International Journal of Analysis and Applications

Typical Sequence of Real Numbers From the Unit Interval Has All Distribution Functions

József Bukor^{1,*}, Kálmán Liptai², János T. Tóth²

¹Department of Informatics, J. Selye University, Komárno, Slovakia ²Department of Mathematics, J. Selye University, Komárno, Slovakia

*Corresponding author: bukorj@ujs.sk

Abstract. This note is devoted to the study of typical properties (in Baire category sense) of sequences of real numbers in [0, 1]. We prove that the subset of sequences that have all distribution functions forms a residual set.

1. Introduction

The concept of Baire categories is one of the possibilities to compare sets. Let *S* be a metric space. A subset $A \subseteq S$ is called *meager* (or of first category) if *A* can be written as a countable union of nowhere dense sets. Any set that is not meager is said to be of second category. The complement of a meager set is called *residual*. We say that a typical element *x* has property *P* if the set $A = \{x \in S | x \text{ has property } P\}$ is residual. For more details we refer the reader to Oxtoby [6].

There are analogous results in Baire category sense for the digit sequences of numbers $z \in [0, 1]$ and the sequences of real numbers. We mention some results. For a fixed positive integer *s* the unique, non-terminating, base *s* expansion of a number $z \in [0, 1]$ is

$$z = \frac{d_1(z)}{s} + \frac{d_2(z)}{s^2} + \dots + \frac{d_n(z)}{s^n} + \dots \quad \text{with } d_i(z) \in \{0, 1, \dots, s-1\}.$$

For each digit $i \in \{0, 1, ..., s - 1\}$ let $\Pi_i(z; n)$ denote the frequency of the digit i among the first n digit of z. It was proved by Šalát [7] that for a typical z, we have $\limsup_{n \to \infty} \Pi_i(z; n) = 1$ and $\liminf_{n \to \infty} \Pi_i(z; n) = 1$. Define the frequency of the digits $i \le x$ among the first n digits of z as

$$F_{z,n}(x) = \sum_{i \le x} \prod_i (z; n).$$

Received: Mar. 19, 2024.

²⁰²⁰ Mathematics Subject Classification. 54E52.

Key words and phrases. Baire category; distribution function; sequences.

Let \mathcal{F} denote the set of all distribution functions of discrete random variables that takes on one of the possible values $0, 1, \ldots, s - 1$. Using this notation, we mention Olsen's [4] fundamental result. For a typical number z we have that for any $f \in \mathcal{F}$ there exists an increasing sequence n_1, n_2, \ldots for that $\lim_{k \to \infty} F_{z,n_k}(x) = f$. Roughly speaking, the digit expansion of a typical number z has all distribution functions from \mathcal{F} .

We will consider the metric space S of all sequences of real numbers in [0, 1] with the Fréchet metric

$$\rho(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k - y_k|}{1 + |x_k - y_k|},$$

where $\mathbf{x} = (x_k)$, $\mathbf{y} = (y_k)$. It is known that (S, ρ) is a complete metric space.

In [3] it was proved that the set of all uniformly distributed sequences is a dense subset of the first Baire category in *S*. The same is true for the set of all statistically convergent sequences of real numbers (cf. [8]). The sequence (x_n) is *maldistributed* if for any non-empty interval *I* the set $\{n \in \mathbb{N} : x_n \in I\}$ has upper asymptotic density 1.

Examples of maldistributed sequences are given in [9] and [2]. In [1] the authors proved that a typical real sequence is maildistributed. The maildistribution property can be characterized by one-jump distribution functions [9], so a typical real sequence has all one-jump distribution functions.

The aim of this not to show that a typical real sequence has all distribution function.

1.1. **Basic notations and properties of distribution functions.** We recall some basic notations and results concerning distribution functions of sequences (e.g., see [11] and [10]).

- Let $\mathbf{x} = (x_n)$ be a sequence from unit interval [0, 1].
- Let $\chi_A(x)$ denote the characteristic function of the set *A*.
- Denote by

$$F_N(x) = \frac{\#\{n \le N; x_n \in [0, x)\}}{N} = \frac{1}{N} \sum_{n=1}^N \chi_{[0, x)}(x_n)$$

the step distribution function for $x \in [0, 1)$, and for x = 1 we define $F_N(1) = 1$.

- A non-decreasing function g : [0,1] → [0,1], g(0) = 0, g(1) = 1 is called a *distribution function* (abbreviated d.f.). We shall identify any two d.f.s coinciding at common points of continuity. Denote by *G* the set of all distribution functions.
- A d.f. g(x) is a d.f. of the sequence **x**, if there exists an increasing sequence $n_1 < n_2 < \cdots$ of positive integers such that

$$\lim_{k\to\infty}F_{n_k}(x)=g(x)$$

almost everywhere on [0, 1]. This is equivalent to the weak convergence, i.e., the preceding limit holds for every point $x \in [0, 1]$ of continuity of g(x). Let $G(\mathbf{x})$ denote the set of all d.f.s of \mathbf{x} .

• $c_{\gamma}(x)$ is one-step d.f. for which $c_{\gamma}(x) = 0$ for $x \in [0, \gamma]$ and $c_{\gamma}(x) = 1$ for $x \in (\gamma, 1]$.

• For every sequence **x** there hold that $G(\mathbf{x})$ is closed and $G(\mathbf{x})$ is connected in the weak topology defined by the metric

$$d(g_1, g_2) = \left(\int_0^1 (g_1(x) - g_2(x))^2 \mathrm{d}x\right)^{\frac{1}{2}}.$$
(1.1)

- For given a non–empty set *H* of d.f.s there exists a sequence \mathbf{x} in [0,1) such that $G(\mathbf{x}) = H$ if and only if *H* is closed and connected.
- First Helly theorem. Every sequence $g_n(x)$ of d.f.s contains a subsequence $g_{k_n}(x)$ such that $\lim_{n \to \infty} g_{k_n}(x) = g(x)$ for every $x \in [0, 1]$. Furthermore, the point limit g(x) is d.f. again.

2. Results

First, we show that a typical sequence has distribution function, which in given point has the function value "near" to the prescribed value.

Lemma 2.1. Let $a, b \in (0, 1)$. For a positive number $\gamma < \min\{b, \frac{a}{4}, \frac{1-a}{4}\}$ denote by $\mathcal{A}(a, b, \gamma)$ the set of all $\mathbf{x} = (x_k) \in S$ for which there is an n_0 such that for any $n \ge n_0$ we have

$$\sum_{i=1}^{n} \chi_{[0,a-\gamma)}(x_i) < (b-\gamma)n \quad or \quad \sum_{i=1}^{n} \chi_{[0,a+\gamma)}(x_i) > (b+\gamma)n.$$
(2.1)

Then $\mathcal{A}(a, b, \gamma)$ *is a set of the first Baire category in S.*

Proof. We define continuous functions $h_{a,\gamma}$: $[0,1] \rightarrow [0,1]$ and $t_{a,\gamma}$: $[0,1] \rightarrow [0,1]$ by

$$h_{a,\gamma}(x) = \begin{cases} 1 & \text{for } x \in [0, a - 2\gamma] \\ \frac{a - \gamma - x}{\gamma} & \text{for } x \in [a - 2\gamma, a - \gamma] \text{, } t_{a,\gamma}(x) = \begin{cases} 1 & \text{for } x \in [0, a + \gamma] \\ \frac{a + 2\gamma - x}{\gamma} & \text{for } x \in [a + \gamma, a + 2\gamma] \text{,} \\ 0 & \text{for } x \in [a - \gamma, 1] \end{cases}$$

see Figure 1.



FIGURE 1. Functions $h_{a,\gamma}(x)$ and $t_{a,\gamma}(x)$

For these functions, we have $h_{a,\gamma}(x) \leq \chi_{[0,a)}(x) \leq t_{a,\gamma}(x)$, where $x \in [0,1]$. Using the functions $h_{a,\gamma}$, $t_{a,\gamma}$ we define for $\mathbf{x} \in S$ and fixed *n* the function $f_n : S \to [0,1]$ in the following way:

$$f_n(\mathbf{x}) = \min\left\{1, \left(\frac{\sum\limits_{i=1}^n h_{a,\gamma}(x_i)}{(b-\frac{\gamma}{2})n}\right)^n\right\} \cdot \min\left\{1, \left(\frac{(b+\frac{\gamma}{2})n}{1+\sum\limits_{i=1}^n t_{a,\gamma}(x_i)}\right)^n\right\}$$

Denote $\mathcal{A}^*(a, b, \gamma)$ the set of all $\mathbf{x} \in S$ for which there exists the limit $\lim_{n \to \infty} f_n(\mathbf{x})$. One can easily check that if (2.1) holds for all sufficiently large n, then $f_n(\mathbf{x}) \to 0$ for $n \to \infty$. Therefore $\mathcal{A}(a, b, \gamma) \subset \mathcal{A}^*(a, b, \gamma)$.

Put $f(\mathbf{x}) = \lim_{n \to \infty} f_n(\mathbf{x})$ for $\mathbf{x} \in \mathcal{R}^*(a, b, \gamma)$. We shall prove that (a) the function f_n (n = 1, 2, ...) is a continuous function on S, (b) f is discontinuous at each point of $\mathcal{R}^*(a, b, \gamma)$.

(a) the continuity of the functions f_n follows from the facts that the functions $h_{a,\gamma}$, $t_{a,\gamma}$ are continuous and the convergence in the space *S* is the coordinate convergence.

(b) Let y = (y_k) ∈ A^{*}(a, b, γ). We have the following two possibilities.
(1) f(y) < 1,
(2) f(y) = 1.

In case (1) we choose a positive ε such that $\varepsilon < 1 - f(\mathbf{y})$. It is suffice to prove that in each ball $K(\mathbf{y}, \delta) = \{\mathbf{x} \in \mathcal{A}^*(a, b, \gamma), \rho(\mathbf{x}, \mathbf{y}) < \delta\}$ $(\delta > 0)$ of the subspace $\mathcal{A}^*(a, b, \gamma)$ of *S* there exists an element $\mathbf{x} \in S$ with $|f(\mathbf{x}) - f(\mathbf{y})| > \varepsilon$.

Let $\delta > 0$ is given. Choose a positive integer *m* such that $\sum_{k=m+1}^{\infty} 2^{-k} < \delta$. Choose a d.f. $g(x) \in \mathcal{G}$ which is continuous in x = a and g(a) = b. Then there exists a sequence $z \in S$ for that $G(z) = \{g(x)\}$. Define the sequence x in the following way:

$$x_k = \begin{cases} y_k, & \text{if } k \le m, \\ \frac{a}{2}, & \text{if } k > m \text{ and } z_k \in [0, a), \\ \frac{a+1}{2}, & \text{if } k > m \text{ and } z_k \in [a, 1] \end{cases}$$

Hence $\rho(\mathbf{x}, \mathbf{y}) < \delta$. Furthermore, $\frac{1}{n} \sum_{i=1}^{n} \chi_{[0,a)}(x_i)$, $\frac{1}{n} \sum_{i=1}^{n} h_{a,\gamma}(x_i)$ and $\frac{1}{n} \sum_{i=1}^{n} t_{a,\gamma}(x_i)$ tend to *b* as $n \to \infty$. Then $f_n(\mathbf{x}) = 1$ for all sufficiently large *n* and therefore $f(\mathbf{x}) = \lim_{n \to \infty} f_n(\mathbf{x}) = 1$. Then immediately follows

$$f(\mathbf{x}) - f(\mathbf{y}) = 1 - f(\mathbf{y}) > \varepsilon.$$

In case (2) we have $g(\mathbf{y}) = 1$. Let δ , *m*, **x** have the previous meaning. Put

$$x_{k} = \begin{cases} y_{k}, & \text{if } k \le m, \\ \frac{a+1}{2}, & \text{if } k > m \text{ and } b \ge \frac{1}{2} \\ \frac{a}{2}, & \text{if } k > m \text{ and } b < \frac{1}{2}. \end{cases}$$

Then, clearly $\rho(\mathbf{x}, \mathbf{y}) < \delta$, and for sufficiently large *n* one of the inequalities (2.1) must be true. So, we have $f(\mathbf{x}) = \lim_{n \to \infty} f_n(\mathbf{x}) = 0$, and therefore $f(\mathbf{y}) - f(\mathbf{x}) = 1 - 0 > 0$. Hence the discontinuity of *f* at $\mathbf{y} \in \mathcal{R}^*(I, \gamma)$ has been proved.

The function f is a limit function (on $\mathcal{A}^*(a, b, \gamma)$) of the sequence of continuous functions $(f_n)_{n=1}^{\infty}$ on $\mathcal{A}^*(a, b, \gamma)$. Then the function f is a function in the first Baire class on $\mathcal{A}^*(a, b, \gamma)$. According to the well-known fact that the set of discontinuity points of an arbitrary function of the first Baire class is a set of the first Baire category (cf. [6], p. 32), we see that the set $\mathcal{A}^*(a, b, \gamma)$ is of the first Baire category in $\mathcal{A}^*(a, b, \gamma)$. Thus $\mathcal{A}^*(a, b, \gamma)$ is in S, too. Since $\mathcal{A}(a, b, \gamma) \subset \mathcal{A}^*(a, b, \gamma)$, the assertion follows.

Consequence 2.1. *For any* $a, b \in (0, 1)$ *the set*

$$\mathcal{P} = \{\mathbf{x} \in \mathbf{S} | \text{ there is a } g(\mathbf{x}) \in G(\mathbf{x}) \text{ for that } g(a) = b\}$$

is residual in S.

Proof. By Lemma 2.1 we have that the infinite union $\bigcup_{n=n_0}^{\infty} \mathcal{A}(a, b, \frac{1}{n})$ (where $\frac{1}{n_0} < \min\{b, \frac{a}{4}, \frac{1-a}{4}\}$) is a meager set in *S*. Let **x** be a sequence from the complementary set to the mentioned infinite union. For the step d.f. of **x** we have

$$b - \gamma \le F_N(a - \gamma) \le F_N(a) \le F_N(a + \gamma) \le b + \gamma$$

for any $\gamma = \frac{1}{n}$ ($n \ge n_0$) and infinitely many N. In this case, First Helly theorem implies that there exists a pointwise convergent subsequence with limit $g(x) \in G(\mathbf{x})$ for that g(a) = b. So, $\mathbf{x} \in \mathcal{P}$ and it means that the set \mathcal{P} is residual in S.

Remark. The assertion of Lemma 2.1 holds for the case b = 1, too. For the case b = 0 we only need to consider right-hand side inequality of (2.1).

In what follows, for simplicity, we will use the notation \mathbf{a}_l for a finite sequence $(a_k)_{k=1}^l$ and the notation \mathbf{b}_l for a finite sequence $(b_k)_{k=1}^l$. We extend the assertion of Lemma 2.1 for arbitrary number of finite points.

Lemma 2.2. Let a positive integer l and finite sequences \mathbf{a}_l and \mathbf{b}_l are given, where $0 < a_1 < a_2 \cdots < a_l < 1$ and $0 < b_1 \le b_2 \cdots \le b_l \le 1$. For a positive number

$$\gamma < \min\left\{b_1, \frac{a_1}{4}, \frac{a_2 - a_1}{4}, \frac{a_3 - a_2}{4}, \dots, \frac{a_l - a_{l-1}}{4}, \frac{1 - a_l}{4}\right\}$$

denote by $\mathcal{A}(\mathbf{a}_l, \mathbf{b}_l, \gamma)$ *the set of all* $\mathbf{x} = (x_k) \in S$ *for which there is an* n_0 *such that for any* $n \ge n_0$ *we have that at least one of the inequalities*

$$\sum_{i=1}^{n} \chi_{[0,a_{j}-\gamma)}(x_{i}) < (b_{j}-\gamma)n, \qquad \sum_{i=1}^{n} \chi_{[0,a_{j}+\gamma)}(x_{i}) > (b_{j}+\gamma)n$$
(2.2)

hold (j = 1, 2, ..., n). Then $\mathcal{A}(\mathbf{a}_l, \mathbf{b}_l, \gamma)$ is a set of the first Baire category in S.

Proof. The proof is analogous to the proof of Lemma 2.1. We mention only the differences. The crucial role is played by the function $f_n : S \to [0, 1]$ given by

$$f_n(\mathbf{x}) = \prod_{j=1}^l \left(\min\left\{ 1, \left(\frac{\sum\limits_{i=1}^n h_{a_j, \gamma}(x_i)}{(b_j - \frac{\gamma}{2})n} \right)^n \right\} \cdot \min\left\{ 1, \left(\frac{(b_j + \frac{\gamma}{2})n}{1 + \sum\limits_{i=1}^n t_{a_j, \gamma}(x_i)} \right)^n \right\} \right).$$

Denote $\mathcal{A}^*(\mathbf{a}_l, \mathbf{b}_l, \gamma)$ the set of all $\mathbf{x} \in S$ for which there exists the limit $\lim_{n \to \infty} f_n(\mathbf{x})$. Similarly as before, $\mathcal{A}(\mathbf{a}_l, \mathbf{b}_l, \gamma) \subset \mathcal{A}^*(\mathbf{a}_l, \mathbf{b}_l, \gamma)$ and put $f(\mathbf{x}) = \lim_{n \to \infty} f_n(\mathbf{x})$ for $\mathbf{x} \in \mathcal{A}^*(\mathbf{a}_l, \mathbf{b}_l, \gamma)$.

In case (b) (1) we prove that f is discontinuous in any $\mathbf{y} \in \mathcal{A}^*(\mathbf{a}_l, \mathbf{b}_l, \gamma)$, where $f(\mathbf{y}) < 1$. Let a positive $\varepsilon < 1 - f(\mathbf{y})$ be given. For given $\delta > 0$ we choose m by the same way. Choose a d.f. $g(x) \in \mathcal{G}$ which is continuous in points $x = a_j$ and $g(a_j) = b_j$ (for j = 1, 2, ..., l. Then there exists a sequence $z \in S$ for that $G(z) = \{g(x)\}$. For simplicity, denote by $a_0 = 0$ and $a_{l+1} = 1$. Define the sequence \mathbf{x} in the following way:

$$x_k = \begin{cases} y_k, & \text{if } k \le m, \\ \frac{a_{j-1}+a_j}{2} & \text{if } k > m \text{ and } z_k \in [a_{j-1}, a_j), \ j = 1, 2, \dots, l+1 \\ 1, & \text{if } k > m, \text{ and } z_k = 1. \end{cases}$$

Then $\frac{1}{n}\sum_{i=1}^{n} \chi_{[0,a_j)}(x_i)$, $\frac{1}{n}\sum_{i=1}^{n} h_{a_j,\gamma}(x_i)$ and $\frac{1}{n}\sum_{i=1}^{n} t_{a_j,\gamma}(x_i)$ tend to b_j as $n \to \infty$ (j = 1, 2, ..., l). So, $f_n(\mathbf{x}) = 1$ for all sufficiently large n. Therefore $f(\mathbf{x}) = \lim_{n \to \infty} f_n(\mathbf{x}) = 1$.

In case (b) (2) we have $g(\mathbf{y}) = 1$. Let δ , *m*, **x** have the previous meaning. Put

$$x_{k} = \begin{cases} y_{k}, & \text{if } k \le m, \\ \frac{a_{l}+1}{2}, & \text{if } k > m \text{ and } b_{1} \ge \frac{1}{2} \\ \frac{a_{1}}{2}, & \text{if } k > m \text{ and } b_{1} < \frac{1}{2}. \end{cases}$$

Then, clearly $\rho(\mathbf{x}, \mathbf{y}) < \delta$, and for sufficiently large *n* at least one of the inequalities (2.2) have to be true. So, we have $f(\mathbf{x}) = \lim_{n \to \infty} f_n(\mathbf{x}) = 0$, and therefore $f(\mathbf{y}) - f(\mathbf{x}) = 1 - 0 > 0$. The rest of the proof follows by the same way as the proof of Lemma 2.1.

Theorem 2.2. Let \mathcal{H} be a subset of S with the property: if $\mathbf{x} \in \mathcal{H}$ than for any positive integer l and arbitrary finite rational sequences \mathbf{a}_l and \mathbf{b}_l with the properties $0 < a_1 < a_2 \cdots < a_l < 1$ and $0 < b_1 \le b_2 \cdots \le b_l \le 1$, there is a d.f. $g(\mathbf{x}) \in G(\mathbf{x})$ for that $g(a_j) = b_j$ (j = 1, 2, ..., l). Then \mathcal{H} is residual in S.

Proof. If we take unions of the sets $\mathcal{A}(\mathbf{a}_l, \mathbf{b}_l, \gamma)$ for all positive integers l = 1, 2, ..., for all rational numbers $a_1, ..., a_l$, $b_1, ..., b_l$, $\gamma = \frac{1}{n}$ (n = 1, 2...) satisfying the necessary conditions, we get countable union of meager sets, which is still set of first Baire category in *S*. The complement of this set is residual in *S*.

Theorem 2.3. Let $\mathcal{M} = \{\mathbf{x} \in S \mid G(\mathbf{x}) = \mathcal{G}\}$. Then the set of sequences \mathcal{M} is residual in S.

Proof. Let us consider a sequence $\mathbf{x} \in \mathcal{H}$ and a d.f. $g(x) \in \mathcal{G}$. As g(x) is monotone on [0, 1], then it is Riemann integrable. Thus, in the sense of (1.1) we can approximate g(x) with arbitrary precision with d.f.s from $G(\mathbf{x})$ which have positive rational function values in points of equidistant partition of the unit interval. It means, that $\mathbf{x} \in \mathcal{M}$ and the assertion follows.

Problem 2.1. In [5] it was proved that a typical (in the sense of Baire) point x has the following property: the all higher order Cesàro averages of digits of x have all distribution functions for discrete random variables that takes possible values of the digits. Let us denote by $C(\mathbf{x})$ the Cesàro average of the sequence \mathbf{x} . It seems to be interesting to ask whether the Cesàro average of a typical sequence has all distribution function? More precisely, is the set of sequences

$$\{\mathbf{x}\in S\,|\,G(C(\mathbf{x}))=\mathcal{G}\}$$

also residual in S?

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- J. Bukor, J. T. Tóth, On Topological Properties of the Set of Maldistributed Sequences, Acta Univ. Sapient. Math. 12 (2020), 272–279. https://doi.org/10.2478/ausm-2020-0018.
- [2] P. Kostyrko, M. Mačaj, T. Šalát, O. Strauch, On Statistical Limit Points, Proc. Amer. Math. Soc. 129 (2000), 2647–2654. https://doi.org/10.1090/s0002-9939-00-05891-3.
- [3] V. László, T. Šalát, The Structure of Some Sequence Spaces, and Uniform Distribution (Mod 1), Period Math. Hung. 10 (1979), 89–98. https://doi.org/10.1007/bf02018376.
- [4] L. Olsen, Extremely Non-Normal Numbers, Math. Proc. Camb. Phil. Soc. 137 (2004), 43–53. https://doi.org/10.1017/ s0305004104007601.
- [5] J. Hyde, V. Laschos, L. Olsen, T. Petrykiewicz, A. Shaw, Iterated Cesàro Averages, Frequencies of Digits, and Baire Category, Acta Arith. 144 (2010), 287–293. https://doi.org/10.4064/aa144-3-6.
- [6] J.C. Oxtoby, Measure and Category, Springer, New York, 1996.
- [7] T. Šalát, A Remark on Normal Numbers, Rev. Roumaine Math. Pures Appl. 11 (1966), 53–56.
- [8] T. Šalát, On Statistically Convergent Sequences of Real Numbers, Math. Slovaca, 30 (1980), 139–150. http://dml.cz/ dmlcz/136236.
- [9] O. Strauch, Uniformly Maldistributed Sequences in a Strict Sense, Monatsh. Math. 120 (1995), 153–164. https://doi.org/10.1007/bf01585915.
- [10] O. Strauch, Distribution of Sequences: A Theory, VEDA, Bratislava; Academia, Prague, 2019.
- [11] O. Strauch, Š. Porubský, Distribution of Sequences: A Sampler, Peter Lang, Frankfurt am Main, 2005.