Asymptotic Distribution of Residuals within congruence classes generated by primes

Gregory M. Sobko

Abstract
By using the Dirichlet characters for a finite abelian group \( G_p = \mathbb{Z}_p = \mathbb{Z} / (p \cdot \mathbb{Z}) \), \( p \in \mathbb{P} \), and the corresponding characteristic functions, we discuss asymptotic distribution for sums of residuals \( r = \text{mod}(\nu, p) = [\nu]_p \), \( p \in \mathbb{P} \), where \( \mathbb{P} \) is a set of prime numbers, and \( \nu \) is a random variable with a certain probability distribution on set \( \mathbb{N} \) of natural numbers. We prove that for a sequence \( \nu_1, \nu_2, \ldots, \nu_n, \ldots \) of independent random integers (not necessarily equally distributed), the residuals of sums \( [V^{(n)}]_p = \sum_{i=1}^{n} [\nu_i]_p \) are asymptotically uniformly distributed on \( G_p \), for every \( p \in \mathbb{P} \), (congruence classes generated by primes). Then, we prove that components of the vector of residuals \( \check{r}(\nu) = (r_1, r_2, \ldots, r_{\pi(\nu)}) \) are asymptotically independent random variables.

1. Characteristic functions for residuals of sums \( [V^{(n)}]_p = \sum_{i=1}^{n} [\nu_i]_p \).
Notice that the vector function \( r(n) = \text{mod}(n, \check{p}(n)) \) is periodic with a period \( T = \prod_{p \leq n} p \) since \( \text{mod}(T, p) = 0 \) for any \( p \leq n \). Due to the Chinese Remainder Theorem (CRT) \([22, \text{p.101}]\), a solution \( x \) to the system of equations \( \text{mod}(x, p_i) = r_i \) \( (1 \leq i \leq m) \) exists, and if \( x \) is a solution to the system, then \( y = x + T \) is also a solution to the same system. Considering the ring of all integers \( \mathbb{Z} \), we write \( \mathbb{Z}_m = \mathbb{Z} / (m \cdot \mathbb{Z}) \).
Here $\mathbb{Z}_m$ consists of $m$ congruence classes: $\mathbb{Z}_m = \{C_{m,0}, C_{m,1}, \ldots, C_{m,m-1}\}$ modulo $m$, also called residue classes, denoted as $[0]_m, [1]_m, \ldots, [m-1]_m$ with the addition and multiplication rules expressed as

$$[k]_m + [l]_m = [\text{mod}(k + l, m)]_m \quad \text{and} \quad [k]_m \cdot [l]_m = [\text{mod}(k \cdot l, m)]_m,$$

respectively. For any prime number $p \in \mathbb{P}$, set $\mathbb{Z}_p$ of congruence classes modulo $p$ is a finite abelian group $G_p = \mathbb{Z}_p = \mathbb{Z}/(p \cdot \mathbb{Z})$, of order $p$.

Consider a random sequence $\omega = (\eta_1, \eta_2, \ldots, \eta_n)$ where $\eta_i \in G_p$ $(i = 1, 2, \ldots, n)$ such that random variables $\eta_1, \eta_2, \ldots, \eta_n$ are mutually independent and we can always find the minimal solution to $\text{mod}(x, p_i) = r_i$ $(1 \leq i \leq m)$ among all solutions.

For example, given $\tilde{p} = (5, 11, 17, 23, 29)$ and $\tilde{r} = (0, 8, 13, 7, 1)$, the system

$$\text{mod}(x, p_i) = r_i \quad (1 \leq i \leq 5)$$

has the minimal solution $x = 30$. One of other possible solutions, for instance, is $x = 623675$.

We are interested in probability measures on the direct product $\mathcal{G} = \prod_{p \in \mathbb{P}} G_p$ such that each non-trivial probability distribution is supported by a finite number of components in $\mathcal{G}$.

For a random sequence $\omega = (\eta_1, \ldots, \eta_n)$ of mutually independent random variables $\eta_i$ $(i = 1, 2, \ldots, n)$ with distributions $P\{\eta_i = r | r \in G_{p_i}\} = q^{(i)}_r$ on $G_{p_i}$, we have

$$P\{\eta_i \in B \subseteq G_{p_i}\} = \sum_{r \in B} q^{(i)}_r, \quad \sum_{r = 0}^{p_i - 1} q^{(i)}_r = 1 \quad (i = 1, 2, \ldots, n).$$

and

$$P\left\{\omega \in \prod_{i=1}^{n} B_i\right\} = \prod_{i=1}^{n} P\{\eta_i \in B_i\} \quad \text{for any} \ B_i \subset G_{p_i} \quad (4.1)$$
Further, we use the following notation: $B-r = \{ s \in G_p | s + r \in B, r \in G_p \}$ and for every probability distribution $P$ on $G_p$ define the ‘shifted’ measure $\theta_r P(B) = P(B-r)$.

Obviously the shifted measure $\theta_r P$ is a probability measure on subsets of a finite set $G_p : \theta_r P(G_p) = P(G_p - r) = 1$ because $G_p - r = G_p$ for any $r \in G_p$ since $G_p$ is a group.

Due to CRT, there exist one-to one correspondence between finite sequences of residues $(r_1, r_2, \ldots, r_n)$ and positive integers $n = \prod_{i=1}^{k} p_i^{q_i}$ such that

mod(n, p_i) = r_i \ (i = 1, 2, \ldots, k). \quad \text{If} \quad \mod(m, p_i) = s_i \ \text{for some number} \ m, \ \text{then}

\mod(n + m, p_i) = \mod(r_i + s_i, p_i). \quad \text{Consider two independent random integers} \ \nu \ \text{and} \ \mu \ \text{with probability measures} \ P^\nu \ \text{and} \ P^\mu, \ \text{and their residuals} \ [\nu]_p, [\mu]_p \ \text{modulo} \ p, \ \text{respectively.} \ \text{We are interested in probability distribution} \ P^{[\nu + \mu]}_r \ \text{of the sum}

\[ [\nu]_p + [\mu]_p = [\nu + \mu]_p. \]

For any set $B \subseteq G_p$ we have

\[ P\left([\nu + \mu]_p \in B\right) = \sum_{(r+s)\in B} P\left([\nu]_p = r\right) \cdot P\left([\mu]_p = s\right) = \sum_{t \in B} P\left([\nu]_p = t-s\right) \cdot P\left([\mu]_p = s\right) \]

and we denote $P^{[\nu + \mu]}_r (B) = P\left([\nu + \mu]_p \in B\right)$ as

\[ P^{[\nu + \mu]}_r (B) = P^{[\nu]}_r \ast P^{[\mu]}_r (B), \]

so that

\[ P^{\nu + \mu} (B) = P^\nu \ast P^\mu (B) = \sum_{t \in B} P\left([\nu]_p = t-s\right) \cdot P\left([\mu]_p = s\right) \]

The measure $P^{\nu + \mu} (B) = P^\nu \ast P^\mu (B)$ is called a \textit{convolution of measures} $P^\nu$ and $P^\mu$. One of interesting questions is an asymptotic distribution of sums of independent random integers $\nu^{(n)} = \nu_1 + \nu_2 + \cdots + \nu_n$ and their corresponding residuals

\[ [\nu^{(n)}]_p = [\nu_1]_p + [\nu_2]_p + \cdots + [\nu_n]_p \]

which are also sums of independent random variables $[\nu_i]_p \ (i = 1, 2, \ldots, n)$. 

3
The answer to the question about the limit distribution of $\nu^{(n)}$ depends in general on the distributions of the terms $\nu_i$ in the sum. Meanwhile the limit behavior of residuals $\left[ \nu^{(n)} \right]_p$ does not depend (under very simple and natural conditions) on the distribution of each term $\left[ \nu_i \right]_p$. In what follows we use the well-known general facts from Probability Theory regarding characteristic functions of probability distributions and their convolutions.

Let $p^\xi$ be a probability measure defined on all finite subsets of $\mathbb{N}$. This means that for every $n \in \mathbb{N}$ there exists $P^\xi(n) = P\left\{ \xi = n \right\} \geq 0$ such that $\sum_{n \in \mathbb{N}} P^\xi(n) = 1$.

Characteristic function $\Phi^\xi$ is defined by the formula

$$\Phi^\xi(t) = E e^{it\xi} = \sum_{n \in \mathbb{N}} e^{i\xi n} P^\xi(n).$$

For a finite abelian additive group $G_p = \mathbb{Z}_p$ we consider a homomorphism $\chi$ of $G_p$ into multiplicative group $C^*$ of complex numbers $\chi: G_p \to C^*$.

A homomorphism $\chi: G_p \to C^*$ is also called a character.

Since any element $[k]_p \in G_p \ (k = 0,1,\ldots,p-1)$ has order $p$, that is $p \cdot [k]_p = [0]_p$, we have $1 = \chi([0]_p) = \chi(p \cdot [k]_p) = \left( \chi([k]_p) \right)^p$. This means that any character value $\chi([k]_p)$ is a $p$-th root of unity.

We can define $p$ such character values: $\chi_r([k]_p) = e^\frac{2\pi i (r k)}{p} \ (r = 0,1,2,\ldots,p-1)$.

Denote $\chi_{rk} = e^\frac{2\pi i (r k)}{p} \ (r,k = 0,1,2,\ldots,p-1)$. Character $\chi_0([k]_p) = 1$ for all $k = 0,1,\ldots,p-1$, and $\chi_0$ is called a principal character.

Consider a square matrix $\chi = [\chi_{rk}] \ (0 \leq r,k \leq p-1)$ of size $p$. All characters are orthogonal to each other in terms of scalar products of rows of matrix $\chi$.
\[ \langle \chi_r, \chi_s \rangle = \sum_{t=0}^{p-1} \chi_{rt} \cdot \bar{\chi}_{st} = \sum_{t=0}^{p-1} e^{\frac{2\pi i (r-s)t}{p}} \cdot e^{-\frac{2\pi i (r-s)t}{p}} = \sum_{t=0}^{p-1} e^{\frac{2\pi i (r-s)t}{p}} = 1 - e^{\frac{2\pi i (r-s)}{p}} \]

Characteristic function \( \Phi^{[\xi]} \) for residual \([\xi]_p\) is given by the formula

\[ \Phi^{[\xi]} (r) = E e^{i \chi \cdot [\xi]_p} = \sum_{k=0}^{p-1} p^{[\xi]}(k) e^{\frac{2\pi i (r-k)}{p}} = \sum_{k=0}^{p-1} \chi_{rk} \cdot P^{[\xi]}(k) = \left[ \chi \cdot P^{[\xi]} \right](r) \]

Since the matrix \( \chi = [\chi_{rk}]_p \) \((0 \leq r, k \leq p-1)\) is orthogonal, the inverse matrix \( \chi^{-1} \)
exists and the probability distribution \( P^{[\xi]}_p \) can be uniquely recovered as

\[ P^{[\xi]}_p = \chi^{-1} \cdot \Phi^{[\xi]}_p \]
given its characteristic function \( \Phi^{[\xi]}_p \).

There is one-to-one correspondence between finite probability distributions and the corresponding characteristic functions.

**2. Convergence of probability distributions of residuals** mod\( (\nu^{(n)}, p) \) as \( n \rightarrow \infty \)

for sums \( \nu^{(n)} = \sum_{i=1}^{n} \nu_i \) \((n = 1, 2, \ldots)\) to uniform distributions on \( G_p \), for every \( p \in \mathbb{P} \)

A probability distribution \( P^{[\xi]}(k) \) \((k = 1, 2, \ldots, n)\) defined on a finite set \( X = \{x_1, x_2, \ldots, x_n\} \)
can be identified with the \( n \)-dimensional vector \( P^{[\xi]} = (p_1, p_2, \ldots, p_n) \) where

\[ p_k = P^{[\xi]} \{ \xi = k \}, 1 \leq k \leq n \]

If we have a sequence of probability distributions \( P^{[\xi_m]} \) \((m = 1, 2, \ldots)\) such that \( P^{[\xi_m]} \rightarrow P \)
in a sense of vector convergence in \( n \)-dimensional vector space to probability measure \( P \) on \( X \), then we can expect the convergence for the sequences of corresponding characteristic functions: \( \Phi^{[\xi_m]} \rightarrow \Phi \), where \( \Phi \) is a characteristic function of some limit random variable \( \xi_\infty \) on \( X \), and vice versa.
One of the most important properties of characteristic functions is that for any two independent random variables $\xi_1, \xi_2$ we have $\Phi^{\xi_1 + \xi_2} = \Phi^{\xi_1} \cdot \Phi^{\xi_2}$, so that $\Phi^{\sum_{i=1}^n \xi_i} = \prod_{i=1}^n \Phi^{\xi_i}$ for independent $\xi_1, \xi_2, \ldots, \xi_n$.

**Theorem 4.1**

For any random integers $\nu$ its residual $[\nu]_p$ for a prime $p \in \mathbb{P}$ has a characteristic function $\Phi^{[\nu]}$ such that $\Phi^{[\nu]}(0) = 1$ and $\left| \Phi^{[\nu]}(r) \right| < 1$, if $0 < r \leq p-1$.

**Proof.**

If a random integer $\lambda$ is such that $[\lambda]_p$ has a uniform distribution on $G_p$, that is

$$P\left([\lambda]_p = k\right) = \frac{1}{p} \text{ for all } k = 0, 1, \ldots, p-1,$$

then $\Phi^{[\lambda]}(r) = \begin{cases} 1, & \text{if } r = 0 \\ 0, & \text{if } r \neq 0 \end{cases}$

We prove this by the direct calculations:

$$\Phi^{[\lambda]}(r) = \sum_{k=0}^{p-1} \chi_{\lambda k} \cdot P^{[\lambda]}(k) = \sum_{k=0}^{p-1} \chi_{\lambda k} \cdot \frac{1}{p} \langle \chi_r, \chi_0 \rangle = \begin{cases} 1, & \text{if } r = 0 \\ 0, & \text{if } r \neq 0 \end{cases}$$

We have $\Phi^{[\nu]}(r) = \sum_{k=0}^{p-1} \chi_{\nu k} \cdot P^{[\nu]}(k) = \left[ \chi \cdot P^{[\nu]} \right](r)$. This implies $\left| \Phi^{[\nu]}(r) \right| \leq 1$.

We have $\Phi^{[\nu]}(0) = 1$. Assume that there exist $r \neq 0 \mod p$ such that $\Phi^{[\nu]}(r) = 1$.

Then, $\Phi^{[\nu]}(r) = \sum_{k=0}^{p-1} P^{[\nu]}(k) e^{\frac{2\pi i}{p} (r \cdot k)} = 1$ and, equivalently,

$$\sum_{k=0}^{p-1} \left( 1 - \cos \left( \frac{2\pi}{p} (r \cdot k) \right) \right) \cdot P^{[\nu]}(k) = 0.$$ 

Since $1 - \cos(\alpha) \geq 0$ for any $\alpha$, and $P^{[\nu]}(k) > 0$ for all $k$, we have $r \cdot k = 0 \mod p$ for $k = 0, 1, 2, \ldots, p-1$, which is possible only if $r = 0 \mod p$.

Q.E.D.
Now, we can answer the question about convergence of probability distributions of residuals \( \text{mod}(v^{(n)}, p) \) as \( n \to \infty \) for sums \( v^{(n)} = \sum_{i=1}^{n} v_i \) \((n = 1, 2, \ldots)\) of independent random integers by the following statement.

**Theorem 4.2**

Let \( v_1, v_2, \ldots, v_n, \ldots \) be a sequence of independent random integers (not necessarily equally distributed) such that for every prime \( p \in \mathbb{P} \) the residuals \([v_i]_p\) \((i = 1, 2, \ldots)\) have probability distributions \( P^{[v_i]}_p(k) > 0 \) for all \( 0 \leq k \leq p-1 \).

We assume that \( \sup_{1 \leq i \leq n, r \neq 0} \left| \Phi^{[v_i]}_p(r) \right| = M < 1 \) for \( r \neq 0 \). Then, the residuals of sums \([v^{(n)}]_p = \sum_{i=1}^{n} [v_i]_p\) are asymptotically uniformly distributed on \( G_p \), for every \( p \in \mathbb{P} \).

**Proof.**

We need to prove that \( \lim_{n \to \infty} P^{(n)}_p = P^\lambda \), or simply that \( [v^{(n)}]_p = \sum_{i=1}^{n} [v_i]_p \to [\lambda]_p \) (in probability) as \( n \to \infty \), where \([\lambda]_p\) is uniformly distributed on \( G_p \).

We have \( \Phi^{(n)} = \prod_{i=1}^{n} \Phi^{v_i} \) and \( \left| \Phi^{v^{(n)}}_p(r) \right| = \prod_{i=1}^{n} \left| \Phi^{v_i}_p(r) \right| \leq M^r \to 0 \) as \( n \to \infty \), for each \( r \neq 0 \).

This implies that \( \lim_{n \to \infty} \Phi^{[v^{(n)}]}_p(r) = \Phi^{[\lambda]}_p(r) = \begin{cases} 1, & \text{if } r = 0 \\ 0, & \text{if } r \neq 0 \end{cases} \), so that \( [v^{(n)}]_p = \sum_{i=1}^{n} [v_i]_p \to [\lambda]_p \).

Thus, random variables \([v^{(n)}]_p\) are asymptotically uniformly distributed on \( G_p = \mathbb{Z}_p \) as \( n \to \infty \).

**Q.E.D.**

For a random variable \( \nu \in \mathbb{N} \) we are interested in the vector of residuals \( \tilde{r}(\nu) = (r_1, r_2, \ldots, r_{\pi(\nu)}) \), where \( \pi(\nu) \) stands for number of primes \( p \leq \nu \).
Here \( \lfloor \nu \rfloor_{p_i} = r_i = \text{mod}(\nu, p_i) \quad (i = 1, 2, \ldots, \pi(\nu)) \) for all \( p_i \leq \nu \).

The asymptotic independence of residuals \( \lfloor \nu \rfloor_{p_i} = r_i = \text{mod}(\nu, p_i) \quad (i = 1, 2, \ldots, \pi(\nu)) \) is addressed in the following statement.

**Theorem 4.3.**

All components of the vector of residuals \( \tilde{r}(\nu) = (r_1, r_2, \ldots, r_{\pi(\nu)}) \) are asymptotically independent random variables.

**Proof.**

Notice that the vector function \( \text{mod}(n, \tilde{p}(\nu)) = \tilde{r}(\nu) = (r_1, r_2, \ldots, r_{\pi(\nu)}) \), where \( \tilde{p}(\nu) = (p_1, p_2, \ldots, p_{\pi(\nu)}) \), is periodic with a period \( T(\nu) = \prod_{p \leq \nu} p \) since \( \text{mod}(T(\nu), p) = 0 \) for any \( p \leq \nu \). This implies that if \( x \) is a solution to the system of equations \( \text{mod}(x, p_i) = r_i \quad (1 \leq i \leq \pi(\nu)) \), then \( \nu = x + T(\nu) \) is also a solution to the same system. We set \( \nu = k(\nu) \cdot T(\nu) + r \), where \( r = \text{mod}(\nu, T(\nu)) \). Then, \( \text{mod}(\nu, p_i) = \text{mod}(r, p_i) = r_i \) and since the combination of residual values \( \tilde{r}(\nu) = (r_1, r_2, \ldots, r_{\pi(\nu)}) \) occurs \( k(\nu) \) times in \( \nu \) trials, then for the relative frequency

\[
f(\nu, \tilde{r}(\nu)) = \frac{k(\nu)}{\nu}, \quad \text{we have:} \quad \left| \frac{k(\nu)}{\nu} - \prod_{i=\pi(\nu)} \frac{1}{p_i} \right| = \left| \frac{1}{T(\nu)} + \frac{r}{k(\nu) T(\nu)} \right| \rightarrow 0 \quad \text{as} \quad \nu \rightarrow \infty.
\]

**Q.E.D.**
REFERENCES


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

Copyright © 2020 Gregory M. Sobko
email: greg.beholder@gmail.com (Gregory Sobko)