SINGLY AND DOUBLY EVEN MULTIPLES OF 6 AND STATISTICAL BIASES IN THE DISTRIBUTION OF PRIMES

WALDEMAR PUSZKARZ

Abstract. Computer experiments show that singly even multiples of 6 surrounded by prime pairs exhibit a larger ratio of nonsquarefree to squarefree multiples than generic singly even multiples of 6, a bias of ca 10.6% measured against the expected value. The same bias occurs for isolated primes next to singly even multiples of 6; here the deviation from the expected value is ca 3.3% of this value. The expected value of the ratio of singly even to doubly even nonsquarefree multiples of 6 also differs from values found experimentally for prime pairs centered on such multiples or isolated primes next to them. For pairs, this ratio exceeds its unbiased value by ca 6.2%, for isolated primes by ca 2.0%. The values cited are for the first 10^{10} primes, the largest range we investigated. This paper broadens our recent study of a newly found bias in the distribution of primes by examining singly and doubly even multiples of 6. In particular, it shows that for primes centered on or next to singly even multiples of 6, the statistical biases in question are more pronounced than in the general case studied by us before.

1. Introduction

In our recent work [1] we discovered that primes, that, with the exception of 2 and 3, always occur by multiples of 6, tend to be drawn in excess to nonsquarefree multiples of 6 compared to their squarefree counterparts. More precisely, which is key here, in excess to what would be expected in a non-biased distribution. The effect is more pronounced for prime pairs than for isolated primes, but it is fairly large in both cases and appears to be persistent and steady over a large range of primes, up to the 10^{10} first primes that we examined.

This persistence is particularly striking for this kind of phenomenon, being in sharp contrast, for instance, to the Chebyshev effect, another bias in the distribution of primes, much better known for it was discovered already in 1853 and has been studied ever since. The Chebyshev bias (see, e.g., [4]) manifest itself in that primes of the form $4k + 3$ are slightly more common than those of the competing form, $4k + 1$. The absolute difference between the numbers of primes in these two classes grows with the number of primes, but too slowly for this effect to remain steady in relative terms.

We found out our bias by examining the ratio of the number of nonsquarefree to the number of squarefree multiples of 6. The ratio exact value can be obtained theoretically. Comparing this theoretical value to the one found experimentally for the multiples of 6 that attract primes, we noticed that they differ quite markedly:

Date: December 21, 2021.

2010 *Mathematics Subject Classification.* 11N05.

Key words and phrases. primes, twin primes, squarefree numbers, nonsquarefree numbers, even numbers, statistical bias.

the actual ratio was bigger, indicating that primes favor nonsquarefree multiples of 6 over their squarefree counterparts.

In this paper, we study the same ratio too, but in the context of singly even multiples of 6. Surprisingly, in this case, the bias, measured in terms of this ratio, is even greater than the one uncovered before [1] for generic multiples of 6. In this situation, the ratio in question too can be established theoretically (or found numerically with very good accuracy) and compared to the one that is actually manifested by singly even multiples of 6 occurring next to primes.

To be more exact, we study two ratios in this paper, or, more exactly still, two sets of ratios: for two classes of primes, twins and isolated ones, as explained below. However, the one mentioned above can be considered of principal interest.

The other ratio studied here is novel: it is that of the number of singly even to the number of doubly even nonsquarefree multiples of 6. It turns out that the value of this ratio determined experimentally in samples of primes near multiples of 6 is also noticeably larger than the one established theoretically for this case.

As in our previous work [1], we concentrate on prime pairs and isolated primes separately. As was the case there, here too we see a marked difference in the effect magnitude, measured in the ratio deviation from its expected value, between these two classes of primes. The difference in magnitudes is about 3 to 1 for the prime pairs, just as in the other effect. There is little doubt that these effects are related, which means that the present work can be seen as an extension of our 2018 study. It is advisable to read [1] to better understand this work.

The rest of the paper is organized as follows. The next section presents the data describing the phenomena studied here along with a computer code that can be used to generate it. The data analysis, with its main focus on the ratios discussed above, is conducted in yet another section, which is followed by conclusion, where some comparisons of our effect to other effects of this kind studied in the literature on the distribution of primes are drawn.

2. Data and Code

In this section, we present and explain the data from computer experiments we conducted using a free number theory software, PARI/GP [2], whose sample code that generates the data is also provided below. The data is collected in four tables. Tables 1 and 3 collect the data for twin primes, while tables 3 and 4 for isolated primes.

Data

Number of singly even multiples of 6 (squarefree and nonsquarefree) inside prime pairs						
						Legend
range exp n	6	7	8	9	10	(1)
sf	25113	215732	1897137	16944418	153121114	(2)
nsf	17896	153723	1351742	12086572	109243899	(3)
exper nsf/sf ratio	0.7126	0.7126	0.7125	0.7133	0.7134	(4)
theor nsf/sf ratio	0.6449	0.6449	0.6449	0.6449	0.6449	(5)
pct deviation	10.49	10.49	10.48	10.6	10.62	(6)
Table 1						

Number of singly even multiples of 6 (squarefree and nonsquarefree) next to isolated primes of the form 6k+1 and 6k+5						
						Legend
range exp n	6	7	8	9	10	(1)
sfnext1	124533	1279931	13060628	132637107	1342701571	(2)
sfnext5	124538	1280277	13062981	132638438	1342703372	(3)
nsfnext1	82548	850810	8690958	88334358	894944193	(4)
nsfnext5	82512	850412	8688616	88331784	894934925	(5)
total sf	249071	2560208	26123609	265275545	2685404943	(6)
total nsf	165060	1701222	17379574	176666142	1789879118	(7)
exper nsf/sf ratio	0.6627	0.6645	0.6653	0.666	0.6665	(8)
theor nsf/sf ratio	0.6449	0.6449	0.6449	0.6449	0.6449	(9)
pct deviation	2.76	3.03	3.16	3.26	3.35	(10)
Table 2						

Number of singly even and doubly even nonsquarefree multiples of 6 inside prime pairs						
						Legend
range exp n	6	7	8	9	10	(1)
se	17896	153723	1351742	12086572	109243899	(2)
de	43018	369142	3248528	29016190	262368498	(3)
exper se/de ratio	0.416	0.4164	0.4161	0.4165	0.4164	(4)
theor se/de ratio	0.3921	0.3921	0.3921	0.3921	0.3921	(5)
pct deviation	6.106	6.213	6.131	6.242	6.199	(6)
Table 3						

Number of singly even and doubly even nonsquarefree multiples of 6 next to isolated primes of the form 6k+1 and 6k+5						
						Legend
range exp n	6	7	8	9	10	(1)
senext1	82548	850810	8690958	88334358	894944193	(2)
senext5	82512	850412	8688616	88331784	894934925	(3)
denext1	206722	2130167	21749745	220979430	2237613820	(4)
denext5	207093	2131209	21752258	220984523	2237635097	(5)
total se	165060	1701222	17379574	176666142	1789879118	(6)
total de	413815	4261376	43502003	441963953	4475248917	(7)
exper se/de ratio	0.3989	0.3992	0.3995	0.3997	0.4	(8)
theor se/de ratio	0.3921	0.3921	0.3921	0.3921	0.3921	(9)
pct deviation	1.73	1.82	1.9	1.95	2.01	(10)
Table 4						

The last column of each table is labeled Legend. It contains references - (1) through (6) for twin primes and (1) through (10) for isolated primes - to the contents of tables' rows. Using these references, we will now briefly explain these contents, starting from the common to all tables.

The top data row in each table is the same. Referenced by (1), it holds the exponents n of powers of 10 that represent the total number of consecutive primes, a certain range, from which data samples were collected. The smallest sample draws the data from the first 10^6 primes, with 6 being the exponent in question, the largest one goes up to the first 10^{10} primes.

The last three bottom rows are also very much the same in that they contain the final results. More precisely, the values of some ratios, both theoretical (expected) and experimental, the latter based on the data in higher rows. We choose to keep the numerical values of these ratios up to 4 decimal digits, but in the analysis in the next section we round them off them to first three leading digits.

The last row contains the measure of deviation, the percent deviation, which informs how much the experimental ratio in question differs from its expected value. It is the ratio of the difference between the values of the experimental ratio and the theoretical one to the latter, this fraction expressed in percents. This number is also the measure of bias(es) that we study in the paper.

Other rows of the tables are specific to each table, so let us now explain them for each table individually, using the row references. The whole numbers they collect per each individual data sample (prime range) differ as the ranges expand, but the ratios in final rows change rather little.

Table 1.
(2) - sf - number of squarefee multiples of 6 that attract twin primes,
(3) - nsf - number of nonsquarefee multiples of 6 that attract twin primes.

Table 2.
(2) - sfnext1 - number of squarefree multiples of 6 that attract isolated primes of the form $6k + 1$,
(3) - sfnext5 - number of squarefree multiples of 6 that attract isolated primes of the form $6k + 5$,
(4) - nsfnext1 - number of nonsquarefree multiples of 6 that attract isolated primes of the form $6k + 1$,
(5) - nsfnext5 - number of nonsquarefree multiples of 6 that attract isolated primes of the form $6k + 5$,
(6) - total sf - total number of squarefree multiples of 6 in (2) - (3),
(7) - total nsf - total number of nonsquarefree multiples of 6 in (4) - (5).

Table 3.
(2) - se - number of singly even multiples of 6 that attract twin primes,
(3) - de - number of doubly multiples of 6 that attract twin primes.

Table 4.
(2) - senext1 - number of singly even multiples of 6 that attract isolated primes of the form $6k + 1$,
(3) - senext5 - number of singly even multiples of 6 that attract isolated primes of the form $6k + 5$,
(4) - denext1 - number of doubly even multiples of 6 that attract isolated primes of the form $6k + 1$,
(5) - denext5 - number of doubly even multiples of 6 that attract isolated primes of the form $6k + 5$,
(6) - total se - total number of singly even multiples of 6 in (2) - (3),
(7) - total de - total number of doubly multiples of 6 in (4) - (5).

Code

As mentioned, the above data was obtained using PARI/GP [2]. To verify it, we employed Wolfram Mathematica [3]. For the sake of completeness, a PARI/GP example code that one can use to produce the data for prime pairs (tables 1 and 3) is included below.

The first piece of the code calculates the numbers of squarefree and nonsquarefree singly even multiples of 6 that attract prime pairs, while the other the numbers of singly and doubly even nonsquarefree multiples of 6 that do so, as also reflected in the code comments.

For Table 1:

```
k=6; c1=0; c2=0; forprime(n=3, prime(10^k), isprime(n+2)&&
valuation(n+1,2)==1&&((issquarefree(n+1)&&c1++)||c2++));
 print1(c1, ", ", c2) \\c1, c2 = #'s of squarefree, nonsquarefree
```

For Table 3:

```
k=6; c1=0; c2=0; forprime(n=3, prime(10^k), isprime(n+2)&&
!issquarefree(n+1)&&((valuation(n+1,2)==1&&c1++)||c2++));
 print1(c1, ", ", c2) \\c1, c2 = #'s of singly, doubly even
```

For isolated primes, the code is more complex and a bit cumbersome to present here.[1] In principle, however, it is as easy as the above pieces that can be copied and pasted into your PARI/GP installation to produce the data for the first 10^k primes (k being 6 above) in virtually no time.

3. Analysis

In [1], we introduced a certain ratio, denoted by R, that allowed us to detect an anomalous distribution of primes with respect to the multiples of 6. To recall, R stands for the ratio of the number of nonsquarefree to the number of squarefree multiples of 6. In what follows we will refer to it as R^g, with the superscript g meaning generic. In this paper, we examine two ratios of this kind. We will denote them by R and S.

The former, as the notation suggests, is of the same nature as in [1], but not exactly the same. Not in principle and, as we will see below, not even close numerically. It cannot be the same as we now limit ourselves to a certain class of multiples of 6: singly even multiples of 6. It is only for this kind of multiples (as opposed to the doubly even ones), that we can calculate this ratio. For the other kind, it is not well defined as there are no doubly even multiples of 6 that would also be squarefree. Hence here, to spell it out, R is the ratio of the number of nonsquarefree to the number of squarefree singly even multiples of 6.

Recall that in the previously studied generic case, the expected (theoretical) value of this ratio was $R_0^g = \pi^2/3 - 1$ and its numerical value was **2.290** for practical purposes, by which we meant primarily rounding it off to the 3rd decimal digit. We use the subscript 0 to denote that we mean the expected value of this ratio. The real (experimental) value can, in principle, be and actually is different as we have found out. It is also the case in this study.

In the situation studied here, $R_0 = \pi^2/6 - 1$ and its numerical value is **0.645**, using the same practical rule for rounding it off.

The other ratio S is the ratio of the number of singly even to the number of doubly even nonsquarefree multiples of 6. Here we are concerned with nonsquarefree multiples of 6, comparing them for singly and doubly even such multiples. As noted above, when discussing R, we cannot form such a ratio for squarefree multiples of 6.

Its expected value is $S_0 = 1 - 6/\pi^2$ and rounded off the same way yields the value of **0.392**.

[1]We plan to provide both the data and the full code in a separate file to be available online.

It is against these values that we compare the distribution of primes with respect to singly (the case of R) as well as both singly and doubly even (the case of S) multiples of 6.

Let us focus on R first. Our findings for this case are presented in tables 1 and 2, the former containing both the data and results for prime pairs while the latter for isolated primes.

As we see from Table 1, there is quite a bit of difference between the experimental and theoretical values of R. Moreover, this difference grows steadily, as the range (number) of primes studied grows too. The growth seems asymptotic for it clearly decelerates. For the largest range we examined (the first 10^{10} primes), the largest value of R turned out to be **0.713**, leading to the relative difference between the experimental and theoretical values of R of 10.62%, quite a difference indeed, bigger than in the generic case where it is only 6.0%.

From Table 2, a similar pattern emerges, although the growth in R here is visibly stronger for it does not seem to show clear signs of slowing down. Yet, the difference between the experimental and theoretical values is smaller than in the case of twin primes: only **0.667** for the first 10^{10} primes, which amounts to a 3.35% difference between the experimental and theoretical values. However, this difference is, again, bigger than in the generic case where it is only 1.9%.

Let us now examine S, for which our findings are collected in tables 3 and 4, the former containing both the data and results for prime pairs while the latter for isolated primes.

The prime pairs present a more interesting case of the two in that S is not growing steadily here, but shows pretty regular oscillations [2]: its values are bigger for prime ranges with odd exponents. Yet, when prime ranges are considered separately with respect to their odd and even exponents, we see a steady growth again. The largest value of S is **0.392**, attained for the first 10^9 primes. This value is practically 6.2% larger than the expected value of S. It would be interesting to see if this pattern holds for another two range exponents. It is of course possible that this effect is spurious being entirely due to statistical fluctuations; a more thorough statistical analysis as well as more data is needed to fully resolve this issue.

The case of isolated primes is more in line with other cases studied here in that a steady growth is also quite apparent. For the largest prime range, we get **0.400** for the experimental value of S, which contrasted with S_0 gives rise to a 2.01% difference, also quite large by typical measures of biases studied in number theory.

Let us summarize our findings now.

What the data from our experiments reveals is that singly even multiples of 6 surrounded by prime pairs exhibit a larger ratio of the number of nonsquarefree to squarefree multiples than generic singly even multiples of 6. It can be as large as ca 10.6% for prime pairs and ca 3.3% for isolated primes measured against the expected value of this ratio, either case exhibiting a more pronounced effect than in the generic situation studied previously [1].

Moreover, the ratio of the number of the singly even to doubly even nonsquarefree multiples of 6 also differs markedly from those found experimentally for prime pairs centered on such multiples or isolated primes next to them. For pairs, this ratio exceeds its unbiased value by ca 6.2%, for isolated primes by ca 2.0%.

[2] Note that the identical values of S in two different places in Table 3 are the result of rounding off. They are really not equal as demonstrated by the values of the percent deviation.

4. Conclusion

Studying biases in the distribution of prime numbers has a pretty long tradition that can be traced back to at least the mid 19-th century when Pafnuty Chebyshev discovered his eponymous effect that we have alluded to in the introductory section. This effect, at first treated as not more than a mere curiosity, has over the years attracted a considerably greater attention, even from first rank mathematicians [4], [5].

In connection with this, let us direct our attention to Table 2. One sees in it that there are more primes of the form $6k + 5$ than of the form $6k + 1$ next to squarefree multiples of 6. Whereas here it is so only for singly even such multiples, it is also true in the generic case of all primes, which is another manifestation of the Chebyshev effect, but for another residue class than originally studied by Chebyshev. Notice, however, that the case of nonsquarefree multiples of 6 is totally reversed: the dominant species are primes of the form $6k + 1$. Now, while this is rather surprising, one observes this kind of squarefree vs nonsquarefree aspect also in the original Chebyshev bias, although not as dramatic (no reversal). This effect may not have been noted in the literature before, and we mention it here in passing only; a more thorough discussion of it is planned for another publication.

Related to the Chebyshev bias, more in spirit than in exact nature, is a phenomenon recently discovered by Lemke Oliver and Soundararajan [6]. While the Chebyshev effect deals with primes of the form $4k + 1$ and $4k + 3$, this effect is concerned with last digits of primes, i.e., with primes congruent to a number that typically represents the last digit of a prime modulo 10, and it is sometimes popularly presented as the repulsion of consecutive primes with the same last digits.

In this paper we have presented the results of research that examined yet another bias, a different kind than the two mentioned above, that may already deserve the status of classics among prime number biases, but similar to the one studied by us previously [1]. As was the case there, here too we have found large biases, which, apart from their size, share with the other bias one more feature - that of persistence. They are not only strong, but their relative magnitude does not decay as is the case with the biases for primes in congruence classes. It may actually grow, if slightly, as we noted in the previous section.

More research, especially from an analytical point of view, is certainly needed to understand this kind of phenomena in the distribution of prime numbers, still very novel having been discovered only in 2018. This is something to be addressed in our future papers on this topic.

Acknowledgements. The author is grateful to the developers of PARI/GP and Wolfram Mathematica whose software was indispensable to this research.

References

[1] Waldemar Puszkarz, *Statistical Bias in the Distribution of Prime Pairs and Isolated Primes*. arXiv:1807.00406 [math.NT].
[2] The PARI Group, PARI/GP, Univ. Bordeaux, `http://pari.math.u-bordeaux.fr/`.
[3] Mathematica, Wolfram Research Inc., `http://www.wolfram.com`.
[4] M. Rubinstein and P. Sarnak, *Chebyshev's bias*, Experiment. Math., 3(3):173–197, (1994).
[5] A. Granville and G. Martin, *Prime number races*, Amer. Math. Monthly, 113(1):1–33, (2006).

[6] R. J. Lemke Oliver and K. Soundararajan, *Unexpected biases in the distribution of consecutive primes*, Proc. Natl. Acad. Sci. USA 113, E4446–E4454, (2016) (also arXiv:1603.03720 [math.NT]).

Email address: psi_bar@yahoo.com