ON THE NOTION OF UNIFORM DISTRIBUTION MOD 1

Rita Giuliano Antonini*
Universita di Pisa, Via Buonarroti, 2, 56100 Pisa, Italy
(Submitted August 1989)

0. Introduction

The notion of a uniformly distributed sequence mod 1 is a classical tool of number theory (see, e.g., [1], [2]), but it is well known that there exist sequences which are not uniformly distributed; it turns out that this kind of sequence is more conveniently treated by notions other than the classical ones.

In this paper one such notion is used, which enables us to study the sequence formed by the fractional parts of decimal logarithms of the integers (it is well known that this sequence is not uniformly distributed in the classical sense; see, e.g., [1]).

With our result, we obtain a simple solution of the so-called first digit problem.

1. Preliminary Results

In this section we list some definitions and results used in the sequel. We begin with the definition of uniform distribution with respect to a measure on $\mathbb{N}^*$.

Definition 1.1: Let μ be a measure on $\mathbb{N}^*$, which we assume to be positive and unbounded; for each integer n, write

$$S_n = \mu([1, n]).$$

Now let $(x_n)_{n \geq 1}$ be a sequence of real numbers in $[0, 1]$. We say that $(x_n)_{n \geq 1}$ is μ-uniformly distributed in $[0, 1]$ if, for each function f in $C([0, 1])$, we have

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} \mu\{k\} f(x_k)}{S_n} = \int_0^1 f(x)\,dx.$$

Remark 1.2: It is easily seen that we may replace (S_n) by any equivalent sequence.

Remark 1.3: The notion of uniform distribution in the sense of Definition (1.1) has been introduced by other authors, although they used different names and symbols.

It is also clear that it can be expressed by saying that the sequence of measures $(\nu_n)_{n \geq 1}$ on $[0, 1]$ defined by

$$\nu_n = \frac{\mu\{k\}}{S_n} \varepsilon_{x_k}$$

weakly converges to the Lebesgue measure on $[0, 1]$ (see, e.g., [8]).

In what follows, we shall use the following proposition, a direct consequence of well-known results concerning weak convergence; note that it is a straightforward generalization of a classical theorem in number theory (see [1], [2]).

*Lavoro svolto nell'ambito del GNAFA e con finanziamento del MPI.

Proposition 1.4: The following conditions are equivalent:

(a) $(x_n)_{n \geq 1}$ is μ-uniformly distributed in [0, 1];
(b) for every interval $[a, b[$ in [0, 1], we have

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} \mu\{k\} 1_{[a,b[}(x_k)}{S_n} = b - a,$$

where $1_{[a,b[}$ stands for the indicator function of $[a, b[$.

For each integer n, write

(1.5) $H_n = \exp S_n$.

We shall assume that the sequence $(H_n)_{n \geq 1}$ is obtained by restriction to $\mathbb{N}^*$ of a function H defined on $\mathbb{R}^+$ and having the following property:

(1.6) There exists a positive constant ℓ and an increasing, slowly varying function L such that

$$H(y) = y^{\ell} L(y).$$

(We recall that L varies slowly at infinity if, for every $x > 0$, we have

$$\lim_{y \to +\infty} \frac{L(xy)}{L(y)} = 1.$$

For further properties, see [5].)

To handle the case $\ell = 0$, we make an additional assumption:

(1.7) For each (a_1, a_2, a_3, a_4) in $\mathbb{R}^4$, where a_1, a_2, a_3, a_4 are strictly positive numbers such that $a_1 \neq a_2$, $a_3 \neq a_4$, we have

$$\frac{L(a_1 y) - L(a_2 y)}{\log(a_1 a_2^{-1})} \sim \frac{L(a_3 y) - L(a_4 y)}{\log(a_3 a_4^{-1})}$$

as y converges to infinity.

We prove the following proposition.

Proposition 1.8:

(a) For every $x > 0$, we have $\displaystyle\lim_{y \to +\infty} \frac{L(x + y)}{L(y)} = 1$.

(b) In the case $\ell = 0$, for each (b_1, b_2, b_3, b_4) in $\mathbb{R}^4$, where b_1, b_2, b_3, b_4 are positive numbers, we have

$$\frac{L(a_1 y + b_1) - L(a_2 y + b_2)}{\log(a_1 a_2^{-1})} \sim \frac{L(a_3 y + b_3) - L(a_4 y + b_4)}{\log(a_3 a_4^{-1})},$$

as y converges to infinity.

Proof: Part (a) follows from the inequalities

$$1 \leq \frac{L(x + y)}{L(y)} \leq \frac{L(2y)}{L(y)},$$

the second of which holds for y sufficiently large.

The assertion of part (b) is proved by noting that, for every $\varepsilon > 0$, we have, for y sufficiently large

$$\frac{L(a_1 y) - L((a_2 + \varepsilon)y)}{L((a_3 + \varepsilon)y) - L(a_4 y)} \leq \frac{L(a_1 y + b_1) - L(a_2 y + b_2)}{L(a_3 y + b_3) - L(a_4 y + b_4)} \leq \frac{L((a_1 + \varepsilon)y) - L(a_2 y)}{L(a_3 y) - L((a_4 + \varepsilon)y)}.$$

Definition 1.9: Let μ be a measure on $\mathbb{N}^*$; we say that μ has property P if (1.6) holds [in the case $\ell = 0$, if (1.6) and (1.7) hold].

We shall also use some results concerned with the notion of density on $\mathbb{N}^*$, which is studied, for example, in [6].

Definition 1.10: Let μ be a measure on $\mathbb{N}^*$, and (S_n) its distribution function as defined in Definition (1.1).

Consider the density on $\mathbb{N}^*$ generated by the sequence of measures, $(\mu)_{n \geq 1}$, defined as follows:

$$\mu_n = \frac{1}{S_n} 1_{[1, n]} \cdot \mu;$$

this density will be called the μ-density.

Definition 1.11: For each $t > 0$, let $\tilde{\mu}_t$ be the measure on $\mathbb{N}^*$ defined by

$$\tilde{\mu}_t = \sum_{k \geq 1} [\exp(-tS_k) - \exp(-tS_{k+1})]\varepsilon_k.$$

The density generated by $(\tilde{\mu}_t)$ will be called the exponential density with respect to μ (or, more briefly, the μ-exponential density).

We state the following result, the proof of which is given in [6].

Proposition 1.12: Assume that the sequence $(\mu\{n\})_{n \geq 1}$ is bounded. Then the μ-density and the μ-exponential density agree everywhere.

The following theorem, proved in [7], gives a practical method for calculating an exponential density.

Theorem 1.13: Let $(\ell_n)_{n \geq 1}$, $(m_n)_{n \geq 1}$ be two sequences of positive real numbers, such that

(i) $\lim\limits_{n \to \infty} \ell_n = \lim\limits_{n \to \infty} m_n = +\infty$ and $\ell_n \leq m_n \leq \ell_{n+1}$ for every integer n;

(ii) the sequence $(m_n - \ell_n)_{n \geq 1}$ is bounded;

(iii) we have $m_n \sim m_{n+1}$, $\ell_n \sim \ell_{n+1}$ as n converges to infinity.

Last, let A be a real number, with $0 \leq A \leq 1$; then the following conditions are equivalent:

$$\text{(a)} \lim_{n \to \infty} \frac{\sum_{k=1}^{n} (\mu_k - \ell_k)}{m_n} = A; \qquad \text{(b)} \lim_{n \to \infty} \frac{\sum_{k=1}^{n} (\mu_k - \ell_k)}{\ell_n} = A;$$

$$\text{(c)} \lim_{n \to 0^+} \sum_{k \geq 1} [\exp(-t\ell_k) - \exp(-tm_k)] = A.$$

2. The Theorem of Uniform Distribution

We shall prove the following result.

Theorem 2.1: Let μ be a measure on $\mathbb{N}^*$, with property P. Then the sequence $(\{\log_{10} n\})_{n \geq 1}$ is μ-uniformly distributed in [0, 1].

Proof: Proposition (1.4) applies, so we can show that, for every interval $[a, b[$ in [0, 1], we have

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} \mu\{k\} 1_{[a, b[}(\{\log_{10} k\})}{\log H_n} = b - a.$$

We can write

$$1_{[a,b[}(\{\log_{10}\}k) = 1_E(k),$$

where E is the subset of $\mathbb{N}^*$ the elements of which are the integers k satisfying the relation

$$10^{n+a} \le k < 10^{n+b}$$

for some integer n; hence, calculating the limit above amounts to finding the μ-density of E (in the sense of Definition 1.10).

First we note that, because of the relations

$$10^{n+b} - 10^{n+a} \ge 1 \quad \text{and} \quad 10^{n+a+1} - 10^{n+b} \ge 1,$$

which hold for n sufficiently large, E is neither finite nor cofinite. Denote by $(p_n)_{n\ge1}$, $(q_n)_{n\ge1}$ the two sequences of integers such that

$$E = \bigcup_{n\ge1} [p_n,\ q_n[.$$

Moreover, for every $x > 0$, write

$$\vartheta(x) = \begin{cases} x & \text{if } x \text{ is an integer} \\ [x] + 1 & \text{otherwise,} \end{cases}$$

so that we have the obvious relations

$$p_n = \vartheta(10^{n+a}); \quad q_n = \vartheta(10^{n+b}).$$

Because of our hypotheses on μ, the sequence $(\mu\{n\})_{n\ge1}$ is bounded; hence, Proposition 1.12 applies, and our goal is equivalent to finding the μ-exponential density of E, that is, we calculate the limit

$$\lim_{t\to0^+} \sum_{n\ge1} [\exp(-t \log H_{p_n}) - \exp(-t \log H_{q_n})];$$

we do this by means of Theorem 1.13, where we put

$$\ell_n = \log H_{p_n}; \quad m_n = \log H_{q_n}.$$

The inequalities $x \le \vartheta(x) \le x + 1$, together with Proposition 1.8, give

$$\lim_{n\to\infty} \frac{m_n - \ell_n}{m_n - m_{n-1}} = b - a;$$

now, a well-known theorem of Cesaro gives the same value for the limit we considered in Theorem 1.13(a).

Remark 2.2: Paper [3] treats, using different techniques, the particular case of the preceding theorem where $\mu\{n\} = 1/n$ (so that $S_n \sim \log n$). Paper [3] also contains a reference to another paper [4] in which the same particular case is studied. The same result is extended in a different direction in Theorem 7.16 on page 64 of [1].

Now, let r be an integer, with $1 \le r \le 9$ and, in the proof of Theorem 2.1, take $a = \log_{10} r$, $b = \log_{10}(r + 1)$; then E turns out to be the set of integers the decimal expansion of which has r as the first digit and the preceding proof gives

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \mu\{k\} 1_E(k)}{\mu([1,\ n])} = \log_{10} \frac{r + 1}{r}.$$

This simple remark may be rephrased as follows:

Corollary 2.3: Let E be the set of integers the decimal expansion of which has r as the first digit; if μ is a measure on $\mathbb{N}^*$ satisfying the property P, then the μ-density of E is $\log_{10}(r+1)/r$.

References

1. L. Kuipers & H. Niederreiter. *Uniform Distribution of Sequences*. New York: Wiley, 1974.
2. I. P. Cornfeld, S. V. Fomin, & Ya. G. Sinai. *Ergodic Theory*. Berlin & New York: Springer, 1982.
3. J. D. Vaaler. "A Tauberian Theorem Related to Weyl's Criterion." *J. Number Theory* *9* (1977):71-78.
4. M. Tsuji. "On the Uniform Distribution of Numbers Mod 1." *J. Math. Soc. Japan* *4* (1952):313-22.
5. W. Feller. *An Introduction to Probability Theory and Its Applications*, vol. II. New York: Wiley, 1971.
6. R. Giuliano Antonini. "Comparaison de densités arithmetiques." *Rend. Acc. Naz. dei XL, Memorie di Mat.* *104* (1986):X, 12, 153-63.
7. R. Giuliano Antonini. "Construction et comparaison de densités arithmétiques." *Rend. Acc. Naz. dei XL, Memorie di Mat.* *106* (1988): XII, 9, 117-23.
8. P. Billingsley. *Convergence of Probability Measures*. New York: Wiley, 1968.

*****