22. Uniform Distribution of Some Special Sequences

By Kazuo GOTO*) and Takeshi KANO**)

(Communicated by Shokichi IYANAGA, M. J. A., March 12, 1985)

We know that $(f(p_n))_1^\infty$, where p_n is n-th prime number, is uniformly distributed mod 1 if $f(x)$ is a polynomial with real coefficients and at least one of the coefficients of $f(x)-f(0)$ is irrational [6 or 2: Theorem 3.2], or if $f(x)$ is an entire function which is not a polynomial [4].

In this note we first consider some sufficient conditions that the sequence $(f(p_n))_1^\infty$ is uniformly distributed mod 1 where $f(x)$ is a kind of log-type function, and next we prove that $(\log n!)_1^\infty$ is uniformly distributed mod 1. In fact we give an estimate for the discrepancy of each sequence.

1. Definition (discrepancy). Let $a_1, a_2, \cdots, a_N$ be a finite sequence of real numbers. Then we define the discrepancy by

$$D_N = D_N(a_1, a_2, \cdots, a_N) = \sup_{0\le\alpha<\beta\le 1} |A([\alpha,\beta):N)/N - (\beta-\alpha)|,$$

where $A([\alpha,\beta):N)$ is the number of terms a_n, $1\leqq n\leqq N$, for which $\{a_n\} \in [\alpha,\beta)$. $\{x\}$ is the fractional part of x.

Lemma 1 (Erdös-Turán [2: p. 114]). *For any finite sequence $x_1, x_2, \cdots, x_N$ of real numbers and positive integer m, we have*

$$D_N \ll (1/m) + \textstyle\sum_{h=1}^m (1/h) |(1/N) \sum_{n=1}^N e^{2\pi i h x_n}|.$$

Theorem 1. *Let $f(x)$ be a continuously differentiable function with $f(x)\to\infty$ $(x\to\infty)$. If $f'(x)\log x$ is monotone, $n|f'(n)|\to\infty$ as $n\to\infty$, and*

$$f(n)/(\log n)^l \to 0 \qquad (n\to\infty) \text{ for some } l>1,$$

then $(\alpha f(p_n))_1^\infty$ is uniformly distributed mod 1, where $\alpha(\neq 0)$ is any real constant.

Proof. First we prove that the discrepancy D_N of $f(p_n)$, $n=1, 2, \cdots, N$, satisfies

$$(1) \qquad D_N \ll \sqrt{f(p_N)/(\log p_N)^l} + 1/N \max(1, 1/(\log p_N)|f'(p_N)|).$$

By Euler's summation formula

$$\begin{aligned} S_N = \sum_{n=1}^N e^{2\pi i h f(p_n)} &= \pi(p_N) e^{2\pi i h f(p_N)} - 2\pi i h \int_2^{p_N} \pi(t) f'(t) e^{2\pi i h f(t)} dt \\ &= \pi(p_N) e^{2\pi i h f(p_N)} - 2\pi i h \int_2^{p_N} (\mathrm{Li}(t) + R(t)) f'(t) e^{2\pi i h f(t)} dt \\ &= N e^{2\pi i h f(p_N)} - \int_2^{p_N} \mathrm{Li}(t)\, d(e^{2\pi i h f(t)}) - 2\pi i h \int_2^{p_N} R(t) f'(t) e^{2\pi i h f(t)} dt \\ &= N e^{2\pi i h f(p_N)} - [\mathrm{Li}(t) e^{2\pi i h f(t)}]_2^{p_N} + \int_2^{p_N} \frac{e^{2\pi i h f(t)}}{\log t} dt - 2\pi i h \int_2^{p_N} R(t) f'(t) e^{2\pi i h f(t)} dt, \end{aligned}$$

where $\pi(x)$ is the number of primes $\leqq x$ and

*) Sakuyo Junior College, Tsuyama, Okayama, Japan.
**) Okayama University, Tsushima, Okayama, Japan.

$$\mathrm{Li}\,(x)=\int_2^x \frac{dt}{\log t}, \qquad R(x)=\pi(x)-\mathrm{Li}\,(x).$$

So we have

$$S_N=(N-\mathrm{Li}\,(p_N))e^{2\pi i h f(p_N)}+\int_2^{p_N} \frac{e^{2\pi i h f(t)}}{\log t}dt-2\pi i h\int_2^{p_N} R(t)f'(t)e^{2\pi i h f(t)}dt.$$

Call these terms I_1, I_2, I_3, respectively. By the P. N. T. of Hadamard and de la Vallée Poussin, we have

$$R(x)\ll x/(\log x)^k \qquad \text{for any } k>1,$$

and

$$|I_1|\ll p_N/(\log p_N)^k.$$

By [5: Lemma 4.3 p. 61] and assumption,

$$|I_2|=\left|\int_2^{p_N} \frac{e^{2\pi i h f(t)}}{\log t}dt\right|\ll\frac{1}{|h|}\max\left(\frac{1}{|f'(2)\log 2|},\ \frac{1}{|f'(p_N)\log p_N|}\right).$$

$$|I_3|\leqq 2\pi h\int_2^{p_N} R(t)|f'(t)|dt\ll\frac{|h|p_N}{(\log p_N)^k}f(p_N).$$

Hence, for $h>0$,

$$\left|\sum_{n=1}^{N} e^{2\pi i h f(p_n)}\right|\ll\frac{1}{h}\max\left(1,\ \frac{1}{|f'(p_N)|\log p_N}\right)+\frac{hp_N f(p_N)}{(\log p_N)^k}.$$

Using Lemma 1 and $p_N\sim N\log N$,

$$D_N\ll\frac{1}{m}+\sum_{h=1}^{m}\frac{1}{hN}\left\{\frac{1}{h}\max\left(1,\ \frac{1}{|f'(p_N)|\log p_N}\right)+\frac{hp_N}{(\log p_N)^k}f(p_N)\right\}$$
$$\ll\frac{1}{m}+\frac{1}{N}\max\left(1,\ \frac{1}{(\log p_N)|f'(p_N)|}\right)+\frac{p_N f(p_N)}{N(\log p_N)^k}m.$$

If we put $m=[\sqrt{N(\log p_N)^k/p_N f(p_N)}]$, then

$$D_N\ll\sqrt{f(p_N)/(\log p_N)^{k-1}}+(1/N)\max\,(1,\ 1/(\log p_N)|f'(p_N)|).$$

Thus we obtain (1). Since by assumption,

$$1/N\cdot 1/(\log p_N)f'(p_N)\ll 1/p_N|f'(p_N)|\to 0 \qquad (N\to\infty),$$

we have $D_N\to 0$ as $N\to\infty$. q.e.d.

Theorem 2. *Let $f(x)$ be a continuously differentiable function with $f'(t)>0$ and $f''(t)>0$. If $t^2f''(t)\to\infty$ as $t\to\infty$ and*

$$f(n)/(\log n)^l\to 0\ \ (n\to\infty)\quad \text{for some } l>1,$$

then $(\alpha f(p_n))_1^\infty$ is uniformly distributed mod 1, where α ($\neq 0$) is any real constant.

Proof. The proof runs along the same lines as that of Theorem 1. Using [7: Lemma 10.2, p. 225], we have

$$|I_2|=\left|\int_2^{p_N}\frac{e^{2\pi i h f(t)}}{\log t}dt\right|\ll\max_{t\in[2,p_N]}\frac{1}{\log t}\left|\frac{1}{hf''(t)}\right|^{1/2}.$$

By Lemma 1, we obtain

$$D_N\ll 1/m+1/N\max_{t\in[2,p_N]}(1/(\log t)\sqrt{f''(t)})+(p_N f(p_N)/N(\log p_N)^k)m.$$

Putting $m=[\sqrt{N(\log p_N)^k/p_N f(p_N)}]$ and using $p_N\sim N\log N$, we have

$$D_N\ll f(p_N)/(\log p_N)^l+\log p_N/(p_N\max_{2\le t\le p_N}(\log t)\sqrt{f''(t)}).$$

Now we consider

$$J:=\frac{\log p_N}{p_N\max_{2\le t\le p_N}(\log t)\sqrt{f''(t)}}=\frac{\log p_N}{p_N\max_{2\le t\le p_N}((\log t)/t)t\sqrt{f''(t)}}.$$

Since $(\log t)/t$ is monotonely decreasing for $t>e$,

$$J \leqq \frac{\log p_N}{p_N(\log p_N/p_N)\max_{2\leq t\leq p_N} t\sqrt{f''(t)}} = \frac{1}{\max_{2\leq t\leq p_N}\sqrt{t^2 f''(t)}}$$

$$\leqq \frac{1}{\sqrt{p_N^2 f''(p_N)}} \to 0 \qquad (N\to\infty).$$

So we have $D_N\to 0$ as $N\to\infty$. q.e.d.

2. The sequence $(\log n!)$, $n=1, 2, \cdots$, is uniformly distributed mod 1, which means that Benford's law holds for the sequence $(n!)$, $n=1, 2, \cdots$, [3]. Now we prove a more precise result by estimating the discrepancy of $(\log n!)$.

Theorem 3. *The discrepancy D_N of* $(\log n!)$, $n=1, 2, \cdots, N$, *satisfies for any* $\varepsilon>0$,

$$D_N \ll N^{-1/2+\varepsilon}.$$

Proof. By Stirling's formula [1 : p. 129], for any positive integer n we have

$$\log n! = \textstyle\sum_{j=1}^n \log j = (n+1/2)\log(n+1)-(n+1)+k+R(n+1),$$

where k is a constant and $R(t)$ is defined by

$$R(t)=\int_0^\infty \frac{p(x)}{t+x}dx, \quad \text{where } p(x)=[x]-x+(1/2), \quad t>0.$$

Then

$$R'(t)=\frac{d}{dt}\int_0^\infty \frac{p(x)}{t+x}dx=\int_0^\infty \frac{-p(x)}{(t+x)^2}dx,$$

$$R''(t)=\int_0^\infty \frac{p(x)}{t+x}dx=\sum_{n=0}^\infty \int_n^{n+1}\frac{2p(x)}{(t+x)^3}dx$$

$$=\sum_{n=0}^\infty \left\{\left[-\frac{n-x+(1/2)}{(t+x)^2}\right]_n^{n+1}-\int_n^{n+1}\frac{dx}{(t+x)^2}\right\}$$

$$=\sum_{n=0}^\infty \left[\frac{1}{2}\left\{\frac{1}{(t+x)^2}+\frac{1}{(t+n+1)^2}+\frac{1}{t+n+1}-\frac{1}{t+n}\right]\right.$$

$$=\sum_{n=0}^\infty \frac{1}{2(t+n)^2(t+n+1)^2}>0.$$

Now we consider

$$S_N=\sum_{n=1}^N e^{2\pi i h\log n!}$$

$$=\sum_{n=1}^N e^{2\pi i h[(n+(1/2))\log(n+1)-(n+1)+k+R(n+1)]}$$

$$=e^{2\pi i h k}\sum_{n=1}^N e^{2\pi i h[(n+(1/2))\log(n+1)+R(n+1)]}.$$

We set for any integers a and b, $S(a,b)=S_b-S_a$, and for $h\in Z-\{0\}$,

$$f(u)=h[(u+(1/2))\log(u+1)+R(u+1)].$$

So we obtain

$$|R'(t)|\leqq 1/2, \quad R''(t)\geqq 0 \quad \text{and} \quad (1/h)f''(t)\geqq(1/(t+1)).$$

Hence by [7 : Lemma 4.6, p. 198],

$$|S(a,b)|<|e^{2\pi i h k}|[|f'(b)-f'(a)|+2]((4\sqrt{b+1})/\sqrt{|h|}+3)$$

$$\ll |h|[\log((b+1)/(a+1))+(1/2)(1/(a+1))-(1/(b+1))+3]\sqrt{b/|h|}.$$

If $a<b\leqq 2a$, then

(2) $$|S(a,b)|\ll\sqrt{|h|}\sqrt{b}.$$

For any given N, we choose a such that $2^a\leqq N<2^{a+1}$. Then we have

$$S_N=S(0,1)+S(1,2)+\cdots+S(2^{a-1},2^a)+S(2^a,N),$$

and by (2),

$$|S_N|\ll\sqrt{|h|}\sum_{m=1}^{a}\sqrt{2^m}+\sqrt{|h|}\sqrt{N}\ll\sqrt{|h|}\sqrt{N}.$$

Thus for any function $g(n)$ which tends monotonically to infinity, we get

$$\lim_{N\to\infty} S_N/\sqrt{N}g(N)=0,$$

which implies that there exists an $N_0(h)$ such that

$$|S_N|<\sqrt{N}g(N) \qquad \text{for all } N\geqq N_0(h).$$

Hence by Lemma 1, we obtain

$$D_N\ll(1/m)+\sum_{h=1}^{m}(1/hN)\sqrt{N}g(N)$$
$$\ll(1/m)+(1/\sqrt{N})g(N)\cdot\log m.$$

If we choose $m=[N^{(1/2)-\varepsilon}]$, then for all $N\geqq\max(N_0,(h+1)^{2/(1-2\varepsilon)})$ we have

$$D_N\ll N^{-(1/2)+\varepsilon}+1/\sqrt{N}g(N)\cdot\log N\ll N^{-(1/2)+\varepsilon}g(N)\ll N^{-(1/2)+\varepsilon}$$

because of the definition of $g(n)$. q.e.d.

Corollary. *The sequence* ($\log n!$), $n=1, 2, \cdots$, *is uniformly distributed mod* 1.

References

[1] H. Cramér: Mathematical Method of Statistics. Princeton (1946).
[2] L. Kuipers and H. Niederreiter: Uniform Distribution of Sequences. Wiley (1974).
[3] K. Nagasaka: On Benford's law. Ann. Inst. Stat. Math., vol. 36, no. 2, part A, pp. 337–352 (1984).
[4] G. Rhin: Répartition modulo 1 de $f(p_n)$ quand f est une série entière. Lecture Notes in Math., vol. 475, Springer, pp. 176–244 (1975).
[5] E. C. Titchmarsh: The Theory of the Riemann Zeta Function. Oxford (1951).
[6] I. M. Vinogradov: On an estimate of trigonometric sums with prime numbers. Izv. Akad. Nauk SSSR, Ser. Mat., **12**, 225–248 (1948) (in Russian).
[7] A. Zygmund: Trigonometric Series. vol. 1, Cambridge (1959).