SOME PROBABILITIES CONCERNING PRIME GAPS

VLADIMIR SHEVELEV

Abstract. With help of the Cramér-like model, we give exact calculation of some probabilities concerning prime gaps.

1. Introduction

As well known, the Bertrand's postulate (1845) states that, for $x > 1$, always there exists a prime in interval $(x, 2x)$. This postulate very quickly-five years later- became a theorem due to Russian mathematician P.L.Chebyshev (cf., e.g., [9, Theorem 9.2]). In 1930 Hoheisel[3] proved that, for $x > x_0(\varepsilon)$, the interval $(x, x + x^{1-\frac{1}{33000}+\varepsilon}]$ always contains a prime. After that there were a large chain of improvements of the Hoheisel's result. Up to now, probably, the best known result belongs to Baker, Harman abd Pintz[1], who showed that even the interval $(x, x + x^{0.525})$ contains a prime. Their result is rather close to the best result which gives the Riemann hypothesis: $p_{n+1} - p_n = O(\sqrt{p_n} \ln p_n)$ (cf. [4, p.299]), but still very far from the Cramér's 1937 conjecture which states that already the interval $(x, x + (1+\varepsilon)\ln^2 x]$ contains a prime for sufficiently large x.

Everywhere below we understand that p_n is the n-th prime and $\mathbb{P}$ is the class of all increasing infinite sequences of primes. If $A \in \mathbb{P}$ then we denote $\mathcal{A}$ the event that prime p is in A. In particular, an important role in our constructions play the following sequences from $\mathbb{P}$: $A^{(i)}$ is the sequence of those primes p_k, for which the interval $(2p_k, 2p_{k+1})$ contains exactly i primes, $i = 1, 2, ...$; A_i is the sequence of those primes p_k, for which the interval $(2p_k, 2p_{k+1})$ contains at least i primes, $i = 1, 2, ...$. So

$$\mathcal{A}_1 = \sum_{i \geq 1} \mathcal{A}^{(i)}, \ \ \mathcal{A}_2 = \sum_{i \geq 2} \mathcal{A}^{(i)}, \ \ etc.$$

Let p be an odd prime. Let, furthermore, $p_n < p/2 < p_{n+1}$. According to the Bertrand's postulate, between $p/2$ and p there exists a prime. Therefore, $p_{n+1} \leq p$. Again, by the Bertrand's postulate, between p and $2p$ there exists a prime. More subtle question, that we study in this paper, is the following.

Problem 1. *Consider the sequence S of primes p possessing the property: if $p/2$ lies in the interval $(p_n, \ p_{n+1})$ then there exists a prime in the interval*

1991 Mathematics Subject Classification. Primary 11A41; secondary 11B05.

$(p, 2p_{n+1})$. *With what probability a random prime* q *belongs to* S *(or the event* $\mathcal{S}$ *does occur)?*

Two words about the structure of the paper. In Sections 2-5 and 8 we create the base for research Problem 1. In Section 6 we construct a sieve for selecting sequence S from all primes. In Section 7 we obtain a lower estimate for the probability of Problem 1. In Section 9 we prove a theorem on precise symmetry in the distribution of primes. Furthermore, in Section 10 we obtain our main results. Finally, in Section 11 we research in a similar style a generalization of Problem 1 when 2 is replaced by arbitrary real number $m > 1$.

2. Independent tests of large integers

Consider the Cramér model in a little modified form (cf. [13]). The principle, based on the fact that an odd number of size about n has two in $\ln n$ chances of being prime, is this:

The indicator function for the set of primes (that is, the function whose value at odd n *is 1 or 0 depending on whether* n *is prime or not) behaves roughly like a sequence of independent, Bernoulli random variable* $X(n)$ *with parameters* $2/\ln n (n \geq 9)$. *In other words, for* $n \geq 9$, *the random variable* $X(n)$ *takes the value 1 (n is 'prime') with probability* $2/\ln n$, *and* $X(n)$ *takes the value 0 (n is 'composite') with probability* $1 - 2/\ln n$. *For* completeness, *let us set* $X(1) = 0, X(3) = 1, X(5) = 1, X(7) = 1$.

As noticed Soundararajan [13], "this must be taken with a liberal dose of salt: a number is either prime or composite; probability does not enter the picture! Nevertheless, the Cramér model is very effective in predicting answers." Thus the Cramér approach consists of a possibility of application of his model to the prediction of the "usual" probability of an event $\mathcal{A}$ (we are writing $\mathrm{P}(\mathcal{A})$). Let us use the Cramér model to predict, for large k, the probability of the event that the interval $(2p_k, 2p_{k+1})$ is free from primes. This probability is:

$$(1 - \frac{2}{\ln(2p_k + 1)})(1 - \frac{2}{\ln(2p_k + 3)})...(1 - \frac{2}{\ln(2p_{k+1} - 1)}).$$

Therefore, the probability of the event $\mathcal{A}_1(k)$ that p_k is in sequence $A_1 \in \mathbb{P}$, i.e. that the interval $(2p_k, 2p_{k+1})$ contains at least one prime, is

$$\mathrm{P}(\mathcal{A}_1(k)) \approx 1 - (1 - \frac{2}{\ln(2p_k + 1)})(1 - \frac{2}{\ln(2p_k + 3)})...(1 - \frac{2}{\ln(2p_{k+1} - 1)}).$$

If it is known, that the interval $(2p_k, 2p_{k+1})$ contains already a prime v, what is the probability that this interval is free from the different from v primes? If $v = 2p_k + 2i + 1,\ 0 \le i \le p_{k+1} - 2p_k - 1$, then this probability is:

$$(1 - \frac{2}{\ln(2p_k + 1)})...(1 - \frac{2}{\ln(2p_k + 2i - 1)})\cdot$$
$$\cdot(1 - \frac{2}{\ln(2p_k + 2i + 3)})...(1 - \frac{2}{\ln(2p_{k+1} - 1)}).$$

Therefore, we have

$$\mathrm{P}_{\mathcal{A}_1(k)}(\mathcal{A}_2(k)) \approx 1 - (1 - \frac{2}{\ln(2p_k + 1)})...(1 - \frac{2}{\ln(2p_k + 2i - 1)})\cdot$$
$$\cdot(1 - \frac{2}{\ln(2p_k + 2i + 3)})...(1 - \frac{2}{\ln(2p_{k+1} - 1)}).$$

Now we find

$$\mathrm{P}_{\mathcal{A}_1(k)}(\mathcal{A}_2(k)) - \mathrm{P}(\mathcal{A}_1(k)) \approx (1 - \frac{2}{\ln(2p_k + 1)}) \cdot ...$$
$$\cdot(1 - \frac{2}{\ln(2p_k + 2i - 1)})(1 - \frac{2}{\ln(2p_k + 2i + 3)})...(1 - \frac{2}{\ln(2p_{k+1} - 1)})\cdot$$
$$\cdot(1 - (1 - \frac{2}{\ln(2p_{k+1} + 2i + 1)})) \approx$$
$$(1 - \mathrm{P}_{\mathcal{A}_1(k)}(\mathcal{A}_2(k)))(\frac{2}{\ln(2p_{k+1} + 2i + 1)}),$$

i.e.

$$\mathrm{P}_{\mathcal{A}_1(k)}(\mathcal{A}_2(k)) = \mathrm{P}(\mathcal{A}_1(k)) + O(1/\ln k).$$

This means that

$$\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}_2) = \mathrm{P}(\mathcal{A}_1). \tag{2.1}$$

Multiplying (2.1) by $\mathrm{P}(\mathcal{A}_1)$, we have

$$\mathrm{P}(\mathcal{A}_1)\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}_2) = \mathrm{P}^2(\mathcal{A}_1)$$

and since $\mathcal{A}_1\mathcal{A}_2 = \mathcal{A}_2$, then

$$\mathrm{P}(\mathcal{A}_2) = \mathrm{P}^2(\mathcal{A}_1). \tag{2.2}$$

Furthermore, if it is known, that the interval $(2p_k, 2p_{k+1})$ contains already two primes $u < v$, what is the probability that this interval is free from the different from u and v primes? Let $u = 2p_k + 2i + 1,\ \ v = 2p_k + 2j + 1,\ \ 0 \le i < j \le p_{k+1} - 2p_k - 1$. Suppose that $2i+3 < 2j-1$ (the cases $2i+3 = 2j-1$, and $2i+3 = 2j+1$ are considered quite analogously). Then this probability is:

$$(1-\frac{2}{\ln(2p_k+1)})...(1-\frac{2}{\ln(2p_k+2i-1)})\cdot$$
$$\cdot(1-\frac{2}{\ln(2p_k+2i+3)})...(1-\frac{2}{\ln(2p_k+2j-1)})\cdot$$
$$\cdot(1-\frac{2}{\ln(2p_k+2j+3)})...(1-\frac{2}{\ln(2p_{k+1}-1)}).$$

Therefore, denoting the event $\mathcal{A}_3(k)$ that p_k is in sequence A_3, we have

$$\mathrm{P}_{\mathcal{A}_2(k)}(\mathcal{A}_3(k)) \approx 1-(1-\frac{2}{\ln(2p_k+1)})...(1-\frac{2}{\ln(2p_k+2i-1)})\cdot$$
$$\cdot(1-\frac{2}{\ln(2p_k+2i+3)})...(1-\frac{2}{\ln(2p_k+2j-1)})\cdot$$
$$\cdot(1-\frac{2}{\ln(2p_k+2j+3)})...(1-\frac{2}{\ln(2p_{k+1}-1)})$$

and again we find

$$\mathrm{P}_{\mathcal{A}_2(k)}(\mathcal{A}_3(k))-\mathrm{P}(\mathcal{A}_1(k)) \approx (1-\frac{2}{\ln(2p_k+1)})\cdot ...$$
$$\cdot(1-\frac{2}{\ln(2p_k+2i-1)})(1-\frac{2}{\ln(2p_k+2i+3)})...$$
$$\cdot(1-\frac{2}{\ln(2p_k+2j+3)})...(1-\frac{2}{\ln(2p_{k+1}-1)})\cdot$$
$$\cdot(1-(1-\frac{2}{\ln(2p_{k+1}+2i+1)})(1-\frac{2}{\ln(2p_{k+1}+2j+1)})) \approx$$
$$(1-P_{\mathcal{A}_2(k)}(\mathcal{A}_3(k)))(\frac{2}{\ln(2p_{k+1}+2i+1)}+\frac{2}{\ln(2p_{k+1}+2j+1)} -$$
$$\frac{4}{\ln(2p_{k+1}+2i+1)\ln(2p_{k+1}+2j+1)}).$$

Thus together with (2.1), (2.2) we have

$$P_{\mathcal{A}_2}(\mathcal{A}_3)=P(\mathcal{A}_1). \tag{2.3}$$

Multiplying (2.3) by $\mathrm{P}(\mathcal{A}_2)$, we have

$$\mathrm{P}(\mathcal{A}_2)\mathrm{P}_{\mathcal{A}_2}(\mathcal{A}_3)=\mathrm{P}(\mathcal{A}_2)\mathrm{P}(\mathcal{A}_1)$$

and since $\mathcal{A}_2\mathcal{A}_3=\mathcal{A}_3$, then using (2.2) we find

$$\mathrm{P}(\mathcal{A}_3)=\mathrm{P}^3(\mathcal{A}_1). \tag{2.4}$$

Note also, that

$$\mathrm{P}(\mathcal{A}_1)\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}_3)=\mathrm{P}(\mathcal{A}_1\mathcal{A}_3)=\mathrm{P}(\mathcal{A}_3)=\mathrm{P}^3(\mathcal{A}_1)$$

and thus

$$\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}_3)=\mathrm{P}^2(\mathcal{A}_1). \tag{2.5}$$

Continuing these arguments, and noticing that

$$\mathrm{P}(\mathcal{A}^{(h)}) = \mathrm{P}(\mathcal{A}_h) - \mathrm{P}(\mathcal{A}_{h-1}),$$
$$\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}^{(h)}) = \mathrm{P}_{\mathcal{A}_1}(\mathcal{A}_h) - \mathrm{P}_{\mathcal{A}_1}(\mathcal{A}_{h-1}),$$

we obtain the following results.

Proposition 1. *For every integer $h \geq 2$, we have*

$$\mathrm{P}(\mathcal{A}_h) = \mathrm{P}^h(\mathcal{A}_1);$$
$$P_{\mathcal{A}_1}(\mathcal{A}_h) = P^{h-1}(\mathcal{A}_1);$$
$$\mathrm{P}(\mathcal{A}^{(h)}) = (1 - \mathrm{P}(\mathcal{A}_1))\mathrm{P}^h(\mathcal{A}_1).$$
$$\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}^{(h)}) = (1 - \mathrm{P}(\mathcal{A}_1))\mathrm{P}^{h-1}(\mathcal{A}_1).$$

Let $A \in \mathbb{P}$. The primes from A we call $A-$*primes*. Let A has the counting function $\pi_A(x)$ of its terms not exceeding x and suppose that there exists the limit

$$\mathrm{P}(A) := \lim_{n\to\infty} \frac{\pi_A(n)}{\pi(n)}. \tag{2.6}$$

By a Cramér-like principle, the indicator function for the set of $A-$primes respectively the set of all primes (that is, the function whose value at prime n is 1 or 0 depending on whether p_n is $A-$prime or not) behaves roughly like a sequence of independent, Bernoulli random variable $X_A(n)$ with parameter $\mathrm{P}(A)$. In other words, the random variable $X_A(n)$ takes the value 1 (p_n is $A-$prime) with probability $\mathrm{P}(A)$, and $X_A(n)$ takes the value 0 (p_n is not $A-$prime) with probability $1 - \mathrm{P}(A)$. Therefore, (2.3) one can consider as the equality

$$\mathrm{P}(A) = \mathrm{P}(\mathcal{A}), \tag{2.7}$$

where $\mathrm{P}(\mathcal{A})$ is the probability of the event that a large random prime p is $A-$prime.

Example 1. *Let $A \in \mathbb{P}$ be the sequence of primes from the arithmetic progression $\{an+b\}_{n\geq 0}$ with relatively prime integers a and b.*

It is well known that in this case $\pi_A(x) \sim \pi(x)/\varphi(a)$, as $x \to \infty$, where $\varphi(x)$ is the Euler's totient function. Thus in this case we have $p(A) = 1/\varphi(a)$.

Remark 1. *If sequence $A \in \mathbb{P}$ contains a subsequence A^* for which* $\mathrm{P}(A^*)$ *exists but it is unknown whether does exist* $\mathrm{P}(A)$, *then we use the notion of "lower probability"* ($\underline{\mathrm{P}}(\mathcal{A})) = \underline{\mathrm{P}}(A)$) *for estimate of the form*

$$\underline{\mathrm{P}}(\mathcal{A})=\underline{\mathrm{P}}(A):=\liminf_{n\to\infty}\frac{\pi_A(n)}{\pi(n)}\geq\lim_{n\to\infty}\frac{\pi_{A^*}(n)}{\pi(n)}=\mathrm{P}(\mathcal{A}^*).$$

3. Equivalence of two conditions for odd primes

Consider the following two conditions for primes:

Condition 1. *Let $p=p_n$, with $n>1$. Then all integers $(p+1)/2,(p+3)/2,...,(p_{n+1}-1)/2$ are composite numbers.*

Condition 2. *Let, for an odd prime p, we have $p_m<p/2<p_{m+1}$. Then the interval $(p,\ 2p_{m+1})$ contains a prime.*

Lemma 1. *Conditions 1 and 2 are equivalent.*

Proof. If Condition 1 is valid, then $p_{m+1}>(p_{n+1}-1)/2$, i.e. $p_{m+1}\geq(p_{n+1}+1)/2$. Thus $2p_{m+1}>p_{n+1}>p_n=p$, and Condition 2 is valid; conversely, if Condition 2 satisfies, i.e. $p_{m+1}>p/2$ and $2p_{m+1}>p_{n+1}>p=p_n$. If k is the least positive integer, such that $p_m<p_n/2<(p_n+k)/2<(p_{n+1}-1)/2$ and $(p_n+k)/2$ is prime, then $p_{m+1}=(p_n+k)/2$ and $p_{n+1}-1>p_n+k=2p_{m+1}>p_{n+1}$. Contradiction shows that Condition 1 is valid. ■

4. Ramanujan primes

In 1919 S. Ramanujan [7]-[8] unexpectedly gave a new short and elegant proof of the Bertrand's postulate. In his proof appeared a sequence of primes

$$2,11,17,29,41,47,59,67,71,97,101,107,127,149,151,167,... \tag{4.1}$$

For a long time, this important sequence was not presented in the Sloane's OEIS [9]. Only in 2005 J. Sondow published it in OEIS (sequence A104272).

Definition 1. *(J. Sondow*[10]*)For $n\geq1$, the* nth Ramanujan prime *is the smallest positive integer (R_n) with the property that if $x\geq R_n$, then $\pi(x)-\pi(x/2)\geq n$.*

In [11], J. Sondow obtained some estimates for R_n and, in particular, proved that, for every $n>1$, $R_n>p_{2n}$. Further, he proved that for $n\to\infty$, $R_n\sim p_{2n}$. From this, denoting $R\in\mathbb{P}$ the sequence of the Ramanujan primes, we have $R_{\pi_R(x)}\sim 2\pi_R(x)\ln\pi_R(x)$. Since $R_{\pi_R(x)}\leq x<R_{\pi_{R(x)}+1}$, then $x\sim p_{2\pi_R(x)}\sim 2\pi_R(x)\ln\pi_R(x)$, as $x\to\infty$, and we conclude that

$$\pi_R(x)\sim\frac{x}{2\ln x},\ \ \textit{or}\ \ \mathrm{P}(R)=1/2. \tag{4.2}$$

It is interesting that quite recently S. Laishram (see [10], comments to A104272) has proved a Sondow conjectural inequality $R_n < p_{3n}$ for every positive n.

5. Ramanujan primes satisfy conditions 1 and 2

Lemma 2. *If p is an odd Ramanujan prime, then Conditions 1 and 2 satisfy.*

Proof. In view of Lemma 1, it is sufficient to prove that Condition 1 satisfies. If Condition 1 does not satisfy, then suppose that $p_m = R_n < p_{m+1}$ and k is the least positive integer, such that $q = (p_m + k)/2$ is prime not more than $(p_{m+1} - 1)/2$. Thus

$$R_n = p_m < 2q < p_{m+1} - 1. \tag{5.1}$$

From Definition 1 it follows (cf.[12]) that, $R_n - 1$ is the maximal integer for which the equality

$$\pi(R_n - 1) - \pi((R_n - 1)/2) = n - 1 \tag{5.2}$$

holds. However, according to (5.1), $\pi(2q) = \pi(R_n - 1) + 1$ and in view if the minimality of the prime q, in the interval $((R_n - 1)/2, q)$ there are not any prime. Thus $\pi(q) = \pi((R_n - 1)/2) + 1$ and

$$\pi(2q) - \pi(q) = \pi(R_n - 1) - \pi((R_n - 1)/2) = n - 1.$$

Since, by (5.1), $2q > R_n$, then this contradicts to the property of the maximality of R_n in (4). ■

Note that, there are non-Ramanujan primes which satisfy Conditions 1,2. We call them *pseudo-Ramanujan* primes, denoting the sequence of such primes R^*. The first R^*-primes are:

$$109, 137, 191, 197, 283, 521, ... \tag{5.3}$$

Definition 2. *We call a prime p an* $\mathbf{R}-$prime *if p satisfies Condition 1 (or, equivalently, Condition 2).*

Thus $\mathbf{R}-$prime is either Ramanujan or pseudo-Ramanujan prime. Then in Problem 1

$$S = \mathbf{R}. \tag{5.4}$$

Give a simple criterion for $\mathbf{R}-$primes.

Proposition 2. *p_n is $\mathbf{R}$-prime if and only if $\pi(\frac{p_n}{2}) = \pi(\frac{p_{n+1}}{2})$.*

Proof. 1) Let $\pi(\frac{p_n}{2})=\pi(\frac{p_{n+1}}{2})$ is valid. From this it follows that if $p_k < p_n/2 < p_{k+1}$, then between $p_n/2$ and $p_{n+1}/2$ there are not exist primes. Thus $p_{n+1}/2 < p_{k+1}$ as well. Therefore, we have $2p_k < p_n < p_{n+1} < 2p_{k+1}$, i.e. p_n is **R**-prime. Conversely, if p_n is **R**-prime, then $2p_k < p_n < p_{n+1} < 2p_{k+1}$, and $\pi(\frac{p_n}{2})=\pi(\frac{p_{n+1}}{2})$ is valid. ■

6. A sieve for selection R−primes from all primes

In this section we build a sieve for selection $\mathbf{R}-$primes from all primes. Recall that the Bertrand sequence $\{b(n)\}$ is defined as $b(1)=2$, and, for $n\geq 2,\ b(n)$ is the largest prime less than $2b(n-1)$ (see A006992 in [10]):

$$2,3,5,7,13,23,43,... \tag{6.1}$$

Put

$$B_1=\{b^{(1)}(n)\}=\{b(n)\}. \tag{6.2}$$

Further we build sequences $B_2=\{b^{(2)}(n)\}, B_3=\{b^{(3)}(n)\}, ...$ according the following inductive rule: if we have sequences $B_1,...,B_{k-1}$, let us consider the minimal prime $p^{(k)}\notin\bigcup_{i=1}^{k-1}B_i$. Then the sequence $\{b^{(k)}(n)\}$ is defined as $b^{(k)}(1)=p^{(k)}$, and, for $n\geq 2,\ b^{(k)}(n)$ is the largest prime less than $2b^{(k)}(n-1)$. So, we obtain consequently:

$$B_2=\{11,19,37,73,...\} \tag{6.3}$$

$$B_3=\{17,31,61,113,...\} \tag{6.4}$$

$$B_3=\{29,53,103,199,...\} \tag{6.5}$$

etc., such that, putting $p^{(1)}=2$, we obtain the sequence

$$\{p^{(k)}\}_{k\geq 1}=\{2,11,17,29,41,47,59,67,71,97,101,107,109,127,...\} \tag{6.6}$$

Sequence (6.6) coincides with sequence (4.1) of the Ramanujan primes up to the 12-th term, but the 13-th term of this sequence is 109 which is the first term of sequence (5.3) of the pseudo-Ramanujan primes.

Theorem 1. *For $n\geq 1$, we have*

$$p^{(n)}=\mathbf{R}_n \tag{6.7}$$

where $\mathbf{R}_n$ is the n-th $\mathbf{R}-$prime.

Proof. The least omitted prime in (6.1) is $p^{(2)}=11=\mathbf{R}_2$; the least omitted prime in the union of (6.2) and (6.3) is $p^{(3)}=17=\mathbf{R}_3$. We use the induction. Let we have already built primes

$$p^{(1)}=2,p^{(3)},...,p^{(n-1)}=\mathbf{R}_{n-1}.$$

Let q be the least prime which is omitted in the union $\bigcup_{i=1}^{n-1} B_i$, such that $q/2$ is in interval (p_m, p_{m+1}). According to our algorithm, q which is dropped should not be the largest prime in the interval $(p_{m+1}, 2p_{m+1})$. Then there are primes in the interval $q, 2p_{m+1})$; let r be one of them. Then we have $2p_m < q < r < 2p_{m+1}$. This means that q, in view of its minimality between the dropping primes more than $\mathbf{R}_{n-1} = p^{(n-1)}$, is the least $\mathbf{R}-$prime more than $\mathbf{R}_{n-1}$ and the least prime of the form $p^{(k)}$ more than $p^{(n-1)}$. Therefore, $q = p^{(n)} = \mathbf{R}_n$.■

Unfortunately the research of this sieve seems much more difficult than the research of the Eratosthenes one for primes. For example, the following question remains open.

Problem 2. *With help of the sieve of Theorem 1 to find a formula for the counting function of* $\mathbf{R}-$*primes not exceeding* x.

Therefore, we choose another way.

7. Lower estimate of probability of Problem 1

We start with the following result. For generality, we use the designations of Remark 1.

Lemma 3.

$$\underline{\mathrm{P}}(\mathbf{R}) \geq \frac{1}{2}.$$

Proof. Using (3), we have

$$\mathrm{P}(\mathbf{R}) \geq \lim_{n\to\infty} \pi_R(n)/\pi(n) = 1/2. \blacksquare$$

D. Berend [2] gave another very elegant proof of this lemma.

Second proof of Lemma 3. We saw that if the interval $(2p_m, 2p_{m+1})$ with odd p_m contains a prime p, then the interval $(p, 2p_{m+1})$ contains in turn a prime if and only if p is an $\mathbf{R}$-prime. Let $n \geq 7$. In the range from 7 up to n there are $\pi(n) - 3$ primes. Put

$$h = h(n) = \pi(n/2) - 2. \tag{7.1}$$

Then $p_{h+2} \leq n/2$ and interval $(p_{h+2}, n/2]$ is free from primes. Look at h intervals:

$$(2p_2, 2p_3), (2p_3, 2p_4), ..., (2p_{h+1}, 2p_{h+2}). \tag{7.2}$$

Our $\pi(n) - 3$ primes are somehow distributed in these h intervals. Suppose $k = k(n)$ of these intervals contain at least one prime and $h - k$ contain no primes. Then for exactly k primes there is no primes between them and the next $2p_j$, and for the other $\pi(n)-3-k$ there is. Hence, among $\pi(n)-3$ primes exactly $\pi(n)-3-k$ are **R**-primes and exactly k non-RPR-primes. Therefore, since $k(n) \leq h(n) \leq \pi(n/2)$, then for the desired lower probability that there is a prime we have:

$$\underline{\mathrm{P}}(\mathbf{R}) = \liminf_{n\to\infty}\frac{\pi_{\mathbf{R}}(n)}{\pi(n)-3} = \liminf_{n\to\infty}\frac{\pi(n)-k(n)}{\pi(n)} \geq 1/2. \tag{7.3}$$

■

Remark 2. *One cannot confuse* **R***-primes and those primes* p_n *for which the interval* $(2p_n,\ 2p_{n+1})$ *contains an* **R***-prime* ($\mathbf{R}^*$*-primes). It is easy to see that* $\mathrm{P}(\mathbf{R}^*) = \mathrm{P}(\mathbf{R})\mathrm{P}(\mathcal{A}_1)$.

8. A symmetrical case of the left intervals

It is clear that for the symmetrical problem of the existence a prime in the left interval $(2p_n, p)$ (for the same condition $p_n < p/2 < p_{n+1}$) we have similar results. Indeed, now in our construction the role of the Ramanujan primes play other primes which appear in OEIS [10] earlier (2003) than the Ramanujan primes due to E. Labos (see sequence A080359):

$$2, 3, 13, 19, 31, 43, 53, 61, 71, 73, 101, 103, 109, 113, 139, 157, 173, ... \tag{8.1}$$

Definition 3. (cf. [9, A080359]) *For* $n \geq 1$, *the* nth Labos prime *is the smallest positive integer* (L_n) *for which* $\pi(L_n) - \pi(L_n/2) = n$.

The sequence (L) of such primes we call the *Labos primes.* Note that, since ([11])

$$\pi(R_n) - \pi(R_n/2) = n, \tag{8.2}$$

then, by the Definition 3, we have

$$L_n \leq R_n. \tag{8.3}$$

As in Section 3, one can prove the equivalence of the following conditions on primes:

Condition 3. *Let* $p = p_n$ *with* $n \geq 3$. *Then all integers* $(p-1)/2, (p-3)/2, ..., (p_{n-1}+1)/2$ *are composite numbers.*

Condition 4. *Let* $p_m < p/2 < p_{m+1}$. *Then the interval* $(2p_m,\ p)$ *contains a prime.*

Furthermore, by the same way as for Lemma 2, one can prove that if p is a Labos prime, then Conditions 3 and 4 satisfy. But again there are non-Labos primes which satisfy Conditions 3,4. We call them *pseudo-Labos* primes, or L^*-primes. The first terms of sequence L^* of such primes are:

$$131, 151, 229, 233, 311, 571, ... \tag{8.4}$$

Definition 4. *We call a prime p a* **L**-prime *if p satisfies Condition 3 (or, equivalently, Condition 4).*

From the above it follows that a **L**-prime is either Labos or pseudo-Labos prime. Besides, quite analogously to Proposition 2 we obtain the following criterium for **L**-primes.

Proposition 3. *p_n is* **L**-*prime if and only if* $\pi(\frac{p_{n-1}}{2}) = \pi(\frac{p_n}{2})$.

Suppose that the probability $\mathrm{P}(\mathcal{S})$ exists. Consider now the probability $\mathrm{P}(\mathcal{S}^*)$ of the event $\mathcal{S}^*$ that the left interval $(2p_n, p)$ contains a prime. From the symmetry (which is in the full concordance with the structure of the second proof of Lemma 3) we should conclude that $\mathrm{P}(\mathcal{S}) = \mathrm{P}(\mathcal{S}^*)$. Note that for the **L**-primes one can build a sieve with help of the Sloane's primes (see A055496 [10]) and the corresponding generalizations of them (cf. constructing in Section 6).

9. A theorem on precise symmetry in distribution of primes and proof of Theorem 2

Now we prove a much stronger statement about the symmetry, which connected with the mutual behaviors of sequences **R** and **L**, which satisfy to Conditions 1,2 and 3,4 correspondingly.

Theorem 2. *Let* $\mathbf{R}_n$ ($\mathbf{L}_n$) *denote the n-th term of the sequence* **R** (**L**). *Then we have*

$$\mathbf{R}_1 \le \mathbf{L}_1 \le \mathbf{R}_2 \le \mathbf{L}_2 \le ... \le \mathbf{R}_n \le \mathbf{L}_n \le ... \tag{9.1}$$

Proof. It is clear that the intervals of considered form, containing not more than one prime, contain neither **R**-primes nor **L**-primes. Moving such intervals, consider the first from the remaining ones. The first its prime is an **R**-prime ($\mathbf{R}_1$). If it has only two primes, then the second prime is an **L**-prime ($\mathbf{L}_1$), and we see that $(\mathbf{R}_1) < (\mathbf{L}_1)$; on the other hand if it has k primes, then beginning with the second one and up to the $(k-1)$-th we have **RL**-primes, i.e. primes which are simultaneously **R**-primes and **L**-primes. Thus, taking into account that the last prime is only **L**-prime , we have

$$\mathbf{R}_1 < \mathbf{L}_1 = \mathbf{R}_2 = \mathbf{L}_2 = \mathbf{R}_3 = ... = \mathbf{L}_{k-1} = \mathbf{R}_{k-1} < \mathbf{L}_k.$$

The second remaining interval begins with an **R**-prime and the process repeats. ■

It is interesting that this property, generally speaking, does not satisfy for proper Ramanujan and Labos primes, and the pseudo-Ramanujan and pseudo-Labos primes appear precisely in those places when this property of the enveloping is broken.

In connection with the considered problem and the corresponding "left" problem, it is natural to consider the following classification of primes: two first primes 2,3 form a separate set of primes; if $p \geq 5$ is a $\mathbf{R}\overline{\mathbf{L}}$-prime, then we call p a *right prime*; if $p \geq 5$ is a $\mathbf{L}\overline{\mathbf{R}}$-prime then we call p a *left prime*, while **RL**-primes we call the *central primes* . Finally, the rest primes it is natural to call *isolated primes* .

10. Solution of Problem 1 and calculating some close probabilities

Greg Martin [5] conjectured that $\mathrm{P}(\mathbf{R}) = 2/3$ and proposed the following heuristic arguments for that: "Imagine the following process: start from p and examine the numbers $p+1, p+2, ...$ in turn. If the number we're examining is odd, check if it's a prime: if so, we "win". If the number we're examining is twice an odd number (that is,2 (mod 4)), check if it's twice a prime: if so, we "lose". In this way we "win" if and only if there is a prime in the interval $(p, 2p_{n+1})$, since we either find such a prime when we "win" or else detect the endpoint $2p_{n+1}$, when we "lose".

Now if the primes were distributed totally randomly, then the probability of each odd number being prime would be the same(roughly $1/lnp$), while the probability of a 2 (mod 4) number being twice a prime would be roughly $1/ln(p/2)$, which for p large is about the same as $1/lnp$. However, in every block of 4 consecutive integers, we have two odd numbers that might be prime and only one 2 (mod 4) number that might be twice a prime. Therefore we expect that we "win" twice as often as we "lose", since the placement of primes should behave statistically randomly in the limit; in other words, we expect to "win" 2/3 of the time." His computions what happens for p among the first million primes show that the probability of "we win" has a steadily increasing trend as p increases, and among the first million primes about 61.2% of them have a prime in the interval $(p, 2p_{n+1})$.

Nevertheless, now we show that $P(\mathbf{R})$ *is some less than 2/3.*

Put

$$\lambda = \mathrm{P}(\mathcal{A}_1). \tag{10.1}$$

We need three lemmas.

Lemma 4.

$$\lambda = 2(1 - \mathrm{P}(\mathcal{S})). \tag{10.2}$$

Proof. Note that in the terms of the second proof of Lemma 3 we have

$$\lambda = \lim_{n\to\infty}\frac{\pi_{A_1}(n)}{\pi(n)} = \lim_{n\to\infty}\frac{k(n)}{h(n)}$$

and, moreover, from this proof, taking into account that $h(n) \sim \pi(n)/2$, we find

$$\lambda/2 = \lim_{n\to\infty}\frac{k(n)}{\pi(n)} = \lim_{n\to\infty}(\pi(n) - \pi_S(n))/\pi(n) = 1 - P(\mathcal{S}). \ \blacksquare$$

Corollary 1. *We have*

$$\mathrm{P}(\mathcal{S}) = \mathrm{P}(\mathbf{R}) = 1 - \frac{\lambda}{2}. \tag{10.3}$$

On the other hand, we prove the following.

Lemma 5. *The desired probability is*

$$\mathrm{P}(\mathbf{R}) = 1 + \frac{1-\lambda}{\lambda}\ln(1-\lambda). \tag{10.4}$$

Proof. Remove the intervals of the form $(2p_n,\ 2p_{n+1})$ which contain no primes. Let a random prime p lies in interval $(2p_n,\ 2p_{n+1})$. By Proposition 1, the conditional probability of the event that interval $(2p_n,\ 2p_{n+1})$ contains exactly k primes in the condition that it already contains at least one prime equals to

$$\mathrm{P}_{\mathcal{A}_1}(\mathcal{A}^{(k)}) = (1-\lambda)\lambda^{k-1},\ \ k = 2, 3, ... \tag{10.5}$$

Note that, in interval of the form $(2p_n,\ 2p_{n+1})$ every prime, except of the last prime, is $\mathbf{R}$-prime. Therefore, the desired probability is

$$\mathrm{P}(\mathbf{R}) = \sum_{k\geq 2} \frac{k-1}{k} \mathrm{P}_{\mathcal{A}_1}(\mathcal{A}^{(k)}) =$$

$$(1-\lambda) \sum_{k\geq 2} \frac{k-1}{k} \lambda^{k-1} = 1 + \frac{1-\lambda}{\lambda} \ln(1-\lambda). \blacksquare$$

Comparing (10.3) and (10.4), we obtain the following equation for λ :

$$(1-\lambda)\ln(1-\lambda) = -\frac{\lambda^2}{2}. \tag{10.6}$$

Lemma 6.

$$P(\mathbf{RL}) = 2(1-\lambda). \tag{10.7}$$

Proof. Indeed, the event $\mathcal{A}^{(k)}$, $k \geq 3$, contributes exactly $k-2$ equalities in (9.1), i.e. from k primes we have exactly $k-2$ central primes. Therefore, as in proof of the previous lemma, we have

$$\mathrm{P}(\mathbf{RL}) = (1-\lambda) \sum_{k\geq 3} \frac{k-2}{k} \lambda^{k-1},$$

and we easily follow that

$$\mathrm{P}(\mathbf{RL}) = 2 - \lambda + 2\frac{1-\lambda}{\lambda} \ln(1-\lambda).$$

Now, using (10.6), we obtain the lemma. $\blacksquare$

Using Corollary 1 and Lemma 6, we find the right (left) primes probability is

$$\mathrm{P}(\mathbf{R}\overline{\mathbf{L}}) = \mathrm{P}(\mathbf{L}\overline{\mathbf{R}}) = \mathrm{P}(\mathbf{R}) - \mathrm{P}(\mathbf{RL}) = 1 - \frac{\lambda}{2} - 2(1-\lambda) = \frac{3}{2}\lambda - 1.$$

Finally, if $\mathbf{I}$ denotes the sequence of isolated primes then

$$\mathrm{P}(\mathbf{I}) = 1 - \mathrm{P}(\mathbf{RL}) - 2\mathrm{P}(\mathbf{R}\overline{\mathbf{L}}) = 1 - 2(1-\lambda) - 2(\frac{3}{2}\lambda - 1) = 1 - \lambda.$$

Thus we obtain the following results.

Theorem 3.

$$\mathrm{P}(\mathcal{S}) = \mathrm{P}(\mathbf{R}) = 1 - \frac{\lambda}{2},$$

$$\mathrm{P}(\mathbf{R}\overline{\mathbf{L}}) = \mathrm{P}(\mathbf{L}\overline{\mathbf{R}}) = \frac{3}{2}\lambda - 1,$$

$$\mathrm{P}(\mathbf{RL}) = 2(1-\lambda), \mathrm{P}(\mathbf{I}) = 1 - \lambda.$$

where $\lambda \approx 0.8010$ *is only positive root of equation (10.6).*

Thus we have

$$\mathrm{P}(\mathcal{S}) = \mathrm{P}(\mathbf{R}) \approx 0.5995, \ \ \mathrm{P}(\mathbf{A}_1) \approx 0.8010;$$
$$\mathrm{P}(\mathbf{R}\overline{\mathbf{L}}) = \mathrm{P}(\mathbf{L}\overline{\mathbf{R}}) \approx 0.2015, \mathrm{P}(\mathbf{RL}) \approx 0.3980, \mathrm{P}(\mathbf{I}) \approx 0.1990.$$

Note that, according to Remark 2, $\mathrm{P}(\mathbf{R}^*) = \mathrm{P}(\mathbf{R})\mathrm{P}(\mathbf{A}_1) \approx 0.4802$. It is interesting also to note that the proportion of Ramanujan primes among **R**-primes is approximately 0.8340. Finally note that, by proof of Proposition 1, we see that a good numerical confirmation we expect to get for sufficiently large $\ln n$. Therefore, a Martin's numerical result 0.612 for $n = 10^6$ when $\ln n \approx 13.81$, we can accept as a very good confirmation of $\mathrm{P}(\mathbf{R}) \approx 0.5995$ with error a little more than 2%.

11. A generalization

In this section we consider a natural generalization of Problem 1 and the corresponding "left" problem.

Problem 3. *Given a real $m > 1$, consider the sequence S_m (S_m^*) of primes p possessing the property: if p/m lies in the interval $(p_n, \ p_{n+1})$ $((p_{n-1}, \ p_n))$ then there exists a prime in the interval (p, mp_{n+1}) $((mp_{n-1}, p))$. With what probability a random prime q belongs to $S_m(S_m^*)$?*

To study this problem, we introduce a natural generalization of Ramanujan primes.

Definition 5. *For real $m > 1$, we call a Ramanujan $m-$prime $R_n^{(m)}$ the smallest integer with the property that if $x \geq R_n^{(m)}$, then $\pi(x) - \pi(x/m) \geq n$.*

It is easy to see (cf. [11]) that $R_n^{(m)}$ is indeed a prime. Moreover, as in [11], one can prove that

$$R_n^{(m)} \sim p_{((m/(m-1))n)},$$

as n tends to the infinity. Denoting $R_m (\in \mathbb{P})$ the sequence of $m-$Ramanujan primes, we have (cf. (4.2))

$$\pi_R^{(m)}(x) \sim (1 - 1/m)\pi(x) \ \ or \ \ \mathrm{P}(\mathcal{R}^m) = 1 - 1/m. \tag{11.1}$$

Consider the corresponding "m-conditions" on primes.

Condition 5. *Let $p = p_n, \ \ n > 1$. Then the interval $(\lceil (p+1)/m \rceil, \ \ \lfloor (p_{n+1} - 1)/m \rfloor)$ is free from primes.*

Condition 6. *Let, for an odd prime p, we have $p_n < p/m < p_{n+1}$. Then the interval $(p, \ \ mp_{n+1})$ contains a prime.*

The following two lemmas are proved by the same way as Lemmas 1 and 2.

Lemma 7. *Conditions 3, 4 are equivalent.*

Lemma 8. *If p is an m-Ramanujan prime, then Condition 5 (or, equivalently, Condition 6) satisfies.*

Some later we prove the following statement.

Proposition 4. *For every $m > 1$ there exists an infinite sequence of nonm-Ramanujan primes which satisfy Condition 6.*

Such primes we call *pseudo-m-Ramanujan primes.* Since we cannot obtain empirically even the first pseudo-m-Ramanujan primes for every $m > 1$, then, in connection with this, it is interesting to study the following problem.

Problem 4. *For every $m > 1$ to estimate the smallest pseudo-m-Ramanujan prime.*

Definition 6. *We call a prime p an m-**R**-prime if p satisfies to Condition 6.*

Note that, as in Section 6, we could construct a sieve for selecting m-**R**-primes from all primes, using a Bertrand-like sequences $B_n^{(m)}$ (cf. (6.2)-(6.5)). The following lemma, as lemma 3, is proved by two ways. The second proof with the Berend's idea is especially important and we give it entirely.

Lemma 9. *We have*

$$\underline{\mathrm{P}}(\mathcal{S}_m) \geq 1 - \frac{1}{m}.$$

Second proof. Choose of the minimal prime $p = p_{t(m)}$ which more than $3m$. Now in the range from $p_{t(m)}$ up to n there are $\pi(n) - \pi(3m)$ primes. Put

$$h_m = h_m(n) = \pi(n/m) - 2. \tag{11.2}$$

Then $p_{h_m+2} \leq n/m$ and interval $(p_{h_m+2}, n/m]$ is free from primes. Furthermore, considering intervals

$$(mp_2, mp_3), (mp_3, mp_4), ..., (mp_{h_m+1}, mp_{h_m+2}). \tag{11.3}$$

Our $\pi(n) - \pi(3m)$ primes are somehow distributed in these h_m intervals. Suppose $k_m = k_m(n)$ of these intervals contain at least one prime and

$h_m - k_m$ contain no primes. Then for exactly k_m primes there is no primes between them and the next mp_j, and for the other $\pi(n)-\pi(3m)-k_m$ there is. Hence, among $\pi(n)-\pi(3m)$ primes exactly $\pi(n)-\pi(3m)-k_m$ are m-**R**-primes and exactly k_m non-m-**R**-primes. Therefore, since $k_m(n) \le h_m(n) \le \pi(n/m)$, then for the desired lower probability, that there is a prime, we have:

$$\underline{\mathrm{P}}(\mathcal{S}_m) = \liminf_{n\to\infty} \frac{\pi_{\mathbf{R}_m}(n)}{\pi(n)-\pi(3m)} =$$

$$\liminf_{n\to\infty} \frac{\pi(n)-k_m(n)}{\pi(n)} \ge \liminf_{n\to\infty} \frac{\pi(n)-\pi(n/m)}{\pi(n)} = 1 - 1/m. \tag{11.4}$$

■

Let $A_{1,m} \in \mathbb{P}$ be the sequence of primes $\{p_{n_k}\}$ for which every interval (mp_{n_k}, mp_{n_k+1}) contains a prime. Then in the terms of the second proof of Lemma 9 we have

$$\mathrm{P}(\mathcal{A}_{1,m}) = \lim_{n\to\infty} \frac{\pi_{A_m}(n)}{\pi(n)} = \lim_{n\to\infty} \frac{k_m(n)}{h_m(n)}$$

and, moreover, from this proof, taking into account that $h_m(n) \sim \pi(n)/m$, we find

$$\mathrm{P}(\mathcal{A}_{1,m}) = m\lim_{n\to\infty} \frac{k_m(n)}{\pi(n)} = m\lim_{n\to\infty}(\pi(n) - \pi_{\mathcal{A}_{1,m}}(n))/\pi(n) = m(1-P(\mathcal{S}_m)).$$

Therefore, $\mathrm{P}(\mathcal{A}_{1,m})$ exists if and only if $P(\mathcal{S}_m)$ exists, and we have

$$\mathrm{P}(\mathcal{A}_{1,m}) = m(1 - \mathrm{P}(\mathcal{S}_m)). \tag{11.5}$$

Note that quite analogously, as in Section 9, one can introduce generalized m-Labos primes and m-**L**-primes, that is the union of m-Labos primes and pseudo-m-Labos primes.

Now, by the same way as in Section 9, we prove the following theorem.

Theorem 4. *For every real $m > 1$ we have*

$$\mathrm{P}(\mathcal{S}_m) = \mathrm{P}(\mathbf{R}_m) = 1 - \frac{\lambda_m}{m},$$

$$\mathrm{P}(\mathbf{R}_m\overline{\mathbf{L}}_m) = \mathrm{P}(\mathbf{L}_m\overline{\mathbf{R}}_m) = (1 + \frac{1}{m})\lambda_m - 1,$$

$$\mathrm{P}(\mathbf{R}_m\mathbf{L}_m) = 2 - (1 + \frac{2}{m})\lambda_m, \ \ P(\mathbf{I}_m) = 1 - \lambda_m,$$

where q_m is only positive root of equation

$$(1-\lambda_m)\ln(1-\lambda_m) = -\frac{{\lambda_m}^2}{m}.$$

It is clear, why $\mathrm{P}(\mathbf{R}_m)$ and $\mathrm{P}(\mathbf{R}_m\mathbf{L}_m)$ tend to 1, as $m \to \infty$, while $\mathrm{P}(\mathbf{R}_m\overline{\mathbf{L}}_m)$ and $\mathrm{P}(\mathbf{I}_m)$ tend to 0.

Remark 3. *(On the singularity value $m = 1$). By Proposition 1, a numerical confirmation with error less than ρ requires a choice $\ln n > N_m(\rho)$. It is natural to conjecture that $N_m(\rho)$ has a form $N_m(\rho) = \frac{C_m(\rho)}{m-1}$, such that for $m = 1$ a confirmation will never come. On the other hand, it is natural conjecture that, for $m > 2$, $C_m < C_2$.*

Proof of Proposition 4. For a fixed $m > 1$, distinguish two cases:
1) Limit (2.6) for sequence $A = S_m$ exists. In this case, since $\mathrm{P}_{\mathcal{S}_m}(\mathcal{R}^{(m)}) < 1$, the theorem is evident.

2)Limit (2.6) for sequence $A = S_m$ does not exist. Now, if to suppose that there exists not more than a finite set of non$-m-$Ramanujan primes which satisfy Condition 6, then, using (11.1) we have

$$\pi_{\mathbf{R}}^{(m)}(n) \sim \pi_R^{(m)}(n) \sim (1 - 1/m)\pi(x).$$

But this means that limit (2.6) for sequence $A = S_m$ exists which contradicts to the condition. ■

Acknowledgment. The author is grateful to Daniel Berend (Ben Gurion University, Israel) for important private communication; Greg Martin (University of British Columbia, Canada) and Jonathan Sondow (USA) for their helpful comments which promoted a significant improvement of the paper.

References

[1] . R. C. Baker, G. Harman and J. Pintz, *The difference between consecutive primes, II,* Proc. London Math. Soc. (3) 72 (1996), 261–280.
[2] . D. Berend, *Private communication.*
[3] . G. Hoheisel, *Primzahlprobleme in der Analysis,* Sitz. Preuss. Akad. Wiss. 33 (1930), 3–11.
[4] . A. Ivić, *The Riemann Zeta-Function,* John Wiley&Sons, New York, 1985.
[5] . G. Martin, *Private communication.*
[6] . K. Prachar, *Primzahlverteilung,* Springer-Verlag, 1957.
[7] . S. Ramanujan, *A proof of Bertrand's postulate,* J. Indian Math. Soc. 11 (1919), 181–182.
[8] . S. Ramanujan, *Collected Papers of Srinivasa Ramanujan* (Ed. G. H. Hardy, S. Aiyar, P. Venkatesvara and B. M. Wilson), Amer. Math. Soc., Providence (2000), 208–209.
[9] . D. Redmond, *Number Theory, An Introduction,* Marcel Dekker, inc., New York-Basel-Hong Kong ,1996.
[10] . N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences* ($http : //www.research.att.com/\sim njas$)
[11] . J. Sondow, *Ramanujan primes and Bertrand's postulate,* Amer. Math. Monthly,116 (2009) 630–635.

[12] . J. Sondow, *Ramanujan primes* in Eric Weisstein's On-Line World of Mathematics; $http://home.earthlink.net/\sim jsondow/$

[13] . K. Soundararajan, *Small gaps between prime numbers: the work of Goldston-Pintz-Yildirim*, Bulletin (New Series) of the Amer. Math. Soc., 44, no.1 (2007), 1–18.

Departments of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. e-mail:shevelev@bgu.ac.il