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

József Bukor\textsuperscript{1,∗}, Kálmán Liptai\textsuperscript{2}, János T. Tóth\textsuperscript{2}

\textsuperscript{1}Department of Informatics, J. Selye University, Komárno, Slovakia
\textsuperscript{2}Department 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

$$z = \frac{d_1(z)}{s} + \frac{d_2(z)}{s^2} + \cdots + \frac{d_n(z)}{s^n} + \cdots$$

with $d_i(z) \in \{0, 1, \ldots, s-1\}$.

For each digit $i \in \{0, 1, \ldots, s-1\}$ 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 $i \leq x$ among the first $n$ digits of $z$ as

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

Received: Mar. 19, 2024.
2020 Mathematics Subject Classification. 54E52.
Key words and phrases. Baire category; distribution function; sequences.

https://doi.org/10.28924/2291-8639-22-2024-72
ISSN: 2291-8639 © 2024 the author(s).
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 

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

where $x = (x_k), y = (y_k)$. It is known that $(S, \rho)$ 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 $x = (x_n)$ be a sequence from unit interval $[0, 1]$.
- Let $\chi_A(x)$ denote the characteristic function of the set $A$.
- Denote by 

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

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

- A non-decreasing function $g : [0, 1] \to [0, 1]$, $g(0) = 0$, $g(1) = 1$ 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 $G(x)$ denote the set of all d.f.s of $x$.

- $c_\gamma(x)$ 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 \( G(x) \) is closed and \( G(x) \) is connected in the weak topology defined by the metric

\[
d(g_1, g_2) = \left( \int_0^1 (g_1(x) - g_2(x))^2 \, dx \right)^{\frac{1}{2}}.
\]

For given a non-empty set \( H \) of d.f.s there exists a sequence \( x \) in \([0, 1]\) such that \( G(x) = H \) 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

\[
\lim_{n \to \infty} g_{k_n}(x) = g(x) \quad \text{for every } x \in [0, 1].
\]

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 \( \gamma < \min\{b, \frac{a}{4}, 1 - \frac{a}{4}\} \) denote by \( \mathcal{A}(a, b, \gamma) \) the set of all \( x = (x_k) \in S \) for which there is an \( n_0 \) such that for any \( n \geq n_0 \) we have

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

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

**Proof.** We define continuous functions \( h_{a,\gamma}(x) : [0, 1] \to [0, 1] \) and \( t_{a,\gamma} : [0, 1] \to [0, 1] \) by

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

see Figure 1.

![Figure 1. Functions \( h_{a,\gamma}(x) \) and \( t_{a,\gamma}(x) \)](image-url)

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 \( x \in S \) and fixed \( n \) the function \( f_n : S \to [0, 1] \) in the following way:
\[ f_n(x) = \min \left\{ 1, \left( \frac{\sum_{i=1}^{n} h_{a,y}(x_i)}{(b - \frac{\gamma}{2})n} \right)^{\frac{1}{n}}, \frac{(b + \frac{\gamma}{2})n}{1 + \sum_{i=1}^{n} t_{a,y}(x_i)} \right\}. \]

Denote \( \mathcal{A}^* (a, b, \gamma) \) the set of all \( x \in S \) for which there exists the limit \( \lim_{n \to \infty} f_n(x) \).

One can easily check that if (2.1) holds for all sufficiently large \( n \), then \( f_n(x) \to 0 \) for \( n \to \infty \). Therefore \( \mathcal{A}(a, b, \gamma) \subset \mathcal{A}^* (a, b, \gamma) \).

Put \( f(x) = \lim_{n \to \infty} f_n(x) \) for \( x \in \mathcal{A}^* (a, b, \gamma) \). We shall prove that

(a) the function \( f_n \) \( (n = 1, 2, \ldots) \) is a continuous function on \( S \),

(b) \( f \) is discontinuous at each point of \( \mathcal{A}^* (a, b, \gamma) \).

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

(b) Let \( y = (y_k) \in \mathcal{A}^* (a, b, \gamma) \). We have the following two possibilities.

(1) \( f(y) < 1 \),
(2) \( f(y) = 1 \).

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

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 \( \mathcal{G}(z) = \{ g(x) \} \).

Define the sequence \( x \) in the following way:

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

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

\[ f(x) - f(y) = 1 - f(y) > \epsilon. \]

In case (2) we have \( g(y) = 1 \). Let \( \delta, m, x \) have the previous meaning. Put

\[ x_k = \begin{cases} 
    y_k, & \text{if } k \leq m, \\
    \frac{a+1}{2}, & \text{if } k > m \text{ and } b \geq \frac{1}{2}, \\
    \frac{a}{2}, & \text{if } k > m \text{ and } b < \frac{1}{2}.
\end{cases} \]
Then, clearly $\rho(x, y) < \delta$, and for sufficiently large $n$ one of the inequalities (2.1) must be true. So, we have $f(x) = \lim_{n \to \infty} f_n(x) = 0$, and therefore $f(y) - f(x) = 1 - 0 > 0$. Hence the discontinuity of $f$ at $y \in \mathcal{A}'(I, \gamma)$ 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

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

is residual in $S$.

**Proof.** By Lemma 2.1 we have that the infinite union $\bigcup_{n=0}^{\infty} \mathcal{A}(a, b, \frac{1}{n})$ (where $\frac{1}{n} < \min[\frac{n}{4}, \frac{1-a}{4}]$) 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

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

for any $\gamma = \frac{1}{n} (n \geq n_0)$ and infinitely many $N$. In this case, First Helly theorem implies that there exists a pointwise convergent subsequence with limit $g(x) \in G(x)$ for that $g(a) = b$. So, $x \in \mathcal{P}$ 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 $a_i$ for a finite sequence $(a_k)_{k=1}^{l}$ and the notation $b_i$ 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 $a_i$ and $b_i$ are given, where $0 < a_1 < a_2 \cdots < a_l < 1$ and $0 < b_1 \leq b_2 \cdots \leq b_l \leq 1$. For a positive number

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

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

$$\sum_{i=1}^{n} x(0, a_i - \gamma)(x_i) < (b_i - \gamma)n, \quad \sum_{i=1}^{n} x(0, a_i + \gamma)(x_i) > (b_i + \gamma)n$$

hold ($j = 1, 2, \ldots, n$). Then $\mathcal{A}(a_i, b_i, \gamma)$ 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

\[
f_n(x) = \prod_{j=1}^{l} \left( \min \left\{ 1, \left( \frac{\sum_{i=1}^{n} h_{a_j, \gamma}(x_i)}{(b_j - \frac{\gamma}{2}) n} \right)^n \right\} \right).
\]

Denote \( \mathcal{A}^*(a_l, b_l, \gamma) \) the set of all \( x \in S \) for which there exists the limit \( \lim_{n \to \infty} f_n(x) \). Similarly as before, \( \mathcal{A}^*(a_l, b_l, \gamma) \subset \mathcal{A}^*(a_l, b_l, \gamma) \) and put \( f(x) = \lim_{n \to \infty} f_n(x) \) for \( x \in \mathcal{A}^*(a_l, b_l, \gamma) \). In case (b) (1) we prove that the function is discontinuous in any \( y \in \mathcal{A}^*(a_l, b_l, \gamma) \) where \( f(y) < 1 \). Let a positive \( \varepsilon < 1 - f(y) \) be given. For given \( \delta > 0 \) we choose \( m \) by the same way. Choose a d.f. \( g(x) \in G \) which is continuous in points \( x = a_j \) and \( g(a_j) = b_j \) for \( j = 1, 2, \ldots, l \). 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 \( x \) in the following way:

\[
x_k = \begin{cases} y_{k_{r}}, & \text{if } k \leq m, \\
\frac{a_{j+1} + a_j}{2}, & \text{if } k > m \text{ and } z_k \in (a_{j-1}, a_j), \ j = 1, 2, \ldots, l + 1 \\
1, & \text{if } k > m \text{ and } z_k = 1.
\end{cases}
\]

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

In case (b) (2) we have \( g(y) = 1 \). Let \( \delta, m, x \) have the previous meaning. Put

\[
x_k = \begin{cases} y_{k_{r}}, & \text{if } k \leq m, \\
\frac{a_{j+1} + a_j}{2}, & \text{if } k > m \text{ and } b_1 \geq \frac{1}{2} \\
\frac{a_j}{2}, & \text{if } k > m \text{ and } b_1 < \frac{1}{2}.
\end{cases}
\]

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

\[\Box\]

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

**Proof.** If we take unions of the sets \( \mathcal{A}(a_l, b_l, \gamma) \) for all positive integers \( l = 1, 2, \ldots, \), for all rational numbers \( a_1, \ldots, a_l, b_1, \ldots, b_l, \gamma = \frac{1}{n} \) \( (n = 1, 2, \ldots) \) 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 \( S \).

\[\Box\]

**Theorem 2.3.** Let \( M = \{x \in S | G(x) = G\} \). Then the set of sequences \( M \) is residual in \( S \).
Proof. Let us consider a sequence \( x \in \mathcal{H} \) 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 \( \mathcal{G}(x) \) which have positive rational function values in points of equidistant partition of the unit interval. It means, that \( x \in \mathcal{M} \) and the assertion follows. \( \square \)

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 \( C(x) \) the Cesàro average of the sequence \( x \). 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

\[ \{ x \in S | G(C(x)) = \mathcal{G} \} \]

also residual in \( S \)?

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

References


