Ruibiao Zou

Hausdorff dimension of the maximal run-length in dyadic expansion

Czechoslovak Mathematical Journal, Vol. 61 (2011), No. 4, 881–888

Persistent URL: http://dml.cz/dmlcz/141793

Czechoslovak Mathematical Journal, 61 (136) (2011), 881–888

HAUSDORFF DIMENSION OF THE MAXIMAL RUN-LENGTH IN DYADIC EXPANSION

Ruibiao Zou, Changsha

(Received April 6, 2010)

Abstract. For any $x \in [0,1)$, let $x = [\varepsilon_1, \varepsilon_2, \ldots,]$ be its dyadic expansion. Call $r_n(x) := \max\{j \geqslant 1\colon \varepsilon_{i+1} = \ldots = \varepsilon_{i+j} = 1,\ 0 \leqslant i \leqslant n-j\}$ the n-th maximal run-length function of x. P. Erdös and A. Rényi showed that $\lim\limits_{n\to\infty} r_n(x)/\log_2 n = 1$ almost surely. This paper is concentrated on the points violating the above law. The size of sets of points, whose run-length function assumes on other possible asymptotic behaviors than $\log_2 n$, is quantified by their Hausdorff dimension.

Keywords: run-length function, Hausdorff dimension, dyadic expansion

MSC 2010: 11K55, 28A78, 28A80

1. Introduction

Let $\mathbf{X}^{(k)}(t) = (X_1(t), \ldots, X_k(t))$ denote a k-vector of i.i.d. random variables, each taking the values 1 or 0 with respective probabilities p and $1-p$. A lot of classical results in probability theory, for instance the strong law of large numbers, the law of iterated logarithm, and so on, concern almost-sure properties of sequences $\{X_n\}$ of i.i.d. random variables. As a process indexed by non-negative t, I. Benjamini et al. proved that $\mathbf{X}^{(k)}(t)$ is strong Markov with invariant measure $((1-p)\delta_0 + p\delta_1)^k$. For the dynamical walk $S_n(t) = X_1(t) + \ldots + X_n(t)$ ($t \geqslant 0$, $n \geqslant 1$), they proved that the law of large numbers and the law of iterated logarithm are dynamically stable while run tests are dynamically sensitive; also, they obtain multi-fractal analysis of exceptional times for run lengths and for prediction [2]. Subsequently, Davar Khoshnevisan et al. showed that in the case that $X_i(0)$'s are standard normal, the classical integer test is not dynamically stable [4]. Then in [5], they extended a result

This work is supported by the HNSN under 09JJ3001, HNKSPP (2010J05) and HNED under 11C0671.

of [2] by proving that if $X_i(0)$'s are lattice, mean-zero and variance-one, and process $(2+\varepsilon)$ finite absolute moments for some $\varepsilon > 0$, then the recurrence of the origin is dynamically stable. Also, they studied some properties of the set of times t when $n \mapsto S_n(t)$ exceeds a given envelope infinitely often, they proved that the infinite-dimensional process $t \mapsto S_{\lfloor n\bullet \rfloor}(t)/\sqrt{n}$ converges weakly in $\mathcal{D}[0,1]$. At the same time, the Bescovitch-Hausdorff dimension of the of set of those points which violate the corresponding law of the iterated logrithm were investigated. In [6], D. Khoshnevisan, D. A. Levin estimated the probability that $X_1(t)+\ldots+X_k(t) = k-l$ for some $t \in F$, where $F \subseteq [0,1]$ is nonrandom and compact.

The run-length function r_n was introduced for the first time in a mathematical experiment of cion tossing, which measures the length of consecutive terms of 'heads' in n times' experiment. The run-length function has been extensively studied and used in probability theory and other subjects, such as in the DNA string machine [1]. For a brief introduction of the run-length function, one can refer to P. Révész's book [8] and references therein.

It is also well known that every $x \in [0,1)$ corresponds to a unique infinite sequence $[\varepsilon_1, \varepsilon_2, \ldots]$ with $\varepsilon_n \in \{0,1\}$ for all $n \geqslant 1$ and $\varepsilon_n = 0$ for infinitely many n's, in the sense that

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_n}{2^n}$$

is the dyadic expansion of x. Naturally, the maximal run-length function $r_n(x)$, for $x \in [0,1)$, can be defined as the length of the longest run of 1's in $[\varepsilon_1(x), \ldots, \varepsilon_n(x)]$, that is

$$r_n(x) = \max\{j \geqslant 1\colon\ \varepsilon_{i+1} = \ldots = \varepsilon_{i+j} = 1,\ 0 \leqslant i \leqslant n-j\}.$$

For the asymptotic behavior of r_n, P. Erdös and A. Rényi showed that, almost surely,

$$\lim_{n\to\infty} \frac{r_n(x)}{\log_2 n} = 1.$$

Nevertheless, the points that violate the above law are visible, in the sense that they carry full Hausdorff dimension [7]. But the above results provide no information about whether there exist points whose run-length function can obey other asymptotic behavior than $\log_2 n$. This motivates us to investigate the set of points with other given asymptotic characters of their run-length function.

Given a nondecreasing integer sequence $\{\delta_n\}_{n=1}^{\infty}$, set

$$E(\{\delta_n\}_{n=1}^{\infty}) = \Big\{x \in [0,1)\colon\ \lim_{n\to\infty} \frac{r_n(x)}{\delta_n} = 1\Big\},$$
$$F(\{\delta_n\}_{n=1}^{\infty}) = \Big\{x \in [0,1)\colon\ \limsup_{n\to\infty} \frac{r_n(x)}{\delta_n} = 1\Big\}.$$

It is natural to ask whether $E(\{\delta_n\}_{n=1}^{\infty})$ and $F(\{\delta_n\}_{n=1}^{\infty})$ are always nonempty. Unexpectedly, it is not the case for $E(\{\delta_n\}_{n=1}^{\infty})$, even if $\{\delta_n\}_{n=1}^{\infty}$ satisfies $0 \leqslant \delta_{n+1} - \delta_n \leqslant 1$ for all $n \geqslant 1$ (See Section 2). So, to guarantee $E(\{\delta_n\}_{n=1}^{\infty}) \neq \emptyset$, some extra conditions must be assumed on $\{\delta_n\}_{n=1}^{\infty}$.

Since the sets in question are all of null Lebesgue measure, Hausdorff dimension is used to quantify their size. In this note, we in particular prove

Theorem 1.1. *Let $\{\delta_n\}_{n=1}^{\infty}$ be a nondecreasing integer sequence with $\delta_n \to \infty$ as $n \to \infty$ and $\lim\limits_{n\to\infty} \delta_{n+\delta_n}/\delta_n = 1$. Then $\dim_H E(\{\delta_n\}_{n=1}^{\infty}) = 1$.*

Theorem 1.2. *Let $\{\delta_n\}_{n=1}^{\infty}$ be an integer sequence with $\delta_n \to \infty$ as $n \to \infty$. Then $\dim_H F(\{\delta_n\}_{n=1}^{\infty}) = \max\{0, 1 - \liminf\limits_{n\to\infty} \delta_n/n\}$.*

At the end, we give some examples of $\{\delta_n\}_{n=1}^{\infty}$ which can fulfil the assumptions of Theorem 1.1:

- $\delta_n = \beta (\log n)^{\gamma}$, $\beta > 0$, $\gamma > 0$,
- $\delta_n = \beta n^{\gamma}$, $\beta > 0$, $0 < \gamma < 1$,
- $\delta_n = \beta n/(\log n)^{\gamma}$, $\beta > 0$, $\gamma > 0$.

We also note that in the set $E(\{\delta_n\}_{n=1}^{\infty})$, δ_n cannot take a large value such as $\delta_n = n$ (see Proposition 2.2). The paper is organized as follows. In Section 2, some intrinsic properties on r_n are established, which will give reasons for the assumption on δ_n in Theorem 1.1. Section 3 and 4 are devoted to presenting Theorem 1.1 and Theorem 1.2 respectively.

2. Properties on run-length function

In this section, an intrinsic property shared by the run-length function is presented. We will see that the assumption in Theorem 1.1 has close relations to this essential feature of r_n. Evidence is also given indicating that not all sequences can serve as the asymptotic function of the run-length function.

Proposition 2.1. *For any $x \in [0,1)$, $r_{n+r_n(x)}(x) = r_n(x)$ holds for infinitely many n's. Consequently,*

$$\liminf_{n\to\infty} \frac{r_{n+r_n}}{r_n} = 1. \tag{2.1}$$

P r o o f. For any $x \in [0,1)$, write $r_n = r_n(x)$ for brevity. By the requirement of uniqueness of the dyadic expansion, we know that $\varepsilon_n(x) = 0$ for infinitely many n's.

However, when $\varepsilon_n(x) = 0$, then

$$r_{n+r_n} = \max\{r_n(\varepsilon_1, \ldots, \varepsilon_n), r_{r_n}(\varepsilon_{n+1}, \ldots, \varepsilon_{n+r_n})\} = \max\{r_n, r_n\} = r_n.$$

Thus we have, for any $x \in [0, 1)$, $r_{n+r_n} = r_n$ for infinitely many n's. □

Proposition 2.2. *For any* $0 < \beta \leqslant 1$,

$$\widetilde{E}(\beta) := \Big\{x \in [0, 1)\colon \lim_{n\to\infty} \frac{r_n(x)}{n} = \beta\Big\} = \emptyset.$$

Proof. (i) $\beta = 1$. For any $x \in \widetilde{E}(\beta)$ and $0 < \varepsilon < 1/4$, there exists $N \geqslant 2$ such that for any $n \geqslant N$, $r_n(x) > (1 - \varepsilon)n$. We will show that $\varepsilon_n(x) = 1$ for all $n \geqslant N$. If this is not the case, we assume that $\varepsilon_n(x) = 0$, then $r_{2n}(x) \leqslant n$. This leads to a contradiction. Since there are infinitely many 0's in the expansion of each $x \in [0, 1)$, we have $\widetilde{E}(\beta) = \emptyset$.

(ii) $0 < \beta < 1$. Let $k = \frac{1}{2}(\frac{1}{1-\beta} + 1)$ and $\varepsilon < \min\{\frac{(k-1)\beta}{k+1}, \frac{\beta(1-\beta)}{2-\beta}\}$, which gives

$$k(\beta - \varepsilon) > \beta + \varepsilon \quad \text{and} \quad k - 1 < k(\beta - \varepsilon).$$

For any $x \in \widetilde{E}(\beta)$, there exists $N \in \mathbb{N}$ such that for any $n \geqslant N$,

$$(\beta - \varepsilon)(n + 1) < r_n(x) < (\beta + \varepsilon)n.$$

We claim that $\varepsilon_n(x) = 1$ for all $n \geqslant N$. If this is not the case for some $n \geqslant N$, then

$$\begin{aligned} r_{[kn]} &= \max\{r_n(\varepsilon_1, \ldots, \varepsilon_n),\ r_{[kn]-n}(\varepsilon_{n+1}, \ldots, \varepsilon_{[kn]})\} \\ &\leqslant \max\{(\beta + \varepsilon)n, kn - n\} < (\beta - \varepsilon)kn < (\beta - \varepsilon)([kn] + 1), \end{aligned}$$

which leads to a contradiction. So, we get $\widetilde{E}(\beta) = \emptyset$. □

3. Proof of theorem 1.2

Recall that

$$F(\{\delta_n\}_{n=1}^{\infty}) = \Big\{x \in [0, 1)\colon \limsup_{n\to\infty} \frac{r_n(x)}{\delta_n} = 1\Big\},$$

where $\{\delta_n\}_{n=1}^{\infty}$ is an integer sequence with $\delta_n \to \infty$ as $n \to \infty$. Write $\beta = \liminf_{n\to\infty} \delta_n/n$ for simplicity.

Lemma 3.1. $\dim_H F(\{\delta_n\}_{n=1}^{\infty}) \leqslant \max\{0, 1-\beta\}$.

P r o o f. When $\beta > 1$, then $F(\{\delta_n\}_{n=1}^{\infty}) = \emptyset$. So we restrict ourselves to $0 \leqslant \beta \leqslant 1$. To get the desired result, it suffices to show that, for any $\varepsilon > 0$ and $s > 1-(1-\varepsilon)\beta$, $\dim_H F(\{\delta_n\}_{n=1}^{\infty}) \leqslant s$.

Note that, for any $\varepsilon > 0$,

$$F(\{\delta_n\}_{n=1}^{\infty}) \subset \{x \in [0,1)\colon\ r_n(x) \geqslant (1-\varepsilon)\delta_n,\ \text{i.o. } n\}.$$

So, for each $N \geqslant 1$,

$$\bigcup_{n \geqslant N} \bigcup_{(\varepsilon_1,\ldots,\varepsilon_n) \in D_n(\varepsilon)} I_n(\varepsilon_1,\ldots,\varepsilon_n)$$

is a cover of $F(\{\delta_n\}_{n=1}^{\infty})$, where

$$D_n(\varepsilon) = \{(\varepsilon_1,\ldots,\varepsilon_n) \in \{0,1\}^n\colon\ r_n(\varepsilon_1,\ldots,\varepsilon_n) \geqslant (1-\varepsilon)\delta_n\}.$$

Then for any $s > 1-(1-\varepsilon)\beta$,

$$\begin{aligned} H^s(F(\{\delta_n\}_{n=1}^{\infty})) &\leqslant \liminf_{N\to\infty} \sum_{n=N}^{\infty} \sum_{(\varepsilon_1,\ldots,\varepsilon_n) \in D_n(\varepsilon)} |I_n(\varepsilon_1,\ldots,\varepsilon_n)|^s \\ &= \liminf_{N\to\infty} \sum_{n=N}^{\infty} \sharp D_n(\varepsilon) \frac{1}{2^{ns}} \\ &\leqslant \liminf_{N\to\infty} \sum_{n=N}^{\infty} n 2^{n-(1-\varepsilon)\delta_n} \frac{1}{2^{ns}} \leqslant 1, \end{aligned}$$

where the last assertion follows from the fact that whenever $s > 1-(1-\varepsilon)\beta$, then $1-(1-\varepsilon)\delta_n/n < s$ for all n large enough. Hence $\dim_H F(\{\delta_n\}_{n=1}^{\infty}) \leqslant s$. $\square$

Lemma 3.2.

$$\dim_H F(\{\delta_n\}_{n=1}^{\infty}) = \begin{cases} 0, & \textit{when } \beta = 1; \\ 1, & \textit{when } \beta = 0. \end{cases}$$

P r o o f. The first assertion follows from Lemma 3.1. When $\beta = 0$, note that

$$\Big\{x \in [0,1)\colon\ \sup_{n \geqslant 1} r_n(x) < \infty\Big\} \subset F(\{\delta_n\}_{n=1}^{\infty}).$$

For any $M \geqslant 3$, set

$$\mathcal{F} = \Big\{f_{\varepsilon_2,\ldots,\varepsilon_{M-1}}(x) = \sum_{n=2}^{M-1} \frac{\varepsilon_n}{2^n} + \frac{x}{2^M},\ \varepsilon_n \in \{0,1\},\ 1 < n < M\Big\}.$$

Let F_M be the attractor of the self-similar IFS $\mathcal{F}$. It is easy to see that

$$\dim_H F_M = \frac{\log 2^{M-2}}{\log 2^M} = \frac{M-2}{M}.$$

Evidently, $F_M \subset \{x \in [0,1) \colon \sup\limits_{n\geqslant 1} r_n(x) < \infty\}$. □

In the sequel, we restrict ourselves to $0 < \beta < 1$. Let β_k be a sequence of rationals decreasing to β. Choose a subsequence N_k of $\mathbb{N}$ satisfying, for each $k \geqslant 1$,

$$N_k \geqslant \frac{8}{\beta_k^2}, \quad N_{k+1} \geqslant (k+1)N_k, \quad \lim_{k\to\infty} \frac{\delta_{N_k}}{N_k} = \beta,$$
$$\beta_k \cdot N_k \in \mathbb{N}, \quad t_k := \frac{N_{k+1} - \beta_{k+1}N_{k+1} - N_k}{\beta_k N_k} \in \mathbb{N}.$$

Set

$$\mathcal{L} = \{N_k + j_k\beta_k N_k,\ 0 \leqslant j_k < t_k,\ \text{and } N_{k+1} - \beta_{k+1}N_{k+1} + 1,$$
$$N_{k+1} - \beta_{k+1}N_{k+1} + 2, \ldots, N_{k+1} - 1,\ k \geqslant 1\}.$$

Define a sequence $\{a_n\}_{n\in\mathcal{L}}$ given as follows. When $i \leqslant N_1$, set $a_i = 0$. When $k \geqslant 1$ and $0 \leqslant j_k \leqslant t_k$, set

$$a_{N_k + j_k\beta_k N_k} = 0,\ a_{N_{k+1}-\beta_{k+1}N_{k+1}+1} = \ldots = a_{N_{k+1}-1} = 1.$$

For any $n \geqslant 1$, define

$$D_n = \{(\varepsilon_1, \ldots, \varepsilon_n) \in \{0,1\}^n \colon\ \varepsilon_k = a_k,\ \text{for } k \in \mathcal{L}\ \text{ and } 1 \leqslant k \leqslant n\}.$$

Define

$$E = \bigcap_{n=1}^{\infty} \bigcup_{(\varepsilon_1,\ldots,\varepsilon_n)\in D_n} I_n(\varepsilon_1, \ldots, \varepsilon_n).$$

Proposition 3.1. $E \subset F(\{\delta_n\}_{n=1}^{\infty})$.

Proof. Fix $x \in E$. For any $n \geqslant N_1$, let $k \geqslant 1$ be the integer such that $N_k \leqslant n < N_{k+1}$.

Case (i). $N_k \leqslant n < N_{k+1} - \beta_{k+1}N_{k+1}$. In this case, $r_n(x) = \beta_k N_k - 1$. Thus,

$$\frac{r_n(x)}{\delta_n} \leqslant \frac{\beta_k N_k - 1}{\delta_{N_k}}.$$

C a s e (ii). $N_{k+1} - \beta_{k+1} N_{k+1} \leqslant n < N_{k+1}$. Thus by the definition of E, we have $r_n(x) = \max\{\beta_k N_k - 1, n - N_{k+1} + \beta_{k+1} N_{k+1}\}$. Thus

$$\begin{aligned}\frac{r_n(x)}{\delta_n} &\leqslant \max\left\{\frac{\beta_k N_k - 1}{\delta_{N_k}}, \frac{n - N_{k+1} + \beta_{k+1} N_{k+1}}{n} \frac{n}{\delta_n}\right\}\\ &\leqslant \max\left\{\frac{\beta_k N_k - 1}{\delta_{N_k}}, \frac{N_{k+1} - N_{k+1} + \beta_{k+1} N_{k+1}}{N_{k+1}} \frac{n}{\delta_n}\right\}.\end{aligned}$$

Thus, in general, for any $x \in E$, we have $\limsup\limits_{n\to\infty} r_n(x)/\delta_n \leqslant 1$.

While, on the other hand, for any $x \in E$ and $k \geqslant 2$ we have $r_{N_k}(x) = \beta_k N_k - 1$, thus, $\limsup\limits_{n\to\infty} r_n(x)/\delta_n \geqslant 1$. □

Lemma 3.3. $\dim_H E = 1 - \beta$.

P r o o f. We show $\dim_H E \geqslant 1 - \beta$ only. First define a mass distribution supported on E. For any $n \geqslant 1$ and $(\varepsilon_1, \ldots, \varepsilon_n) \in D_n$, set

$$\mu(I(\varepsilon_1, \ldots, \varepsilon_n)) = \frac{1}{\sharp D_n}.$$

Then by Kolomogrov's consistency theorem, μ can be extended to a probability measure supported on E. In what follows, we estimate the measure $\mu(I_n(x))$ for any $x \in E$. Assume that $N_k \leqslant n < N_{k+1}$.

C a s e (i). $N_k + j_k \beta_k N_k \leqslant n < N_k + (j_k + 1)\beta_k N_k$. In this case,

$$\mu(I_n(x)) = \left(\prod_{i=1}^{k-1} 2^{N_{i+1} - \beta_{i+1} N_{j+1} - N_i - t_i} \cdot 2^{n - N_k - j_k}\right)^{-1}.$$

Thus,

$$\begin{aligned}\frac{\log \mu(I_n(x))}{-n \log 2} &\geqslant \frac{n - N_k - j_k + \sum_{i=1}^{k-1}(N_{i+1} - \beta_{i+1} N_{i+1} - N_i - t_i)}{n}\\ &\geqslant 1 - \frac{N_k + j_k - \sum_{i=1}^{k-1}(N_{i+1} - \beta_{i+1} N_{i+1} - N_i - t_i)}{N_k + j_k \beta_k N_k}\\ &\geqslant 1 - \frac{N_k - \sum_{i=1}^{k-1}(N_{i+1} - \beta_{i+1} N_{i+1} - N_i - t_i)}{N_k}\\ &\qquad\qquad \text{(increasing with respect to } j_k\text{)}\\ &\to 1 - \beta, \quad \text{as } k \to \infty.\end{aligned}$$

C a s e (ii). $N_{k+1} - \beta_{k+1} N_{k+1} \leqslant n < N_{k+1}$. In this case,

$$\mu(I_n(x)) = \left(\prod_{i=1}^{k} 2^{N_{i+1} - \beta_{i+1} N_{i+1} - N_i - t_i}\right)^{-1}.$$

Thus,

$$\begin{aligned}\frac{\log\mu(I_n(x))}{-n\log 2} &= \frac{\sum_{i=1}^{k}(N_{i+1}-\beta_{i+1}N_{i+1}-N_i-t_i)}{n}\\ &\geqslant \frac{\sum_{i=1}^{k}(N_{i+1}-\beta_{i+1}N_{i+1}-N_i-t_i)}{N_{k+1}}\\ &\to 1-\beta, \quad \text{as } k\to\infty.\end{aligned}$$

In general, we have

$$\liminf_{n\to\infty}\frac{\log\mu(I_n(x))}{\log|I_n(x)|}\geqslant 1-\beta.$$

An application of Billingsley' Theorem (see [3], p. 141, Theorem 14.1) yields $\dim_H E \geqslant 1-\beta$.

Acknowledgment. The author expresses thanks to the editor for helpful and detailed suggestions.

References

[1] *R. Arratia, L. Gordon, M. S. Waterman*: The Erdös-Rényi law in distribution, for coin tossing and sequence matching. Ann. Stat. *18* (1990), 539–570.
[2] *I. Benjamini, O. Häggström, Y. Peres, J. E. Steif*: Which properties of a random sequence are dynamically sensitive? Ann. Probab. *31* (2003), 1–34.
[3] *P. Billingsley*: Ergodic Theory and Information,. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons, 1965.
[4] *D. Khoshnevisan, D. A. Levin, P. J. Méndez-Hernández*: On dynamical Gaussian random walks. Ann. Probab. *33* (2005), 1452–1478.
[5] *D. Khoshnevisan, D. A. Levin, P. J. Méndez-Hernández*: Exceptional times and invariance for dynamical random walks. Probab. Theory Relat. Fields. *134* (2006), 383–416.
[6] *D. Khoshnevisan, D. A. Levin*: On dynamical bit sequences. arXiv:0706.1520v2.
[7] *J.-H. Ma, S.-Y. Wen, Z.-Y. Wen*: Egoroff's theorem and maximal run length. Monatsh. Math. *151* (2007), 287–292.
[8] *P. Révész*: Random Walk in Random and Non-Random Enviroments. Singapore. World Scientific, 1990.

Author's address: Ruibiao Zou, Hunan Agriculture University, Changsha, Hunan, China P. R., 410128, e-mail: rbzou@163.com.