Proceedings of the Edinburgh Mathematical Society (2018) **61**, 363–369
doi:10.1017/S0013091517000116

A NOTE ON RÉNYI'S 'RECORD' PROBLEM AND ENGEL'S SERIES

LULU FANG[1]*AND MIN WU[2]

[1]*School of Mathematics, Sun Yat-sen University, Guangzhou 510275, P.R. China* (fanglulu@mail.sysu.edu.cn)

[2]*Department of Mathematics, South China University of Technology, Guangzhou 510640, P.R. China* (wumin@scut.edu.cn)

(Received 27 April 2016)

Abstract In 1973, Williams [D. Williams, On Rényi's 'record' problem and Engel's series, *Bull. London Math. Soc.* **5** (1973), 235–237] introduced two interesting discrete Markov processes, namely C-processes and A-processes, which are related to record times in statistics and Engel's series in number theory respectively. Moreover, he showed that these two processes share the same classical limit theorems, such as the law of large numbers, central limit theorem and law of the iterated logarithm. In this paper, we consider the large deviations for these two Markov processes, which indicate that there is a difference between C-processes and A-processes in the context of large deviations.

Keywords: large deviations; Markov processes; record times; Engel's series

2010 *Mathematics subject classification:* Primary 60F10; 60J10
Secondary 62G30; 11A67

1. Introduction and main results

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and we assume that the following processes both define on it.

1.1. C-processes

A stochastic process $\{C_n : n \geqslant 1\}$ taking values in $\mathbb{N}$ will be called a *C-process* if the sequence $C_0, C_1, C_2, \ldots$ where $C_0 = 1$, forms a time-homogeneous Markov chain with one-step transition probabilities:

$$p_{ij} = \frac{i}{j(j-1)} \quad \text{for any } j \geqslant i+1; \quad p_{ij} = 0 \quad \text{for any } j \leqslant i. \tag{1.1}$$

Next we introduce two interesting examples of C-processes in number theory and statistics respectively.

* Corresponding author.

© 2018 The Edinburgh Mathematical Society

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press

Example 1.1 (modified Engel's series). In particular, $\Omega = (0,1)$, $\mathcal{F}$ is the Borel σ-algebra on $(0,1)$ and $\mathbf{P}$ denotes the Lebesgue measure on $(0,1)$. In [11], Rényi studied the *modified Engel's series* in the form

$$x = \frac{1}{d_1} + \frac{1}{(d_1-1)d_2} + \cdots + \frac{1}{(d_1-1)\cdots(d_{n-1}-1)d_n} + \cdots,$$

where $d_n \geqslant 2$ is an integer and $d_{n+1} \geqslant d_n + 1$ for any $n \in \mathbb{N}$. Then the digit sequence $\{d_n : n \geqslant 1\}$ is a C-process (see Rényi [11, Equation (5.3)]).

Example 1.2 (Record times). Let $\{X_n : n \geqslant 1\}$ be a sequence of independent and identically distributed real-valued random variables with a common continuous distribution function. The *record times* L_n are defined by $L_0 = 1$ and

$$L_{n+1} = \min\{j : X_j > X_{L_n}, j > L_n\} \quad \text{for all } n \geqslant 0.$$

Then $\{L_n : n \geqslant 1\}$ is a C-process (see Rényi [12, Equation (20)]). For more properties and applications of record times, see [1, 7, 10, 14] and the references therein.

1.2. A-processes

We say a stochastic process $\{A_n : n \geqslant 1\}$ taking values in $\mathbb{N}$ an *A-process* if it is identical in law to the Engel's series (see Erdős *et al.* [3]). That is to say, an A-process can be constructed as follows: an irrational number x in $(0,1)$ is chosen at random according to the uniform distribution and expressed uniquely as

$$x = \frac{1}{A_1} + \frac{1}{A_1A_2} + \cdots + \frac{1}{A_1A_2\cdots A_n} + \cdots,$$

where A_n are integers and $A_n \geqslant A_{n-1}$ for any $n \in \mathbb{N}$ with the convention that $A_0 = 2$. Erdős *et al.* [3, Equation (1.3)] proved that the digit sequence of Engel's series forms a time-homogeneous Markov chain with one-step transition probabilities:

$$p_{ij} = \frac{i-1}{j(j-1)} \quad \text{for any } j \geqslant i \geqslant 2; \quad p_{ij} = 0 \quad \text{for any } j < i. \tag{1.2}$$

Since A-processes are identical in law to the Engel's series, we know that an A-process also forms a time-homogeneous Markov chain with one-step transition probabilities defined as Equation (1.2).

In his paper [15], Williams pointed out that the C-process and A-process fulfil the following representations respectively:

$$C_0 = 1, \quad C_n = [C_{n-1}\exp(W_n)] + 1, \quad a_0 = 1, \quad a_n = [a_{n-1}\exp(W_n)], \quad A_n = a_n + 1,$$

where $[x]$ denotes the greatest integer not exceeding x and $W_1, W_2, \ldots$ are independent and identically distributed exponential random variables with mean 1. Here denoted by $B_n^* = W_1 + \cdots + W_n$, Williams also showed that the sequences $\{B_n^* - \log C_n\}_{n\geqslant 1}$ and $\{B_n^* - \log A_n\}_{n\geqslant 1}$ are almost surely bounded, which implies that $(\log C_n)/n$ and $(\log A_n)/n$ are asymptotically close to B_n^*/n, i.e., the empirical mean of independent and identically distributed exponential random variables with mean 1. As a consequence, the

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press

strong law of large numbers, the central limit theorem and the law of iterated logarithm for C-processes and A-processes can be read off from the classical results for $\{B_n^*, n \geqslant 1\}$. However, our following results indicate that the approximation of $(\log C_n)/n$ by B_n^*/n is enough while the approximation of $(\log A_n)/n$ by B_n^*/n is not enough in the the context of large deviations. This is a difference between these two Markov processes.

1.3. Main results

It is worth remarking that the classical limit theorems basically concern that the averages taken over large samples converge to expected values in some sense, but say little or nothing about the rate of convergence. One way to address this is the theory of large deviations in modern probability theory. Let $\{Y_n : n \geqslant 1\}$ be a sequence of the real-valued random variables defined on $(\Omega, \mathcal{F}, \mathbf{P})$. A function $I : \mathbb{R} \to [0, \infty]$ is called a *good rate function* if it is lower semi-continuous and has compact level sets. We say that the sequence $\{Y_n : n \geqslant 1\}$ satisfies a *large deviation principle* (LDP for short) with speed n and good rate function I, if for any Borel set Γ,

$$-\inf_{x\in\Gamma^\circ} I(x) \leqslant \liminf_{n\to\infty} \frac{1}{n}\log \mathbf{P}(Y_n \in \Gamma) \leqslant \limsup_{n\to\infty} \frac{1}{n}\log \mathbf{P}(Y_n \in \Gamma) \leqslant -\inf_{x\in\overline{\Gamma}} I(x),$$

where Γ° and $\overline{\Gamma}$ denote the interior and the closure of Γ respectively. For instance, we can study the probability that the empirical mean of a sequence of random variables deviates away from its ergodic mean. These probabilities are exponentially small in general and follow a large deviation principle. For an introduction to the theory of large deviations, we refer the reader to Dembo and Zeitouni [**2**] and Touchette [**13**].

Theorem 1.3. *Let $\{C_n : n \geqslant 1\}$ be a C-process. Then $\{(\log C_n - n)/n : n \geqslant 1\}$ satisfies an LDP with speed n and good rate function*

$$I_C(x) = \begin{cases} x - \log(x+1), & \text{if } x > -1; \\ +\infty, & \text{if } x \leqslant -1. \end{cases} \tag{1.3}$$

As an application of Theorem 1.3, we immediately obtain the large deviations for modified Engel's series and record times by Examples 1.1 and 1.2.

Corollary 1.4. *The sequence $\{(\log d_n - n)/n : n \geqslant 1\}$ or $\{(\log L_n - n)/n : n \geqslant 1\}$ satisfies an LDP with speed n and good rate function defined as Equation (1.3).*

Remark 1.5. It is worth pointing out that Gut [**8**, **9**] has considered the following tail probabilities for record times

$$\mathbf{P}(|\log L_n - n| \geqslant n\varepsilon),$$

where $\varepsilon > 0$ and $n \in \mathbb{N}$. Moreover, he also showed that these tail probabilities are exponentially small (see [**8**, Theorem 6]). However, the Corollary 1.4 indicates that the probability of the rate event that $(\log L_n)/n$ deviates away from 1 decays to zero exponentially with an explicit rate function.

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press

Since an A-process is identical in law to the Engel's series, the large deviation principle for A-processes can be stated as follows (see Zhu [16, Theorem 1.2]).

Theorem 1.6 ([16]). *Let $\{A_n : n \geqslant 1\}$ be an A-process. Then $\{(\log A_n - n)/n : n \geqslant 1\}$ satisfies an LDP with speed n and good rate function*

$$I_A(x) = \begin{cases} x - \log(x+1), & \text{if } x > -\frac{1}{2}; \\ -x - 1 + \log 2, & \text{if } -1 \leqslant x \leqslant -\frac{1}{2}; \\ +\infty, & \text{if } x < -1. \end{cases} \tag{1.4}$$

Remark 1.7. As we have seen in Theorem 1.3, the rate function I_C completely coincides with the rate function for B_n^*/n (see [13, Example 3.2]). However, the rate function I_A in Theorem 1.6 coincides with the rate function for B_n^*/n only when $x > -1/2$. That is to say, the approximation of $(\log A_n)/n$ by the empirical mean of independent and identically distributed exponential random variables with mean 1 is not enough in the context of large deviations (see Zhu [16, Remark 1.4]).

2. Proof of main result

We use the notation $\mathbf{E}(\xi)$ to denote the expectation of a random variable ξ with respect to the probability measure $\mathbf{P}$. In this section, we will give the proofs of Theorems 1.3 and 1.6. In fact, the large deviations for stochastic processes related Markov chain occurring in number theory have been discussed by Zhu [16], Fang [4, 5] and Fang and Shang [6]. More precisely, Zhu [16] studied the large deviations for Engel's series whose digit sequence is non-decreasing. In [6], the authors considered the large deviations for Engel continued fractions, whose digit sequence is also non-decreasing but does not form a Markov chain. In [4, 5], the author mainly investigated large deviations for two expansions of real numbers whose digit sequence both are strictly increasing. So we will use the methods in [4, 5] to prove our main result. To do this we need an elementary lemma.

Lemma 2.1. *Let $\theta < 1$ be a real number. For any $j \geqslant 1$, we have*

$$\left(1 + \frac{1}{j}\right)^{\theta-1} \cdot \frac{1}{1-\theta} \leqslant \sum_{k=j+1}^{\infty} \frac{j}{k(k-1)} \left(\frac{k}{j}\right)^{\theta} \leqslant \left(1 + \frac{1}{j}\right) \cdot \frac{1}{1-\theta}.$$

Proof. The proof is similar to the proof of Lemma 3.1 in [5]. □

Recall that $C_0, C_1, C_2, \ldots$ is a strictly increasing stochastic process and also forms a time-homogeneous Markov chain. Lemma 2.1 and the methods of Lemma 3.2 in [5] immediately yield the following.

Lemma 2.2.

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{E}(C_n^{\theta}) = \begin{cases} \log \dfrac{1}{1-\theta}, & \text{if } \theta < 1; \\ +\infty, & \text{if } \theta \geqslant 1. \end{cases}$$

Now we are ready to prove Theorem 1.3.

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press

Proof of Theorem 1.3. For any $\theta \in \mathbb{R}$, we define the logarithmic moment generating function (see Dembo and Zeitouni [**2**]) for C-processes as

$$\Lambda_C(\theta) = \lim_{n\to\infty} \frac{1}{n} \log \mathbf{E}\left(\exp\left(\frac{\log C_n - n}{n} \cdot n\theta\right)\right)$$

if the limit exists. In view of Lemma 2.2, we know that

$$\Lambda_C(\theta) = -\theta + \lim_{n\to\infty} \frac{1}{n} \log \mathbf{E}(C_n^\theta) = \begin{cases} -\theta - \log(1-\theta), & \text{if } \theta < 1; \\ +\infty, & \text{if } \theta \geqslant 1. \end{cases}$$

By the Gärtner-Ellis theorem (see [**2**, Theorem 2.3.6]), we obtain that the sequence $\{(\log C_n - n)/n : n \geqslant 1\}$ satisfies an LDP with speed n and good rate function, i.e., the Legendre transform of $\Lambda_C(\theta)$,

$$I_C(x) = \sup_{\theta\in\mathbb{R}}\{\theta x - \Lambda_C(\theta)\} \text{ for all } x \in \mathbb{R}.$$

Now we will show that $I_C(x)$ is as defined in Theorem 1.3. Since $\Lambda_C(\theta) = +\infty$ if $\theta \geqslant 1$, then $I_C(x)$ is alternatively given by

$$I_C(x) = \sup_{\theta<1}\{\theta x + \theta + \log(1-\theta)\} \text{ for all } x \in \mathbb{R}.$$

For any $x \in \mathbb{R}$, let $f(\theta) = \theta x + \theta + \log(1-\theta)$ for all $\theta < 1$. It is worth pointing out that f is strictly concave and that $\theta = x/(x+1)$ is the unique maximal point of it. When $x \leqslant -1$, we know that $f(\theta) = \theta(x+1) + \log(1-\theta)$ is decreasing with respect to θ on the interval $(-\infty, 1)$ and hence that $I_C(x) = +\infty$. If $x > -1$, we have that $x/(x+1) < 1$. That is to say, the unique maximal point of f locates in its domain. So, $I_C(x) = x - \log(x+1)$ for any $x > -1$. Thus, we determine the rate function $I_C(x)$ defined as

$$I_C(x) = \begin{cases} x - \log(x+1), & \text{if } x > -1; \\ +\infty, & \text{if } x \leqslant -1. \end{cases}$$ □

Proof of Theorem 1.6. Since the proof is similar to the proof of Theorem 1.3, we confine ourselves to providing a sketch. Although A-processes are non-decreasing (not strictly increasing) Markov chains, the methods in the proof of Theorem 1.3 are still valid for A-processes if we properly modify the corresponding proofs of Lemmas 2.1 and 2.2. In fact, we can obtain

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{E}(A_n^\theta) = \begin{cases} \max\left\{-\log 2, \log\dfrac{1}{1-\theta}\right\}, & \text{if } \theta < 1; \\ +\infty, & \text{if } \theta \geqslant 1. \end{cases}$$

Replacing C by A in the relevant objects defined as in the proof of Theorem 1.3, we actually deduce that

$$\Lambda_A(\theta) = \begin{cases} -\theta - \log 2, & \text{if } \theta \leqslant -1; \\ -\theta - \log(1-\theta), & \text{if } -1 < \theta < 1; \\ +\infty, & \text{if } \theta \geqslant 1. \end{cases}$$

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press

By the Gärtner-Ellis theorem, we know the sequence $\{(\log A_n - n)/n : n \geqslant 1\}$ satisfies an LDP and its rate function I_A given by Equation (1.4) is the Legendre transform of $\Lambda_A(\theta)$. It is worth remarking that $\log 2$ appears in this formula for A-processes by comparing with the result of Lemma 2.2. This fact eventually leads to the difference of rate functions for C-processes and A-processes. □

3. Further remarks

This section is devoted to giving some further remarks about the differences between C-processes and A-processes.

1. It is clear to see $I_C(x) \geqslant I_A(x)$ for any $x \in \mathbb{R}$. Strict inequality holds for the interval $[-1, -1/2)$. This shows that the decay speed of the probability for which $(\log C_n)/n$ deviates away from 1 is faster than that of $(\log A_n)/n$. In fact, to understand the difference between I_C and I_A, we take a value of x of between -1 and $-1/2$ and then $A_n = 2$ implies $(\log A_n - n)/n \leqslant x$ for sufficiently large n. By the definition of A-processes, we have $\mathrm{P}(A_n = 2) = 2^{-n}$ and hence that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathrm{P}\left(\frac{\log A_n - n}{n} \leqslant x\right) \geqslant -\log 2,$$

which implies $I_A(-1) \leqslant \log 2$, whereas this type of behaviour is ruled out for the C-process.

2. Now we consider the large deviation problems for C-processes with $C_0 = c$ ($c \in \mathbb{N}$) and A-processes with $A_0 = a$ ($a \geqslant 2$ and are integers), our methods in this paper imply that this class of C-processes satisfy an LDP and its rate function is still $I_C(x)$ defined as Equation (1.3) while these A-processes satisfy an LDP depending on the initial value $A_0 = a$, more precisely, the good rate function is

$$I_{A,a}(x) = \begin{cases} x - \log(x+1), & \text{if } x > -1 + 1/a; \\ (1-a)(x+1) + \log a, & \text{if } -1 \leqslant x \leqslant -1 + 1/a; \\ +\infty, & \text{if } x < -1, \end{cases}$$

see Zhu [**16**, Remark 1.3]. Clearly, $I_A = I_{A,2}$ and $I_{A,a} \to I_C$ as $a \to +\infty$. This indicates that the initial value does not impact the large deviation result for C-processes, but it will change the rate function of large deviations for A-processes.

Acknowledgements. The author wishes to thank the referee for helpful comments, and Professor Jonathan Warren and Professor Sandy Davie for useful discussions and suggestions.

References

1. K. Chandler, The distribution and frequency of record values, *J. Roy. Statist. Soc. Ser. B.* **14** (1952), 220–228.
2. A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd edition (Springer, New York, 1998).

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press

3. P. Erdős, A. Rényi and P. Szüsz, On Engel's and Sylvester's series, *Ann. Univ. Sci. Budapest. Eötvös. Sect. Math.* **1** (1958), 7–32.
4. L. Fang, Large and moderate deviations for modified Engel continued fractions, *Statist. Probab. Lett.* **98** (2015), 98–106.
5. L. Fang, Large and moderate deviation principles for alternating Engel expansions, *J. Number Theory* **156** (2015), 263–276.
6. L. Fang, M. Wu and L. Shang, Large and moderate deviation principles for Engel continued fractions, *J. Theor. Probab.*, DOI:10.1007/s10959-016-0715-3.
7. J. Galambos and E. Seneta, Record times, *Proc. Amer. Math. Soc.* **50** (1975), 383–387.
8. A. Gut, Convergence rates for record times and the associated counting process, *Stochastic Process. Appl.* **36** (1990), 135–151.
9. A. Gut, Precise asymptotics for record times and the associated counting process, *Stochastic Process. Appl.* **101** (2002), 233–239.
10. D. Pfeifer, On a joint strong approximation theorem for record and inter-record times, *Probab. Theory Related Fields* **75** (1987), 213–221.
11. A. Rényi, A new approach to the theory of Engel's series, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **5** (1962), 25–32.
12. A. Rényi, Théorie des éléments saillants d'une suite d'observations, *Ann. Fac. Sci. Univ. Clermont-Ferrand* **8** (1962), 7–13 (in French).
13. H. Touchette, The large deviation approach to statistical mechanics, *Phys. Rep.* **478** (2009), 1–69.
14. G. Wergen, A. Hense and J. Krug, Record occurrence and record values in daily and monthly temperatures, *Climate Dynamics* **42** (2014), 1275–1289.
15. D. Williams, On Rényi's 'record' problem and Engel's series, *Bull. London Math. Soc.* **5** (1973), 235–237.
16. L. Zhu, On the large deviations for Engel's, Sylvester's series and Cantor's products, *Electron. Commun. Probab.* **19** (2014), 1–9.

https://doi.org/10.1017/S0013091517000116 Published online by Cambridge University Press