# A Probabilistic Approach to some Additive and Multiplicative Problems of Number Theory

Gregory M. Sobko

**Abstract.** We suggest here a probabilistic approach that helps to address some classical questions and problems of Number Theory, like the Goldbach Conjecture [1], distributions of twin- and $d$-primes and primes among arithmetic sequences and many others.
The problem mentioned above will be addressed in publications that follow later.
In this paper we discuss the concepts of 'randomness' and 'independence' relevant to number-theoretic problems and interpret the basic concepts of divisibility of natural number in terms of probability spaces and appropriate probability distributions on classes of congruence.
We analyze and demonstrate the importance of Zeta probability distribution and prove theorems stating the equivalence of probabilistic independence of divisibility of random integers by coprime factors, and the fact that random variables with the property of independence of coprime factors must have Zeta probability distribution. The idea to use Zeta distribution is motivated by the fact that it provides the validity of the probabilistic Cramér's model for asymptotic prime number distribution, in a full agreement with the Prime Number Theorem. Multiplicative and additive models with recurrent equations for generating sequences of prime numbers are derived based on the reduced Sieve of Eratosthenes Algorithm. This allows to interpret such sequences as realizations of random walks on set $\mathbb{N}$ of natural numbers and on multiplicative semigroups $S(\mathbb{P})$ generated by sets of prime numbers $\mathbb{P}$, representing paths of stochastic dynamical systems. The H. Cramér's model for probability distribution of primes is modified as a *generalized predictable non-stationary Bernoulli process with unequally distributed terms* that are *asymptotically pairwise independent*. This model is applied then to analyze the sequences of primes generated by appropriate random walks. With an intense use of Zeta probability distribution it seems possible by using the modified Cramér's model to approximate the probability distribution of various arithmetic function.

*"...Mathematics is the art of giving the same name to different things...The only facts worthy of our attention are those which introduce order into this complexity and so make it accessible to us".*

(Henry Poincaré, The Value of Science, Random House, Inc., 2001.

## 1. Stochastic Predictable Sequences, Prime Numbers and Zeta Probability Distribution

Let $\mathbb{N}$ denote the set of natural numbers and $\mathbb{P}$ the set of all primes. Our major assumption follows the amazing Cramér's idea [9] to represent a deterministic sequence of prime numbers as realizations of binary random variables $\xi_k$ in the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ with an appropriate choice of their probability distributions.

Pursuing this idea we address two problems:

1) the choice of an adequate probability distributions $P_k$ for each $\xi_k$;

2) stochastic relationship among all $\xi_k$ in the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$.

We need several definitions [7].

**Definition 1.1**

Let $\left\{v_n \mid n\in\mathbb{N}\right\}$ be random variables $v_n:\Omega\to\mathbb{N}$ defined on probablity space $\left(\Omega,\mathcal{F},P\right)$ and $\mathcal{F}_n=\sigma\left\{v_k \mid 1\le k\le n\right\}$ a $\sigma$-algebra generated by all events created by random variables $\left\{v_k \mid 1\le k\le n\right\}$. We have: $\mathcal{F}_1\subseteq\cdots\mathcal{F}_n\subseteq\mathcal{F}_{n+1}\subseteq\cdots\subseteq\mathcal{F}$, and for each $n\in\mathbb{N}$ random variable $v_n$ is $\mathcal{F}_n$ -measurable. Then, the sequence $\left(v_n,\mathcal{F}_n\right)_{n\in\mathbb{N}}$ is called a *stochastic sequence*.

A stochastic sequence $\left(v_n,\mathcal{F}_n\right)_{n\in\mathbb{N}}$ is called *predictable* if for each $n\in\mathbb{N}$ there exists $k=k(n)<n$ such that $v_n$ is $\mathcal{F}_{k(n)}$-measurable. A pedictable sequence we can write as $\left(v_n,\mathcal{F}_{k(n)}\right)_{n\in\mathbb{N}}$.

*Predictability of a stochastic sequence* $\left(v_n,\mathcal{F}_n\right)_{n\in\mathbb{N}}=\left(v_n,\mathcal{F}_{k(n)}\right)_{n\in\mathbb{N}}$ means that for each $n\in\mathbb{N}$ the probability distribution $P_n$ of $v_n$ given the entire prehistory $\mathcal{F}_{n-1}$ is compeletly determined by the condition $\mathcal{F}_{k(n)}$, that is depends on values taken by some (or all) variables $v_1,v_2,\ldots,v_{k(n)}$ , where $k(n) < n$. So, in terms of conditional probabilities,

$$P_n\left\{\nu_n \in A \mid \mathcal{F}_{n-1}\right\} = P_n\left\{\nu_n \in A \mid \mathcal{F}_{k(n)}\right\} \text{ for all } A \in \mathcal{F}_n . \tag{1.1}$$

Notice that general stochastic sequences include classes of sequences of independent as well as dependent random variables like martingales, Markov chains, etc.

A sequence of mutually independent random variables $\left(\nu_n, \mathcal{F}_n\right)_{n\in\mathbb{N}}$ is *unpredictable* since probability distribution of each $\nu_n$ is determined only by events from $\mathcal{F}_n$ and does not depend on condition given by 'previous' events from $\mathcal{F}_k$ $(k < n)$. Markov chains and martingales are examples of predictable stochastic sequences.

In Number Theory we are interested in recursively defined sequences of numbers, generated by certain recurrent relations, mostly nonlinear. From probabilistic point of view, such recurrent relations generate sequences of dependent random variables. The problem of dependence of events and random variables in the framework of Number Theory had been dicussed in some detail in the monograph of Mark Kac [4]. As M. Kac underlined in [4], the concept of independence "though of central importance in probability theory, is not a purely mathematical notion", and it appears quite naturally in Statistical Physics. He mentioned that "the rule of multiplication of probabilities of independent events is an attempt to formalize this notion and to build a calculus arount it". By using informal language, the concept of independence is stated in [14] as follows: "Two events are said to be independent if they have 'nothing to do' with each other". To decide whether a 'randomly choosen' (odd) integer $\nu > 2$ is a prime number, we subject $\nu$ to a divisibility test, by using the Eratosthenes algorithm. If event $A = \left\{p_i \setminus \nu\right\}$ (' $p_i$ divides $\nu$ ') does not tell us anything about event $B = \left\{p_j \setminus \nu\right\}$ (' $p_j$ divides $\nu$ ') for $i \neq j$ , we can say that $A$ and $B$ do not depend on each either logically or statistically, and should be considered as independent events for a 'reasonable' choice of probability distribution of random variable $\nu$. Meantime, events $\left\{\nu \in \mathbb{P}\right\}$ and $\left\{(\nu+1) \in \mathbb{P}\right\}$ are dependent events since they exclude each other for $\nu > 2$, because only one of them holds true at a time.

More sofisticated example of dependent events represent $\left\{\nu \in \mathbb{P}\right\}$ and $\left\{(\nu+2) \in \mathbb{P}\right\}$, which are both true for twin primes, and false otherwise.

We demonstrate below that with an appropriate choice of probability distribution for random variable $\nu$ events $A=\{p_i \setminus \nu\}$and $B=\{p_j \setminus \nu\}$ are independent for any choice of prime numbers $p_i \neq p_j$. Such a choice is provided by Zeta probability distribution

$$P\{\nu_m = n\} = \frac{n^{-s}}{\zeta(s)} \quad (s>1), \quad n \in \mathbb{N} \tag{1.2}$$

Both dependence and independence of 'events' in Number Theory are results of complicated recurrent nonlinear relations between terms of numeric sequences, which can generate 'dynamical chaos', imitating pseudo-randomness in a long run behavior of such deterministic sequences. The precise prediction of behavior of terms in the sequences demands for 'big' numbers almost infeasible calculations caused by the expanding memory of prehistory of their evolution. To make a study feasible and overcome "the curse of dependence" researchers in this area typically suggest heuristic assumptions that terms in a $(\nu_n)_{n\in\mathbb{N}}$ are independent, or asymptotically independent, or uncorrelated, or 'weakly' dependent, in a certain sense.

**Proposition**

The basic fact is that the set of prime numbers $\mathbb{P}$ is a *recursive set* [17].

**Proof.**

We can prove this by using an indicator function $I_{\mathbb{P}} : \mathbb{N} \to \{0,1\}$ of set $\mathbb{P}$. We need to show that the function $I_{\mathbb{P}}$ is *recursively defined*.

(1) *Initial step*: let $I_{\mathbb{P}}(2)=1$, $I_{\mathbb{P}}(3)=1$.

(2) *Inductive step*: if $n>3$ is the smallest number such that $k \nmid n$ for each $k \le \sqrt{n}$ (symbol $\nmid$ means 'does not divide'), then $I_{\mathbb{P}}(n)=1$, otherwise $I_{\mathbb{P}}(n)=0$.

Notice that such number $n$ exists since $\mathbb{N}$ is a well-ordered set so that any nonempty subset of $\mathbb{N}$ has the least element (the smallest number).

(3) *Closure step*: Only numbers $n$ obtained in steps (1) and (2) satisfy condition $I_{\mathbb{P}}(n)=1$.

It holds true that if a function is recursively defined then it is unique [17].

We can explain the above statement concerning the recursive definition of prime numbers as follows. Occurrence of a prime number $n = p \in \mathbb{P}$ in the sequence of consecutive natural numbers $n = \{2,3,4,\ldots\}$ depends on the values of reminders $r = \mathrm{mod}(n,p)$ for all primes $p \le n$, due to the Sieve of Eratosphenes Algorithm [5]. This requirement can be relaxed:

we need to consider only divisibility of $n$ by all primes $p \le \sqrt{n}$.

The proof of this statement (attributed to Fibonacci) follows below.

**Lemma 1.1**

A natural number $n \ge 5$ is prime if and only if $n$ is not divisible by of any prime numbers $p \le \sqrt{n}$, or, equivalently, if $r = \mathrm{mod}(n,p) \ne 0$ for all primes $p \le \sqrt{n}$.

**Proof.**

If we assume that $n$ is a composite number with no primes $p \le \sqrt{n}$ that divide $n$, then $n$ should be divided by primes $p_1'$ and $p_2'$ both greater than $\sqrt{n}$, and therefore also divided by their product $p_1' \cdot p_2'$. But this would imply that $p_1' \cdot p_2' > n$, which is impossible. This means that if $n$ is not divisible by any of prime numbers $p \le \sqrt{n}$, then $n$ itself must be a prime number.

**Q.E.D.**

The above discussion implies that sequence of consecutive primes can be considered as a realization of a predictable stochastic sequence $\left(\nu_n, \mathcal{F}_{k(n)}\right)_{n\in\mathbb{N}}$, where $k(n) = \left[\sqrt{n}\right]$ for all $n > 3$ ($\left[x\right]$ stands for integer part of x).

One of the most challenging problems of Number Theory is the distribution of primes in the set $\mathbb{N}$ of natural numbers. The sequence of consecutive odd prime numbers $(3,5,7,11,\ldots)$ may look like a path of sporadic walks $\omega : \mathbb{N} \to \mathbb{P}$ given by a random sequence of natural numbers $\omega = \left(\nu_k(\omega) \mid k \in \mathbb{N}\right)$ where randomness of each term $\nu_j$ is determined by the choice of elementary event $\omega \in \Omega$ due to a probability distribution $P$ defined by a probability space $\left(\Omega, \mathcal{F}, P\right)$.

Primes in $\omega = \left(\nu_1, \nu_2 \ldots, \nu_j, \ldots\right)$ for each $\nu_k = k$ can be represented by the indicator function $I_{\mathbb{P}}(k) = \xi_k$ as a sequence of binary-valued variables $\xi_k = \begin{cases} 1, \text{ if } \nu_k(\omega) = k \in \mathbb{P} \\ 0, \text{ otherwise} \end{cases}$, This can be directly observed in the sequence of prime numbers below $100$:

$$\left(2\ 3\ 5\ 7\ 11\ 13\ 17\ 19\ 23\ 29\ 31\ 37\ 41\ 43\ 47\ 53\ 59\ 61\ 67\ 71\ 73\ 79\ 83\ 89\ 97\right)$$

**Table 1.1**

*The sequence* $\left(\xi(n) \mid 1 \le n \le 100\right)$ *of sequential primes among natural numbers from 1 to 100 represented by values of* $n$ *such that* $\xi_k = 1$ if $k$ is prime*:*

0 1 1 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0

0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0

In Number Theory we are interested in recursive sequences of numbers, generated by certain recurrent relations, mostly nonlinear. Let $\left(\Omega, \mathcal{F}, P\right)$ be a probability space, where $\Omega$ is a set of all $\mathbb{N}$-valued sequences, $\mathcal{F}$ is a $\sigma$-algebra generated by the algebra of cylinder sets in $\Omega$, and $P$ is a probability measure on $\left(\Omega, \mathcal{F}\right)$. Unsurmountable challenge is to describe probability distributions on a set $\Omega = \mathbb{N}^{\mathbb{N}}$ of all $\mathbb{N}$-valued sequences $\left(v_k\right)_{k \in \mathbb{N}}$ that include all recursively generated sequences of positive integers with all possible dependences between their terms. A sequence of integers in their natural increasing order $\left\{v_k = k \middle| k \in \mathbb{N}\right\}$ is our main concern.

In the framework of Probability Theory, we consider basic sequences $\left(\xi(n) \mid n \in \mathbb{N}\right)$ as realizations of $\left(0,1\right)$-valued random variables traditionally called *Bernoulli variables*.

To avoid pure heuristic justification of probabilistic conclusions, we try to conduct our discourse entirely in the framework of Probability Theory. This means that, prior to discussion of dependence issues related to sequences like $\omega = \left(v_1, v_2 \ldots, v_j, \ldots\right)$, we should introduce random variables $v_j : \Omega \to \mathbb{N}$ with the corresponding probability distribution $P$ defined on $\sigma$-algebra $\mathcal{F}$ of events $v_j^{-1}(A)$ (generated in our context by all finite subsets $A \subseteq \mathbb{N}$).

We assume that a binary-valued sequence $\left(\xi_k \mid k \in \mathbb{N}\right)$, where $\xi_k(\omega) = \begin{cases} 1, \text{ if } \ \nu_k(\omega) = k \in \mathbb{P} \\ 0, \text{ otherwise} \end{cases}$,

representing primes, is a realization of a non-stationary sequence of possibly *dependent* Bernoulli variables, by postulating probabilities

$$P\left\{\xi_k = 1\right\} = P_k, \quad P\left\{\xi_k = 0\right\} = Q_k = 1 - P_k, \text{ where } 0 \le P_k \le 1. \tag{1.3}$$

The major challenges in the study of such sequences are evaluation of $P_n$ in (1.3) and analysis of dependence of random variables $\left(\xi_k \mid k = 1,2,3,\ldots\right)$ included in the sequence. The problem of dependence of events and random variables in the framework of Number Theory had been discussed in some detail in the monograph of Mark Kac [4]. In number of works authors tried to avoid a standard probabilistic approach based on the concept of sigma-additive probability measures and the corresponding probability spaces, and considered instead so-called 'density' measures, which are additive but not $\sigma$-additive. As M. Kac underlined in [4], the concept of independence "though of central importance in probability theory, is not a purely mathematical notion", as it appears quite naturally in Statistical Physics.
He mentioned that "the rule of multiplication of probabilities of independent events is an attempt to formalize this notion and to build a calculus around it". Moreover, the notions of statistical (probabilistic) independence and dependence of events have been sometimes confused with mathematical (functional) or logical dependence.
Both dependence and independence of "events" in Number Theory are results of complicated recursive nonlinear relations between terms of numeric sequences, which can generate a 'dynamical chaos', imitating pseudo-randomness in the long run behavior of purely 'deterministic' sequences. The precise prediction of behavior on a 'long run' for terms in such sequences demanding tremendous calculations requires expanding memory of prehistory of their evolution. To make a study feasible and overcome 'the curse of dependence', a typically suggested heuristic assumption is that terms in $\left(\xi_k\right)_{k\in\mathbb{N}}$ are 'asymptotically independent', or 'uncorrelated', or 'weakly' dependent in a certain sense.

In the framework of modified H. Cramér's model we show that the sequence of dependent not identically distributed random variables $\left(\xi_k\right)_{k\in\mathbb{N}}$ is *asymptotically pairwise independent* in a sense that we are going to discuss below.

Surprisingly, in many discussions of probabilistic interpretations of Number Theory problems, some authors use 'by default' an approach as in the following sentence:

"Assume that we choose number $X$ at random from $1$ to $n$. Then $\text{Prob}(X \text{ is prime}) = \dfrac{\pi(n)}{n}$ ...".

The above sentence, due to its ambiguity, raises the following critical comments.

When one chooses number $X$ "at random" in the sense of Probability Theory, it is presumed that the probability distribution of $X$ exists and is known (at least theoretically).

The formula $\text{Prob}(X \text{ is prime}) = \dfrac{\pi(n)}{n}$ cited above tells us that the probability distribution is assumed to be uniform on the sequence of integers $\{1,2,3,\dots,n\}$.

Here $\pi(n) = \#\{p \in \mathbb{P} \mid p \le n\}$ denotes a counting function of number of primes not exceeding $n$. If the probability distribution of $X$ is not uniform on the interval of integers $[1,n] = \{1,2,\dots,n\}$, then, in a statistical framework, $\dfrac{\pi(n)}{n}$ can be interpreted not as a probability but rather as an observed relative frequency of occurrences of prime numbers in the interval $[1,n]$.

One of goals in our study is to construct a probabilistic model for the "statistical" distribution of primes given by the observed frequencies $\dfrac{\pi(n)}{n}$. Notice here the obvious fact that a *discrete uniform probability distribution* does not exist on an infinite support, that is on infinite subsets of $\mathbb{N}$ (including $\mathbb{N}$ itself). The following analysis is about divisibility of $\nu$ by a prime $p \le n$.

Denote $p \cdot \mathbb{N}$ a set of all multiples of number $p$. As mentioned above, the probability $P\{\nu \in p \cdot \mathbb{N}\}$ does not exists if $\nu$ is evenly distributed on $\mathbb{N}$. But the problem can be easily resolved if one assigns the probability $P\{\nu \in p \cdot \mathbb{N}\}$ to the class $C_{p,0} = \{n \mid n = k \cdot p,\ k \in \mathbb{N}\}$ of integers in $\mathbb{N}$ congruent $0$ modulo $p$. There are exactly $p$ congruent classes modulo $p$:

$$C_{p,r} = \{n \mid n = k \cdot p + r; 0 \le r \le p-1; k \in \mathbb{N} \cup \{0\}\},$$

which make a partition of $\mathbb{N}$. Then, we can define a probability distribution

$P\left(C_{p,r}\right)=q_{p,r}$ $(r=0,1,2,\ldots,p-1)$ on $\left\{C_{p,0},C_{p,1},\ldots,C_{p,p-1}\right\}$ such that $\sum_{r=0}^{p-1} q_{p,r} = 1$.

Then, $P\left\{\nu \in p\cdot\mathbb{N}\right\}=P\left(C_{p,0}\right)=q_{p,0}$. By assuming equal probabilities to randomly choose a class of congruence for a number $\nu$ given by $P\left(C_{p,r}\right)=q_p$ for all $r:\ 0\le r\le p-1$, we have $P\left(C_{p,r}\right)=\frac{1}{p}$, where $C_{p,0}=p\cdot\mathbb{N}$. Then, $P\left\{\nu\in p\cdot\mathbb{N}\right\}=P\left(C_{p,0}\right)=\frac{1}{p}$. Considering $P\left\{\nu\in p\cdot\mathbb{N}\right\}$, we have nothing but to assume that random variable $\nu$ can take any ('unknown') value within a congruence class $p\cdot\mathbb{N}$. The value of probability can be different from $P\left\{\nu\in p\cdot\mathbb{N}\right\}=\frac{1}{p}$ if we impose some limitations on $\nu$, say, if we assume that $\nu\le n$. For arbitrary $n\in\mathbb{N}$ and a given probability distribution of $\nu$, an event $\left\{\nu\le n\right\}$ may not belong, in general, to the algebra of events created by the partition of $\mathbb{N}$ into $p$ congruence classes $\left\{C_{p,r}\mid r=0,1,2,\ldots,p-1\right\}$, therefore, it would be impossible to assign probability to the event $\left\{\nu\in C_{p,o}\cap[1,n]\right\}$, where we denote $[1,n]=\left\{k\mid k=1,2,\ldots,n\right\}$. Since $[1,n]$ is a finite set, we can define a uniform probability distribution on this set, but the agreement of uniform distribution with the assumption $P\left\{C_{p,0}\cap[1,n]\right\}=\frac{1}{p}$ would depend on the choice of $n$, specifically, on divisibility of $n$ by $p$. For example, if $p=3$ and $n=20$, we have $P\left\{C_{3,0}\cap[1,20]\right\}=\frac{6}{20}=0.3\ne\frac{1}{3}$.

For $p=3$ and $n=21$ we have:

$$C_{3,0}\cap[1,21]=\left\{3,6,9,12,15,18,21\right\},\text{ and } P\left\{C_{3,0}\cap[1,21]\right\}=\frac{7}{21}=\frac{1}{3}\approx 0.333\ldots$$

Independence of divisibility of random number $\nu$ by different primes is determined by choice of probability distribution of $\nu$. As it had been noticed by Mark Kac in [4], "*primes play a game of chance*". He pointed out to the obvious fact that $\nu$ to be divisible by both different primes $p$ and $q$ is equivalent of being divisible by $p\cdot q$. This mean that if $P\{C_{m,0}\} = \frac{1}{m}$ for any positive integer $m$ , then, since $C_{p\cdot q,0} = C_{p,0} \cap C_{q,0}$ , we have

$$P\left\{C_{pq,0}\right\} = P\left\{C_{p,0}\right\} \cdot P\left\{C_{q,0}\right\} \text{ because } \frac{1}{p\cdot q} = \frac{1}{p} \cdot \frac{1}{q}. \quad (1.4)$$

Mark Kac was not able to establish and use the independence of divisibility events in terms of probability theory since he used a density set functions $d(A) = \lim_{n\to\infty} \frac{A(n)}{n}$, where $A(n) = A \cap [0,n],\ A \subset \mathbb{N}$ , which is not a probability measure as it is additive but not $\sigma$-additive.

**Definition. 1.1**

We call a probability distribution $P_f$ on $\mathbb{N}$ of a random variable $\nu$ *multiplicative or completely multiplicative* if for all $A \subseteq \mathbb{N}$ we have:

$$P_f\left\{\nu \in A\right\} = \frac{1}{Z} \sum_{n\in A} f(n) \text{ , where } f : \mathbb{N} \to (0,1], \quad (1.5)$$

is a *multiplicative*, *or respectively, completely multiplicative function,* such that $Z = \sum_{n\in\mathbb{N}} f(n)$ is a convergent series.

As we show below, independence of divisibility of random number $\nu$ by different primes can be guaranteed if $\nu$ has a multiplicative probability distribution defined above.
Each prime number $p$ determines a partition of the set $\mathbb{N}$ into $p$ classes of congruence modulo $p$: $C_{p,r}$ , where $r \in \{0,1,2,\ldots,p-1\}$. We show below that a randomly chosen value $\nu$ with multiplicative distribution $P_f$ is divisible by natural $m$ with probability $f(m)$ .

For $f(n) = \frac{1}{n^s}$ $(s > 1)$ and $Z = \zeta(s)$(where $\zeta(s)$ is Zeta function), the probability $P_f$ on $\mathbb{N}$ is *Zeta probability distribution*

$$P_{\zeta(s)}\left\{\nu = n\right\} = \frac{n^{-s}}{\zeta(s)},\ n \in \mathbb{N}, \text{ for any choice of } s > 1$$

and random $\nu$ with Zeta distribution is divisible by a prime number $p$ with probability

$\frac{1}{p^s}$ , so that for each $p \in \mathbb{P}$,

$$P_s\{\nu \in C_{p,0}\} = \frac{1}{p^s} \;,\; P_s\{\nu \notin C_{p,0}\} = 1 - \frac{1}{p^s} \tag{1.6}$$

Each natural $n$, due to the Fundamental Theorem of Arithmetic, can be represented in the unique form

$$n = p_1^{a_1} \cdot p_2^{a_2} \cdots p_k^{a_k} = \prod_{i=1}^{k} p_i^{a_i} \tag{1.7}$$

where $p_1, p_2, \ldots, p_k$ are distinct primes, and $a_1, a_2, \ldots, a_k$ are natural numbers.

The formula (1.7) is called a *canonical representation* of $n$, where $a_1, a_2, \ldots, a_k$ are called *multiplicities* of prime factors of $n$.

If $n$ is a realization of a random variable $\nu$, that is $\nu(\omega) = n$ , then (1.7) can be written in the form

$$\nu = \prod_{p \in \mathbb{P}} p^{\alpha(\nu,p)} = \prod_{k=1}^{\kappa(\nu)} p_k^{\alpha_k(\nu)} \tag{1.8}$$

as a canonical presentation of a random variable $\nu$ . Here $\alpha(\nu, p_k) = \alpha_k(\nu)$ .

Thus, (1.8) implies that the probability that $p$ does not divide $\nu$ equals $P_s\{\alpha(\nu, p) = 0\} = 1 - \frac{1}{p^s}$ .

In general, the event $\{\alpha(\nu, p) = k\}$ in (1.8) means that $p^k$ divides $\nu$ but $p^{k+1}$ does not divide $\nu$. From independence of these events, it follows:

$$P_s\{\alpha(\nu, p) = k\} = \left(\frac{1}{p^s}\right)^k \cdot \left(1 - \frac{1}{p^s}\right),\; k = 0, 1, 2, 3, \ldots , \tag{1.9}$$

This shows that each $\alpha(p_j, \nu)$ has a *geometric distribution* with parameter $\frac{1}{p_j}$

and we have $E\alpha(\nu, p) = \frac{p^{-s}}{1 - p^{-s}} = \frac{1}{p^s - 1}$; $Var\big(\alpha(\nu, p)\big) = \frac{p^{-s}}{\left(1 - p^{-s}\right)^2} = \left(\frac{p^s}{p^s - 1}\right) \cdot \left(\frac{1}{p^s - 1}\right)$.

Sum $\varphi(\nu) = \sum_{p \in \mathbb{P}} \alpha(\nu, p)$ counts the total number of prime factors (with their multiplicities) in the prime factorization of $\nu = \prod_{p \in \mathbb{P}} p^{\alpha(\nu,p)}$. Here are parameters of $\varphi(\nu)$:

$$E_s\left[\varphi(\nu)\right] = \sum_{p \in \mathbb{R}} \frac{1}{p^s - 1},\ Var_s\left[\varphi(\nu)\right] = \sum_{p \in \mathbb{R}} \left(\frac{p^s}{p^s - 1}\right) \cdot \left(\frac{1}{p^s - 1}\right).$$

Assume now that there is a vector $\vec{p} = \left(p_1, p_2, \ldots, p_N\right)$, which components are $N$ different consecutive prime numbers, and we consider a multiplicative semigroup $S(\vec{p})$ with unity, generated by components of vector $\vec{p}$ and number $1$.

For any $n \in S(\vec{p})$ we have $n = \prod_{i=1}^{k} p_i^{a_i}$ where $a_i > 0$ for all $i$ $(1 \le i \le k)$, $k \le N$.

Notice that by using computer simulation, we can generate $N$ pseudo-random variables $\alpha_j = \alpha(p_j, \nu), 1 \le j \le N$ , where each $\alpha(p_j, \nu)$ has a geometric distribution with parameter $\frac{1}{p_j}$

and then, simulate a 'pseudo-random' number $\nu = \prod_{j=1}^{k} p_j^{\alpha_j}$ with $k = k(\nu) \le N$ .

Further we consider a multiplicative semigroup $S(\mathbb{P}_N)$ generated by all primes $\mathbb{P}_N$ not exceeding $N \in \mathbb{N}$ , that is $\mathbb{P}_N = \left\{p \le N \mid p \in \mathbb{P}\right\}$.

**THEOREM 1.1.**

If $P_f$ is a *multiplicative probability distribution* on $\mathbb{N}$ and $\nu$ is a random variable such that

$$P_f\left\{\nu \in A\right\} = \frac{1}{Z} \cdot \sum_{n \in A} f(n) \text{ where } A \subseteq \mathbb{N},\ f : \mathbb{N} \to (0,1],$$

then

1) For any natural $m \ge 2$ random event $E$ of occurrence of a random number $\nu$ divisible by $m$ has probability $P_f(E) = P_f(C_{m,0}) = f(m)$.
2) for any two mutually prime numbers $m_1$ and $m_2$, random events $E_1$ and $E_2$ of occurrence of $\nu$ divisible by both $m_1$ and by $m_2$, respectively, are $P_f$ -independent events: $P_f(E_1 \cap E_2) = P_f(E_1) \cdot P_f(E_2)$.

Since $E_1 = C_{m_1,0}$ , $E_2 = C_{m_2,0}$ and $E_1 \cap E_2 = C_{m_1 \cdot m_2,0}$ we have, equivalently,

$$P_f(C_{m_1,0} \cap C_{m_2,0}) = P_f(C_{m_1,0}) \cdot P_f(C_{m_2,0})$$

**Proof.**

For $m = m_1 \cdot m_2$ we have:

$P_f\left(C_{m,0}\right) = \frac{1}{Z}\sum_{k\in\mathbb{N}} f(m\cdot k) = \frac{1}{Z}\sum_{k\in\mathbb{N}} f(m)\cdot f(k) = f(m) = f(m_1)\cdot f(m_2)$ since $\frac{1}{Z}\sum_{k\in\mathbb{N}} f(k) = 1$,

and $P_f\left(C_{m_i,0}\right) = \frac{1}{Z}\sum_{k\in\mathbb{N}} f(m_i \cdot k) = f(m_i)\ (i = 1,2)$. Then, $C_{m_1\cdot m_2} = C_{m_1} \cap C_{m_2}$ implies

$$P_f(C_{m_1,0} \cap C_{m_2,0}) = P_f(C_{m_1\cdot m_2,0}) = P_f(C_{m_1,0}) \cdot P_f(C_{m_2,0})$$

**Q.E.D.**

The following theorem states that the assumption that the probability distribution $P_f$ on $\mathbb{N}$ is ‘complete multiplicative’ (with an appropriate choice of function $f$) is a necessary and sufficient condition for such distribution $P_f$ to be Zeta probability distribution.

**THEOREM 1.2.**

Let $\nu$ be a random variable with values in $\mathbb{N}$ with probability distribution

$$P_f\left\{\nu \in A\right\} = \frac{1}{Z}\sum_{n\in A} f(n) \ , \tag{1.10}$$

where $f:\mathbb{N} \to \left[0,1\right]$, $A \subseteq \mathbb{N}$ and $Z = \sum_{n=1}^{\infty} f(n)$ is a convergent series.

The series $Z = \sum_{n=1}^{\infty} f(n)$ takes a form of the ‘Euler product of the series’ [12, p.230]:

1) if $f$ in (1.5) is multiplicative, then $Z = \sum_{n=1}^{\infty} f(n) = \prod_{p\in\mathbb{P}}\left[1 + f(p) + f(p^2) + \cdots\right]$;

2) if $f$ in (1.5) is a completely multiplicative function such that $0 < f(p) < 1$ for all $p \in \mathbb{P}$, then

$$Z = \sum_{n=1}^{\infty} f(n) = \prod_{p\in P} \frac{1}{1 - f(p)} \ ;$$

3) the probability distribution $P_f$ is a Riemann Zeta distribution

$P_{\zeta(s)}\left\{\nu=n\right\}=\dfrac{n^{-s}}{\zeta(s)}$, $n\in\mathbb{N}$, for any choice of $s>1$. Further we denote $P_{\zeta(s)}=P_s$ .

**Proof.**

1) Let $S(\mathbb{P}_N)$ be a semigroup of all integers generated by $\mathbb{P}_N\cup\{1\}$: $\mathbb{P}_N=\{p\mid p\le N, p\in\mathbb{P}\}$.

Due to the Fundamental Theorem of Arithmetic,

$$n=\prod_{p\in\mathbb{P}^*}p^{\alpha(n,p)},\text{ where }\alpha(n,p)\ge 0,\ \alpha(n,p)=\begin{cases}a_j>0\text{ if }p^{a_j}\mid n\text{ and }p^{a_j+1}\nmid n\\ 0,\text{ otherwise}\end{cases}$$

Then, if $f$ is a multiplicative function, we have

$$Z=\sum_{n=1}^{\infty}f(n)=\sum_{n=1}^{\infty}\left[\prod_{p\in\mathbb{P}}f(p^{\alpha(n,p)})\right]=\prod_{p\in\mathbb{P}}\left[\sum_{k=0}^{\infty}f(p^k)\right]=\prod_{p\in\mathbb{P}}\left[1+f(p)+f(p^2)+\cdots\right].$$

2) In the proof above we have used the multiplicative property of function $f$.

If $f$ is completely multiplicative, we have $f(p^k)=\left(f(p)\right)^k$. Then, we can write

$1+f(p)+\left(f(p)\right)^2+\left(f(p)\right)^3+\cdots=\dfrac{1}{1-f(p)}$ and the above equality takes a form:

$$Z=\sum_{n=1}^{\infty}f(n)==\prod_{p\in\mathbb{P}}\left[\sum_{k=0}^{\infty}\left(f(p)\right)^k\right]=\prod_{p\in\mathbb{P}}\frac{1}{1-f(p)}$$

Notice that the right-hand sides of the above equalities are convergent infinite products, since the left-hand side is given by the convergent series.

3) Notice that for any $n\in S(\mathbb{P}_N)$ we have $n^s=\prod_{p\in\mathbb{P}}p^{a(n,p)\cdot s}=\prod_{p\in\mathbb{P}}p^{a(p)\cdot s}$,

where $a(n,p)=a(p)\in\mathbb{N}\cup\{0\}$ .

Since $\dfrac{1}{1-\dfrac{1}{p^s}}=\sum_{k=0}^{\infty}\left(\dfrac{1}{p^s}\right)^k$, we have $\zeta_{\mathbb{P}_N}(s)=\prod_{p\le N}\left[\sum_{k=0}^{\infty}p^{-sk}\right]=\sum_{\{a(p),p\le N\}}\prod_{p\le N}p^{-s\cdot a(p)}=\sum_{n\in S(\mathbb{P}_N)}n^{-s}$

Denote $\xi_p(\nu)=p^{\alpha(\nu,p)}$, $\xi=p$ .Then, $P\left\{\xi_{\nu,p}=p^{\alpha(\nu,p)}\right\}=\left[P\left\{\xi=p\right\}\right]^{\alpha(\nu,p)}$

For any natural $m$ we write the event "$m$ divides $\nu$" as $E=\{m\setminus\nu\}$ and the opposite event

" $m$ does not divide $\nu$" as $\bar{E}=\{m\nmid\nu\}$. The probability that a prime number $p$ divides $\nu$ is $P\{p\backslash\nu\}=f(p)$ and the probability that $p$ does not divide $\nu$ is $P\{p\nmid\nu\}=1-f(p)$. The probability that the number $\nu$ divides $p^k$ and does not divide $p^{k+1}$ is given by the formula

$$P\left\{\left(p^k\backslash\nu\right)\cap\left(p^{k+1}\nmid\nu\right)\right\}=\left(f(p)\right)^k\cdot\left(1-f(p)\right)$$

Then, by virtue of Theorem 1.1 and the canonical factorization of $n$, we have

$$\begin{aligned}P\left\{\nu=n\right\}&=\prod_{p\in\mathbb{P}}P\left\{\left(p^{a(n,p)}\,|\,\nu\right)\cap\left(p^{a(n,p)+1}\nmid\nu\right)\right\}\\&=\prod_{p\in\mathbb{P}}\left[\left(f(p^{a(n,p)})\right)\cdot\left(1-f(p)\right)\right]=\prod_{p\in\mathbb{P}}\left[f(p)\right]^{a(\nu,p)}\cdot\prod_{p\in\mathbb{P}}\left[1-f(p)\right]\end{aligned}\qquad(1.11)$$

Summation of both sides of (1.11) results in the formula:

$$1=\sum_{n\in\bullet}P\{\nu=n\}=\prod_{p\in\mathrm{P}}(1-f(p))\cdot\sum_{\nu\in\bullet}\prod_{p\in\mathrm{P}}[f(p)]^{\alpha(\nu,p)}$$, which implies:

$$\prod_{\nu\in\mathbb{N}}\frac{1}{1-f(p)}=\sum_{\nu\in\mathbb{N}}\prod_{p\in\mathbb{P}}f(p^{\alpha(n,p)})=\sum_{\nu\in\mathbb{N}}f\left(\prod_{p\in\mathbb{P}}p^{\alpha(n,p)}\right)=\sum_{\nu\in\mathbb{N}}f(n)=Z\qquad(1.12)$$

provided that $f(n)$ is such that the infinite product and the infinite sum in the above formulas are both convergent. Completely multiplicative function $f:\mathbb{N}\to(0,1]$ satisfies the functional equation $f(x\cdot y)=f(x)\cdot f(y)$, known as one of 'fundamental' Cauchy functional equations. Due to Theorem 3, p.41 in [19] for positive $x,y$, it has the most general solution of the form $f(x)=e^{c\cdot\ln x}=x^c$. Obviously, in our context $f(n)=n^{-s}(s>1)$ is a completely multiplicative arithmetic function and for this choice of $f$, $Z(f)=\zeta(s)$ is Zeta function which generates Zeta probability distribution

$$P_{\zeta(s)}\left\{\nu=n\right\}=\frac{1}{n^s\cdot\zeta(s)},\ n\in\mathbb{N}.$$

**Q.E.D.**

**Remark 1.1.**

The problem with the choice $f(n)=\dfrac{1}{n}$ for $s=1$ is that it leads to the divergent harmonic series $\zeta(1)=\sum_{n=1}^{\infty}\frac{1}{n}$. To avoid the situation with the series divergence, we follow the steps of Euler [3] by restricting values of $s$ to $s>1$. Zeta function $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$ is well known to be directly related to the probability distribution of prime numbers. This motivates the choice of Zeta distribution. Due to the property of independence of divisibility for Zeta distribution, if $p$ divides $\nu$, then $\nu = p\cdot\nu'$ while the quotient $\nu'=\frac{\nu}{p}$ is again distributed over $p$ classes of congruence $C_{p,r}$, and so on. A number $\nu = n$ is prime if and only if it does not divide all primes less than or equal to $\sqrt{n}$:

$$P_s\left\{\nu\in\mathbb{P}\,|\,\nu=n\right\}=P_s\left\{\bigcap_{p\le\sqrt{n}}\left[\alpha(\nu,p)=0\right]|\,\nu=n\right\}=\prod_{p\le\sqrt{n}}\left(1-\frac{1}{p^s}\right) \tag{1.13}$$

For $p\in\mathbb{P}$ we have $P_s\left\{\alpha(\nu,p)=k\right\}=p^{-sk}\cdot\left(1-p^{-s}\right)$ for all $k=0,1,2,\ldots$.

In particular, $P_s\left\{\alpha(\nu,p)=1\right\}=p^{-s}\cdot\left(1-p^{-s}\right)$, and $P_s\left\{\alpha(\nu,p)=0\right\}=1-p^{-s}$.

Then, the probability of $\left\{\nu=p_j\in\mathbb{P}\right\}$ is calculated as

$$P\left\{\nu=p_j\right\}=P_s\left\{\alpha(\nu,p_1)=0,\ldots,\alpha(\nu,p_{j-1})=0,\alpha(\nu,p_j)=1,\alpha(\nu,p_{j+1})=0,\ldots\right\}$$

$$=\left(\frac{1}{p_j^s}\right)\cdot\left(1-\frac{1}{p_j^s}\right)\cdot\prod_{k\ne j}P_s\left\{\alpha(\nu,p_k)=0\right\}=\left(\frac{1}{p_j^s}\right)\cdot\prod_{k=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)=\frac{p_j^{-s}}{\zeta(s)}$$

Probability of $\left\{\nu=1\right\}$, due to the canonical presentation (1.8), can be expressed as

$$P_s\left\{\nu=1\right\}=P_s\left\{\bigcap_{p\in\mathbb{P}}\left\{\alpha(1,p)=0\right\}\right\}=\prod_{p\in\mathbb{P}}\left(1-\frac{1}{p^s}\right)=\frac{1}{\zeta(s)}$$

In general, for any natural number $\nu=n=\prod_{p\in\mathbb{P}}p^{\xi_p}=p_1^{k_1}\cdot p_2^{k_2}\cdots p_m^{k_m}\cdots$, we have

$$P_s\left\{\nu=n\right\}=\prod_{p\in\mathrm{P}}\left[\left(\frac{1}{p^s}\right)^{\alpha(n,p)}\cdot\left(1-\frac{1}{p^s}\right)\right]=\prod_{p\in\mathrm{P}}\left(\frac{1}{p^s}\right)^{\alpha(n,p)}\cdot\prod_{p\in\mathrm{P}}\left(1-\frac{1}{p^s}\right)=\frac{1}{n^s}\cdot\zeta^{-1}(s)\,,$$

that is

$$P_s\left\{\nu=n\right\}=\frac{n^{-s}}{\zeta(s)} \tag{1.14}$$

Formula (1.14) may provide some probabilistic interpretations of Riemann Zeta function. If $\nu$ has Zeta probability distribution, then the probability that $\nu(\omega)$ for certain $\omega$ results in a prime number is evaluated as

$$P_s\{\nu \in \mathbb{P}\} = \frac{1}{\zeta(s)} \sum_{p\in\mathbb{P}} p^{-s} \tag{1.15}$$

Notice that formula (1.13) does not provide 'reasonable' values of probabilities for specific realizations of $\nu$ . For example, it is not equal to zero for any composite value of $\nu$, say for even $\nu$. Actually, $P_s\{\, \nu \text{ is prime} \mid \nu = n\} = \prod_{p\le\sqrt{n}} \left(1-\frac{1}{p^s}\right)$ evaluates a relative frequency for occurrence of prime numbers $p \le n$ . As we show further, formula (1.13) gives satisfactory predictions of asymptotic values of probability $P_s\{\, \nu \text{ is prime} \mid \nu = n\}$ as $n \to \infty$.

Since $\{\nu \le n\} = \bigcup_{i=1}^{n} \{\nu = i\}$, we have

$$P_s\{\nu \le n\} = \frac{\sum_{k=1}^{n} k^{-s}}{\zeta(s)} \text{ and } P_s\{(\nu \in \mathbb{P}) \cap (\nu \le n)\} = \frac{\sum_{p\in\mathbb{P}_n} p^{-s}}{\zeta(s)}, \text{where } \mathbb{P}_n = \{p \in \mathbb{P} \mid p \le n\} .$$

We could compare the last probability with the frequency estimate $\frac{\pi(n)}{n}$ or with (1.13) and with the Cramér's model prediction $\frac{1}{\ln n}$ , though, dependence of probability $P_s$ on parameter $s > 1$ makes the above formulas harder to interpret. As we know, one can circumvent divergence of $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ for $s \le 1$ by using the analytic continuation of $\zeta(z)$ on the complex plane $\mathbb{C}$, as suggested by B. Riemann. Meanwhile, as we have mentioned above, the use of *Incomplete Product Zeta* function (*IPZ*) $\zeta_{\mathbb{P}_N}(s)$ defined as a partial product of $\zeta(s)$, $\zeta_{\mathbb{P}_N}(s) = \prod_{p\le N} \frac{1}{1-\frac{1}{p^s}}$ , provides another opportunity to deal with the divergence of $\zeta(1) = \prod_{p\in\mathbb{P}} \frac{1}{1-\frac{1}{p^s}}$ for $s = 1$ .

**Lemma 1.2**

Let $S(\mathbb{P}_N)$ be a semigroup of all integers generated by $\mathbb{P}_N \cup \{1\}$,

$\mathbb{P}_N = \{p \mid p \le N, p \in \mathbb{P}\}$.

Then,

$$\zeta_{\mathbb{P}_N}(s) = \sum_{n \in S(\mathbb{P}_N)} n^{-s} \cdot$$

**Proof.**

Notice that for any $n \in S(\mathbb{P}_N)$ we have $n^s = \prod_{p \in \mathbb{P}} p^{\alpha(p) \cdot s}$, where $\alpha(p) \in \mathbb{N} \cup \{0\}$.

Since $\dfrac{1}{1-\dfrac{1}{p^s}} = \sum_{k=0}^{\infty} \left(\dfrac{1}{p^s}\right)^k$, we have $\zeta_{\mathbb{P}_N}(s) = \prod_{p \le N}\left[\sum_{k=0}^{\infty} p^{-sk}\right] = \sum_{\{a(p), p \le N\}} \prod_{p \le N} p^{-s \cdot a(p)} = \sum_{n \in S(\mathbb{P}_N)} n^{-s}$.

**Q.E.D.**

**Lemma 1.2**

If $\nu$ follows Zeta distribution $P_s$, then

$$P_s\{\nu = n \in \mathbb{P}\} = \prod_{p < \sqrt{n}} \left(1 - \frac{1}{p^s}\right). \tag{1.16}$$

**Proof.**

By using the recursive property of the sequence of prime numbers $p \le \nu$ with the memory size $\sqrt{\nu}$ and the property of independence of divisibility for Zeta distribution, we have $P\{\nu = n \in \mathbb{P}\} = P\left\{\bigcap_{p \le \sqrt{n}} \{p \nmid \nu\}\right\} = \prod_{p \le \sqrt{n}} P_s\{p \nmid \nu\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p^s}\right)$,

which means that $\nu = n$ is prime if and only if any prime $p \le \sqrt{n}$ does not divide $n$. Formally, $\nu = n$ is prime if and only $r = \mathrm{mod}(n, p') \ne 0$ for all primes $p' \le \sqrt{n}$.

This implies that if $\nu > 5$ follows Zeta probability distribution then

$$P_s\{\,\nu \text{ is prime }\} = \prod_{p \le \sqrt{\nu}} \left(1 - \frac{1}{p^s}\right).$$

**Q.E.D.**

**Lemma 1.3**

Let $S(\mathbb{P}_N)$ be a semigroup of all integers generated by $\mathbb{P}_N \cup \{1\}$,

$\mathbb{P}_N = \{p \mid p \le N, p \in \mathbb{P}\}$.

Then,

$$\zeta_{\mathbb{P}_N}(s) = \sum_{n \in S(\mathbb{P}_N)} n^{-s}.$$

**Proof.**

Notice that for any $n \in S(\mathbb{P}_N)$ we have $n^s = \prod_{p \in \mathbb{P}} p^{\alpha(p)\cdot s}$, where $\alpha(p) \in \mathbb{N} \cup \{0\}$.

Since $\dfrac{1}{1-\dfrac{1}{p^s}} = \sum_{k=0}^{\infty}\left(\dfrac{1}{p^s}\right)^k$, we have $\zeta_{\mathbb{P}_N}(s) = \prod_{p \le N}\left[\sum_{k=0}^{\infty} p^{-sk}\right] = \sum_{\{a(p), p \le N\}} \prod_{p \le N} p^{-s\cdot a(p)} = \sum_{n \in S(\mathbb{P}_N)} n^{-s}$.

**Q.E.D.**

## 2. Multiplicative and Additive Recurrent models for Primes

The famous Harald Cramér's model [2,3] describes the occurrence of prime numbers as a *sequence of independent Bernoulli variables* with probabilities

$$P\{\xi_n = 1\} = \frac{1}{\ln n},\ \ P\{\xi_n = 0\} = 1 - \frac{1}{\ln n},\ \text{ where } n \ge 2. \tag{2.1}$$

Notice that similar to (1.16), formula (2.1) is valuable only asymptotically for distribution of primes. In what follows, we provide rigorous arguments in support of Cramér's model, related to the values of probabilities $P_n = \dfrac{1}{\ln n}$, and then analyze dependence of $\xi_n$ in the sequence $(\xi_n \mid n = 1,2,\ldots)$. As we have discussed above, appearance of a prime $\nu_k = k$ in the sequence $\{\nu_k = k \mid k \in \mathbb{N}\}$ are dependent events determined by the prehistory $\mathcal{F}_{\sqrt{n}} = \sigma\{\nu_k \mid 1 \le k \le \sqrt{n}\}$. Obviously, if $\nu_k = p \in \mathbb{P}$, then $\nu_{k+1} = p + 1 \notin \mathbb{P}$ since $p+1$ is an even number. Even if we restrict values of $\nu_k$ to odd numbers $2k+1$, still divisibility of $\nu_k = 2k+1$ by the previously occurred primes would depend on the prehistory $\mathcal{F}_{\sqrt{2k+1}}$ .

Therefore, the sequence of consecutive primes and the corresponding Bernoulli variables $\xi_k$ cannot be interpreted as occurrence of independent events in the sequence, or as a realization of a Markov chain with a constant size of 'memory', because for each $\nu_k = k$ the size $\left[\sqrt{k}\right]$ of the 'memory' $\mathcal{F}_{\sqrt{k}}$ increases in the sequence with $k$.

We analyze the sequence of prime numbers $\left\{\nu_k = p \mid p \in \mathbb{P}, k \in \mathbb{N}\right\}$ by using *multiplicative and additive models*. In any kind of a model, we will be using the equivalent canonical realizations

$$\left(\Omega, \mathcal{F}, P\right) = \left(X^T, \mathcal{B}^T, P_X\right) \text{ so that } \nu(\omega, t) = \nu(t) \ .$$

The transformations $\theta_t : X^T \to X^T$, $t \in T$, are $\mathcal{B}^T / \mathcal{B}^T$ -measurable.

We define the transformations by $\theta_s \nu(t) = \nu(t+s),$ for $s, t \in T$.

A *multiplicative model* is based on the *canonical representation of primes* [5, p.18]:

$$n = \prod_{p \in \mathbb{P}} p^{\alpha(n,p)} \text{ where } \alpha(n,p) = \begin{cases} \alpha_p > 0 \text{ if } p \text{ divides } n \\ 0, \text{ otherwise} \end{cases} \tag{2.2}$$

and is concerned with the questions of divisibility of integer-valued random variables by integers, and with their connection to Zeta probability distribution:

$$P_s\left\{\nu \in A\right\} = \frac{1}{\zeta(n)} \cdot \sum_{n \in A} \frac{1}{n^s} \text{ , for any subset } A \subseteq \mathbb{N}. \tag{2.3}$$

For the multiplicative model of the dynamical system representing (2.2), where $\nu = n$, we define

$$\theta_i \nu = \nu_i; \;\; \nu_0 = 1, \theta_{i+1} \nu = \theta_i \nu \cdot \eta_{i+1},$$

$$\text{where } \eta_{i+1} = p_{i+1}^{\alpha_{i+1}(\nu)} (i = 0,1,2,\ldots,\kappa(\nu)-1). \tag{2.4}$$

*Additive models* are useful in problems related to counting of various types of integers in $\mathbb{N}$.

In additive models dynamical systems are defined by the equations:

$$\theta_i \nu = \nu(i); \nu_0 = 0, \theta_{i+1} \nu = \theta_i \nu + \xi_{i+1}, \tag{2.5}$$

where definition of the 'updating' term $\xi_{i+1}$ determines the specifics of the model, as illustrated below.

First, we consider the function $\pi(x)$, counting the number of primes less than or equal to $x$.

Second, for all $m \ge 3$ we consider the number $G(2m)$ of *Goldbach m-primes,* or $G_m$*-primes*, which are such primes $p$ that a difference $2m - p$ is again a prime number.

In the first situation we use recurrent equations:

$$\begin{cases} \pi(1) = 0 \\ \pi(k+1) = \pi(k) + \xi_{k+1}, \ k \in \mathbb{N} \end{cases} \tag{2.6}$$

It is well-known that the connections between additive and multiplicative properties of numbers are extraordinarily complicated, and this leads to various difficult problems in Number Theory. We start from the *division algorithm* [5, p.19]. Given integers $n$ and $m > 0$ there exists a unique pair of integers $k$ and $r$ such that $n = mk + r$, with $0 \le r \le m$. In this equation, $r = 0$ if and only if $m$ divides $n$. We derive here a recursive formula generating a sequence of prime numbers: $2, 3, 5, 7, \ldots$ For any prime number $p \in \mathbb{P}$ and a natural number $n \ge 2$, consider a function $\mathrm{mod}(n, p) = r$ of residuals (remainders) such that $n = m \cdot p + r,\ 0 \le r < p,$ where $m \in \mathbb{N} \cup \{0\}$ and vector of consecutive prime numbers $\vec{p}(n) = (p_1, p_2, \ldots, p_k)$ such that $p_k \le n$ and $p_{k+1} > n$. Index $k$ determines here the value $\pi(n) = k$ for the number of primes less than or equal to $n$, so that $\vec{p}(n) = (p_1, p_2, \ldots, p_{\pi(n)})$. For each coordinate $p_i$ of vector the $\vec{p}(n)$ we determine the residual value $r_i = \mathrm{mod}(n, p_i),\ i = 1, 2, \ldots, \pi(n)$, and vector of residuals $\vec{r}(n) = (r_1, r_2, \ldots, r_{\pi(n)})$.

Notice that, due to Sieve Algorithm and Lemma 1.1, for an integer $n > 2$ to be prime it is necessary and sufficient that all coordinates $r_i$ of the 'reduced' vector of residuals $\vec{r}(n)$ such that $1 \le i \le \pi(\sqrt{n})$ be different from zero. Thus, the events

$$\left\{ \min_{i \le \pi(\sqrt{\nu})} \{ r_i \mid r_i = \mathrm{mod}(\nu, p_i) \} > 0 \right\} \text{ and } \{ \nu \in \mathbb{P} \}$$

are equivalent. See calculations below in the Table 2.1.

**Table 2.1.** *The recursive sequence of primes driven by their residuals*

| $n$ | $\pi(n)$ | $\vec{p}(n)=\left(p_1,p_2,\ldots,p_{\pi(n)}\right)$ | $\vec{r}(n)=\text{mod}\left(n,\vec{p}(n)\right)=\left(r_1,r_2,\ldots,r_{\pi(n)}\right)$ |
|---|---|---|---|
| 2 | 1 | $(2)$ | $(0)$ |
| 3 | 2 | $(2,3)$ | $(1,0)$ |
| 4 | 2 | $(2,3)$ | $(0,1)$ |
| 5 | 3 | $(2,3,5)$ | $(1,2,0)$ |
| 6 | 3 | $(2,3,5)$ | $(0,0,1)$ |
| 7 | 4 | $(2,3,5,7)$ | $(1,1,2,0)$ |
| 8 | 4 | $(2,3,5,7)$ | $(0,2,3,1)$ |
| 9 | 4 | $(2,3,5,7)$ | $(1,0,4,2)$ |
| 10 | 4 | $(2,3,5,7)$ | $(0,1,0,3)$ |
| 11 | 5 | $(2,3,5,7,11)$ | $(1,2,1,4,0)$ |
| 12 | 5 | $(2,3,5,7,11)$ | $(0,0,2,5,1)$ |
| 13 | 6 | $(2,3,5,7,11,13)$ | $(1,1,3,6,2,0)$ |
| … | … | … | … |
| 30 | 10 | $(2,3,5,7,11,13,17,19,23,29)$ | $(0,0,0,2,8,4,13,11,7,1)$ |
| 31 | 11 | $(2,3,5,7,11,13,17,19,23,29,31)$ | $(1,1,1,3,9,5,14,12,8,2,0)$ |

We evaluate $P\{\nu\in\mathbb{P}\,|\,\nu=n\}$ assuming that a random integer $\nu$ follows Zeta probability distribution. To assign a probability value to a set $m\cdot\mathbb{N}$ ("all multiples of number $m$"), we should refer it to the class $C_{m,0}=\{n\,|\,n=k\cdot m,\,k\in\mathbb{N}\}$ of integers in $\mathbb{N}$ congruent $0$ modulo $m$ so that $C_{m,0}=m\cdot\mathbb{N}$ .

There are exactly $m$ congruent classes modulo $m$: $C_{m,r}=\{n\,|\,n=r+k\cdot n,\,k\in\mathbb{N}\cup\{0\}\},\,0\le r\le m-1$, which make a finite partition of $\mathbb{N}$. Then, for each integer $m>1$ we can define a probability distribution on $\{C_{m,0},C_{m,1},\ldots,C_{m,m-1}\}$ :

$$P\left\{\nu \in C_{m,r}\right\} = q_{m,r} \geq 0,\, 0 \leq r \leq m-1 \quad \text{and} \quad \sum_{r=0}^{m-1} q_{m,r} = 1,\ m = 2,3,4\ldots$$

**Theorem 2.1**

Let $\nu$ be a random variable with Zeta probability distribution $P_s$ and

its canonical representation.

Variables $\alpha_k(\nu) = \alpha(\nu, p_k)$ are independent for all primes $p_k$ $(k = 1,2,\ldots,\kappa(\nu))$ as well as factors $p_k^{\alpha(\nu,p_k)}$ and $p_j^{\alpha(\nu,p_j)}$ for all $k \neq j$ in the canonical factorization $\nu = \prod_{p\in\mathbb{P}} p^{\alpha(\nu,p)}$.

Then, each random variable $\alpha(\nu, p)$ in (2.8) has geometric probability distribution with a parameter $u = \frac{1}{p^s}$ $(0 < u < 1)$ :

$$P_s\left\{\alpha(\nu,p) = a\right\} = u^a \cdot (1-u) = \left(\frac{1}{p^s}\right)^a \cdot \left(1 - \frac{1}{p^s}\right), \quad a = 0,1,2,\ldots \tag{2.9}$$

***Proof.***

Notice that (2.8) implies $\nu^s = \prod_{p\in\mathbb{P}} p^{s\cdot\alpha(\nu,p)} = \prod_{k=1}^{\kappa(\nu)} p_k^{s\cdot\alpha_k(\nu)}$ where $s > 1$.

Denote $a \setminus b$ and $a \nmid b$ events '$a$ divides $b$' and '$a$ does not divide $b$', respectively.

Event $E_m = C_{m,0}$ stands for $\left\{m \setminus \nu\right\}$ ('$m$ divides $\nu$'). Then, for $m = m_1 \cdot m_2$ we have

$$P_{\zeta(s)}\left(E_m\right) = P_{\zeta(s)}\left(C_{m,0}\right) = \sum_{k\geq 11} \frac{\left(m\cdot k\right)^{-s}}{\zeta(s)} = \frac{1}{m^s}\sum_{k\geq 1}\frac{k^{-s}}{\zeta(s)} = \frac{1}{m^s} = \frac{1}{m_1^{\,s}\cdot m_2^s},$$

$$P_{\zeta(s)}\left(C_{m_i,0}\right) = \sum_{k\geq 11} \frac{\left(m_i\cdot k\right)^{-s}}{\zeta(s)} = \frac{1}{m_i^s}\sum_{k\geq 1}\frac{k^{-s}}{\zeta(s)} = \frac{1}{m_i^s}\ (i = 1,2).$$

Similar,

$$\begin{aligned} P_{\zeta(s)}\left(C_{m_1\cdot m_2,0}\right) &= \sum_{k\geq 11} \frac{\left(m_1\cdot m_2\cdot k\right)^{-s}}{\zeta(s)} = \sum_{k\geq 1}\frac{k^{-s}}{\zeta(s)}\cdot\frac{1}{m_1^s\cdot m_2^s} = \frac{1}{m_1^s\cdot m_2^s}. \\ &= P_{\zeta(s)}\left(C_{m_1,0}\right)\cdot P_{\zeta(s)}\left(C_{m_2,0}\right) \end{aligned}$$

If $m_1$ and $m_2$ are co-prime numbers, then $C_{m_1\cdot m_2,0} = C_{m_1} \cap C_{m_2}$, that is $E_{m_1\cdot m_2} = E_{m_1} \cap E_{m_2}$, and $P_s\left(E_{m_1}\cap E_{m_2}\right)=P_s\left(E_{m_1}\right)\cdot P_s\left(E_{m_2}\right)$, which holds true for any two different primes $m_1=p_1$ and $m_2=p_2$ . This proves independence of $\alpha(p,\nu)$ for different primes $p$, as well as independence of factors $p_i^{\alpha(\nu,p_i)}$ and $p_j^{\alpha(\nu,p_i)}$ for all $i\neq j$ in the canonical factorization $\nu=\prod_{p\in\mathbb{P}}p^{\alpha(p,\nu)}$. We have:

$$P_s\left\{\left(p^k\setminus\nu\right)\cap\left(p^{k+1}\nmid\nu\right)\right\}=P_s\left\{p^k\setminus\nu\right\}-P_s\left\{p^{k+1}\setminus\nu\right\}\ \text{ since }\left\{p^{k+1}\setminus\nu\right\}\subset\left\{p^k\setminus\nu\right\}.$$

Notice that $P_s\left\{p^k\setminus\nu\right\}=P_s\left\{\nu\in p^k\cdot\mathbb{N}\right\}=\frac{1}{\zeta(s)}\cdot\sum_{m\in\mathbb{N}}\frac{1}{\left(p^k\cdot m\right)^s}=\left(\frac{1}{p^s}\right)^k\cdot\frac{1}{\zeta(s)}\cdot\sum_{m\in\mathbb{N}}\frac{1}{m^s}=\left(\frac{1}{p^s}\right)^k$.

Since

$$\left\{\left(p^{\alpha(\nu,p)}\setminus\nu\right)\cap\left(p\nmid\frac{\nu}{p^{\alpha(\nu,p)}}\right)\right\}=P\left\{\left(p^{\alpha(\nu,p)}\setminus\nu\right)\cap\left(p^{\alpha(\nu,p)+1}\nmid\nu\right)\right\}=\left(\frac{1}{p}\right)^{\alpha(\nu,p)}\cdot\left(1-\frac{1}{p}\right),$$

we have

$$P_s\left\{\alpha(\nu,p)=a\right\}=P_s\left\{\left(p^a\setminus\nu\right)\cap\left(p\nmid\frac{\nu}{p^a}\right)\right\}=\left(\frac{1}{p^s}\right)^a\cdot\left(1-\frac{1}{p^s}\right)$$

**Q.E.D**.

**Theorem 2.2**

Random variables $\nu_k$, $k=0,1,2,3,...$ with Zeta distribution

$$P_s\left\{\nu_k=n\right\}=\frac{n^{-s}}{\zeta(s)},\ s>0,\ \ n\in\mathbb{N}$$

represents a random walk $\left\{\nu_k\,|\,0\le k\le\kappa(\nu)\right\}$ on a multiplicative semigroup $S(\mathbb{P}^*)$ generated by the extended set of primes $\mathbb{P}^*=\mathbb{P}\cup\left\{1\right\}$.

The walk on $\mathbb{P}^*$ is defined recursively as follows:

$$\begin{cases}\nu_1=\nu_0\cdot\eta_1,\ \text{ where }\nu_0=1,\ \eta_1=p_1^{\alpha_1(\nu)}\\ \nu_{i+1}=\nu_i\cdot\eta_{i+1},\ \text{ where }\ \eta_{i+1}=\ p_{i+1}^{\alpha_{i+1}(\nu)}(i=0,1,2,\ldots,\kappa(\nu)-1)\end{cases}\qquad(2.10)$$

The sequence $\left\{\nu_i\,|\,0\le i\le\kappa(\nu)\right\}$ is a finite walk on $S(\mathbb{P}^*)$ with independent multiplicative increments $\eta_i = p_i^{\alpha_i(\nu)}$ such that $P\left\{\eta_i = p_i^{a_i}\right\} = \left(\frac{1}{p_i^s}\right)^{a_i} \cdot \left(1 - \frac{1}{p_i^s}\right)$,

and $\kappa(\nu) \le \log_{p_{\min}} \nu = \frac{\ln \nu}{\ln p_{\min}}$, where $p_{\min}$ is the least prime number that divides $\nu$.

**Proof.**

Formulas (1.7) and (1.9) imply: $\nu = \prod_{p \in \mathbb{P}} p^{\alpha(\nu,p)} = \left(\prod_{p:\,\alpha(\nu,p)=0} 1\right) \cdot \left(\prod_{p:\,\alpha(\nu,p)>0} p^{\alpha(\nu,p)}\right) = \prod_{k=1}^{\kappa(\nu)} p_i^{\alpha_i}$

Since $\xi_i = p_i^{\alpha_i}$ and all $\alpha_i = \alpha(\nu, p_i)$, due to Theorem 1, are independent random variables each with

geometric distribution, we have $P_s\left\{\eta(i) = p_i^{a_i}\right\} = \left(\frac{1}{p_i^s}\right)^{a_i} \cdot \left(1 - \frac{1}{p_i^s}\right)$,

were $i = 1, 2, \ldots, n$, so that $\nu_n = \prod_{i=1}^{n} \eta_i$ for all $n$: $1 \le n \le \kappa(\nu)$ and $\nu(n) = \nu$ if $n = \kappa(\nu)$.

Thus, $P\{\nu = m\} = \prod_{i=1}^{\kappa(m)} \left(\frac{1}{p_i^s}\right)^{\alpha_i} \cdot \prod_{i=1}^{\infty} \left(1 - \frac{1}{p_i^s}\right) = \frac{1}{m^s} \cdot \frac{1}{\zeta(s)}$ since $m = \prod_{i=1}^{\kappa(m)} p_i^{\alpha_i}$ .

Then, $m = \prod_{i=1}^{\kappa(m)} p_i^{\alpha_i} \ge \left(p_{\min}\right)^{\kappa(m)}$, where $p_{\min} \le p_i$ for all $i$: $1 \le i \le \kappa(m)$, implies: $\kappa(m) \le \log_{p_{\min}} m$

**Q.E.D.**

**Theorem 2.3**

Let $h: R \to \{0,1\}$ be the Heaviside function $h(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases}$ , $\vec{r}(\nu) = \left(r(\nu_i) \,|\, 1 \le i \le \pi\left(\sqrt{\nu}\right)\right)$

a vector of residuals $r(\nu_i) = \operatorname{mod}(\nu, p_i)$, and $\rho(\nu) = \min\left(\vec{r}(\nu)\right) = \min_i \left(r(\nu_i) \,|\, 1 \le i \le \pi\left(\sqrt{\nu}\right)\right)$.

If a random variable $\nu$ has Zeta probability distribution and $\xi(n) = h\left(\rho(n)\right)$, then for each $n \in \mathbb{N}$ the following statements hold true:

(1) $P_s\{\nu = n \in \mathbb{P}\} = P_s\left\{h\left(\rho(\nu)\right) = 1\right\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p^s}\right)$ (2.11)

(2) $P_s\left\{\xi(n+1)=\pi(n+1)-\pi(n)=1\right\}=P_s\left\{h(\rho(\nu)=1\middle|\nu=n+1\right\}=\prod_{p\le\sqrt{n+1}}\left(1-\frac{1}{p^s}\right)$

**Proof.**

Theorem 1 implies

$$P_s\left\{\nu\in\mathbb{P}\middle|\nu=n\right\}=P_s\left\{\bigcap_{p\le\sqrt{n}}\{p\nmid\nu\}\middle|\nu=n\right\}=\prod_{p\le\sqrt{n}}P_s\left\{\{p\nmid\nu\}\mid\nu=n\right\}=\prod_{i=1}^{\pi(\sqrt{n})}\left(1-\frac{1}{p_i}\right)$$

Notice that the event $\left\{\bigcap_{p\le\sqrt{n}}\{p\nmid\nu\}\middle|\nu=n\right\}$ can be expressed in the form of conditions

$$\left\{\bigcap_{p\le\sqrt{n}}\left\{\left[\operatorname{mod}(\nu,p)>0\middle|p\in\mathbb{P},\nu=n\right]\right\}\right\}=\left\{\bigcap_{1\le i\le\sqrt{\pi(n)}}\{r_i>0\}\right\}=\left\{\min\left[r_i\middle|1\le i\le\sqrt{\pi(n)}\right]>0\right\}. \qquad (2.12)$$

By using the Heaviside function $h(x)=\begin{cases}1 & \text{if } x>0\\ 0 & \text{if } x\le 0\end{cases}$, we can write the recursive equation

for $\pi(n)$ in the form: $\pi(n+1)=\pi(n)+h\left(\min_{p\le\sqrt{n+1}}\left\{\operatorname{mod}(n+1,p)\mid p\in\mathbb{P}\right\}\right)$

or, equivalently,

$$\pi(n+1)=\pi(n)+h\left(\min_{i\le\sqrt{n}}\left\{r_i\mid r_i=\operatorname{mod}(n+1,p_i)\right\}\right)=\pi(n)+h\left(\min(\vec{r}(n+1)\right) \qquad (2.13)$$

which controls the occurrence of prime numbers in the sequence of all integers $n=3,4,5,6,\ldots$ For a random number $\nu$ with Zeta probability distribution, vector of residuals $\vec{r}(\nu)=\left(r_1(\nu),r_2(\nu),\ldots,r_{\kappa(\nu)}(\nu)\right)$ is a vector with independent random components $r_k(\nu)=\operatorname{mod}(\nu,p_k)$ distributed within congruence classes $C_{p_k,r_k(\nu)}$ for all $k:\ 1\le k\le\pi(\nu)$. For $\nu$ to be prime is necessary and sufficient that $\nu$ should not be divisible by all of primes $p\le\sqrt{\nu}$, which means that $\rho(\nu)=\min\left\{r_i(\nu)\mid 1\le i\le\pi\left(\sqrt{\nu}\right)\right\}>0$. Denoting $\xi(n)=h\left(\rho(n)\right)$ $(n=1,2,3,\ldots)$, we have:

$$P_s\left\{\xi(n+1)=\pi(n+1)-\pi(n)=1\right\}=P\left\{h(\rho(n+1)=1\right\}$$

$$= P_s\left\{\min\left\{(\overline{r}(n+1) > 0\right\}\right\} = \prod_{p \le \sqrt{n+1}} \left(1 - \frac{1}{p^s}\right), \tag{2.14}$$

since $P_s\left\{\pi(1) = 0\right\} = \dfrac{1}{\zeta(s)}$ . Therefore, by letting $s \to 1$ , we obtain

$$P_s\left\{\xi(n) = 1 \middle| \pi(1) = 0\right\} \to \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right) \tag{2.15}$$

Probability of random variable $\nu$ with Zeta distribution to be a prime number in the interval $[2, n]$ for all $n \ge 5$ is given by the formulas:

$$P\left\{\nu = n \in \mathbb{P}\right\} = P\left\{h\left(\rho(\nu)\right) = 1\right\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right),$$

$$P\left\{\xi(n) = 1 \middle| \min(r_i \mid 1 \le i \le \pi(\sqrt{n}) > 0\right\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right) \tag{2.16}$$

***Examples.***

1) $\nu = 108 = 1 \cdot 2^2 \cdot 3^3 \cdot 5^0 \cdot 7^0 \cdots$ with $\alpha(108, p) = 0$ for all $p > 3$.

We have: $\alpha(108,2) = 2, \alpha(108,3) = 3; \kappa(108) = 2$

2) $\nu$=110=$2 \cdot 3^0 \cdot 5 \cdot 7^0 \cdot 11 \cdot 13^0 \cdot 17^0 \cdots$ with $\alpha(110, p) = 0$ for $p = 3, 7$, and all $p > 11$

We have: $\alpha(110,2) = 1, \alpha(110,5) = 1, \alpha(110,11) = 1;$ $\kappa(110) = 3$.

In the above setting, the number $108 = \prod_{i=0}^{\infty} \xi(i)$ in example 1) represents the path:

$$1 \to 2^2 \to 3^3 \to 5^0 \to 7^0 \to \cdots$$

The number $110 = \prod_{i=0}^{\infty} \xi(i)$ in example 2) represents the path:

$$1 \to 2 \to 3^0 \to 5 \to 7^0 \to 11 \to 13^0 \to 17^0 \to \ldots$$

By setting $P\left\{\xi(j) = p_j^{\alpha_j}\right\} = \left(\dfrac{1}{p_j^s}\right)^{\alpha_j}$ for all $p_j \in \mathbb{P}$ , we can calculate probability $P\left\{\nu = n\right\}$ of any given value $n \in \mathbb{N}$.

In example 1):

$$P_s\{\nu=108\}=\frac{1}{2^{2s}}\cdot\left(1-\frac{1}{2^s}\right)\cdot\frac{1}{3^{3s}}\cdot\left(1-\frac{1}{3^s}\right)\cdot\left(1-\frac{1}{7^s}\right)\cdot\left(1-\frac{1}{11^s}\right)\cdots\left(1-\frac{1}{p_j^s}\right)\cdots$$

$$=\frac{1}{2^{2s}\cdot 3^{3s}}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)=\frac{1}{108^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)$$

In example 2):

$$P_s\{\nu=110\}=\frac{1}{2^s}\cdot\left(1-\frac{1}{2^s}\right)\cdot\left(1-\frac{1}{3^s}\right)\cdot\frac{1}{5^s}\cdot\left(1-\frac{1}{5^s}\right)\cdot\left(1-\frac{1}{7^s}\right)\cdot\frac{1}{11^s}\cdot\left(1-\frac{1}{11^s}\right)\cdot\left(1-\frac{1}{13^s}\right)\cdots\left(1-\frac{1}{p_j^s}\right)\cdots$$

$$=\frac{1}{2^s\cdot 5^s\cdot 11^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)=\frac{1}{110^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)$$

Notice that, in general, in the formal expression $P_s\{\nu=n\}=\frac{1}{n^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)$

the product involves a set of all prime numbers. In the above expressions the ‘probability’ $P_s\{\nu=n\}$ depends on a parameter $s$:

$$P_s\{\nu=n\}=\frac{1}{n^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)=\frac{1}{n^s\cdot\zeta(s)},\ n\in\mathbb{N}, s>1 \qquad (2.17)$$

To cope with the divergence of the infinite product $\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j}\right)=\prod_{p\in\mathbb{P}}\left(1-\frac{1}{p}\right)=\zeta(1)$,

we consider $\zeta(s)=\sum_{n\geq 1}\frac{1}{n^s}$ for $s>1$, and define the probability $P_s$ as a function of parameter $s$.

Meanwhile, there is another way to cope with divergence of $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$ for $s\leq 1$. We can do so by introducing a sequence of *incomplete* (or *partial*) *Riemann Zeta* functions. We define the *incomplete product Zeta function* $\zeta_{\mathbb{P}_N}(s)$ as a partial product in the multiplicative presentation of $\zeta(s)$ for $s\geq 1$:

$$\zeta_{\mathbb{P}_N}(s)=\prod_{p\leq N}\frac{1}{1-\frac{1}{p^s}} \qquad (2.18)$$

**Remark 2.1.**

Since $\dfrac{1}{1-\dfrac{1}{p^s}} = \sum_{k=0}^{\infty}\left(\dfrac{1}{p^s}\right)^k$, we have a convergent additive partial presentation of $\zeta(s)$:

$$\zeta_{\mathbb{P}_N}(s) = \prod_{p \le N}\left[\sum_{k=0}^{\infty} p^{-sk}\right] = \sum_{n \in S(\mathbb{P}_N)} n^{-s}. \tag{2.19}$$

Here $S(\mathbb{P}_N)$ is a multiplicative semigroup of all integers generated by $\mathbb{P}_N{}^* = \mathbb{P}_N \cup \{1\}$, where $\mathbb{P}_N = \{p \mid p \le N, p \in \mathbb{P}\}$. Notice that $S(\mathbb{P}_N)$ is an infinite set generated by a finite set $\mathbb{P}_N{}^*$. Then, we consider the corresponding probability distribution $P_{s,N}$ , $s > 1$, given by the formula:

$$P_{s,N}\{\nu = n\} = \frac{1}{n^s \cdot \zeta_{\mathbb{P}_N}(s)},\ n \in S(\mathbb{P}_N),\ s > 0,\ N \in \mathbb{N} \tag{2.20}$$

Since $\zeta_{\mathbb{P}_N}(s) = \prod_{p \le N}\left(1-\dfrac{1}{p^s}\right)^{-1} = \sum_{n \in S(\mathbb{P}_N)} \dfrac{1}{n^s}$ , we have $\sum_{n \in S(\mathbb{P}_N)} P_{s,N}\{\nu = n\} = 1$.

The probability $P_{s,N}$ of $\nu$ to be a prime number in the set of numbers $S(\mathbb{P}_N)$, generated by primes not exceeding $N$, can be calculated by the formula:

$$P_{s,N}\{\nu \text{ is prime} | \nu \in S(\mathbb{P}_N)\} = \frac{\sum_{p \in \mathbb{P}_N} p^{-s}}{\zeta_{\mathbb{P}_N}(s)} = \frac{\sum_{p \le N} p^{-s}}{\prod_{p \le N} \frac{1}{1-p^{-s}}} = \left(\sum_{p \le N} p^{-s}\right) \cdot \prod_{p \le N}\left(1-\frac{1}{p^s}\right) \tag{2.21}$$

The convergence of the infinite series $\zeta_{\mathbb{P}_N}(z) = \sum_{n \in S(\mathbb{P}_N)} n^{-s}$ is guaranteed by (2.18) and (2.19). In general, from the probabilistic point of view, every finite path on the monoid set. $S(\mathbb{P}^*) = \mathbb{N}$ can be identified with a randomly chosen natural number $\nu$ by assuming that it has a probability distribution $P\{\nu = n\}$, $n \in \mathbb{N}$, such that

$$\sum_{n=1}^{\infty} P\{\nu = n\} = 1 .$$

## 3. Asymptotics of generalized Bernoulli processes and the Cramér's model of prime numbers distribution

**Definition 3.1**

A sequence of $\{0,1\}$-valued random variables $(\xi_k)_{k\in\mathbb{N}}$ defined on probability space $(\Omega,\mathcal{F},P)$ which terms are not in general independent and identically distributed we call *generalized Bernoulli process*. We have:

$$P\{\xi_k(\omega)=1\}=P_k, P\{\xi_k(\omega)=0\}=Q_k,\ P_k+Q_k=1,\ k\in\mathbb{N}.$$

Probabilitstic approach to distribution of prime numbers in $\mathbb{N}$ is addresed in the Harald Cramér's model [2,3]. The sequence of random variables $(\nu_k)_{k\in\mathbb{N}}$ on the main probability space $(\Omega,\mathcal{F},P)$, such that $\nu_n:\Omega\to\mathbb{N}$, for some $\omega\in\Omega$ has realizations resulted in prime numbers: $\nu_k(\omega)=n\in\mathbb{P}$. The assignments of probabilities

$$P\{\nu_k(\omega)=k\in\mathbb{P}\}=P\{\xi_k=1\}=\frac{1}{\ln k},\ P\{\nu_k(\omega)=k\notin\mathbb{P}\}=P\{\xi_k=0\}=1-\frac{1}{\ln k}, \tag{3.1}$$

in the Cramér's model was originally motivated by the Prime Number Theorem [10, p.133],

where the counting function of primes on $\mathbb{N}$ is given by the asymptotic formula

$$\pi(x)=\sum_{p\in\mathbb{P}\cap[2,x]}1\sim Li(x)=\int_2^x\frac{dt}{\ln t},$$

which leads to the heuristic assumption about the probability $P\{p\in[x-1,x]\}\sim\int_{x-1}^x\frac{dt}{\ln t}\sim\frac{1}{\ln x}.$

The Cramér's model describes the occurrence of prime numbers as a special case of a Bernoulli process given by a *sequence of independent Bernoulli variables* $(\xi_k\mid k\in\mathbb{N},)$ with probabilities defined in (3.1):

$$P\{\xi_k=1\}=\frac{1}{\ln k},\ \ P\{\xi_k=0\}=1-\frac{1}{\ln k}\ ,\ \text{where } n\geq 2.$$

The above formulas (2.1) and (2.16) , due to the Merten's 1st and 2nd theorems [9, p.22], have the asymptotic expression:

$$P\{\xi(n)=1\}=\prod_{p\le\sqrt{n}}\left(1-\frac{1}{p}\right)\approx\frac{e^{-\gamma}}{\frac{1}{2}\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right]=\frac{c}{\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right], \tag{3.2}$$

where $c=\frac{2}{e^{\gamma}}\approx 1.122918968$. In both models A and B we consider all values of $n>N$ in (3.1) by choosing an arbitrary large natural $N$. As we pointed above, the Cramér's assumption about *independence of terms* in the sequence $\left(\xi(n)\,|\,n=1,2,\ldots\right)$ is not accurate for any finite subset of $\mathbb{N}$ . The more adequate approach would be to consider the sequence of consecutive primes represented by $\left(\nu_n\right)_{n\in\mathbb{N}}$ and $\left(\xi(n)\,|\,n=1,2,\ldots\right)$, respectively, as *stochastic predictable sequences of dependent random variables.*

Actually the sequence of random variables in the updated Cramér's model is *asymptotically Bernoullian* (and asymptotically pairwise independent) in a sense of Definition 3.1 given below. Meanwhile, the demand for idependence of terms in the sequence $\left(\nu_n\right)_{n\in\mathbb{N}}$ and in $\left(\xi(n)\right)_{n\in\mathbb{N}}$ could be be relaxed for the Cramér's model, due to the version of the *Central Limit Theorem (CLT) for dependent random variables* in sequences with a sort of '*asymptotically forgetful memory*' [ 7 ]. This version of CLT tracks back to the S.N. Bernstein's ideas [ 22 ]. One of the most general forms of the Central Limit Theoems for dependent variables has been proved for sequences of random walks on differentiable manifolds and Lie groups by the author [24,25].

Never the less, in what follows we use the assumption of independent terms in sequences of random variables $\{\nu_k\}_{k\in\mathbb{N}}$ as the most adequate for the goals of this article and apply here the classical form of the CLT [23].

Let discuss now, following M. Loèv [18], *asymptotic behavior of a generalized Bernoulli process.* We have for $\xi_k$ mathematical expectation $E\{\xi_k\}=P_k$ and variance $V\{\xi_k\}=P_k\cdot Q_k$

Let's denote $X_n = \frac{1}{n}\sum_{k=1}^{n}\xi_n$. Then $E\{X_n\} = \frac{1}{n}\sum_{k=1}^{n} P_k$. Since $(\xi_k)^2 = \xi_k$, we have

$E\{(\xi_k)^2\} = E\{\xi_k\}$, $E\{\xi_k \cdot \xi_l\} = P\{\xi_k \cdot \xi_l = 1\} = P_{kl}$. Then, $(E\{X_n\})^2 = \frac{1}{n^2}\left(\sum_{k=1}^{n} P_k^2 + 2\sum_{1\le k<l<]\le n} P_k \cdot P_l\right)$

and $E\{(X_n)^2\} = \frac{1}{n^2}\left(\sum_{k=1}^{n} P_k + 2\sum_{k<l} P_{kl}\right)$.

This implies:

$$V\{X_n\} = E\{(X_n)^2\} - (E\{X_n\})^2 = \frac{1}{n^2}\left(\sum_{k=1}^{n} P_k Q_k + 2\sum_{1\le k<l\le n}(P_{kl} - P_k \cdot P_l)\right) = \frac{1}{n^2}\sum_{k=1}^{n} V(\xi_k) + D_n \,, \qquad (3.4)$$

where $D_n = \frac{2}{n^2}\sum_{1\le k<l\le n}(P_{kl} - P_k \cdot P_l) = \frac{n(n-1)}{2n^2}\left(\frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_{kl} - \frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_k \cdot P_l\right)$.

If terms in $(\xi(n))_{n\in\mathbb{N}}$ are pairwise independent, then $P_{kl} = E\{\xi_k \cdot \xi_l\} = E\{\xi_k\} \cdot E\{\xi_l\} = P_k \cdot P_l$

and $D_n = 0$, which implies $V\{X_n\} = \frac{1}{n^2}\sum_{k=1}^{n} V\{\xi_k\}$.

Thus, $D_n$ can be viewed as a *cummulative measure of pairwise independence* of terms in a Bernoulli process $(\xi(n))_{n\in\mathbb{N}}$. Denote:

$$\overline{P}_1(n) = \frac{1}{n}\sum_{k=1}^{n} P_k \text{ and } \overline{P}_2(n) = \frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_{kl}.$$

Notice that

$$D_n = \frac{n-1}{2n}\left(\overline{P}_2 - \overline{P}_{1,2}\right) \text{ where } \overline{P}_{12} = \frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_k \cdot P_l.$$

We consider below a slightly different measure $d_n$ that shows how close a Bernoulli process $(\xi(n))_{n\in\mathbb{N}}$ is to a classical Bernoulli sequence of independent equally distributed random variables.

Then, $E\{(X_n)^2\} = \frac{1}{n^2}\left(\sum_{k=1}^{n} P_k + 2\sum_{k<l} P_{kl}\right) = \frac{\overline{P}_1 - \overline{P}_2}{n} + \overline{P}_2$ and $(E\{X_n\})^2 = (\overline{P}_1)^2$

Since $V\{X_n\} = E\{(X_n)^2\} - (E\{X_n\})^2 = \frac{\overline{P}_1 - \overline{P}_2}{n} + \overline{P}_2 - (\overline{P}_1)^2$, we have

$$V\left(X_n\right)=\frac{\bar{P}_1-\bar{P}_2}{n}+\bar{P}_2-\left(\bar{P}_1\right)^2=\frac{\bar{P}_1-\bar{P}_2}{n}+d_n \tag{3.5}$$

where $d_n=\bar{P}_2-\left(\bar{P}_1\right)^2$.

In the classical Bernoulli scheme witn independent identically distributed terms $\left(\xi_k\right)_{k\in\mathbb{N}}$ we have $P_{kl}=E\left\{\xi_k\cdot\xi_l\right\}=P\left\{\xi_k\cdot\xi_l=1\right\}=P\left\{\xi_k=1\right\}\cdot P\left\{\xi_l=1\right\}=P_k\cdot P_l=P^2$, due to independence and equal distribution of terms in the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$, so that $d_n=\bar{P}_2-\left(\bar{P}_1\right)^2=P^2-P^2=0$.

This implies $d_n=0$ and $V\left\{X_n\right\}=\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}$. This means that the value of $d_n$ is a measure of a deviation of the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ from a classical Bernoulli scheme.

**Definition 3.1**

We call a sequence $\left(\xi_n\right)_{n\in\mathbb{N}}$ of $\left\{0,1\right\}$-valued random variables defined on probability space $\left(\Omega,\mathcal{F},P\right)$ *asymptotically pairwise Bernoullian* if $\max_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|\to 0$ as $N\to\infty$. This means that for sufficiently large $N$ variables $\xi_k,\xi_l$ are asymptotically independent for all $l>k>N$.

**Lemma 3.1**

For *asymptotically Bernoullian sequence* $\left(\xi_n\right)_{n\in\mathbb{N}}$ we have $D_n\to 0$ so that

$$\left|V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}\right|\to 0 \text{ as } n\to\infty .$$

**Proof.**

Due to (3.2), $V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}=D_n$.

Since $D_n=\frac{2}{n^2}\sum_{k<l\le n}\left(P_{kl}-P_k\cdot P_l\right)$, and $\left|\sum_{k<l\le n}\left(P_{kl}-P_k\cdot P_l\right)\right|\le\frac{n(n-1)}{2}\max_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|$,

we have $\left|D_n\right|\le\frac{2}{n^2}\cdot\frac{n(n-1)}{2}\cdot\max_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|\to 0.$

This implies $\left|V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}\right|\to 0$.

**Q.E.D.**

Keeping in mind approximation (3.1), we restrict the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ by considering its 'tail' $\left(\xi_k\right)_{k>N}$ of the original seqience for sufficiently large $N$.

**Theorem 3.1**

The sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ in the modified Cramér's model is *asymptotically pairwise Bernoullian,*

that is $\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|=O\left(\dfrac{1}{\ln N}\right)$, where $P\left\{\xi_k=1\right\}=P_k$, $P\left\{\xi_k\cdot\xi_l=1\right\}=P_{kl}$,

and $\left|V\left\{X_n\right\}-\dfrac{1}{n^2}\sum\limits_{k=1}^{n}V\left\{\xi_k\right\}\right|=O\left(\dfrac{1}{\ln N}\right)$ as $D_n=O\left(\dfrac{1}{\ln N}\right)$ for all $n>N$. (3.6)

**Proof.**

Indeed, $P\left\{\xi_k=1\right\}=P_k$, $P\left\{\xi_k\cdot\xi_l=1\right\}=P_{kl}$. Then, since $\left|P_{kl}-P_k\cdot P_l\right|<P_k\le\dfrac{1}{\ln N}$

for all $N<k<l\le n$, we have $\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|\le\dfrac{c}{\ln N}\to 0$ and $D_n=O\left(\dfrac{1}{\ln N}\right)$ as $N\to\infty$.

This implies $\left|V\left\{X_n\right\}-\dfrac{1}{n^2}\sum\limits_{k=1}^{n}V\left\{\xi_k\right\}\right|=O\left(\dfrac{1}{\ln N}\right)$ for all $n>N$.

**Q.E.D.**

In the Cramér's model $\hat{\pi}_N\left(n\right)=\sum\limits_{k=N}^{N+n}\xi(k)$ represents the number of primes among $n$ terrms

in the interval $(N,N+n]$ of the sequence and $\dfrac{\hat{\pi}_N(n)}{n}=\dfrac{1}{n}\sum\limits_{k=N}^{N+n}\xi(k)=\dfrac{\hat{\pi}_N(n)}{n}$ is a relative freqiency

of primes for these terms. predicted by the improved model based on Zeta probability distribution.

In the Table 4 below, we demonstrate how well $E\left\{\dfrac{\hat{\pi}(n)}{n}\right\}$ approximates relative frequencies

of primes $\dfrac{\pi(n)}{n}$ in the Zeta distribution model for $\left(\xi_k\right)_{k\ge 3}$ as $n$ increses from $10^1$ to $10^9$.

**Table 3.1**. *Comparison of probabilities* $P\{\nu \in \mathbb{P} | \nu = n\}$ *and frequencies* $\frac{\pi(n)}{n}$ *of primes in intervals* $[1, n]$

| Natural $n$ | $P\{\nu \in \mathbb{P} \| \nu = n\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right)$ | $\frac{\pi(n)}{n}$ |
|---|---|---|
| $10^1$ | 0.33333333 | 0.400000 00 |
| $10^2$ | 0.22857143 | 0.25000000 |
| $10^3$ | 0.15285215 | 0.16800000 |
| $10^4$ | 0.12031729 | 0.12290000 |
| $10^5$ | 0.09621491 | 0.09592000 |
| $10^6$ | 0.08096526 | 0.07849800 |
| $10^7$ | 0.06957939 | 0.06645790 |
| $10^8$ | 0.06088469 | 0.05761455 |
| $10^9$ | 0.05416682 | 0.05084753 |

Consider now the *Generalized Law of Large Numbers* for a general Bernoulli process as it stated in [18] and apply it then to Zeta distribution model for $(\xi_k)_{k \in \mathbb{N}}$.

**Theorem 3.2**

Let $\xi(k) = \begin{cases} 1 \text{ if } k \in \mathbb{P} \\ 0 \text{ otherwise} \end{cases}$ and $\frac{\hat{\pi}_N(n)}{n} = \frac{1}{n} \sum_{k=N+1}^{N+n} \xi(k)$ be a relative freqiency of primes the interval $[N, N+n]$. Then, the *Generalized Law of Large Numbers* holds true:

$$P\left\{\left|\frac{\hat{\pi}_N(n)}{n} - E\left\{\frac{\hat{\pi}_N(n)}{n}\right\}\right| > \varepsilon\right\} \to 0 \text{ as } N, n \to \infty. \tag{3.7}$$

If $d_{n.N} = \frac{2}{n(n-1)} \sum_{N \le k < l \le N+n} P\{(k \in \mathbb{P}) \cap (l \in \mathbb{P})\} - \left(\frac{1}{n} \sum_{k=N}^{N+n} P\{k \in \mathbb{P}\}\right)^2 = O\left(\frac{1}{n}\right)$,

then the *Generalized Strong Law of Large Numbers* holds true:

$$P\left\{\left|\frac{\hat{\pi}_N(n)}{n} - E\left\{\frac{\hat{\pi}_N(n)}{n}\right\}\right| \to 0\right\} = 1 \text{ as } N, n \to \infty \tag{3.8}$$

**Proof.**

Due to [25], we apply the following *Propositions:*

1. Generalized Bernoulli Theorem that for every $\varepsilon > 0$: $P\left\{\left|X_n - E\left\{X_n\right\}\right| > \varepsilon\right\} \to 0$

holds true for a Bernoulli process $\left(\xi_k\right)_{k\in\mathbb{N}}$ if and only if $d_n = \overline{P}_2 - \left(\overline{P}_1\right)^2 \to 0$ as $n \to \infty$.

2. Generalized Strong form of Bernoulli Theorem that $P\left\{\left|X_n - E\left\{X_n\right\}\right| \to 0\right\} \to 1$

holds true if $d_n = O\left(\frac{1}{n}\right)$.

We show here that these propositions asymptotically hold true for tails $\left(\xi_k\right)_{k\geq N}$.

For tail $\left(\xi_k\right)_{k\geq N}$ in the framework of Cramér's model we have:

$$\overline{P}_{1,N}(n) = E\left\{\frac{\hat{\pi}_N(n)}{n}\right\} = \frac{1}{n}\sum_{k=N+1}^{N+n} E\left\{\xi(k)\right\} = \frac{1}{n}\sum_{k=N+1}^{N+n} P\left\{k\in\mathbb{P}\right\} = \frac{1}{n}\sum_{k=N+1}^{N+n}\frac{1}{\ln k} \sim \int_N^{N+n}\frac{dt}{\ln t} = Li(N+n) - Li(N)$$

$$\overline{P}_{2,N}(n) = \frac{2}{n(n-1)}\sum_{N\leq k<l\leq N+n} E\left\{\xi(k)\cdot\xi(l)\right\} = \frac{2}{n(n-1)}\sum_{N\leq k<l\leq N+n} P\left\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\right\}\cdot$$

Then, $d_{n.N} = \dfrac{2}{n(n-1)}\displaystyle\sum_{N\leq k<l\leq N+n} P\left\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\right\} - \left(\frac{1}{n}\sum_{k=N}^{N+n} P\left\{k\in\mathbb{P}\right\}\right)^2$

Notice that $d_{n.N} \leq \max\limits_{N\leq k<l\leq N+n}\left(P\left\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\right\}\right) < \dfrac{1}{\ln N}$ implies $d_{n,N} \to 0$ as $n, N \to \infty$ .

This implies $d_{n,N} \to 0$ as $N, n \to \infty$.

Then, $P\left\{\left|\dfrac{\hat{\pi}_N(n)}{n} - E\left\{\dfrac{\hat{\pi}_N(n)}{n}\right\}\right| > \varepsilon\right\} = O\left(\dfrac{1}{\ln N}\right)$ as $n, N \to \infty$ and (9) holds true.

In addition, if $d_{n,N} = O\left(\dfrac{1}{n}\right)$, as $n, N \to \infty$, then (10) holds true.

**Q.E.D.**

## 4. Probabilistic proof of the Strong Goldbach Conjecture

According to the conjecture stated by Goldbach in his letter to Euler in 1742, "every even number $2m \geq 6$ is the sum of two odd primes" [1]. Regardless numerous attempts to prove the statement, supported in our days by computer calculations up to $4 \times 10^{18}$ , it remains unproven till now.
In this part we try to solve the 'puzzle' in the framework of Probability Theory, by using the modified Cramér's probabilistic model for distribution of primes in the sequence of natural numbers $\mathbb{N}$.
Strong Goldbach Conjecture (SGC), as one of the oldest notoriously known problems in Number Theory, raises a question, why it seems so difficult to decide whether the equation

$$p + p' = 2m \tag{4.1}$$

where $p$ and $p'$ are prime numbers, has at least one solution for each even number $2m \geq 6$.
Indeed, occurrences of primes look very sporadic, so this is hard to oversee all possible partitions $(p, p')$ for even numbers, like $p + p' = 2m$, especially for 'big' values of $m$.
One of ideas to solve such a combinatorial problem is to apply methods of Probability Theory.
The first obstacle in probabilistic approach is assumed 'randomness' of occurrences of prime numbers, the second – 'independence' of their occurrences. From probabilistic point of view, a sequence of natural numbers is considered as realization of a series of trials, each of which results either in a prime number or in a composite number, occurring with certain probabilities.
Notice that every integer solution $(n, n')$ in primes to the equation (4.1) must satisfy the inequalities: $3 \leq n \leq n' \leq 2m - 3$. For each integer $m \geq 3$ we can populate interval of integers $I_m = [3, 2m-3]$ by randomly and independently chosen numbers $(n, n')$ that belong to this interval, in a hope that a pair $(n, n') = (p, p') \in \mathbb{P}^2$ would satisfy the equation (4.1), if such a pair exists.

A well-known serious objection to this approach in solving SGC problem, pointed out by Hardy and Littlewood [26], is that 'randomly chosen' primes in a pair $(p, p')$ such that $p + p' = 2m$, must be realizations of dependent random variables. This means that, given $m$, the choice of a prime number $p$ in the equation (1) completely determines the choice of $p'$, which should occur with the same probability as $p$. Meanwhile, a useful probabilistic assumption is that each occurrence of a pair of primes $(n, n')$ is a realization of a trial given by two independent random variables $(\nu, \nu')$.

In the Hardy-Littlewood objection, instead of simultaneously 'rolling' two dice at a time for a pair of integer outcome $(n,n')$, Hardy and Littlewood were 'rolling' just one die.

The correct resolution of this issue should be based on a reasonable definition of probability space $\Omega = \Omega_{\nu} \times \Omega_{\nu'}$ (a set of all possible elementary outcomes) of the 'game', generating pairs of prime numbers. The key point is that $\nu$ and $\nu'$ are considered as independent random variables in a pair $(\nu,\nu')$, and among their realizations $(n,n')$ we are interested in those with satisfy the equation (4.1). If there exists a pair $(\nu,\nu')$ with probability distribution that guarantees for every $m \ge 3$ occurrence of pairs $(p,p')$ satisfying (4.1), then we can say that SGC is confirmed. This is an objective of this part of the paper.

**Definition 4.1**

Prime numbers $p \in \mathbb{P}, p' \in \mathbb{P}$ we call $G_m$- *primes* if there exist an even number $2m \ge 6$ such that $2m = p + p'$. The set of all $G_m$-*primes* for a given $m$ we denote $G_m\mathbb{P}$.

For each natural $m \ge 3$ we define Goldbach function $G(2m)$ as a number of primes solving the equation $2m = p + p'$. Thus, $G(2m) = \left|G_m\mathbb{P}\right|$, where $|A|$ is a number of elements in a finite set $A$. As we have pointed above, we have $G_m\mathbb{P} \subseteq I_m = [3, 2m-3]$ for each $m \ge 3$.

In the context of the Strong Goldbach conjecture (SGC) we are interested in evaluation of $G(2m)$ for all even numbers $2m$ in the form $2m = p + p'$, where $(p,p') \in \mathbb{P}^2$, $m \ge 3$.

Evaluation of $G(2m)$ for each natural $m \ge 3$ is a difficult combinatorial problem.

Calculations show that Goldbach function $G(2m)$ asymptotically increases as $m$ increases (rather not in a monotonic way) and becomes larger for the larger values of $m$ (see Figure 4.3), but so far there is no conclusive statements regarding behavior of $G(2m)$ as $m \to \infty$. Examples of sets $G_m\mathbb{P}$ for $2m = 10, 10^2, 10^3$ with the corresponding values of $G(2m)$ are given in the following table.

**Table 4.1.**

| | | |
|---|---|---|
| $G_m$ - *primes in sets* $G_mP$ *and Goldbach function values* | | |
| $2m$ | $G_m\mathbb{P}$ sets | $G(2m)$ |
| 10 | 3 5 7 | 3 |
| 100 | 3 11 17 29 41 47 53 59 71 83 89 97 | 12 |
| 1000 | 3 17 23 29 47 53 59 71 89 113 | |
| | 137 173 179 191 227 239 257 281 317 | |
| | 347 353 359 383 401 431 443 479 491 | |
| | 509 521 557 569 599 617 641 647 653 | |
| | 683 719 743 761 773 809 821 827 863 | |
| | 887 911 929 941 947 953 971 977 983 | |
| | 997 | 56 |

The idea of probabilistic approach in this context is based on presentation of a naturally ordered sequence $(p_i)_{i\in\mathbb{N}}$ of prime numbers as *realizations* $\nu_k(\omega)=k$ of independent random variables in the sequence $(\nu_k)_{k\in\mathbb{N}}$ such that $P\{\nu_k = k \in \mathbb{P}\} = P_k$, $P\{\nu_k = k \notin \mathbb{P}\} = Q_k = 1 - P_k$ .

There are two requirements for adequate presentation of primes by a sequence of random variables $(\nu_k)_{k\in\mathbb{N}}$ that we demand from a probabilistic model:

1) the choice of probability values $P_k$ should provide an accurate asymptotic approximation to the actual distribution of prime numbers in $\mathbb{N}$ for large values of $k$ (that is as $k \to \infty$) Meanwhile, a probabilistic model is not designed to guarantee 'intuitively correct' assignments of probabilities $P\{\nu_k(\omega) = k \in \mathbb{P}\}$ to concrete values of each natural number $k \in \mathbb{N}$.
2) the joint probability distribution of random variables in the sequence $(\nu_k)_{k\in\mathbb{N}}$ should objectively reproduce dependence in occurrence of primes in $(p_i)_{i\in\mathbb{N}}$ if such probabilistic dependence exists, as $\nu_k = k \to \infty$.

To address the conditions mentioned above, we consider two options for the sequence $(v_k)_{k\in\mathbb{N}}$:

A. *The Cramér's model* for occurrences of prime numbers in the sequence of independent random variables $(v_k)_{k\in\mathbb{N}}$. Recall that in the Cramér's model we consider the sequence of prime numbers as realizations of independent random variables $v_k,\, k\in\mathbb{N}$, on a probability space $(\Omega,\mathcal{F},\mathcal{P})$, such that

$$P\{\xi_k=1\}=P\{v_k=k\in\mathbb{P}\}=P_k=\frac{1}{\ln k},$$
$$P\{\xi_k=0\}=P\{v_k=k\notin\mathbb{P}\}=Q_k=1-\frac{1}{\ln k} \qquad (4.1)$$

Here $\xi_k=\begin{cases}1, \text{ if } v_k=k\in\mathbb{P}\\ 0, \text{ otherwise}\end{cases}$, is an indicator function for primes in sequence of realizations $v_k=k$ of random variables $v_k,\, k\in\mathbb{N}$.

B. *Zeta probability distribution model* considers occurrences of prime numbers in the sequence of independent random variables $(v_k)_{k\in\mathbb{N}}$, where each integer $k$ is *a realization* $v_k=k$ *of random variable* $v_k$ on a probability space $(\Omega,\mathcal{F},\mathcal{P})$, following Zeta probability distribution

$$P_s\{v_k=k\}=\frac{k^{-s}}{\zeta(s)},\, s>1, \qquad (4.2)$$

**Remark 4.**1

In both models A and B we consider sample spaces $\Omega$ for sequences of random variables $(v_k)_{k\in\mathbb{N}}$ as $\Omega=\mathbb{N}^{\mathbb{N}}$, respectively, and sample spaces for the corresponding $(1,0)$-valued sequences $\{\xi_k\}_{k\in\mathbb{N}}$ as $\Omega=\{1,0\}^{\mathbb{N}}$. In both models, each $\sigma$-algebra $\mathcal{F}$ of events is generated by all finite subsets of the corresponding sample space $\Omega$. As stated in Theorem 2.3 (formula 2.11), if each $v_k$ follows Zeta probability distribution (4.2),

$$P_s\{\xi_k = 1\} = P_s\{\nu_k = n \in \mathbb{P}\} = P_k = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p^s}\right)$$
$$P_s\{\xi_k = 0\} = P_s\{\nu_k = n \notin \mathbb{P}\} = Q_k = 1 - \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p^s}\right) \quad (4.3)$$

This allows us to think of each pair $(\nu_k, \nu_{k'})$ with $k \neq k'$, taken from sequence $(\nu_k)_{k \in \mathbb{N}}$, as of a pair of independent random variables with probabilities given in (4.3). Since the sequences $(\nu_k)_{k \in \mathbb{N}}$ in both Cramér's model and Zeta distribution model are assumed to consist of independent terms $\nu_k$, pairs $(\nu_k, \nu_{k'})$ inherit the same property of independence for their components.

Notice that both models asymptotically agree with each other as stated below in Lemma 4.3.

**Remark 4.2**

As we mentioned above, realization of sums $\nu_k + \nu_{k'} = 2m$ may cause certain confusion related to possible dependence of events $\{\nu_k \in \mathbb{P}\}$ and $\{\nu_{k'} \in \mathbb{P}\}$. The problem of dependence for primes in the equation $\nu_k + \nu_{k'} = 2m$ allegedly undermines 'heuristic justification' of "a very crude probabilistic argument" [see the article 'Goldbach conjecture', Wikipedia] for evaluation of probability for realizations of a 'random' pair $(\nu_k, \nu_{k'})$ based on a 'product rule':

$$P\{\nu_k = k \in \mathbb{P}, \nu_{k'} = 2m - k \in \mathbb{P}\} = P\{\nu_k = k \in \mathbb{P}\} \cdot P\{\nu_{k'} = 2m - k \in \mathbb{P}\}.$$

This situation has been addressed in 1923 by Hardy and Littlewood in their *Hardy – Littlewood prime tuple conjecture* [26]. Meanwhile, the dependence of variables $\nu_k$ and $\nu_{k'}$ should be considered in the framework of an appropriate probability space. Our approach surmounts this obstacle:every pair of integers $(n, n')$ is considered as *a realization of independent random variables* $(\nu_k, \nu_{k'})$ with $k \neq k'$, each of which follows distributions specified in model A or in model B. We assume that *random variables in the sequence* $(\nu_k)_{k \in \mathbb{N}}$ *are* ***independent***.

To evaluate $G(2m)$, define indicator function:

$$\gamma_m(k, k') = \begin{cases} 1 \text{ if } k \in \mathbb{P} \text{ and } k' \in \mathbb{P} \\ 0, \text{ otherwise} \end{cases} \quad (4,4)$$

Then, $G(2m) = \sum_{k=3}^{2m-3} \gamma_m(k, 2m-k)$. Consider integers $k$ and $k'$ in the given interval $[3, 2m-3]$

as *realizations* $n = v_k$ and $n' = v_{k'}$ *of random variable*s $v_k(\omega)$and $v_{k'}(\omega)$

on probability space $(\Omega, \mathcal{F}, \mathcal{P})$, which follow probability distribution according

to Cramér's model or Zeta probability distribution model.

Realizations $v_k(\omega) = k$ and $v_{k'}(\omega) = k'$ are determined by the choice of elementary events

$\omega \in \Omega$ from the set $\Omega$ of all elementary events. The choice of Zeta distribution is motivated, in particular, by the fact that, due to Theorems 2.1, 2.2, 2.3 and Lemmas 4.1 and 4.3, it provides the validity of the probabilistic Cramér's model for asymptotic prime number distribution, in a full agreement with the *Prime Number Theorem*. This is especially important for SGC since we are interested in the asymptotic behavior of $G(2m)$ as $m \to \infty$.

By substituting $n = v_{mk}(\omega)$ and $n' = v_{k'}(\omega)$ into each deterministic indicator function $\gamma_m(n)$, we obtain 'randomization' of these functions. Thus, each of the 'randomized' functions $\gamma_m(v_k, v_{k'}) = \gamma_m(k, k')$ takes values $1$ or $0$ with probabilities $P\{(v_k, v_{k'}) = (k, k')\}$.

The combinations of $1$ or $0$ values of $\gamma(k, k')$, occurred for all $k, k'$ in the interval $[3, 2m-3]$, determines the counts of numbers in each set $G_m\mathbb{P}$. We have then, $2m = v_k + v_{k'}$, where $v_k \in \mathbb{P}$, $v_{k'} \in \mathbb{P}$. Consider set $\vec{v}_m = \{v_k\}_{k \in I_m}$, where $I_m = [3, 2m-3]$, and define random variable $G(2m, \vec{v}_m)$ as a sum of (2m-5) independent Bernoulli variables

$G(2m, \vec{v}_m) = \sum_{k=3}^{2m-3} \gamma_m(v_k, v_{2m-k})$. We summarize this in the following Lemma

**Lemma 4.1**

The Goldbach function $G(2m)$, $m \geq 3$, represents a sum of $2m-5$ realizations of independent Bernoulian random variables:

$$G(2m, \vec{v}_m) = \sum_{k=3}^{2m-3} \gamma_m(v_k, v_{2m-k}), \tag{4.3}$$

where $\gamma_m(k,k') = \begin{cases} 1 \text{ if } (k,k') \in \mathbb{P}^2 \\ 0 \text{ otherwise} \end{cases}$ , $\vec{v}_m = \{v_k\}_{k \in I_m}$ is a subsequence of the sequence of independent variables $(v_k)_{k\in\mathbb{N}}$ that follows Zeta probability distribution and $P\{v_k = k \in \mathbb{P}\} = P\{\xi_k = 1\}$,

Then,

$$P\{\gamma_m(v_k, v_{2m-k}) = 1\} = \prod_{p \le \sqrt{k}} \left(1 - \frac{1}{p}\right) \cdot \prod_{p \le \sqrt{2m-k}} \left(1 - \frac{1}{p}\right)$$

**Proof.**

Due to (2.11) and (4.5), $P\{\xi_k = 1\} = \prod_{p \le \sqrt{k}} \left(1 - \frac{1}{p}\right)$. Independence of random variables $v_k, v_{2m-k}$ in the subsequence $\vec{v}_m = \{v_k\}_{k \in I_m}$ implies

$$P\{\gamma_m(v_k, v_{2m-k}) = 1\} = P\{v_k \in \mathbb{P}\} \cdot P\{v_{2m-k}\mathbb{P}\}$$

$$= P\{\xi_k = 1\} \cdot P\{\xi_{2m-k} = 1\} = \prod_{p \le \sqrt{k}} \left(1 - \frac{1}{p}\right) \cdot \prod_{p \le \sqrt{2m-k}} \left(1 - \frac{1}{p}\right)$$

**Q.E.D.**

Defining $\gamma_m(k,k')$ as above, we write $P\{\gamma_m(k, 2m-k) = 1\} = P\{v_k \in \mathbb{P}, v_{2m-k} \in \mathbb{P}\}$, and

$$G(2m, \vec{v}_m) = \sum_{k=3}^{2m-3} \gamma_m(k, 2m-k) \tag{4.4}$$

Denote $g(n) = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right)$ and $\beta(m,n) = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right) \cdot \prod_{p \le \sqrt{2m-n}} \left(1 - \frac{1}{p}\right) = g(n) \cdot g(2m-n)$

Notice that the defined above $g(n)$ is a monotonically decreasing function $g: \mathbb{N} \to (0,1)$.

The following Lemma adresses the behavior of $\beta(m,n)$ for $3 \le n \le 2m$ .

**Lemma 4.2**

Let $g: [a, 2m-a] \to (0,1)$ be a monotonically decreasing function of natural $n$ where $a \le n \le 2m-a$. Then $\beta(m,n) = g(n) \cdot g(2m-n)$ gets its minimum value on $[a, 2m]$ at $n = m$, that is $\min_{a \le n \le 2m-1} \beta(m,n) = \beta(m,m) = g^2(m)$ for $n < m$.

**Proof.**

For $n \le m$ we have $\beta(m,n) = g(n) \cdot g(2m-n) \ge g(m) \cdot g(m)$ since $g(n) > g(m)$ and $g(2m-n) \ge g(m)$. Notice that $\beta(m,n) = g(n) \cdot g(2m-n)$, as a function of $n$, has a symmetry about $m$. This implies that for $n \ge m$ we have $\beta(m,n) \ge \beta(m,m) = g^2(m)$.

Thus, $\min\limits_{a \le n \le 2m-a} \beta(m,n) = \beta(m,m) = g^2(m)$.

**Q.E.D.**

Lemma 4.2 implies the inequality

$$\beta(m,n) \ge \beta(m,m) \text{ for } m,n \text{ such that } 3 \le n \le 2m-n \tag{4.5}$$

for functions $g(n) = \prod\limits_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right)$ and $\beta(m,n) = g(n) \cdot g(2m-n)$

**Lemma 4.3.**

Both models A and B asymptotically equivalent, that is $P_A\{\xi(n)\} \sim P_B\{\xi(n)\}$ as $n \to \infty$,

where $P_A\{\xi(n)=1\} = \frac{1}{\ln n}$, $P_B\{\xi(n)=1\} = \prod\limits_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right)$.

**Proof.**

Validity of the choice of probabilities in the Cramer's model (A) and in Zeta probability model (B) is supported by formula the (2.11) in Theorem 2.3, and by Merten's 2nd theorem ('Merten's Formula') [9, p.21-22]. Indeed, by using (2.11) and the Merten's 2-nd theorem (30), we have:

$$P_B\{\xi(n)=1\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\frac{1}{2}\ln(n)} \left[1 + O\left(\frac{1}{\ln(n)}\right)\right] = \frac{c}{\ln(n)} \left[1 + O\left(\frac{1}{\ln(n)}\right)\right] \tag{4.6}$$

where $c = \frac{2}{e^{\gamma}} \approx 1.122918968$ and $\gamma = \lim\limits_{n \to n} \left( \sum\limits_{k=1}^{n} \frac{1}{k} - \ln n \right) \approx 0.577215664$ is Euler's constant.

Then, we have $P_A\{\xi(n)=1\} = E\{\xi(n)\} \sim \frac{c}{\ln(n)}$. This implies $P_A\{\xi(n)\} \sim P_B\{\xi(n)\}$ as $n \to \infty$.

**Q.E.D.**

Due to (4.5, 4.6), we can evaluate

$$\beta(m,n) = c^2 \cdot \frac{1}{\ln(n)} \cdot \frac{1}{\ln(2m-n)} \cdot \left[1 + \frac{C}{\ln(n)}\right] \cdot \left[1 + \frac{C}{\ln(2m-n)}\right] \tag{4.7}$$

for $3 \le n \le 2m-3$ with an certain choice of a constant $C > 0$. Then, (4.7) implies

$\beta(m,m) = \frac{C'}{\ln^2(m)}\left[1 + \frac{C'}{\ln(m)}\right]^2$ for $n = m$ and an appropriate choise of constant $C' > 0$.

Some authors assume the constant $c = \frac{2}{e^{\gamma}}$ in (4.6) appears as a correcting coefficient for the Cramér's model as a compensation for possible pairwise dependence of prime occurences ignored in the model.

Then, since $E\left\{\gamma(v_{mj})\right\} = \beta(m,n)$ and $G(2m,\vec{v}_m) = \sum_{k=3}^{m-3} \gamma_m(v_k, v_{k'})$, we have the following expressions for mathematical expectation and variance, respectively:

$$E\left\{G(2m,\vec{v}_m)\right\} = \sum_{n=3}^{m-3} E\left\{\gamma_m(v_k, v_{k'})\right\} = \sum_{n=3}^{2\cdot m-3} \beta(m,n) \sim \int_3^{2m-3} \beta(m,t)dt, \tag{4.8}$$

Due to independence of $v_k$ in $\vec{v}_m = \left(v_k\right)_{3 \le k \le m-3}$, we have

$$\begin{aligned} Var\left\{G(2m,\vec{v}_m)\right\} &= \sum_{k=3}^{m-3} Var\left\{\gamma_m(v_k, v_{2m-k})\right\} \\ &= \sum_{n=3}^{2\cdot m-3} \left[\beta(m,n)\cdot(1-\beta(m,n)\right] \sim \int_3^{2m-3} \beta(m,t)\cdot\left(1-\beta(m,t)\right)dt \end{aligned} \tag{4.9}$$

Notice that we use the approximations (4.6), (4.7) to prove the following Theorem 4.1 related to the Goldbach Conjecture. Figure 4.1 below illustrate growth of function $E\left\{G(2m,\vec{v}_m)\right\}$ for $m = 10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7$.

**Figure 4.1**

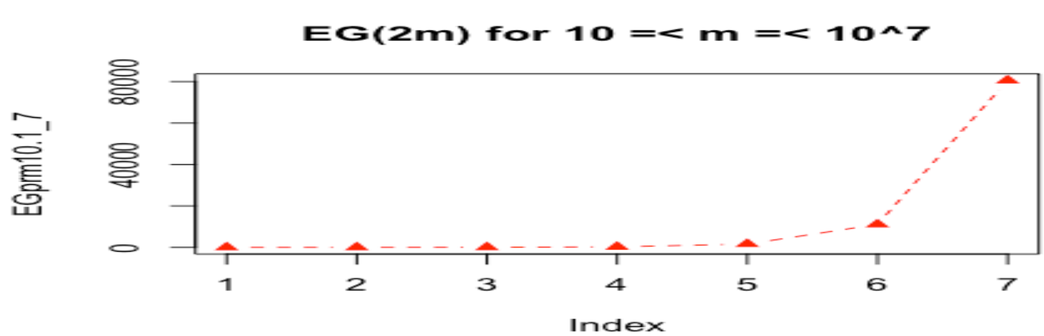
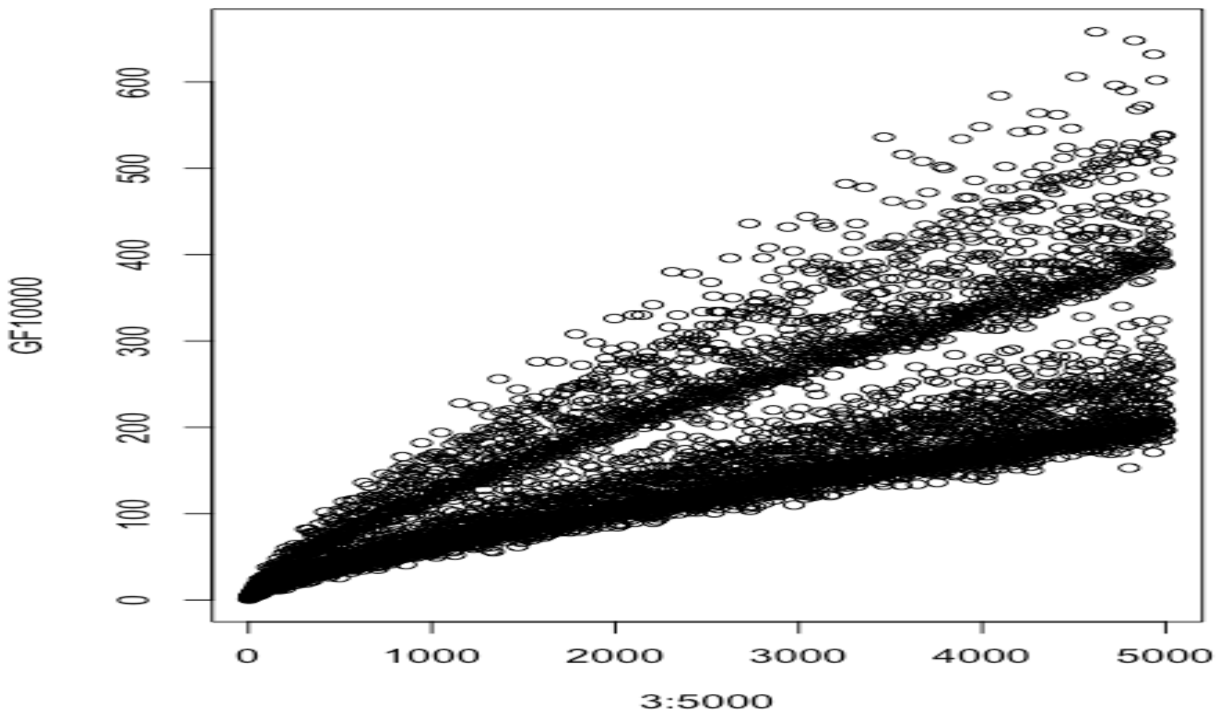



**Figure 4.2**

Plot of Goldbach function $G(2m)$ for the natural sequence of $m$ in the domain $3 \leq m \leq 5000$




The Goldbach Conjecture for large values of $m$ can be stated in the form:

probability that $G(2m,\vec{v}_m)=\sum_{k=3}^{2m-3}\gamma_m(k,2m-k)\}$ tends to 1 for all $m \geq M$ as $M \to \infty$ .

Assumption that $G(2m,\vec{v}_m)=0$ for some arbitrary large value of $m$ is in contradiction with stochastic behavior of $G(2m,\vec{v}_m)$ when $m$ increases, as we demonstrate below.

**Theorem 4.1**

Let $G_m\mathbb{P}$ for $m \geq 3$ be a set of all $G$-primes, that is prime numbers $p, p' \in \mathbb{P}$ such that $p + p' = 2m$. Let each random variable $v_k$ in the sequence of independent random variables $\left(v_k\right)_{k\in\mathbb{P}}$ follow Zeta probability distribution: $P\left\{v_k = n\right\} = \dfrac{n^{-s}}{\zeta(s)}$ $(s>1)$ and $\vec{v}_m = \left\{(v_k\right\}_{3\leq k\leq 2m-3}$ is a subsequence of the sequence of primes $\left(v_k\right)_{k\in\mathbb{P}}$.

$\gamma_m(k,k') = \begin{cases} 1 \text{ if } k \in \mathbb{P} \text{ and } k' \in \mathbb{P} \\ 0 \text{ otherwise} \end{cases}$, where $\gamma_m\left(v_k, v_{2m-k}\right) = \gamma_m\left(k, 2m-k\right)$.

Then $\left\{\gamma_m(v_k, v_{2m-k}\right\}_{3\le k\le 2m-3}$ is a sequence of independent Bernoulli variables and the randomized Goldbach function $G(2m,\vec{v}_m) = \sum_{k=3}^{m-3} \gamma_m(v_k, v_{2m-k})$ has the following properties:

(1) $$P\left\{G(2m,\vec{v}_m)=0\right\} = P\left\{\bigcap_{k=3}^{2m-3}\left\{\gamma_m(v_k,v_{2m-k})=0\right\}\right\} \to 0 \text{ as } m\to\infty.$$

(2) $$\sum_{m=3}^{\infty} P\left\{G(2m,\vec{v}_m)=0\right\} < \sum_{m=3}^{\infty} e^{-C\cdot\frac{2m-6}{\ln^2(m)}} < \infty \quad \left(C>0\right).$$

(3) $$\lim_{M\to\infty} P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,\vec{v}_m) = |G_m\mathbb{P}| \ >0\right\}\right\} \to 1$$

**Proof.**

(1) Independence of the Bernoulli variables in the set $\left\{\gamma_m(v_{k,} v_{2m-k,}) \,|\, 3\le k\le 2m-3\right\}$ follows from assumed independence of $v_k$ in the sequence $\left\{v_k\right\}_{k\in\mathbb{N}}$. This implies:

$$P\left\{G(2m,\vec{v}_m)=0\right\} = P\left\{\bigcap_{k=3}^{2m-3}\left\{\gamma_m(v_k,v_{2m-k})=0\right\}\right\} = \prod_{i=3}^{2m-3} P\left\{\gamma_m(v_k,v_{2m-k})=0\right\} = \prod_{n=3}^{2m-3}\left[1-\beta(m,n)\right].$$

Due to Lemmas 4.1 and formula (4.7), $\beta(m,n) \ge \beta(m,m) \sim \frac{C}{(\ln m)^2}\left[1+\frac{C}{\ln m}\right]^2$ for an appropriate choice of constant $C>0$. From this follows $1-\beta(m,n) \le 1-\beta(m,m)$ and

(2) $$P\left\{G(2m,\vec{v}_m)=0\right\} \le \prod_{n=3}^{2m-3}\left[1-\frac{C}{(\ln m)^2}\cdot\left(1+\frac{C}{\ln m}\right)^2\right] = \left[1-\frac{C}{(\ln m)^2}\cdot\left(1+\frac{C}{\ln m}\right)^2\right]^{2m-5} \sim e^{-\frac{2m-5}{D(m)}}$$

where $D(m) = \frac{C}{(\ln m)^2}\cdot\left(1+\frac{C}{\ln m}\right)^2$ and $e^{-\frac{2m-5}{D(m)}} \to 0$ as $m\to\infty$.

This proves that $P\left\{G(2m,\vec{v}_m)>0\right\} \to 1$ as $m\to\infty$.

(3) A critical question for the Goldbach Conjecture can be stated as follows:
is this true that for 'sufficiently large' values of $m$ such that $m > M \ge 3$ the probability that all sets $G_m\mathbb{P}$ are not empty is equal to $1$:

$$P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,\vec{v}_m)=|G_m\mathbb{P}|\ >0\right\}\right\}\to 1 \text{ as } M\to\infty.$$

Consider the probability of the opposite event: $P\left\{\bigcup_{m=M}^{\infty}\left\{G(2m,\vec{v}_m)\right\}=0\right\}$ and prove that

$$P\left\{\bigcup_{m=M}^{\infty}\left\{G(2m,\vec{v}_m)\right\}=0\right\}\to 0 \text{ as } m\to\infty.$$

This is a probability that for sufficiently large value of $M$ there exists at least one value of $m\geq M$ such that the set $G_m\mathbb{P}$ is empty.

We have: $P\left\{\bigcup_{m=3}^{\infty}\left\{G(2m,\vec{v}_m)=0\right\}\right\}\leq\sum_{m=3}^{\infty}P\left\{G(2m,\vec{v}_m)=0\right\}<\sum_{m=3}^{\infty}e^{-\frac{2m-5}{D(m)}}<\infty.$

Then, $P\left\{\bigcup_{m=M}^{\infty}\left\{G(2m,\vec{v}_m)\right\}=0\right\}\leq\sum_{m=M}^{\infty}P\left\{G(2m,\vec{v}_m)=0\right\}\to 0$ as $M\to\infty$

due to convergence of the series $\sum_{m=3}^{\infty}P\left\{G(2m,\vec{v}_m)=0\right\}$.

**Q.E.D.**

There is another way to evaluate the probability $P\left\{\left|G_m\mathbb{P}\right|<1\right\}$ as demonstrated below.

**Theorem 4.2**

Let $G_m\mathbb{P}$ for $m\geq 3$ be a set of all $G$-primes, that is prime numbers $p,p'\in\mathbb{P}$ such that $p+p'=2m$.

Let each random variable $\nu_k$ in the sequence $\left(\nu_k\right)_{k\in\mathbb{N}}$ follows Zeta probability distribution:

$P\left\{\nu_k=n\right\}=\dfrac{n^{-s}}{\zeta(s)}$ $(s>1)$. Consider $\vec{\nu}_m=\left(\nu_k\right)_{3\leq k\leq 2m-3}$ as a subsequence of $\left(\nu_k\right)_{k\in\mathbb{N}}$.

Let $\gamma_m\left(\nu_k,\nu_{2m-k}\right)=\gamma_m\left(k,2m-k\right)$, where $\gamma_m(k,k')=\begin{cases}1 \text{ if } k\in\mathbb{P} \text{ and } k'\in\mathbb{P}\\ 0 \text{ otherwise}\end{cases}.$

Denote $Y_{mk}=\gamma_m(\nu_k,\nu_{2m-k})-E\left\{\gamma_m(\nu_k,\nu_{2m-k})\right\}=\gamma_m(\nu_k,\nu_{2m-k})-\beta(m,n)$,

where $E\left\{\gamma_m(\nu_k,\nu_{2m-k})\right\}=\beta(m,n)$, and $Y_m=\sum_{k=3}^{2m-3}Y_m$

Then, for $X_m = \frac{Y_m - E\left\{Y_m\right\}}{\sqrt{Var\left\{Y_m\right\}}} = \frac{G(2m,\vec{v}_m) - E\left\{G(2m,\vec{v}_m)\right\}}{\sqrt{Var\left\{G(2m,\vec{v}_m)\right\}}}$ we have

$$P\left\{G(2m,\vec{v}_m) < 1\right\} = P\left\{X_m < x_{cr}(m)\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_{cr}(m)} e^{-\frac{1}{2}t^2}\, dt, \quad \text{where } x_{cr}(m) = \frac{1 - E\left\{G(2m,\vec{v}_m)\right\}}{\sqrt{Var\left\{G(2m,\vec{v}_m)\right\}}} \to -\infty$$

and $\lim_{m\to\infty} P\left\{G(2m,\vec{v}_m) < 1\right\} = 0$ or, equivalently, $\lim_{m\to\infty} P\left\{\left|G_m\mathbb{P}\right| \geq 1\right\} = \lim_{m\to\infty} P\left\{G(2m,\vec{v}_m) \geq 1\right\} = 1$

**Proof.**

We have:

$$Var\left\{Y_{mk}\right\} = Var\left\{\gamma_m(\nu_k,\nu_{2m-k}\right\} = \beta(m,k)\cdot\left(1-\beta(m,k)\right) \geq \beta(m,m)\cdot\left(1-\beta(m,n)\right) \geq \beta(m,m)\cdot\left(1-\beta(m,3)\right).$$

since $\beta(m,k) \geq \beta(m,m)$ for all $k$: $3 \leq k \leq 2m-3$, due to Lemma 4.2.

For $Y_m = \sum_{k=3}^{2m-3} Y_{mk}$ , we write $Var\left\{Y_m\right\} = Var\left\{G(2m,\vec{v}_m)\right\} = \sum_{k=3}^{2m-k} Var\left\{Y_{mk}\right\}$.

Then, $E\left\{\left|Y_{mk}\right|\right\} = E\left\{\left|\gamma_{mk} - \beta(m,k)\right|\right\} \leq \left(1+\beta(m,n)\right) \leq 2$

$$\beta_{3,mk} = E\left\{\left|Y_{mk}\right|^3\right\} = E\left\{\left|\gamma_{mk} - \beta(m,k)\right|^3\right\} = p_{mk}\cdot q_{mk}\cdot(p_{mk}^2 + q_{mk}^2) \leq p_{mk}\cdot q_{mk} = \sigma_{mk}^2,$$

where $p_{mk} = E\left\{\gamma_{mk}\right\} = \beta(m,k),\ q_{mk} = 1 - p_{mk}, \sigma_{mk}^2 = Var\left\{\gamma_{mk}\right\} = p_{mk}\cdot q_{mk}$

Then, $E\left\{Y_m\right\} = 0$, $Var\left\{Y_m\right\} = \sum_{k=3}^{2m-3} Var\left\{Y_{mk}\right\} = \sum_{k=3}^{2m-3} \sigma_{mk}^2 = \sum_{k=3}^{2m-3}\left[\beta(m,k)\cdot\left(1-\beta(k,m)\right)\right] = \sigma_m^2.$

Due to (4.7), we have $\beta(m,m) = \frac{C'}{\ln^2(m)}\left[1 + \frac{C'}{\ln(m)}\right]^2$.

This implies: $\sigma_m^2 = Var\left\{Y_m\right\} = \sum_{k=3}^{2m-3} \sigma_{mk}^2 \geq \left(2m-5\right)\cdot\beta(m,m)\cdot\left(1 - \frac{1}{\ln^2 3}\right) \to \infty$ as $m \to \infty$

so that $\sigma_m^2 = Var\left\{Y_m\right\} \to \infty$ as $m \to \infty$.

All terms $Y_{mk} = \gamma_{mk} - p_{mk}$ in the sum $Y_m = \sum_{k=3}^{2m-3} Y_{mk}$ are uniformly bounded ( $\left|Y_{mk}\right| \leq 1$ for all $m$)

and centered, because $Y_{mk} = \gamma_{mk} - E\left\{\gamma_{mk}\right\} = \gamma_{mk} - p_{mk}$, so that $E\left\{Y_m\right\} = 0$.

We have also $\beta_{3,mk} = E\left\{\left|Y_{mk}\right|^3\right\} \leq \sigma_{mk}^2$. Since $\sigma_m^2 = Var\left\{Y_m\right\} \to \infty$ as $m \to \infty$,

this implies the sufficient Liapunov condition

$$\frac{1}{\sigma_m^3}\sum_{k=3}^{2m-3} E\left\{\left|Y_{mk}\right|^3\right\} = \frac{1}{\sigma_m^3}\sum_{k=3}^{2m-3}\beta_{3,mk} \le \frac{1}{\sigma_m^3}\sum_{k=3}^{2m-3}\sigma_{mk}^2 = \frac{\sigma_m^2}{\sigma_m^3} = \frac{1}{\sigma_m} \to 0 \text{ as } m \to \infty$$

for the *Central Limit Theorem* (called in this situation the *Bounded Liapunov Theorem* [ 18], [23])

for the sequence of normed and centered variables $X_m = \frac{Y_m}{\sigma_m}$, such that $E\{X_m\} = 0, Var\{X_m\} = 1.$

This guarantees the uniform convergence of probability distribution function $F_{X_m}(x)$ of $X_m$

to the standard normal probability distribution $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\,dt$.

Recall that $X_m = \frac{Y_m - E\{Y_m\}}{\sqrt{Var\{Y_m\}}} = \frac{G(2m,\vec{v}_m) - E\{G(2m,\vec{v}_m)\}}{\sqrt{Var\{G(2m,\vec{v}_m)\}}}$

where $E\{G(2m,\vec{v}_m)\} = \sum_{k=3}^{m-3} E\{\gamma_m(v_k, v_{2m-k})\}$ , $G(2m,\vec{v}_m) = \sum_{i=3}^{2m-3}\gamma_m(v_k, v_{2m-k})$.

Then, we have $P\{G(2m,\vec{v}_m) < 1\} = P\{X_m < x_{cr}(m)\} \approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_{cr}(m)} e^{-\frac{1}{2}t^2}\,dt,$

where $x_{cr}(m) = \frac{1 - E\{G(2m,\vec{v}_m)\}}{\sqrt{Var\{G(2m,\vec{v}_m)\}}} = \frac{1 - \sum_{k=3}^{3m-3} p_{mk}}{\sigma_m}$.

Since $\frac{1}{\sigma_m}\sum_{k=3}^{2m-3} p_{mk} \ge \frac{1}{\sigma_m}\sum_{k=3}^{2m-3} p_{mk}q_{mk} = \frac{\sigma_m^2}{\sigma_m} = \sigma_m \to \infty$ as $m \to \infty$, we have $x_{cr}(m) \to -\infty$ as $m \to \infty$.

This implies $\lim_{m\to\infty} P\{G(2m,\vec{v}_m) < 1\} = 0$, which means that $\lim_{m\to\infty} P\{\left|G_m\mathbb{P}\right| \ge 1\} = \lim_{m\to\infty} P\{G(2m,\vec{v}_m) \ge 1\} = 1.$

**Q.E.D.**

The values of $P\{\left|G_m\mathbb{P}\right| < 1\}$ and $x_{cr}(m)$ for $m = 10^3, 10^4, \ldots, 10^8$ are given in the following table.

**Table 4.2**

| $m$ | $10^3$ | $10^4$ | $10^5$ | $10^6$ | $10^7$ | $10^8$ |
|---|---|---|---|---|---|---|
| $x_{cr}(m)$ | -6.866973 | -16.130926 | -40.343498 | -105.469447 | -284.348502 | -783.836910 |
| $P\{\left|G(2m)\right| < 1\}$ | $3.278916\times10^{-12}$ | $7.734173\times10^{-59}$ | 0.0000000 | 0.0000000 | 0.0000000 | 0.0000000 |

## REFERENCES


1. Yuang Wang, *The Goldbach Conjecture.* Second Edition. World Scientific. Singapore, London, 2002.
2. Cramér Harald, *On the order of magnitude of the difference between consecutive prime numbers,* Acta Arithmetica, 2: 23-46.1936.
3. Granville, A., *Harald Cramér and the distribution of prime numbers*, Scandinavian Actuarial Journal, 1:12-28. 1995]
4. Mark Kac, *Statistical Independence in Probability, Analysis and Number Theory*, The Mathematical Association of America, John Wiley and Sons, Inc.,1972.
5. Tom M. Apostol, *Introduction to Analytic Number Theory*, Springer, 2000.
6. Song Y. Yan, *Number Theory for Computing*, Springer, Springer-Verlag, 2000.
7. A.N. Shiryaev, *Probability*, $2^{nd}$ edition, Springer, 1996.
8. V.S. Varadarajan, *Euler Trough Time: A new Look at Old Themes*, AMS, 2006
9. Gérald Tenenbaum and Michael Mendès, *The Prime Numbers and Their Distribution,* AMS, 2000.
10. А. Г. Постников, *Введение в аналитическую теорию чисел*, Москва, Наука, 1971.
11. Gwo Dong Lin and Chin-Yuan Hu, *The Riemann zeta distribution*, Bernoulli 7(5), 2001, pp. 817-828.
12. Philippe Biane, Jim Pitman, and Mark Yor, *Probability Laws related to the Jacobi Theta and Riemann Zeta Functions, and Brownian Excursions*, Bulletin o the American Mathematical Society. Vol. 38, Number 4, p. 435-465, 2001.
13. Solomon W. Golomb, *A Class of Probability Distributions on the Integers*. Journal of Number Theory 2, 1970, p. 189-192.
14. Andrew Granville, *Primes in Intervals of Bounded Length*, Bulletin of the American Mathematical Society, Vol.52, Number 2, April 2015, p.171-222.
15. Bryan Kra, *Poincaré Reccurence and Number Theory: Thisty Years Later*. Bulletin of the American Mathematical Society, Vol.48, Number 4 October 2011, p. 497-501.
16. Harry Fusrstenberg, *Poincaré Reccurence and Number Theory*, Bulletin of the American Mathematical Society, Volume 5, Number 3, November 1981, p. 211-234.
17. V.K. Balakrishnan, Introductory Discrete Mathematics, Prentice-Hall International, Inc.,1991.

18. Michel Loèv, Probability Theory, Third Edition. Dover Publications, Inc., Mineola, New York, 2017.
19. J. Aczél, *Lectures on Functional Equations and Their Applications*, Vol.14, Academic Press, New York and London, 1966.
20. Julia B. *Statistical Theory of Numbers*. In: Luck JM., Moussa P., Waldse Schmidt M. (eds), Number Theory in Physics, Springer Proceedings in Physics, Berlin, Heidelberg. 1990
21. Pete L. Clark, *Probabilistic Ideas and Methods in Analytic Number Theory (Part I),* From a 11/5/08 talk given in the UGA number theory seminar.
22. S. Bernstein, *Sur l'extension du theorem limite du calculi des probabilités aux summes du quantités dependents. Math. Annalen*, 97 (1927), 1– 59.
23. Cramér, H., Random variables and probability distributions, Stockholm, 1937.
24. G.M. Sobko, *The First Diffusion Problem on Differentiable Manifolds*, Theory of Probability and its Applications, v. XVII, n.3, 1972, pp. 521 - 528
25. G. M. Sobko, *A Diffusion Approximation of non-Markov Random Walks on Differentiable Manifolds*, Theory of Probability and its Applications, v. XVIII, n.1, 1973, pp. 41 - 53
26. Hardy, G. H. and Littlewood, J. E. , *Some Problems of 'Partitio Numerorum.' III. On the Expression of a Number as a Sum of Primes. Acta Math.* **44**, 1-70, 1923.

**Key Words**
Prime numbers, Goldbach conjecture,
Goldbach function, Zeta probability distribution,
random walks, Cramér's model, Bernoulli process,
independent variables, semigroups generated by primes,
multiplicative and additive recurrent models.

**2010 Mathematics Subject Classification**
Primary 11Nxx, 11N05, 11A41,11A07, 11A25, 11A51, 11N60,
11N37, 11P32, 11Mxx, 11M06, 11Y05;
Secondary: 11N36, 11Y16, 60-xx, 60Fxx, 60J65, 60G50, 60Bxx, 60B15


## ABOUT THE AUTHOR

Gregory M. Sobko is a retired professor of mathematics at UC San Diego Extension and National University, San Diego. He obtained his MS degree in 1967 and PhD in Mathematics in 1973 from Moscow State University, Russia. His works are in Probability Theory (Limit Theorems for random walks on Lie groups and differentiable manifolds) and applications (Random Fields, Stochastic Differential Equations) and Number Theory.

Email: gbeholder@gmail.com (Gregory Sobko)