Discrete uniform limit law for additive functions on shifted primes

Gediminas Stepanauskas, Laura Žvinytė

Faculty of Mathematics and Informatics, Vilnius University,
Naugarduko str. 24, LT-03225 Vilnius, Lithuania
gediminas.stepanauskas@mif.vu.lt; laura.zvinyte@mif.vu.lt

Received: September 8, 2015 / Revised: January 15, 2016 / Published online: March 7, 2016

Abstract. The sufficient and necessary conditions for a weak convergence of distributions of a set of strongly additive functions \( f_x, x \geq 2 \), the arguments of which run through shifted primes, to the discrete uniform law are obtained. The case when \( f_x(p) \in \{0, 1\} \) for every prime \( p \) is considered.

Keywords: additive function, discrete uniform law, frequency, weak convergence.

1 Introduction

Let \( f_x, x \geq 2 \), be a set of strongly additive functions such that \( f_x(p) \in \{0, 1\} \) for all primes \( p \) and all \( x \geq 2 \). It follows from the strong additivity that for every positive integer \( n \)

\[
f_x(n) = \sum_{p|n} f_x(p) = \sum_{p|n, f_x(p) = 1} 1.
\]

The problem of a weak convergence of distributions

\[
\nu_x(n \leq x, f_x(n) < u) := \frac{1}{[x]} \sum_{n \leq x, f_x(n) < u} 1 \quad \text{as } x \to \infty
\]

is of key importance in probabilistic number theory. There are interesting general conditions of convergence, classes of possible limit distributions, conditions of convergence to particular distributions. A detailed account of particular and general results you can find in the monographs [1, 2, 5].

© Vilnius University, 2016
In the articles [7, 8, 9], the case of the Poisson distribution as a limit law was considered. It was shown there that the Poisson law can occur as a limit one for the distributions:

\[
\nu_x(p \leq x, f_x(p + 1) < u),
\nu_x(n \leq x, f_x(n) + g_x(n + 1) < u),
\nu_x(p \leq x, f_x(p + 1) + g_x(p + 2) < u).
\]

The Bernoulli, geometrical, binomial, discrete uniform distributions as limit ones for \( \nu_x(f_x(n) < u) \) were investigated in [10, 11, 12]. Several results (general enough) can be found in [1, 4, 13, 14].

In this work, we consider the weak convergence of distribution functions \( \nu_x(f_x(p + 1) < u) = \nu_x(p \leq x, f_x(p + 1) < u) \)

\[
:= \frac{1}{\pi(x)} \# \{ p \leq x: f_x(p + 1) < u, p \text{ prime} \}
\] (1)

to the discrete uniform law

\[
U(u, L) := \sum_{k=0,1,...,L-1 \atop k<u} \frac{1}{L},
\] (2)

where the parameter \( L \in \mathbb{N}, L \geq 2 \). Similarly as in [9], we use in the proofs the method of factorial moments and we have to restrict the behaviour of additive functions on large primes (see condition (H)). But the authors think that this condition is not necessary for the weak convergence to the discrete uniform distribution. Maybe, the problem could be solved applying the Kubilius model of probability spaces [1, 2, 5]. But the large primes have to be overcome there as well.

Throughout the paper, we keep the following notation. The values of \( p, p_1, p_2, \ldots \) mean prime numbers; \( c \) is an absolute positive constant not always the same. By the symbol \( \varepsilon(x) \) we denote values vanishing as \( x \to \infty \). The notation \( a \ll b \) is equivalent to the inequality \( |a| \leq cb \). If the constant \( c \) the constant included in \( \ll \), or the vanishing function \( \varepsilon(x) \) depend on a parameter \( a \), we write \( c_a, \ll_a, \varepsilon_a(x) \). The notation \( F_x(u) \Rightarrow F(u) \) means that the distribution functions \( F_x(u) \) converge weakly to the distribution function \( F(u) \) as \( x \to \infty \). The superscript " at the signs of sum or maximum, \( \sum^*, \max^* \), means that the summation or maximum is expanded over the primes \( p \) for which \( f_x(p) = 1 \). The other notation is generally accepted or is later discussed in the text.

2 Main result and auxiliary lemmas

**Theorem 1.** Let \( f_x, x \geq 2, \) be a set of strongly additive functions. Assume that \( f_x(p) \in \{0, 1\} \) for all prime numbers \( p \) and

\[
\lim_{x \to \infty} \log x \sum_{x^* < p \leq x} \frac{1}{p} = 0 \quad (H)
\]
for all $\gamma \in (0, 1)$. The distributions $\nu_x(f_x(p + 1) < u)$ converge weakly to the limit discrete uniform law $\mathcal{L}(u, L)$ as $x \to \infty$ if and only if $L = 2$ and

$$f_x(3) = 1, \quad \lim_{x \to \infty} \sum_{3 < p \leq x}^* \frac{1}{p} = 0. \quad (3)$$

The proof of this main theorem is based on the following three lemmas on the limit behaviour of factorial moments of the distribution $\nu_x(f_x(p + 1) < u)$.

**Lemma 1.** Let $f_x, x \geq 2$, be a set of strongly additive functions such that $f_x(p) \in \{0, 1\}$ for all primes $p$. If distributions (1) converge weakly to some distribution function $F(u)$ with a jump at the point $u = 0$ as $x \to \infty$, then the quantities

$$\beta(l, x) := \frac{1}{\pi(x)} \sum_{p \leq x} f_x(p + 1)(f_x(p + 1) - 1) \cdots (f_x(p + 1) - l + 1),$$

$l = 1, 2, \ldots$, have finite limits

$$\lim_{x \to \infty} \beta(l, x) = g_l, \quad (4)$$

where $g_l$ is the $l$th factorial moment of the limit law.

**Lemma 2.** (See [9, Lemma 2].) If a set of strongly additive functions $f_x$ satisfies the conditions of Theorem 1 and

$$\sum_{p \leq x}^* \frac{1}{p} \ll 1, \quad (5)$$

then

$$\beta(l, x) = \sum_{p_1, p_2, \ldots, p_l \leq x \atop p_i \neq p_j, i \neq j} \frac{1}{(p_1 - 1)(p_2 - 1) \cdots (p_l - 1)} + \varepsilon_l(x),$$

$l = 1, 2, \ldots$.

According to this statement and equality (4) in the case of convergence of $\nu_x(f_x(p + 1) < u)$, we have that

$$\lim_{x \to \infty} \sum_{p_1, p_2, \ldots, p_l \leq x \atop p_i \neq p_j, i \neq j} \frac{1}{(p_1 - 1)(p_2 - 1) \cdots (p_l - 1)} = g_l \quad (6)$$

for each $l \in \{1, 2, \ldots\}$.

**Lemma 3.** Let $f_x, x \geq 2$, be a set of strongly additive functions such that $f_x(p) \in \{0, 1\}$ for all primes $p$ and condition (H) hold. If distributions (1) converge weakly to the distribution $F_{\xi}$ of the random variable $\xi$ with a finite support $\{0, 1, \ldots, L - 1\}$, $L \geq 2$, then there exists some constant $D \geq 2$ such that

$$\limsup_{x \to \infty} \#\{p \leq D: f_x(p) = 1\} \leq L - 1, \quad (7)$$

\[
\lim_{x \to \infty} \sum_{D < p \leq x^{1/L}}^{\ast} \frac{1}{p} = 0, \quad (8)
\]

\[
\lim_{x \to \infty} \sum_{D < p \leq x^{1/L}}^{\ast} \frac{1}{(p_1 - 1)(p_2 - 1) \cdots (p_l - 1)} = g_l, \quad (9)
\]

\[l = 1, 2, \ldots, L - 1.\]

Moreover, the characteristic function of the limit distribution \(F_\xi\) is equal to

\[
1 + \sum_{l=1}^{L-1} \frac{g_l}{l!} (e^{it} - 1)^l.
\]

From Theorem 1 we get the following example.

**Example 1.** Let

\[
f_x(p) = \begin{cases} 
1 & \text{if } p = 3, \\
1 & \text{if } x^\alpha \leq p < x^\alpha + \varepsilon_x,
0 & \text{otherwise},
\end{cases}
\]

where \(\varepsilon_x > 0\) and \(\varepsilon_x \log x \to 0\) as \(x \to \infty\). Then

\[\nu_x(f_x(p) + 1) < U(u, 2) \Rightarrow U(u, 2) \text{ as } x \to \infty.\]

### 3 Proofs of lemmas

**Proof of Lemma 1.** Suppose that distribution functions (1) converge weakly to the limit distribution \(F(u)\) with the jump at the point \(u = 0\). From the weak convergence we have that

\[\lim_{x \to \infty} \nu_x(f_x(p + 1) = 0) = F(0^+) - F(0) \geq c.\]

Using this estimate, it is proved in [9] (see inequality (10)) that

\[\beta(l, x) \ll_1 l, \quad l \geq 1.\]

(10)

According to this,

\[
\frac{1}{\pi(x)} \sum_{p \leq x}^{P_x(p+1)=k} 1 = \frac{1}{\pi(x)} \sum_{p \leq x}^{P_x(p+1)=k} \frac{k(k-1) \cdots (k-l-1)}{k(k-1) \cdots (k-l-1)} \leq \frac{1}{k(k-1) \cdots (k-l-1)} \beta(l+2, x) \ll_1 \frac{1}{k(k-1) \cdots (k-l-1)}
\]

http://www.mii.lt/NA
for every $k \geq l + 2$. Therefore,

$$F(k+1) - F(k) = \lim_{x \to \infty} \nu_x(f_x(p+1) = k) \ll \frac{1}{k(k-1) \cdots (k-l-1)}$$

for every $k \geq l + 2$.

Let us fix $l \in \mathbb{N}$ and choose $K > l + 2$. Using estimate (10), analogously as in [9], we get

$$\beta(l, x) = \frac{1}{\pi(x)} \sum_{p \leq x, 1 \leq f_x(p+1) \leq K} f_x(p+1)(f_x(p+1) - 1) \cdots (f_x(p+1) - l + 1)$$

$$+ \frac{1}{\pi(x)} \sum_{p \leq x, f_x(p+1) > K} f_x(p+1)(f_x(p+1) - 1) \cdots (f_x(p+1) - l + 1) \frac{f_x(p+1) - l}{f_x(p+1) - l}$$

$$= \sum_{k=l}^{K} k(k-1) \cdots (k-l + 1) \frac{1}{\pi(x)} \sum_{f_x(p+1) = k} 1 + O\left(\frac{\beta(l+1, x)}{K-l}\right)$$

$$= \sum_{k=l}^{K} k(k-1) \cdots (k-l + 1)(F(k+1) - F(k)) + \varepsilon_{K,l}(x) + O_l\left(\frac{1}{K-l}\right)$$

$$= g_l + \varepsilon_{K,l}(x) + O_l\left(\sum_{k=K+1}^{\infty} \frac{1}{(k-l)(k-l-1)}\right) + O_l\left(\frac{1}{K}\right)$$

$$= g_l + \varepsilon_{K,l}(x) + O_l\left(\frac{1}{K}\right).$$

Taking the limit in the last equality as $x$ tends to infinity and then as $K$ tends to infinity, we obtain relation (4). Lemma 1 is proved.

**Proof of Lemma 3.** The proof of the lemma almost coincides with the proof of the necessity part of Corollary 4 from [6]. From the conditions of the lemma we have that there is some $k \in \{0, 1, \ldots, L-1\}$ such that

$$\lim_{x \to \infty} \nu_x(f_x(p+1) = k) = F_\xi(k+1) - F_\xi(k) \geq c.$$

Thus, from the inequality (see [3])

$$\nu_x(f_x(p+1) = k) \ll \left(4 + \sum_{p \leq x} \frac{1}{p}\right)^{-1/2}$$

(11)

it follows inequality (5).
Now according to Lemma 2, equality (6) holds. Let \( d \geq 2 \) be a temporarily fixed positive integer. If \( x/d^{L-1} > d \), we have from equality (6) that
\[
\frac{1}{(p_1 - 1)(p_2 - 1) \cdots (p_L - 1)}.
\]
Assuming that
\[
\limsup_{x \to \infty} \# \{ p \leq d : f_x(p) \neq 0 \} \geq L,
\]
we get a contradiction to the condition \( g_L = 0 \). Thus,
\[
\limsup_{x \to \infty} \# \{ p \leq d : f_x(p) \neq 0 \} \leq L - 1
\]
for each fixed positive integer \( d \).

Put
\[
a_d = \limsup_{x \to \infty} \# \{ p \leq d : f_x(p) \neq 0 \}.
\]
The sequence \( a_d \) (\( d \geq 2 \)) is integer-valued, non-decreasing, and bounded. Thus, there exists a positive integer \( D \geq 2 \) such that
\[
\limsup_{x \to \infty} \# \{ p \leq D : f_x(p) \neq 0 \} = a_d
\]
for \( d \geq D \). Since \( a_d \leq L - 1 \) for all positive integers \( d \), we obtain condition (7).

On the other hand, from the reasoning above it follows that
\[
\lim_{x \to \infty} f_x(p) = 0
\]
for each fixed prime \( p > D \). Consequently, we obtain that
\[
\lim_{x \to \infty} \max^* \frac{1}{p} = 0.
\]
Since for every pair \( i, j, 1 \leq i < j \leq L \),
\[
\sum_{D < p_1 \ldots p_L \leq x^{1/L} \atop p_i = p_j} 1
\]
\[
\leq \max^* \frac{1}{p - 1} \left( \sum_{p \leq x} \frac{1}{p - 1} \right)^{L-1} \to 0 \quad (x \to \infty),
\]
equality (6) shows that
\[
g_L \geq \limsup_{x \to \infty} \left( \sum_{D < p \leq x^{1/L}} \frac{1}{p - 1} \right)^L.
\]
Thus, the condition \( g_L = 0 \) implies (8). The last condition (9) of the lemma now follows from (6), equality (8), and condition (H). Lemma 3 is proved. \( \Box \)
4 Proof of Theorem 1

Necessity. Suppose that

$$\nu_x(f_x(p+1) < u) \Rightarrow U(u,L) \quad \text{as} \quad x \to \infty \quad (12)$$

with parameter $L \geq 2$. From Lemma 1 we obtain that

$$\lim_{x \to \infty} \beta(x,l) = \frac{(L-1)!}{(L-1-l)!(l+1)}$$

for $l = 1, 2, \ldots, L-1$. The values of $g_l$ are the factorial moments of the limit distribution. In the case of the uniform distribution, we get that

$g_1 = \frac{L-1}{2}$, \quad $g_2 = \frac{(L-1)(L-2)}{3}$,

$g_k = \frac{(L-1)(L-2)\cdots(L-k)}{k+1}$, \quad $k = 3, 4, \ldots, L-1$,

$g_k = 0$, \quad $k = L, L+1, \ldots$.

From (12) we have that

$$\lim_{x \to \infty} \nu_x(f_x(p+1) = 0) = U(0+,L) - U(0,L) = \frac{1}{L} > 0.$$  

Thus, inequality (5) follows from inequality (11) with $k = 0$. We apply now Lemma 2.

The values of $g_l$ are the factorial moments of the limit distribution. It is clear that

$$g_l \leq g_{l-k}g_k$$

for all $l = 2, 3, \ldots$ and all $k = 1, 2, \ldots, l-1$. In the particular case,

$$g_2 \leq g_1^2.$$

Therefore,

$$\frac{(L-1)(L-2)}{3} \leq \left( \frac{L-1}{2} \right)^2,$$

which implies $L \leq 5$. Further, we examine separately the cases $L = 2, 3, 4, 5$.

Let $L = 2$. In this case,

$$g_1 = \frac{1}{2}, \quad g_2 = g_3 = \cdots = 0.$$

Using Lemma 3, we have that, for some $D \geq 2$,

$$\limsup_{x \to \infty} \# \{p \leq D: f_x(p) = 1\} = \kappa \leq 1,$$

\[
\lim_{x \to \infty} \sum_{D < p \leq x^{1/2}}^* \frac{1}{p^s} = 0, \\
\lim_{x \to \infty} \sum_{p \leq D}^* \frac{1}{p-1} = \frac{1}{2}.
\] (13)

If \( \kappa = 1 \), then we obtain from relations above and condition (H) that there is only one case

\[ f_x(3) = 1 \]

for large \( x \) and

\[ \lim_{x \to \infty} \sum_{3 < p \leq x}^* \frac{1}{p} = 0. \]

If \( \kappa = 0 \), then \( f_x(p) = 0 \) for every fixed \( p \) and sufficiently large \( x \). In this case, equality (13) cannot be satisfied. It follows that the case \( \kappa = 0 \) cannot occur.

Let \( L = 3 \). Then

\[ g_1 = 1, \quad g_2 = \frac{2}{3}, \quad g_3 = g_4 = \cdots = 0. \]

According to Lemma 3, we have that, for some fixed \( D \geq 2 \), the following conditions hold:

\[
\limsup_{x \to \infty} \# \{ p \leq D: f_x(p) = 1 \} = \kappa \leq 2, \\
\lim_{x \to \infty} \sum_{D < p \leq x^{1/3}}^* \frac{1}{p} = 0, \\
\lim_{x \to \infty} \sum_{p \leq D}^* \frac{1}{p-1} = 1, \\
\lim_{x \to \infty} \sum_{p_1, p_2 \leq D}^* \frac{1}{(p_1-1)(p_2-1)} = \frac{2}{3}. \] (14)

(15)

First, we suppose that \( \kappa = 2 \). From (14) and condition (H) we have that

\[ f_x(p_1) = f_x(p_2) = 1 \]

for large \( x \) and

\[ \frac{1}{p_1 - 1} + \frac{1}{p_2 - 1} = 1 \]

for some fixed primes \( p_1 < p_2 \). Since the last equality is impossible for any pair of different primes \( p_1, p_2 \), then the case \( \kappa = 2 \) cannot occur.

From equality (15) it follows that the cases \( \kappa = 1, \kappa = 0 \) cannot occur as well.
Let $L = 4$. Then
\[ g_1 = \frac{3}{2}, \quad g_2 = 2, \quad g_3 = \frac{3}{2}, \quad g_4 = g_5 = \cdots = 0. \]

According to Lemma 3, there exists some constant $D \geq 2$ for which
\[
\limsup_{x \to \infty} \# \{ p \leq D : f_x(p) = 1 \} = \kappa \leq 3, \tag{16}
\]
\[
\lim_{x \to \infty} \sum_{1 \leq p \leq x^{1/4}}^* \frac{1}{p} = 0,
\]
\[
\lim_{x \to \infty} \sum_{p \leq D}^* \frac{1}{p - 1} = \frac{3}{2}, \tag{17}
\]
\[
\lim_{x \to \infty} \sum_{p_1, p_2 \leq D \atop p_1 \neq p_2}^* \frac{1}{(p_1 - 1)(p_2 - 1)} = 2,
\]
\[
\lim_{x \to \infty} \sum_{p_1, p_2, p_3 \leq D \atop p_1 \neq p_j, i \neq j}^* \frac{1}{(p_1 - 1)(p_2 - 1)(p_3 - 1)} = \frac{3}{2}. \tag{18}
\]

It follows from (18) that $\kappa$ cannot be 0, 1, and 2.

Suppose $\kappa = 3$. Then equalities (16) and (17) imply that there exist fixed primes $p_1 < p_2 < p_3$ for which $f_x(p_1) = f_x(p_2) = f_x(p_3) = 1$ for large $x$ and
\[
\frac{1}{p_1 - 1} + \frac{1}{p_2 - 1} + \frac{1}{p_3 - 1} = \frac{3}{2}. \tag{19}
\]
But there are no primes satisfying equality (19). So, the case $\kappa = 3$ is impossible as well.

Let $L = 5$. Then
\[ g_1 = 2, \quad g_2 = 4, \quad g_3 = 6, \quad g_4 = \frac{24}{5}, \quad g_5 = g_6 = \cdots = 0. \]

According to Lemma 3, there exists some $D \geq 2$ for which
\[
\limsup_{x \to \infty} \# \{ p \leq D : f_x(p) = 1 \} = \kappa \leq 4, \tag{20}
\]
\[
\lim_{x \to \infty} \sum_{1 \leq p \leq x^{1/5}}^* \frac{1}{p} = 0,
\]
\[
\lim_{x \to \infty} \sum_{p \leq D}^* \frac{1}{p - 1} = 2,
\]
\[
\lim_{x \to \infty} \sum_{p_1, p_2 \leq D \atop p_1 \neq p_j}^* \frac{1}{(p_1 - 1)(p_2 - 1)} = 4,
\]
$$\lim_{x \to \infty} \sum_{\substack{p_1, p_2, p_3 \leq D \\ p_i \neq p_j, i \neq j}}^{*} \frac{1}{(p_1 - 1)(p_2 - 1)(p_3 - 1)} = 6,$$

$$\lim_{x \to \infty} \sum_{\substack{p_1, p_2, p_3, p_4 \leq D \\ p_i \neq p_j, i \neq j}}^{*} \frac{1}{(p_1 - 1)(p_2 - 1)(p_3 - 1)(p_4 - 1)} = \frac{24}{5}. \quad (21)$$

It follows from equality (21) that $\kappa$ cannot be equal to 0, 1, 2, and 3.

Suppose $\kappa = 4$. Then from (20) we deduce that there exist primes $p_1 < p_2 < p_3 < p_4$ such that $f_x(p_1) = f_x(p_2) = f_x(p_3) = f_x(p_4) = 1$ for large $x$ and

$$\frac{1}{p_1 - 1} + \frac{1}{p_2 - 1} + \frac{1}{p_3 - 1} + \frac{1}{p_4 - 1} = 2.$$

But

$$\frac{1}{p_1 - 1} + \frac{1}{p_2 - 1} + \frac{1}{p_3 - 1} + \frac{1}{p_4 - 1} \leq 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{6} < 2.$$

So, the case $\kappa = 4$ is impossible as well.

**Sufficiency.** Assume that conditions (3) of the theorem together with additional condition (H) are satisfied. Estimate (5) follows from condition (3). Now using Lemma 2, we obtain that

$$\lim_{x \to \infty} \beta(l, x) = \lim_{x \to \infty} \sum_{\substack{p_1, p_2, \ldots, p_l \leq x \\ p_i \neq p_j, i \neq j}}^{*} \frac{1}{(p_1 - 1)(p_2 - 1) \cdots (p_l - 1)} = 0$$

if $l > 1$ and

$$\lim_{x \to \infty} \beta(1, x) = \frac{1}{2}.$$

Put

$$\psi_x(t) = \frac{1}{\pi(x)} \sum_{p \leq x} e^{itf_x(p + 1)}$$

for $x \geq 2$ and $t \in \mathbb{R}$. Since

$$|e^{itr} - 1 - r(e^{it} - 1)| \leq \frac{r(r - 1)}{2} |e^{it} - 1|^2$$

for all $r \in \{0\} \cup \mathbb{N}$, we have

$$\psi_x(t) = 1 + \beta(1, x)(e^{it} - 1) + O(\beta(2, x)).$$

Taking the limit in the last equality, we conclude that

$$\lim_{x \to \infty} \psi_x(t) = \frac{1}{2}(e^{it} + 1).$$

But this is the characteristic function of the uniform distribution $U(1, 2)$.

So, the sufficiency of the theorem follows. Theorem 1 is now proved.
References