] Gauss CF. Third proof of the law of quadratic reciprocity, in Smith D.E. ASource Book in Mathematics, Dover Publications Inc. New York, unabridged republic- ation of the 1rst edition, originally published in 1929 by McGraw-Hill Book Co., Inc; 1959.

[31] Mathews J, Walker RL. – CalTech, Mathematical Methods of Physics ©1964 W.A. Benjamin Inc. New York N.Y.

[32] Morice E. Dizionario di statistica ©1971 by ISEDI. Milano. Dictionnaire de Statistique ©1967 by Dunod, Paris.

[33] Walck C. Hand-book on statistical distributions for experimentalists, Internal Report SUF-PFY/96-01, Stockolm, last modification 10 September 2007, University of Stockolm; 11 December 1996.

[34] Babusci D, Dattoli G, Del Franco M. Lectures on mathematical methods for physics RT/2010/58/ENEA, ENEA-Roma-I; 2010.

[35] Soundararajan K. Small Gaps between Prime Numbers: the Work of Goldston-Pintz- Yildirim Bulletin (New Series) of the American Mathematical Society. January, 2007;44(1):1–18, S 0273-0979(06)01142-6 Article electronically published on September 25, 2006.

[36] Tao T. Small and large gaps in the primes University of California, Los Angeles, Latinos in the Mathematical Sciences Conference; Apr 9, 2015.

[37] Visser R. Large Gaps Between Primes Essay setter: Dr Thomas Bloom Part III Essay, University of Cambridge; 2020.

[38] Matsushita R, Da Silva, S. A Power Law Governing Prime Gaps. Open Access Library Journal. 2016;3:1-6.
DOI:10.4236/oalib.1102989.

[39] Anonymous
Available:http://originlab.com/www/helponline/origin/en/UserGuide/Extreme.html

[40] Fubini A, Alberini R, Lattanzi D. ^{203}Tl(n,γ) Reaction and Level Structure of ^{204}Tl, Il Nuovo Cimento (S.I.F.) Serie, 21 and references therein. Dicembre 1973;11(18):711-725.
Available:http://link.springer.com/article/10.1007%2FBF02727587#

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

[42] Cox CD, Van Der Poorten J. On a sequence of prime numbers, J. Austral. Math. Soc. 1968;8:571-574.

[43] Dahmen SR, Prado SD, Stuermer-Daitx T. Similarity in the statistics of prime numbers Physica A 296 2001;523–528. Elsevier,
Available:www.elsevier.com/locate/physa

[44] Granville A, Martin GG. Prime number races the american mathematical monthly. published by Mathematical Association of America. Jan. 2006;113(1):1-33.

[45] Granville A. Prime Number Patterns, The American Mathematical Monthly. Apr. published by Mathematical Association of America. 2008;115(4):279-296.

[46] Jeong S, Lee G, Kim G. Statistical and structural analysis of the appearance of prime numbers, J. Appl. Math. Comp© Korean Society for Computational and Applied Mathematics 2012. 2013;41:283-299. DOI: 10.1007/s12190-012-0601-9

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/112293