Asymptotic Distribution of Residuals within congruence classes generated by primes

Gregory M. Sobko

Abstract

By using the Dirichlet characters for a finite abelian group $G_p = \mathbb{Z}_p = \mathbb{Z}/\left(p\cdot\mathbb{Z}\right)$, $p\in\mathbb{P}$, and the corresponding characteristic functions, we discuss asymptotic distribution for sums of residuals $r = \mathrm{mod}(\nu, p) = \left[\nu\right]_p$, $p\in\mathbb{P}$, where $\mathbb{P}$ is a set of prime numbers, and ν 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 $\left[\nu^{(n)}\right]_p = \sum_{i=1}^{n}\left[\nu_i\right]_p$ 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 $\vec{r}(\nu) = \left(r_1, r_2, \ldots r_{\pi(\nu)}\right)$ are asymptotically independent random variables.

1. Characteristic functions for residuals of sums $\left[\nu^{(n)}\right]_p = \sum_{i=1}^{n}\left[\nu_i\right]_p$.

Notice that the vector function $r(n) = \mathrm{mod}\left(n, \vec{p}(n)\right)$ is periodic with a period $T = \prod_{p\le n} p$ since $\mathrm{mod}(T, p) = 0$ for any $p \le n$. Due to the Chinese Remainder Theorem (CRT) [22, p.101], a solution x to the system of equations $\mathrm{mod}(x, p_i) = r_i$ $\left(1 \le i \le m\right)$ 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 $\mathbb{Z}_m = \mathbb{Z}/\left(m\cdot\mathbb{Z}\right)$.

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

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

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 $\eta_i \in G_{p_i} (i = 1, 2, \ldots, n)$ such that random variables $\eta_1, \eta_2, \ldots, \eta_n$ are mutually independent and we can always find the minimal solution to $\mathrm{mod}(x, p_i) = r_i \ (1 \le i \le m)$ among all solutions.

For example, given $\vec{p} = (5, 11, 17, 23, 29)$ and $\vec{r} = (0, 8, 13, 7, 1)$, the system $\mathrm{mod}(x, p_i) = r_i \ (1 \le i \le 5)$ 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 $P\{\eta_i = r \mid r \in G_{p_i}\} = q_r^{(i)}$ on G_{p_i}, we have

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

and

$$P\left\{\omega \in \prod_{i=1}^{n} B_i\right\} = \prod_{i=1}^{n} P\{\eta_i \in B_i\} \text{ for any } B_i \subset G_{p_i} \qquad (4.1)$$

Further, we use the following notation: $B - r \equiv \left\{ s \in G_p \middle| s + r \in B, r \in G_p \right\}$ 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 $n = \prod_{i=1}^{k} p_i^{\alpha_i}$ such that $\mathrm{mod}(n, p_i) = r_i \ (i = 1, 2, \ldots, k)$. If $\mathrm{mod}(m, p_i) = s_i$ for some number m, then $\mathrm{mod}(n + m, p_i) = \mathrm{mod}(r_i + s_i, p_i)$. Consider two independent random integers ν and μ with probability measures P^{ν} and P^{μ}, and their residuals $[\nu]_p$, $[\mu]_p$ modulo p, respectively. We are interested in probability distribution $P^{[\nu+\mu]_p}$ of the sum $[\nu]_p + [\mu]_p = [\nu + \mu]_p$. For any set $B \subset G_p$ we have

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

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

$$P^{[\nu+\mu]_p}(B) = P^{[\nu]_p} * P^{[\mu]_p}\left(B\right), \tag{4.2}$$

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

The measure $P^{\nu+\mu}(B) = P^{\nu} * P^{\mu}\left(B\right)$ is called a *convolution* of measures P^{ν} and P^{μ}. 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

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

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

The answer to the question about the limit distribution of $\nu^{(n)}$ depends in general on the distributions of the terms ν_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 $[\nu_i]_p$. In what follows we use the well-known general facts from Probability Theory regarding characteristic functions of probability distributions and their convolutions.

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

Characteristic function Φ^ξ is defined by the formula

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

For a finite abelian additive group $G_p = \mathbb{Z}_p$ we consider a homomorphism χ 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 $1 = \chi\left([0]_p\right) = \chi\left(p\cdot[k]_p\right) = \left(\chi\left([k]_p\right)\right)^p$. This means that any character value $\chi\left([k]_p\right)$ is a p-th root of unity.

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

Denote $\chi_{rk} = e^{\frac{2\pi i}{p}(r\cdot k)}$ $(r,k = 0,1,2,\ldots,p-1)$. Character $\chi_0\left([k]_p\right) = 1$ for all $k = 0,1,\ldots,p-1$, and χ_0 is called a *principal character*.

Consider a square matrix $\chi = \left[\chi_{rk}\right]$ $(0 \leq r,k \leq p-1)$ of size p. All characters are orthogonal to each other in terms of scalar products of rows of matrix χ :

$$\langle \chi_r, \chi_s \rangle = \sum_{t=0}^{p-1} \chi_{rt} \cdot \bar{\chi}_{st} = \sum_{t=0}^{p-1} e^{\frac{2\pi i}{p}(r\cdot t)} \cdot e^{\frac{-2\pi i}{p}(s\cdot t)} = \sum_{t=0}^{p-1} e^{\frac{2\pi i}{p}((r-s)\cdot t)} = \frac{1-e^{2\pi i\cdot(r-s)}}{1-e^{\frac{2\pi i\cdot(r-s)}{p}}} = \begin{cases} p, & \text{if } r = s \\ 0, & \text{if } r \neq s \end{cases}$$

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

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

Since the matrix $\chi = [\chi_{rk}]_p$ $(0 \le r, k \le p-1)$ is orthogonal, the inverse matrix χ^{-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 $\mathrm{mod}(v^{(n)}, p)$ as $n \to \infty$ for sums $v^{(n)} = \sum_{i=1}^{n} v_i$ $(n = 1,2,\ldots)$ to uniform disributions 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 $p_k = P\{\xi = k\}$, $1 \le k \le n$.

If we have a sequence of probability distributions P^{ξ_m} $(m = 1,2,\ldots)$ such that $P^{\xi_m} \to P$ 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: $\Phi^{\xi_m} \to \Phi$, where Φ is a characteristic function of some limit random variable ξ_∞ on X, and vice versa.

One of the most important properties of characteristic functions is that for any two independent random variables ξ_1, ξ_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 ν its residual $[\nu]_p$ for a prime $p \in \mathbb{P}$ has a characteristic function $\Phi^{[\nu]_p}$ such that $\Phi^{[\nu]_p}(0) = 1$ and $\left|\Phi^{[\nu]_p}(r)\right| < 1,$ if $0 < r \le p-1$.

Proof.

If a random integer λ is such that $[\lambda]_p$ has a uniform distribution on G_p , that is $P\left\{[\lambda]_p = k\right\} = \dfrac{1}{p}$ for all $k = 0,1,\ldots,p-1$, then $\Phi^{[\lambda]_p}(r) = \begin{cases} 1, & \text{if } r = 0 \\ 0, & \text{if } r \neq 0 \end{cases}$

We prove this by the direct calculations:

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

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

We have $\Phi^{\nu_i}(0) = 1.$ Assume that there exist $r \neq 0 \bmod p$ such that $\Phi^{\nu_i}(r) = 1.$

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

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

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

Q.E.D.

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

Theorem 4.2

Let $\nu_1, \nu_2, \ldots, \nu_n, \ldots$ be a sequence of independent random integers (not necessarily equally distributed) such that for every prime $p \in \mathbb{P}$ the residuals $[\nu_i]_p \ (i = 1,2,\ldots)$ have probability distributions $P^{[\nu_i]_p}(k) > 0$ for all $0 \le k \le p-1$.

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

Proof.

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

We have $\Phi^{\nu^{(n)}} = \prod_{i=1}^{n} \Phi^{\nu_i}$ and $\left| \Phi^{\nu^{(n)}}(r) \right| = \prod_{i=1}^{n} \left| \Phi^{\nu_i}(r) \right| \le M^n \to 0$ as $n \to \infty$, for each $r \ne 0$.

This implies that $\lim_{n \to \infty} \Phi^{\left[\nu^{(n)}\right]_p}(r) = \Phi^{[\lambda]_p}(r) = \begin{cases} 1, \text{ if } r = 0 \\ 0, \text{if } r \ne 0 \end{cases}$, so that $\left[\nu^{(n)}\right]_p = \sum_{i=1}^{n} [\nu_i]_p \to [\lambda]_p$.

Thus, random variables $\left[\nu^{(n)}\right]_p$ 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 $\vec{r}(\nu) = \left(r_1, r_2, \ldots r_{\pi(\nu)}\right)$, where $\pi(\nu)$ stands for number of primes $p \le \nu$.

Here $[\nu]_{p_i} = r_i = \mathrm{mod}(\nu, p_i)\ (i = 1, 2, \ldots, \pi(\nu))$ for all $p_i \le \nu$.

The asymptotic independence of residuals $[\nu]_{p_i} = r_i = \mathrm{mod}(\nu, p_i)\ (i = 1, 2, \ldots, \pi(\nu))$ is addressed in the following statement.

Theorem 4.3.

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

Proof.

Notice that the vector function $\mathrm{mod}(n, \vec{p}(\nu)) = \vec{r}(\nu) = (r_1, r_2, \ldots r_{\pi(\nu)})$,

where $\vec{p}(\nu) = (p_1, p_2, \ldots, p_{\pi(\nu)})$, is periodic with a period $T(\nu) = \prod_{p \le \nu} p$ since

$\mathrm{mod}(T(\nu), p) = 0$ for any $p \le \nu$. This implies that if x is a solution to the system of equations $\mathrm{mod}(x, p_i) = r_i\ (1 \le i \le \pi(\nu))$, then $y = x + T(\nu)$ is also a solution to the same system. We set $\nu = k(\nu) \cdot T(\nu) + r$, where $r = \mathrm{mod}(\nu, T(\nu))$. Then,

$\mathrm{mod}(\nu, p_i) = \mathrm{mod}(r, p_i) = r_i$ and since the combination of residual values $\vec{r}(\nu) = (r_1, r_2, \ldots r_{\pi(\nu)})$ occurs $k(\nu)$ times in ν trials, then for the relative frequency

$f(\nu, \vec{r}(\nu)) = \frac{k(\nu)}{\nu}$, we have: $\left| \frac{k(\nu)}{\nu} - \prod_{i \le \pi(\nu)} \frac{1}{p_i} \right| = \left| \frac{1}{T(\nu) + \frac{r}{k(\nu)}} - \frac{1}{T(\nu)} \right| \to 0$ as $\nu \to \infty$.

Q.E.D.

REFERENCES

1. Granville, A., *Harald Cramér and the distribution of prime numbers*, Scandinavian Actuarial Journal, 1:12-28. 1995]

2. Mark Kac, *Statistical Independence in Probability, Analysis and Number Theory*, The Mathematical Association of America, John Wiley and Sons, Inc.,1972.

3. Tom M. Apostol, *Introduction to Analytic Number Theory*, Springer, 2000.

4. Edward Beltrami, *What is Random? Chance and Order in Mathematics and Life.* Springer-Verlag New York, Inc., 1999.

5. Song Y. Yan, *Number Theory for Computing*, Springer, Springer-Verlag, 2000.

6. A.N. Shiryaev, *Probability*, 2nd edition, Springer, 1996.

7. V.S. Varadarajan, *Euler Trough Time: A new Look at Old Themes*, AMS, 2006.

8. Gérald Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, 3rd Edition, AMS, 2015.

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. Harry Fusrstenberg, *Poincaré Reccurence and Number Theory*, Bulletin of the American Mathematical Society, Volume 5, Number 3, November 1981, p. 211-234.

12. V.K. Balakrishnan, Introductory Discrete Mathematics, Prentice-Hall International, Inc.,1991.

13. Michel Loèv, Probability Theory, Third Edition. Dover Publications, Inc., Mineola, New York, 2017.

14. G.M. Sobko, *The First Diffusion Problem on Differentiable Manifolds*, Theory of Probability and its Applications, v. XVII, n.3, 1972, pp. 521 – 528.

15.. 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

16. А. Н. Колмогоров, *Основные понятия теории вероятностей*. Москва, “Наука”, 1974.

17. A.N. Shiryaev, *Probability*, 2nd edition, Springer, 1996.

18. S.R.S Varadhan, *Probability Theory*. Courant Institute of Mathematical Studies, American Mathematical Society. 2001.

19. S.R.S Varadhan, *Stochastic Processes*. Courant Institute of Mathematical Studies, American Mathematical Society. 2007.

20. Yu. I. Manin, Alexey A. Panchishkin, *Introduction to Modern Number Theory*. Springer, 2006

21. Jacque Neveu, *Bases Mathématiques du Calcul des Probabilités*, Masson et Cie, Paris, 1964.

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

email: greg.beholder@gmail.com (Gregory Sobko)