PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 38, Fasc. 2 (2018), pp. 407–428
doi:10.19195/0208-4147.38.2.8

# ON THE LONGEST RUNS IN MARKOV CHAINS

BY

ZHENXIA LIU (LINKÖPING) AND XIANGFENG YANG (LINKÖPING)

*Abstract.* In the first $n$ steps of a two-state (success and failure) Markov chain, the longest success run $L(n)$ has been attracting considerable attention due to its various applications. In this paper, we study $L(n)$ in terms of its two closely connected properties: moment generating function and large deviations. This study generalizes several existing results in the literature, and also finds an application in statistical inference. Our method on the moment generating function is based on a global estimate of the cumulative distribution function of $L(n)$ proposed in this paper, and the proofs of the large deviations include the Gärtner–Ellis theorem and the moment generating function.

**2010 AMS Mathematics Subject Classification:** Primary: 60F10, 44A1; Secondary: 60J10, 60G70.

**Key words and phrases:** Longest run, moment generating function, large deviation principle, Markov chain.

## 1. INTRODUCTION

Let $\{X_k\}_{k\geqslant 1}$ be a time-homogeneous two-state (success and failure) Markov chain. We assume that the initial distribution is $\mathbb{P}(X_1 = 0) = p_0$ and $\mathbb{P}(X_1 = 1) = p_1 = 1 - p_0$, with '1' and '0' denoting the 'success' and 'failure', respectively. The transition matrix of $\{X_k\}_{k\geqslant 1}$ is written as

$$T = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}.$$

To avoid triviality, it is assumed throughout the paper that $0 < p_0 < 1$ and $0 < p_{ij} < 1$ for $i, j = 0, 1$, which indicates that the Markov chain is ergodic. In the first $n$ steps of the Markov chain, the longest success run $L(n)$, namely the longest stretch of consecutive successes, has been attracting considerable attention due to its applications in various fields, such as reliability and statistics (cf. [1]). We refer to [4] and [5] for the first few seminal works in the 1970s, and [8]–[11] for the latest progress.

Among various studies on the longest success run $L(n)$, the probability estimating of $L(n)$ for large $n$ (such as large deviations) is an important topic. Part of the reason is that the exact distribution of $L(n)$ (cf. [7]) is intricate despite known explicit formulas, which gives no information as $n$ approaches infinity. Even in the identically independent case (that is, $\{X_k\}_{k\geqslant 1}$ are independent and identically distributed), there is much complexity of the exact distribution of $L(n)$ which can be seen (for instance cf. [8]) as follows:

$$\mathbb{P}(L(n) < k) = \sum_{r=0}^{[\frac{n+1}{k+1}]} (-1)^r p_1^{rk} p_0^{r-1} \left[ \binom{n-rk}{r-1} + p_0 \binom{n-rk}{r} \right],$$

where $[\cdot]$ denotes the integer part of a constant. One topic of this paper is to study the large deviations of $L(n)$ in a Markov chain $\{X_k\}_{k\geqslant 1}$ defined above. To appropriately propose such deviations, recall a law of large numbers (cf. e.g. [14]):

$$\frac{L(n)}{\log_{1/p_{11}} n} \to 1 \quad \text{in probability as } n \to \infty.$$

Such a limit in independent trails is a well-known result (cf. [4], [5], [12]). This suggests to study the large deviation probabilities in the form $\mathbb{P}\big(L(n)/\log_{1/p_{11}} n \in A\big)$, where the set $A$ does not include the most probable point 1. Our first result is formulated as follows.

THEOREM 1.1. *For each* $x > 0$, *we have*

$$\lim_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\bigg( \frac{L(n)}{\log_{1/p_{11}} n} \geqslant 1 + x \bigg) = -x \cdot \ln(1/p_{11}). \tag{1.1}$$

*For each* $0 < x < 1$, *we have*

$$\lim_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \bigg[ -\ln \mathbb{P}\bigg( \frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1 - x \bigg) \bigg] = x \cdot \ln(1/p_{11}). \tag{1.2}$$

Theorem 1.1 tells that the probability $\mathbb{P}\big(L(n)/\log_{1/p_{11}} n \geqslant 1 + x\big)$ decays in a power rate, while the probability $\mathbb{P}\big(L(n)/\log_{1/p_{11}} n \leqslant 1 - x\big)$ decays exponentially fast. If $\{X_k\}_{k\geqslant 1}$ is a sequence of identically independent trails, namely $p_{00} = p_{10} = p_0$ and $p_{01} = p_{11} = p_1$, then the limits (1.1) and (1.2) trivially hold because of a well global estimate (cf. [7] and [9]): for $k = 1, \ldots, n$,

$$(1 - p_1^k)^{n-k+1} \leqslant \mathbb{P}\big(L(n) < k\big) \leqslant (1 - p_0 p_1^k)^{n-k+1}. \tag{1.3}$$

Due to the lack of satisfactory estimates as above (namely (1.3)) for general Markov chains $\{X_k\}_{k\geqslant 1}$, the proof of Theorem 1.1 will be based on a less precise global estimate proposed below (see Lemma 2.1) in this paper. Here we note that essentially the same large deviation probability as (1.1) was claimed to be proved in [14] in the form: for all $x > 0$,

(1.4)
$$\lim_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\big(L(n) - \lfloor \log_{1/p_{11}} n \rfloor \geqslant x \cdot \log_{1/p_{11}} n\big) = -x \cdot \ln(1/p_{11}).$$

Unfortunately, the proof of (1.4) therein contains a mistake stemming from the employed (Stein–Chen) method, which seems to be impossible to be corrected in principle. Section 4 includes detailed explanations on this aspect.

A natural generalization of the limit (1.1) (not (1.2)) is a *large deviation principle* for the family of random variables $L(n)/\log_{1/p_{11}} n$. For identically independent trails $\{X_k\}_{k\geqslant 1}$, large deviation principles were recently derived in [9] based on (1.3). There are also related discussions on the large deviations of $L(n)$ in [7] and [11]. The second result of this paper is to establish a large deviation principle for $L(n)$, which includes (1.1) (or (1.4)) as a special case. To this end, we define a function $\Lambda^*(x)$ as

$$\Lambda^*(x) = \begin{cases} +\infty, & x < 1, \\ (x-1)\ln(1/p_{11}), & x \geqslant 1. \end{cases} \tag{1.5}$$

THEOREM 1.2. *The normalized longest success run $L(n)/\log_{1/p_{11}} n$ satisfies a large deviation principle with a good rate function $\Lambda^*(x)$ given by* (1.5) *and a speed $\log_{1/p_{11}} n$. Namely,*

(i) *for any open set $O \subseteq \mathbb{R}$,*

$$\liminf_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\left(\frac{L(n)}{\log_{1/p_{11}} n} \in O\right) \geqslant -\inf_{x\in O} \Lambda^*(x); \tag{1.6}$$

(ii) *for any closed set $F \subseteq \mathbb{R}$,*

$$\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\left(\frac{L(n)}{\log_{1/p_{11}} n} \in F\right) \leqslant -\inf_{x\in F} \Lambda^*(x). \tag{1.7}$$

It is clear that the special case (1.1) (or (1.4)) comes from Theorem 1.2 with an open set $O = (1+x, \infty)$ and a closed set $F = [1+x, \infty)$. The proof of Theorem 1.2 is given in Section 3.2.

The large deviation principle in Theorem 1.2 is *non-trivial* since the rate function $\Lambda^*(x)$ is not always *zero* or *infinity*. Now an interesting question arises: besides the family of random variables $L(n)/\log_{1/p_{11}} n$, are there other families which admit non-trivial large deviation principles? Note that large deviation principles have very close connections with the corresponding Laplace transforms (or the moment generating functions; see the Gärtner–Ellis theorem [3]), thus the above question leads to the third result of this paper: precise logarithmic asymptotics for the moment generating function of $L(n)$ as formulated in the following theorem, based on which there are (only) two families which admit non-trivial large deviation principles: $\{L(n)/\log_{1/p_{11}} n\}$ and $\{L(n)/n\}$. Throughout the paper, $a(n) \sim b(n)$ as $n \to \infty$ stands for $\lim_{n\to\infty} a(n)/b(n) = 1$.

THEOREM 1.3. *The moment generating function of $L(n)$ has the following logarithmic asymptotics*:

(i) *for* $\lambda < \ln(1/p_{11})$,

$$\ln \mathbb{E}e^{\lambda L(n)} \sim \lambda \log_{1/p_{11}} n;$$

(ii) *for* $\lambda = \ln(1/p_{11})$,

$$\ln \mathbb{E}e^{\lambda L(n)} \sim 2\lambda \log_{1/p_{11}} n;$$

(iii) *for* $\lambda > \ln(1/p_{11})$,

$$\begin{aligned}\lambda - \ln(1/p_{11}) &\leqslant \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{E}e^{\lambda L(n)} \leqslant \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E}e^{\lambda L(n)} \\ &\leqslant \max\left\{\lambda - \ln(1/p_{11}),\ \lambda - \ln \frac{1}{|p_{00} - p_{10}|}\right\},\end{aligned}$$

*and*, *in particular*, *if* $p_{10} \leqslant p_{00} + p_{11}$, *then*

$$\ln \mathbb{E}e^{\lambda L(n)} \sim \lambda - \ln(1/p_{11}).$$

Similar results for the identically independent case have been recently proved in [9], where the condition $p_{10} \leqslant p_{00} + p_{11}$ is automatically fulfilled. Technically speaking, the condition $p_{10} \leqslant p_{00} + p_{11}$ is due to an extra error term $e(n)$ in Lemma 2.2 below. In terms of the structure of the Markov chain, this condition means that the transition probability $p_{10}$ from the state '1' to the state '0' should not exceed the probability that the chain stays still, which is $p_{00} + p_{11}$. Although we think that such a condition can be removed by using a more precise estimate than the one in Lemma 2.2, the current method in this paper cannot get rid of this condition.

Several new difficulties arise in the proof of Theorem 1.3 due to the lack of satisfactory global estimates of the cumulative distribution function of $L(n)$, and we overcome them using suitable non-global estimates included in Section 2.2. To see how Theorem 1.3 yields non-trivial large deviation principles, we first consider the logarithmic moment generating function of $L(n)/\log_{1/p_{11}} n$ (according to (i) and (ii) of Theorem 1.3) defined as $\Lambda_n(\lambda) = \ln \mathbb{E} \exp\{\lambda \cdot L(n)/\log_{1/p_{11}} n\}$ for $\lambda \in \mathbb{R}$, and the *cumulant* defined as $\Lambda(\lambda) := \lim_{n\to\infty} \Lambda_n(\lambda \cdot \log_{1/p_{11}} n)/\log_{1/p_{11}} n$. Then the Gärtner–Ellis theorem (cf. [3], Section 2.3) suggests that there is a non-trivial large deviation principle for the family $L(n)/\log_{1/p_{11}} n$ with a rate function $\Lambda^*$ defined via the Fenchel–Legendre transform of $\Lambda$: $\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}}[\lambda \cdot x - \Lambda(\lambda)]$.

This is verified in detail in Theorem 1.2. Now, according to (iii) of Theorem 1.3, we can also consider the logarithmic moment generating function of $L(n)/n$ as $\widetilde{\Lambda}_n(\lambda) = \ln \mathbb{E} \exp\{\lambda \cdot L(n)/n\}$ for any $\lambda \in \mathbb{R}$, and obtain the cumulant, under the condition $p_{10} \leqslant p_{00} + p_{11}$, in the form

$$\widetilde{\Lambda}(\lambda) := \lim_{n\to\infty} \frac{1}{n} \widetilde{\Lambda}_n(\lambda \cdot n) = \begin{cases} \lambda - \ln(1/p_{11}), & \lambda \geqslant \ln(1/p_{11}), \\ 0, & \lambda < \ln(1/p_{11}). \end{cases}$$

The Gärtner–Ellis theorem again suggests that there is a non-trivial large deviation principle for the family $L(n)/n$ with a rate function $\widetilde{\Lambda}^*(x)$ defined as the Fenchel–Legendre transform of $\widetilde{\Lambda}(\lambda)$:

$$\widetilde{\Lambda}^*(x) = \begin{cases} +\infty, & x < 0, \\ x \ln(1/p_{11}), & 0 \leqslant x \leqslant 1, \\ +\infty, & x > 1. \end{cases} \tag{1.8}$$

This large deviation principle for the family $\{L(n)/n\}$ corresponds to the law of large numbers $L(n)/n \to 0$ which is directly from $L(n)/\log_{1/p_{11}} n \to 1$. We formulate this observation as our last result in the following theorem.

THEOREM 1.4. *If $p_{10} \leqslant p_{00} + p_{11}$, then the normalized longest success run $L(n)/n$ satisfies a large deviation principle with a good rate function $\widetilde{\Lambda}^*(x)$ given by* (1.8) *and a speed $n$. Namely*,

(i) *for any open set $O \subseteq \mathbb{R}$*,

$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \in O\right) \geqslant -\inf_{x\in O} \widetilde{\Lambda}^*(x); \tag{1.9}$$

(ii) *for any closed set $F \subseteq \mathbb{R}$*,

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \in F\right) \leqslant -\inf_{x\in F} \widetilde{\Lambda}^*(x). \tag{1.10}$$

Here we draw the reader's attention that the Gärtner–Ellis theorem will be used to prove the aforementioned two large deviation principles. It should be noted that there are other methods to achieve such large deviation principles, such as the Bryc's Inverse Varadhan Lemma (cf. Section 4.4 in [3]). In [10] the Bryc's Inverse Varadhan Lemma was used to obtain a large deviation principle for $L(n)$ with a general speed in the identically independent case.

The rest of the paper is organized as follows. Section 2 includes global and non-global estimates of the cumulative distribution function of $L(n)$ which will be used throughout the paper. In the first part of Section 3, we give the proof of the main result of the paper: the precise logarithmic asymptotics for the moment generating function (Theorem 1.3). Then we show that the two large deviation principles (Theorems 1.2 and 1.4) follow from Theorem 1.3 and the Gärtner–Ellis theorem, which is included in the second part of Section 3. The last part of Section 3 contains a very concise proof of Theorem 1.1. The use of the Stein–Chen method in estimating the large deviation probabilities of $L(n)$ is briefly described in Section 4, where a mistake of proving (1.4) in [14] is pointed out. Finally, an application of the derived results to statistical inference is presented in Section 5.

## 2. ESTIMATES OF THE DISTRIBUTION FUNCTION

In this section, we first propose a global estimate for the cumulative distribution function of $L(n)$ which will be used throughout the paper. Then we present several special non-global estimates which have more explicit forms.

**2.1. Global estimate.**

LEMMA 2.1. *For all $k = 1, \ldots, n$, we have*

(2.1)
$$1 - p_{11}^{k-1}\,[c_1 \cdot (n-k) + c_2] - c(n,k) \leqslant \mathbb{P}\big(L(n) < k\big) \leqslant (1 - c_3 \cdot p_{11}^{k-1})^{n-k+1},$$

*where*

$$c_1 = \frac{p_{01}p_{10}}{p_{01}+p_{10}} > 0, \qquad c_2 = \frac{p_{01}(p_0p_{01} - p_1p_{10})}{(p_{01}+p_{10})^2},$$

$$c_3 = \min\left\{p_1,\ \frac{c_1 + (p_{01}+p_{10}) \cdot \min\{0,\ c_2,\ c_2(p_{00}-p_{10}),\ c_2(p_{00}-p_{10})^2\}}{1 + c_1/p_{01} + |c_2|(p_{01}+p_{10})/p_{01}}\right\}$$

$(c_3 > 0)$, *and*

$$\begin{aligned} c(n,k) = & \frac{c_1 p_{11}^{k-1}}{1-p_{11}} - \frac{c_2(p_{01}+p_{10})}{p_{01}}(p_{00}-p_{10})^{n-1} \\ & - \frac{c_2(p_{01}+p_{10})}{p_{11}} \cdot \frac{(p_{00}-p_{10})^n - p_{11}^{k-1}(p_{00}-p_{10})^{n-k}}{p_{00}-p_{10}-p_{11}} > 0. \end{aligned}$$

P r o o f. We first note that the exact distribution of $L(n)$ has been known (cf. [7]), but it hardly helps to gain useful information on the asymptotics as $n \to \infty$. The proof of Lemma 2.1 is based on a newly built Markov chain $\{\eta_k\}_{1 \leqslant k \leqslant n}$, where $\eta_k$ is defined as the length of success runs at the end of the $k$-th step, namely

$$\{\eta_k = i\} \text{ is equivalent to } \{X_k = 1, \ldots, X_{k-i+1} = 1, X_{k-i} = 0\}.$$

In this setting, the longest success run $L(n) = \max_{1 \leqslant k \leqslant n} \eta_k$. This enables us to estimate $\mathbb{P}\big(L(n) < k\big)$ a little more explicitly, using the probabilities involving $\eta_k$. This idea was introduced in [6], where the derived results are

$$\mathbb{P}\big(L(n) < k\big) \geqslant 1 - p_{01}p_{11}^{k-1} \sum_{i=k}^{n-1} b(i-k) - c(n,k) \tag{2.2}$$

and

(2.3)

$$\mathbb{P}\big(L(n) < k\big) \leqslant (1 - p_1 p_{11}^{k-1}) \prod_{i=k+1}^{n} \left( 1 - \frac{p_{01} p_{11}^{k-1} b(i-k)}{b(i-1) + p_{01} \sum\limits_{j=1}^{k-1} p_{11}^{j-1} b(i-j-1)} \right)$$

with

$$b(i) = p_0 (p_{00} - p_{10})^{i-1} + \frac{p_{10}\big(1 - (p_{00} - p_{10})^{i-1}\big)}{1 - p_{00} + p_{10}}.$$

To achieve the upper bound in (2.1) from (2.3), we note that $p_1 \geqslant c_3$, and

$$\frac{p_{01} b(i-k)}{b(i-1) + p_{01} \sum\limits_{j=1}^{k-1} p_{11}^{j-1} b(i-j-1)} \geqslant \frac{p_{01} \min_j b(j)}{\max_j b(j) + 1},$$

since $p_{01} \sum_{j=1}^{k-1} p_{11}^{j-1} b(i-j-1) = \mathbb{P}(\eta_i = 0, 1, \dots, k-1) \leqslant 1$. To estimate two quantities $\min_j b(j)$ and $\max_j b(j)$, we rewrite $b(j)$ as

$$b(j) = \alpha + \beta \cdot (p_{00} - p_{10})^{j-1}, \qquad \text{where } \alpha = \frac{p_{10}}{p_{01} + p_{10}} \text{ and } \beta = \frac{p_0 p_{01} - p_1 p_{10}}{p_{01} + p_{10}}.$$

It then follows that

$$\max_j b(j) \leqslant \alpha + |\beta|,$$

and

$$\min_j b(j) \geqslant \min\{\alpha, \alpha + \beta, \alpha + \beta(p_{00} - p_{10}), \alpha + \beta(p_{00} - p_{10})^2\}.$$

Therefore,

$$\frac{p_{01} \min_j b(j)}{\max_j b(j) + 1} \geqslant c_3,$$

which implies the upper bound in (2.1).

To obtain the lower bound in (2.1) from (2.2), we see that the sum in (2.2) is

$$\begin{aligned} \sum_{i=k}^{n-1} b(i-k) &= (n-k)\alpha + \beta \cdot \frac{1 - (p_{00} - p_{10})^{n-k}}{1 - (p_{00} - p_{10})} \\ &\leqslant (n-k)\alpha + \beta \cdot \frac{1}{p_{01} + p_{10}}, \end{aligned}$$

which gives the lower bound. To see the positivity of $c(n,k)$, we note that

$$c(n,k) = \mathbb{P}(\eta_n \in \{k, k+1, \dots, n\}) > 0. \quad \blacksquare$$

**2.2. Non-global estimates.** One might be interested in comparing the global estimate (2.1) in Lemma 2.1 with the i.i.d. case (1.3). They actually look alike under suitable conditions, which will be summarized as follows.

LEMMA 2.2. *If* $n > k := k(n) \geqslant 1 + \log_{1/p_{11}} \left(\frac{n(c_1/(1-p_{11})+|c_2|)}{2}\right)$, *then we have*

$$(1 - c_5 \cdot p_{11}^k)^{n-k+1} - e(n) \leqslant \mathbb{P}\big(L(n) < k\big) \leqslant (1 - c_4 \cdot p_{11}^k)^{n-k+1} \tag{2.4}$$

*for large* $n$, *where* $c_4$ *and* $c_5$ *are two* (*uniform*) *positive constants*, *and* $e(n)$ *is a term which converges to zero exponentially fast as* $n \to \infty$ (*note that* $e(n) = 0$ *when* $p_{00} = p_{10}$).

P r o o f. In (2.4) the claimed upper bound $\mathbb{P}\big(L(n)<k\big) \leqslant \big(1 - c_4 \cdot p_{11}^k\big)^{n-k+1}$ comes directly from the upper bound of (2.1) by setting $c_4 = c_3/p_{11}$, uniformly in $k$. To achieve the lower bound of (2.4), we first rewrite the lower bound of (2.1) as follows:

$$\begin{aligned}
&\mathbb{P}\big(L(n) < k\big) \geqslant 1 - p_{11}^{k-1}\left[c_1 \cdot (n-k) + c_2\right] - c(n,k) \\
&= 1 - p_{11}^{k-1}\left[c_1 \cdot (n-k) + c_2 + \frac{c_1}{1-p_{11}}\right] + \frac{c_2(p_{01}+p_{10})}{p_{01}}(p_{00}-p_{10})^{n-1} \\
&\quad + \frac{c_2(p_{01}+p_{10})}{p_{11}} \cdot \frac{(p_{00}-p_{10})^n - p_{11}^{k-1}(p_{00}-p_{10})^{n-k}}{p_{00}-p_{10}-p_{11}} \\
&=: 1 - p_{11}^{k-1}\left[c_1 \cdot (n-k) + c_2 + \frac{c_1}{1-p_{11}}\right] + e(n).
\end{aligned}$$

It is clear that the term $e(n)$ converges to zero exponentially fast for all $k$, and $e(n) = 0$ if $p_{00} = p_{10}$. If we define $c_* = c_1/(1-p_{11}) + |c_2|$, then (with $k < n$)

$$1 - p_{11}^{k-1}\left[c_1 \cdot (n-k) + c_2 + \frac{c_1}{1-p_{11}}\right] \geqslant 1 - p_{11}^{k-1} \cdot c_* \cdot (n-k).$$

In order to estimate $1 - p_{11}^{k-1} \cdot c_* \cdot (n-k)$, we set $N = n-k+1$, $a = p_{11}^{k-1} \cdot c_*$, and obtain

$$(1-a)^N \leqslant 1 - (N-1)a(1-a)^{N-2}\left[\frac{N}{N-1}(1-a) - Na/2\right].$$

Since $n > k(n) \geqslant 1 + \log_{1/p_{11}} \left(\frac{n(c_1/(1-p_{11})+|c_2|)}{2}\right)$ and $n$ is large, $a$ is small. Therefore,

$$(1-a)^{N-2} = [(1-a)^{1/a}]^{a(N-2)} \geqslant [(1-a)^{1/a}]^{c_* \cdot p_{11}^{c-1}} \geqslant (e/2)^{c_* \cdot p_{11}^{c-1}}$$

with $c = 1 - \log_{1/p_{11}}(2/c_*)$,

$$\frac{N}{N-1}(1-a) \geqslant 1+\delta$$

for some small $\delta > 0$, and

$$Na/2 \leqslant 1.$$

In summary, we have

$$(1-a)^N \leqslant 1 - (N-1)a \cdot \delta(e/2)^{c_* \cdot p_{11}^{c-1}},$$

which gives

$$(1 - c_* \cdot p_{11}^{k-1})^{n-k+1} \leqslant 1 - (n-k) \cdot c_* \cdot p_{11}^{k-1} \cdot \delta(e/2)^{c_* \cdot p_{11}^{c-1}}.$$

Replacing $c_*$ by $c_*/\delta(e/2)^{c_* \cdot p_{11}^{c-1}}$ proves the lower bound of (2.4). ∎

In Lemma 2.2, if $k$ is exactly the size $\alpha \cdot \log_{1/p_{11}} n$ with $\alpha > 1$, then we have the following more explicit estimate.

LEMMA 2.3. *If $x > 0$ and $k(n) = [(1+x)\log_{1/p_{11}} n]$, then*

$$c_6 \cdot n^{-(1+x)}\,(n-k) \leqslant \mathbb{P}\big(L(n) > k\big) \leqslant c_7 \cdot n^{-(1+x)}\,(n-k)$$

*for large $n$, where $c_6$ and $c_7$ are two (uniform) positive constants.*

P r o o f. To see the lower bound, we infer from Lemma 2.2 that

$$\begin{aligned}
\mathbb{P}\big(L(n) > k\big) &= 1 - \mathbb{P}\big(L(n) \leqslant k\big) \\
&\geqslant 1 - (1 - c_4 \cdot p_{11}^{k+1})^{n-k} \\
&= 1 - [(1 - c_4 \cdot p_{11}^{k+1})^{1/(c_4 \cdot p_{11}^{k+1})}]^{c_4 \cdot p_{11}^{k+1}(n-k)} \\
&= -[(1 - c_4 \cdot p_{11}^{k+1})^{1/(c_4 \cdot p_{11}^{k+1})}]^{\theta_n} \cdot \ln\big((1 - c_4 \cdot p_{11}^{k+1})^{1/(c_4 \cdot p_{11}^{k+1})}\big) \\
&\quad \times c_4 \cdot p_{11}^{k+1}(n-k) \\
&\geqslant \text{const} \cdot p_{11}^{k}(n-k) \geqslant \text{const} \cdot n^{-(1+x)}(n-k),
\end{aligned}$$

where $\theta_n \in [0,\ c_4 \cdot p_{11}^{k+1}(n-k)]$. The upper bound can be similarly handled by noticing that

$$e(n) \sim \text{const} \cdot \exp\left\{-n \cdot \ln\ \frac{1}{|p_{00} - p_{10}|}\right\} \leqslant \text{const} \cdot n^{-(1+x)}\,(n-k)\,. \quad ∎$$

The next estimate is the case when $k$ is of size $\alpha \cdot \log_{1/p_{11}} n$ with $\alpha < 1$.

LEMMA 2.4. *If* $0 < x < 1$ *and* $k(n) = [(1-x)\log_{1/p_{11}} n]$, *then*

$$c_8 \cdot n^x \leqslant \ln \mathbb{P}\big(L(n) < k\big) \leqslant c_9 \cdot n^x$$

*for large* $n$, *where* $c_8$ *and* $c_9$ *are two* (*uniform*) *negative constants.*

Proof. With $k(n) = [(1-x)\log_{1/p_{11}} n]$, it follows from Lemma 2.1 that

$$\begin{aligned}&\mathbb{P}\big(L(n) < k\big) \geqslant 1 - p_{11}^{k-1}\,[c_1 \cdot (n-k) + c_2] - c(n,k)\\ &\geqslant 1 - \text{const}_1 \cdot p_{11}^k (n-k) - \text{const}_2 \cdot |p_{00} - p_{10}|^n - \text{const}_3 \cdot p_{11}^k |p_{00} - p_{10}|^{n-k}.\end{aligned}$$

If we apply the inequality $\ln(1-a) \geqslant -2a$ for $0 < a < 1/2$, then

$$\begin{aligned}&\ln \mathbb{P}\big(L(n) < k\big)\\ &\geqslant -2\,\text{const}_1\, p_{11}^k (n-k) - 2\,\text{const}_2 |p_{00} - p_{10}|^n - 2\,\text{const}_3\, p_{11}^k |p_{00} - p_{10}|^{n-k}\\ &\geqslant \text{const} \cdot n^{-x}.\end{aligned}$$

The upper bound is similarly proved with the help of the arguments in the proof of Lemma 2.3. ■

## 3. MOMENT GENERATING FUNCTION AND LARGE DEVIATIONS

In this section, we first give a proof of Theorem 1.3 regarding the precise logarithmic asymptotics for the moment generating function, which is the main result of the paper. Then, using this proved result, we derive two large deviation principles (Theorems 1.2 and 1.4) with the help of the Gärtner–Ellis theorem. At the end, a very concise proof of Theorem 1.1 is included.

### 3.1. Proof of Theorem 1.3.

Step 1. The following estimate holds for all $\lambda \in \mathbb{R}$:

$$\liminf_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \geqslant \lambda.$$

The case when $\lambda = 0$ is trivial. If $\lambda > 0$, then

$$\begin{aligned}&\frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\}\\ &\quad\geqslant \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E}\bigg(\exp\{\lambda \cdot L(n)\}, \bigg\{\bigg|\frac{L(n)}{\log_{1/p_{11}} n} - 1\bigg| \leqslant \varepsilon\bigg\}\bigg)\\ &\quad\geqslant \frac{1}{\log_{1/p_{11}} n} \ln \exp\{\lambda \cdot (1-\varepsilon)\log_{1/p_{11}} n\} \cdot \mathbb{P}\bigg(\bigg|\frac{L(n)}{\log_{1/p_{11}} n} - 1\bigg| \leqslant \varepsilon\bigg)\\ &\quad= \lambda \cdot (1-\varepsilon) + \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\bigg(\bigg|\frac{L(n)}{\log_{1/p_{11}} n} - 1\bigg| \leqslant \varepsilon\bigg).\end{aligned}$$

Since $L(n)/\log_{1/p_{11}} n$ converges to one almost surely, we have

$$\lim_{\varepsilon\to 0^+} \liminf_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \geqslant \lim_{\varepsilon\to 0^+} \lambda \cdot (1-\varepsilon) = \lambda.$$

If $\lambda < 0$, a similar argument as above yields

$$\lim_{\varepsilon\to 0^+} \liminf_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \geqslant \lim_{\varepsilon\to 0^+} \lambda \cdot (1+\varepsilon) = \lambda.$$

S t e p 2. The following estimate holds for $\lambda < \ln(1/p_{11})$:

$$\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \leqslant \lambda.$$

To see this, we first rewrite

$$\begin{aligned} &\ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \\ &= \ln \mathbb{E}\bigg( \exp\{\lambda \cdot L(n)\}, \bigg\{ \bigg| \frac{L(n)}{\log_{1/p_{11}} n} - 1 \bigg| \leqslant \varepsilon \bigg\} \cup \bigg\{ \bigg| \frac{L(n)}{\log_{1/p_{11}} n} - 1 \bigg| > \varepsilon \bigg\} \bigg). \end{aligned}$$

Therefore,

$$\begin{aligned} (3.1)\quad &\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \\ &= \max\bigg\{ \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E}\bigg( \exp\{\lambda \cdot L(n)\}, \bigg\{ \bigg| \frac{L(n)}{\log_{1/p_{11}} n} - 1 \bigg| \leqslant \varepsilon \bigg\} \bigg), \\ &\qquad \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E}\bigg( \exp\{\lambda \cdot L(n)\}, \bigg\{ \bigg| \frac{L(n)}{\log_{1/p_{11}} n} - 1 \bigg| > \varepsilon \bigg\} \bigg) \bigg\}. \end{aligned}$$

It is clear that the first limit satisfies

$$\begin{aligned} (3.2)\quad &\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E}\bigg( \exp\{\lambda \cdot L(n)\}, \bigg\{ \bigg| \frac{L(n)}{\log_{1/p_{11}} n} - 1 \bigg| \leqslant \varepsilon \bigg\} \bigg) \\ &\qquad\qquad \leqslant \begin{cases} \lambda(1+\varepsilon), & \lambda > 0, \\ \lambda(1-\varepsilon), & \lambda < 0. \end{cases} \end{aligned}$$

The second limit is more complicated, and the assumption $\lambda < \ln(1/p_{11})$ is needed. We rewrite

$$
\begin{aligned}
&\ln \mathbb{E}\left( \exp\{\lambda \cdot L(n)\}, \left\{ \left| \frac{L(n)}{\log_{1/p_{11}} n} - 1 \right| > \varepsilon \right\} \right) \\
&= \ln \mathbb{E}\left( \exp\{\lambda \cdot L(n)\}, \left\{ \frac{L(n)}{\log_{1/p_{11}} n} - 1 > \varepsilon \right\} \cup \left\{ \frac{L(n)}{\log_{1/p_{11}} n} - 1 < -\varepsilon \right\} \right).
\end{aligned}
$$

On the first part $\left\{ \frac{L(n)}{\log_{1/p_{11}} n} - 1 > \varepsilon \right\}$, if $\lambda < 0$, then similar things can be done as above. But if $\lambda > 0$, then we need to make the following separation:

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E}\left( \exp\{\lambda \cdot L(n)\}, \left\{ \frac{L(n)}{\log_{1/p_{11}} n} - 1 > \varepsilon \right\} \right) \\
&= \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \\
&\quad \times \ln \mathbb{E}\left( \exp\{\lambda \cdot L(n)\}, \bigcup_{k=1}^{\infty} \left\{ 1 + k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1 + (k+1)\varepsilon \right\} \right) \\
&\leqslant \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \left( \sum_{k=1}^{\infty} e^{\lambda[1+(1+k)\varepsilon] \log_{1/p_{11}} n} \cdot \mathbb{P}\left( 1 + k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \right) \right) \\
&= \lambda(1+\varepsilon) + \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \left( \sum_{k=1}^{\infty} e^{\lambda k\varepsilon \log_{1/p_{11}} n} \cdot \mathbb{P}\left( 1 + k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \right) \right).
\end{aligned}
$$

It now follows from Lemma 2.3 that

$$
\mathbb{P}\left( 1 + k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \right) = 1 - \mathbb{P}\left( \frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1 + k\varepsilon \right) \leqslant \text{const} \cdot n^{-k\varepsilon},
$$

which gives

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E}\left( \exp\{\lambda \cdot L(n)\}, \left\{ \frac{L(n)}{\log_{1/p_{11}} n} - 1 > \varepsilon \right\} \right) \\
&= \lambda(1+\varepsilon) \\
&\quad + \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \left( \sum_{k=1}^{\infty} e^{\lambda k\varepsilon \log_{1/p_{11}} n} \cdot \mathbb{P}\left( 1 + k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \right) \right) \\
&\leqslant \lambda(1+\varepsilon) + \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \left( \sum_{k=1}^{\infty} e^{\lambda k\varepsilon \log_{1/p_{11}} n} \cdot n^{-k\varepsilon} \right) \\
&= \lambda(1+\varepsilon) + \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \left( \sum_{k=1}^{\infty} n^{-(1 - \frac{\lambda}{\ln(1/p_{11})})k\varepsilon} \right) \\
&\leqslant \lambda(1+\varepsilon),
\end{aligned}
$$

where the last step follows from the fact that $\lambda < \ln(1/p_{11})$. Namely, we have proved that

(3.3)
$$\limsup_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\ln\mathbb{E}\bigg(\exp\{\lambda\cdot L(n)\},\bigg\{\frac{L(n)}{\log_{1/p_{11}}n}-1>\varepsilon\bigg\}\bigg)\leqslant\lambda(1+\varepsilon).$$

On the second part $\Big\{\frac{L(n)}{\log_{1/p_{11}}n}-1<-\varepsilon\Big\}$, the case when $\lambda>0$ can be similarly handled. For the case $\lambda<0$, we can do a similar separation to that in the proof of (3.3), but the argument here is a little different. We have

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\ln\mathbb{E}\bigg(\exp\{\lambda\cdot L(n)\},\bigg\{\frac{L(n)}{\log_{1/p_{11}}n}-1<-\varepsilon\bigg\}\bigg)\\
&=\limsup_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\\
&\quad\times\ln\mathbb{E}\bigg(\exp\{\lambda\cdot L(n)\},\ \bigcup_{k=1}^{[1/\varepsilon]-1}\bigg\{1-(k+1)\varepsilon<\frac{L(n)}{\log_{1/p_{11}}n}\leqslant1-k\varepsilon\bigg\}\bigg)\\
&\leqslant\limsup_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\\
&\quad\times\ln\bigg(\sum_{k=1}^{[1/\varepsilon]-1}e^{\lambda[1-(k+1)\varepsilon]\log_{1/p_{11}}n}\cdot\mathbb{P}\bigg(1-(k+1)\varepsilon<\frac{L(n)}{\log_{1/p_{11}}n}\leqslant1-k\varepsilon\bigg)\bigg).
\end{aligned}$$

Since there are only finite terms in the summation, we can simplify the above quantity, noticing that it is less than or equal to

$$\begin{aligned}
&\max_{1\leqslant k\leqslant[1/\varepsilon]-1}\bigg\{\lambda[1-(k+1)\varepsilon]+\limsup_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\ln\mathbb{P}\bigg(\frac{L(n)}{\log_{1/p_{11}}n}<1-k\varepsilon\bigg)\bigg\}\\
&\qquad=\max_{1\leqslant k\leqslant[1/\varepsilon]-1}\{\lambda[1-(k+1)\varepsilon]-\infty\}=-\infty,
\end{aligned}$$

where the '$-\infty$' appears because of Lemma 2.4. Therefore,

(3.4)
$$\limsup_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\ln\mathbb{E}\bigg(\exp\{\lambda\cdot L(n)\},\bigg\{\frac{L(n)}{\log_{1/p_{11}}n}-1<-\varepsilon\bigg\}\bigg)=-\infty.$$

Now the proof is done by taking the estimates (3.2), (3.3) and (3.4) back into (3.1).

S t e p 3. If $\lambda=\ln(1/p_{11})$, then

$$\lim_{n\to\infty}\frac{1}{\log_{1/p_{11}}n}\ln\mathbb{E}\exp\{\lambda\cdot L(n)\}=2\lambda.$$

On the one hand, it follows from Lemma 2.3 that, for every $\varepsilon>0$,

$$
\begin{aligned}
&\frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp \{\lambda \cdot L(n)\} \\
&\geqslant \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp \left\{\lambda \cdot L(n), \left\{\frac{L(n)}{\log_{1/p_{11}} n} > 1+\varepsilon\right\}\right\} \\
&= \frac{1}{\log_{1/p_{11}} n} \\
&\quad \times \ln \mathbb{E} \exp \left\{\lambda \cdot L(n), \bigcup_{k=1}^{q} \left\{1 + k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1 + (k+1)\varepsilon\right\}\right\} =: K,
\end{aligned}
$$

where (and in the sequel) we put

$$
q = \left[\frac{1}{\varepsilon}\left(\frac{n}{\log_{1/p_{11}} n} - 1\right)\right].
$$

Now we have

$$
\begin{aligned}
K &\geqslant \frac{1}{\log_{1/p_{11}} n} \ln \sum_{k=1}^{q} \exp\{(1+k\varepsilon)(\log_{1/p_{11}} n) \cdot \ln(1/p_{11})\} \\
&\quad \times \left(\mathbb{P}\left\{\frac{L(n)}{\log_{1/p_{11}} n} > 1 + k\varepsilon\right\} - \mathbb{P}\left\{\frac{L(n)}{\log_{1/p_{11}} n} > 1 + (k+1)\varepsilon\right\}\right) \\
&\geqslant \ln(1/p_{11}) + \frac{1}{\log_{1/p_{11}} n} \ln \sum_{k=1}^{q} n^{k\varepsilon}\Big(c_6 \cdot n^{-(1+k\varepsilon)}\big(n - (1+k\varepsilon)\log_{1/p_{11}} n\big) \\
&\qquad - c_7 \cdot n^{-(1+(k+1)\varepsilon)}\Big(n - \big(1 + (k+1)\varepsilon\big)\log_{1/p_{11}} n\Big)\Big) \\
&= \ln(1/p_{11}) + \frac{1}{\log_{1/p_{11}} n} \ln \sum_{k=1}^{q} \Big(\frac{c_6}{n}\big(n - (1+k\varepsilon)\log_{1/p_{11}} n\big) \\
&\qquad - \frac{c_7}{n^{1+\varepsilon}}\Big(n - \big(1 + (k+1)\varepsilon\big)\log_{1/p_{11}} n\Big)\Big) \\
&\sim \ln(1/p_{11}) + \frac{1}{\log_{1/p_{11}} n} \ln \left[\frac{c_6}{n} \cdot \frac{n^2}{2\varepsilon \log_{1/p_{11}} n} - \frac{c_7}{n^{1+\varepsilon}} \cdot \frac{n^2}{2\varepsilon \log_{1/p_{11}} n}\right] \\
&\sim \ln(1/p_{11}) + \frac{1}{\log_{1/p_{11}} n} \ln \left[\frac{c_6}{n} \cdot \frac{n^2}{2\varepsilon \log_{1/p_{11}} n}\right] \\
&\sim \ln(1/p_{11}) + \ln(1/p_{11}) = 2\ln(1/p_{11}).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp \{\lambda \cdot L(n)\} \\
&= \max\left\{\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp \left\{\lambda \cdot L(n), \left\{\frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1+\varepsilon\right\}\right\},\right. \\
&\qquad \left.\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp \left\{\lambda \cdot L(n), \left\{\frac{L(n)}{\log_{1/p_{11}} n} > 1+\varepsilon\right\}\right\}\right\}.
\end{aligned}
$$

The first limit is estimated as

$$\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\left\{\lambda \cdot L(n), \left\{\frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1+\varepsilon\right\}\right\}$$
$$\leqslant \limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \exp\{(1+\varepsilon)\log_{1/p_{11}} n \cdot \ln(1/p_{11})\} \mathbb{P}\left\{\frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1+\varepsilon\right\}$$
$$= (1+\varepsilon)\ln(1/p_{11}).$$

The second limit is estimated as

$$\frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\left\{\lambda \cdot L(n), \left\{\frac{L(n)}{\log_{1/p_{11}} n} > 1+\varepsilon\right\}\right\}$$
$$= \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\left\{\lambda \cdot L(n), \bigcup_{k=1}^{q} \left\{1+k\varepsilon < \frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1+(k+1)\varepsilon\right\}\right\}$$
$$\leqslant \frac{1}{\log_{1/p_{11}} n} \ln \sum_{k=1}^{q} \exp\left\{\big(1+(k+1)\varepsilon\big)\log_{1/p_{11}} n \cdot \ln(1/p_{11})\right\}$$
$$\times \mathbb{P}\left\{\frac{L(n)}{\log_{1/p_{11}} n} > 1+k\varepsilon\right\}$$
$$\leqslant (1+\varepsilon)\ln(1/p_{11})$$
$$+ \frac{1}{\log_{1/p_{11}} n} \ln \sum_{k=1}^{q} n^{k\varepsilon} \cdot c_7 \cdot n^{-(1+k\varepsilon)}\big(n-(1+k\varepsilon)\log_{1/p_{11}} n\big)$$
$$\sim (1+\varepsilon)\ln(1/p_{11}) + \frac{1}{\log_{1/p_{11}} n} \ln \frac{c_7}{n} \cdot \frac{n^2}{\varepsilon \log_{1/p_{11}} n}$$
$$\sim (1+\varepsilon)\ln(1/p_{11}) + \ln(1/p_{11}).$$

Therefore,

$$\limsup_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \leqslant (1+\varepsilon)\ln(1/p_{11}) + \ln(1/p_{11}),$$

which completes the proof.

S t e p 4. In order to study the asymptotic behavior of $\mathbb{E}\exp\{\lambda \cdot L(n)\}$ when $\lambda > \ln(1/p_{11})$, we need to consider a large deviation probability which may be of independent interest.

LEMMA 3.1. *For a fixed* $0 < x < 1$, *we have*

$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \geqslant x\right) \geqslant -x\ln(1/p_{11})$$

*and*

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \geqslant x\right) \leqslant \max\left\{-x\ln(1/p_{11}),\ -\ln\frac{1}{|p_{00}-p_{10}|}\right\}.$$

*In particular*, *if* $p_{10} \leqslant p_{00} + p_{11}$, *then*

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \geqslant x\right) = -x \ln(1/p_{11}).$$

P r o o f o f L e m m a 3.1. We apply Lemma 2.2 with $k(n) = [nx]$ and obtain the following:

$$1 - (1 - c_4 \cdot p_{11}^k)^{n-k+1} \leqslant \mathbb{P}\left(\frac{L(n)}{n} \geqslant x\right) \leqslant 1 - (1 - c_5 \cdot p_{11}^k)^{n-k+1} + e(n).$$

The lower bound can be handled as

$$\begin{aligned} &1 - (1 - c_4 \cdot p_{11}^k)^{n-k+1} \\ &= 1 - [(1 - c_4 \cdot p_{11}^k)^{1/(c_4 \cdot p_{11}^k)}]^{c_4 \cdot p_{11}^k (n-k+1)} \\ &= -[(1 - c_4 \cdot p_{11}^k)^{1/(c_4 \cdot p_{11}^k)}]^{\theta_n} \ln\left((1 - c_4 \cdot p_{11}^k)^{1/(c_4 \cdot p_{11}^k)}\right) \cdot c_4 \cdot p_{11}^k (n-k+1), \end{aligned}$$

where $\theta_n \in [0, c_4 \cdot p_{11}^k (n-k+1)]$. Therefore, for big enough $n$, the lower bound satisfies

$$1 - (1 - c_4 \cdot p_{11}^k)^{n-k+1} \geqslant c_4 \cdot (1 - \delta) p_{11}^k (n-k+1)$$

for some small $\delta > 0$, which proves the lower bound. The upper bound can be handled similarly except for the extra term $e(n)$. In this case,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \ln |e(n)| \\ &\qquad \leqslant \limsup_{n\to\infty} \frac{1}{n} \ln[\mathrm{const}_1 \cdot |p_{00} - p_{10}|^n + \mathrm{const}_2 \cdot p_{11}^k |p_{00} - p_{10}|^{n-k}] \\ &\qquad \leqslant \max\left\{-\ln \frac{1}{|p_{00} - p_{10}|},\ -x \ln(1/p_{11})\right\}, \end{aligned}$$

from which the upper bound follows. ■

S t e p 5. If $\lambda > \ln(1/p_{11})$, then

$$\begin{aligned} \lambda - \ln(1/p_{11}) \leqslant \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{E} e^{\lambda L(n)} &\leqslant \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E} e^{\lambda L(n)} \\ &\leqslant \max\left\{\lambda - \ln(1/p_{11}),\ \lambda - \ln \frac{1}{|p_{00} - p_{10}|}\right\}. \end{aligned}$$

It follows from Lemma 3.1 that, for any $0 < x < 1$,

$$
\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \\
&\geqslant \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\Big[ \exp\{\lambda \cdot L(n)\}, \Big\{ \frac{L(n)}{n} > x \Big\}\Big] \\
&\geqslant \lambda x + \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\Big( \frac{L(n)}{n} > x \Big) \\
&= \lambda x - x \ln(1/p_{11}) = \lambda - \ln(1/p_{11}) \quad \text{ as } x \to 1.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E} \exp\{\lambda \cdot L(n)\} \\
&\quad = \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\Big( \exp\{\lambda \cdot L(n)\}, \Big\{ \frac{L(n)}{n} \leqslant \varepsilon \Big\} \cup \Big\{ \frac{L(n)}{n} > \varepsilon \Big\}\Big) \\
&\quad = \max\Big\{ \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\Big( \exp\{\lambda \cdot L(n)\}, \Big\{ \frac{L(n)}{n} \leqslant \varepsilon \Big\}\Big), \\
&\qquad\qquad \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\Big( \exp\{\lambda \cdot L(n)\}, \Big\{ \frac{L(n)}{n} > \varepsilon \Big\}\Big)\Big\}.
\end{aligned}
$$

The first limit is

$$
\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E} \left( \exp\{\lambda \cdot L(n)\}, \ \left\{ \frac{L(n)}{n} \leqslant \varepsilon \right\} \right) \leqslant \lambda\varepsilon.
$$

The second limit is handled as follows:

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E} \left( \exp\{\lambda \cdot L(n)\}, \ \left\{ \frac{L(n)}{n} > \varepsilon \right\} \right) \\
&= \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\Big( \exp\{\lambda \cdot L(n)\}, \ \bigcup_{k=1}^{[1/\varepsilon]-1} \Big\{ k\varepsilon < \frac{L(n)}{n} \leqslant (k+1)\varepsilon \Big\}\Big) \\
&= \max_{1\leqslant k\leqslant[1/\varepsilon]-1} \Big\{ \lambda(k+1)\varepsilon + \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\Big( k\varepsilon < \frac{L(n)}{n} \Big)\Big\} \\
&\leqslant \max_{1\leqslant k\leqslant[1/\varepsilon]-1} \Big\{ \lambda(k+1)\varepsilon + \max\Big\{ -k\varepsilon \ln(1/p_{11}), \ -\ln \frac{1}{|p_{00}-p_{10}|} \Big\}\Big\} \\
&= \max_{1\leqslant k\leqslant[1/\varepsilon]-1} \Big\{ \lambda \cdot \varepsilon + k\varepsilon\big(\lambda - \ln(1/p_{11})\big), \ \lambda(k+1)\varepsilon - \ln \frac{1}{|p_{00}-p_{10}|} \Big\} \\
&= \max\Big\{ \lambda - \ln(1/p_{11}) + \lambda \cdot \varepsilon, \ \lambda - \ln \frac{1}{|p_{00}-p_{10}|} \Big\}.
\end{aligned}
$$

The condition $\lambda > \ln(1/p_{11})$ is used when the maximum is attained with $k = [1/\varepsilon] - 1$. The proof now follows by taking $\varepsilon \to 0^+$.

**3.2. Proofs of Theorems 1.2 and 1.4.** Using the proved Theorem 1.3, we are now ready to prove Theorems 1.2 and 1.4 with the help of the Gärtner–Ellis theorem. The proofs of Theorems 1.2 and 1.4 are essentially the same, and here we only show the details for the one of Theorem 1.4. Let us define the logarithmic moment generating function of $L(n)/n$ as

$$\widetilde{\Lambda}_n(\lambda) = \ln \mathbb{E} \exp\{\lambda \cdot L(n)/n\}, \quad \lambda \in \mathbb{R},$$

and the *cumulant* as

$$\widetilde{\Lambda}(\lambda) := \lim_{n\to\infty} \frac{1}{n}\widetilde{\Lambda}_n(\lambda \cdot n) = \begin{cases} \lambda - \ln(1/p_{11}), & \lambda \geqslant \ln(1/p_{11}), \\ 0, & \lambda < \ln(1/p_{11}), \end{cases}$$

where the last limit is from Theorem 1.3, under the condition $p_{10} \leqslant p_{00} + p_{11}$. Then the large deviation upper bound (1.10) follows directly from the Gärtner–Ellis theorem (cf. [3], Section 2.3) with the rate function $\widetilde{\Lambda}^*$ in (1.8) defined by the Fenchel–Legendre transform of $\widetilde{\Lambda}$ as $\widetilde{\Lambda}^*(x) = \sup_{\lambda\in\mathbb{R}}[\lambda \cdot x - \widetilde{\Lambda}(\lambda)]$.

For the large deviation lower bound (1.9), it suffices to prove that for a fixed point $0 < y < 1$,

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \in B_{y,\delta}\right) \geqslant -y\ln(1/p_{11}), \tag{3.5}$$

where $B_{y,\delta}$ is the open ball centered at $y$ with a radius $\delta$. To achieve (3.5), we write

$$\mathbb{P}\left(\frac{L(n)}{n} \in B_{y,\delta}\right) = \mathbb{P}\left(\frac{L(n)}{n} > y - \delta\right) - \mathbb{P}\left(\frac{L(n)}{n} \geqslant y + \delta\right),$$

and apply an inequality in the form $\ln(a-b) \geqslant \ln(a) - \frac{b}{a-b}$ for $a > b > 0$ to show that

$$\begin{aligned} \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \in B_{y,\delta}\right) &\geqslant \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n}\Bigg( \ln\left[\mathbb{P}\left(\frac{L(n)}{n} > y - \delta\right)\right] \\ &\quad - \frac{\mathbb{P}\big(L(n)/n \geqslant y+\delta\big)}{\mathbb{P}\big(L(n)/n > y-\delta\big) - \mathbb{P}\big(L(n)/n \geqslant y+\delta\big)}\Bigg). \end{aligned} \tag{3.6}$$

Lemma 3.1 implies that the first limit is, under the assumption $p_{10} \leqslant p_{00} + p_{11}$,

$$\begin{aligned} \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \ln \left[\mathbb{P}\left(\frac{L(n)}{n} > y - \delta\right)\right] \\ = \lim_{\delta\to 0} -(y-\delta)\ln(1/p_{11}) = -y\ln(1/p_{11}). \end{aligned} \tag{3.7}$$

For the second ratio term, applying Lemma 3.1 twice gives

$$\begin{aligned}(3.8)\quad &\frac{\mathbb{P}\big(L(n)/n \geqslant y+\delta\big)}{\mathbb{P}\big(L(n)/n > y-\delta\big) - \mathbb{P}\big(L(n)/n \geqslant y+\delta\big)} \\ &\qquad = \frac{1}{\mathbb{P}\big(L(n)/n > y-\delta\big)/\mathbb{P}\big(L(n)/n \geqslant y+\delta\big) - 1} \\ &\qquad \leqslant \frac{1}{e^{(2\delta\ln(1/p_{11})-\varepsilon)n} - 1} \to 0,\end{aligned}$$

as $n \to \infty$, for sufficiently small $\varepsilon > 0$ with $2\delta \ln(1/p_{11}) - \varepsilon > 0$. Then (3.5) follows by taking (3.7) and (3.8) back into (3.6).

**3.3. Proof of Theorem 1.1.** The limit (1.1) comes directly from Lemma 2.3. For the limit (1.2), we apply Lemma 2.4 for each $0 < x < 1$ and obtain

$$\ln\left[-c_9 \cdot n^x\right] \leqslant \ln\left[-\ln \mathbb{P}\left(\frac{L(n)}{\log_{1/p_{11}} n} \leqslant 1 - x\right)\right] \leqslant \ln[-c_8 \cdot n^x].$$

Then the proof follows directly by taking the limit $\lim_{n\to\infty} 1/\log_{1/p_{11}} n$.

## 4. THE STEIN–CHEN METHOD

The aim of this section is to introduce the use of the Stein–Chen method in estimating the large deviation probabilities of $L(n)$ in [14], and point out a mistake in the proof of (1.4). It turns out that the employed Stein–Chen method is insufficient to prove such large deviation probabilities. Let us recall the limit (1.4): for all $x > 0$,

$$\lim_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\big(L(n) - \lfloor \log_{1/p_{11}} n\rfloor \geqslant x \cdot \log_{1/p_{11}} n\big) = -x \cdot \ln(1/p_{11}).$$

The idea used in the proof of (1.4) in [14] is to approximate the large deviation probabilities $\mathbb{P}\big(L(n) - \lfloor \log_{1/p_{11}} n\rfloor \geqslant x \cdot \log_{1/p_{11}} n\big)$ by the ones involving Poisson random variables, and then to control the error term using the Stein–Chen method.

By setting $k = \big\lfloor \lfloor \log_{1/p_{11}} n\rfloor + x \cdot \log_{1/p_{11}} n\big\rfloor + 1$, it was proved on p. 1947 of [14] that

$$\begin{aligned}(4.1)\quad &\Big|\mathbb{P}\big(L(n) - \lfloor \log_{1/p_{11}} n\rfloor \geqslant x \cdot \log_{1/p_{11}} n\big) \\ &\qquad - \big(1 - \exp\{-n\pi_1(1-p_{11})p_{11}^{k-1} + o(1)\}\big)\Big| \\ &\qquad \leqslant \text{Error}\Big(W(n), Po\big(\lambda(n)\big)\Big),\end{aligned}$$

where $\pi_1$ is a constant, $W(n)$ is a random variable depending on $n$, defined on p. 1941, and $Po\big(\lambda(n)\big)$ is a Poisson random variable whose intensity $\lambda(n)$, also depending on $n$, was defined on p. 1942. It was then proved that

$$\big(1 - \exp\{-n\pi_1(1-p_{11})p_{11}^{k-1} + o(1)\}\big) = O(1)n^{-x}.$$

The error term was estimated via the Stein–Chen method as

$$\mathrm{Error}\Big(W(n), Po\big(\lambda(n)\big)\Big) = O\bigg(\frac{\ln(n)}{n}\bigg).$$

It is then obvious true that if $0 < x < 1$, then the limit (1.4) holds since the error term (which is of order $O\big(\frac{\ln(n)}{n}\big)$) is smaller than $n^{-x}$. But the problem occurs when $x > 1$, since in this case the error term is much bigger than the target $n^{-x}$, and the limit in (1.4) is unclear. Therefore, while employing this method, the limit (1.4) is true only for $0 < x < 1$. Furthermore, the Stein–Chen method seems to be impossible to remove the restriction $0 < x < 1$ since it gives an error of power orders, while the target term $n^{-x}$ is also of power order which can be any size depending on $x$.

## 5. AN APPLICATION IN CONFIDENCE INTERVALS

Given simulations of the Markov chain $\{X_k\}_{1\leqslant k\leqslant n}$ with the transition matrix

$$\begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix},$$

the aim of this section is to make statistical inferences on the transition probabilities $p_{ij}$. Since our interest throughout the paper is the longest success run, we will apply Theorem 1.2 to study the confidence intervals of $p_{11}$.

Theorem 1.2 implies that for each $x \geqslant 1$,

$$\lim_{n\to\infty} \frac{1}{\log_{1/p_{11}} n} \ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p_{11}} n} \geqslant x\bigg) = -(x-1)\cdot \ln(1/p_{11}).$$

If $x = 1 - \ln(\alpha)/\ln(n)$ with a given small $\alpha > 0$, then it holds true asymptotically that $\mathbb{P}(p_{11} < e^{-(\ln(n)-\ln(\alpha))/L(n)}) = \alpha$. This suggests a $100(1-\alpha)\%$ lower confidence bound of $p_{11}$ as follows:

$$I_{p_{11}} = \bigg(\exp\bigg\{-\frac{\ln(n)-\ln(\alpha)}{\hat{L}(n)}\bigg\}, 1\bigg),$$

where $\hat{L}(n)$ is a point estimate of $L(n)$. A reasonable point estimate of $L(n)$ is the observed longest success run. We can also obtain a point estimate $\hat{p}_{11}$ of $p_{11}$ using the observed (state '1'$\rightarrow$ state '1') proportion. For estimating the transition probabilities in terms of confidence intervals, there are many existing (more complicated) methods (cf. [2] and [13] for instance), but the advantage of our method is that the lower confidence bound is very simple and neat involving only one observation $\hat{L}(n)$.

Below in Table 1 we have simulations for different transition matrices. Although the point estimate $\hat{p}_{11}$ does not work well, the derived lower confidence bound $I_{p_{11}}$ works really good. We chose the $p$ which is close to 1, since $\hat{p}_{11}$ is only a lower confidence bound. As the other transition probabilities change (see $T_2$ and $T_3$), the confidence interval $I_{p_{11}}$ does not change much. This is as expected since the observed longest success run $\hat{L}(n)$ is not supposed to change when the other transition probabilities change. Meanwhile, the point estimates $\hat{p}_{11}$ are quite different due to the fact that the Markov chain with $T_3$ will have more chance to stay at the state '0' when it is at '0' now.

TABLE 1. $100(1-\alpha)\%$ lower confidence bound of $p_{11}$.

| | | | | |
|---|---|---|---|---|
| | $T_1 = \begin{bmatrix} 0.4 & 0.6 \\ 0.05 & 0.95 \end{bmatrix}$ | $n = 1000$ | $\alpha = 0.05$ | |
| $\hat{p}_{11} = 0.8810$<br>$\hat{L}(n) = 111$<br>$I_{p_{11}} = (0.9146, 1)$ | $\hat{p}_{11} = 0.8650$<br>$\hat{L}(n) = 102$<br>$I_{p_{11}} = (0.9075, 1)$ | $\hat{p}_{11} = 0.8780$<br>$\hat{L}(n) = 190$<br>$I_{p_{11}} = (0.9492, 1)$ | $\hat{p}_{11} = 0.8900$<br>$\hat{L}(n) = 99$<br>$I_{p_{11}} = (0.9048, 1)$ | $\hat{p}_{11} = 0.8630$<br>$\hat{L}(n) = 127$<br>$I_{p_{11}} = (0.9250, 1)$ |
| | $T_2 = \begin{bmatrix} 0.4 & 0.6 \\ 0.02 & 0.98 \end{bmatrix}$ | $n = 1000$ | $\alpha = 0.05$ | |
| $\hat{p}_{11} = 0.9510$<br>$\hat{L}(n) = 302$<br>$I_{p_{11}} = (0.9677, 1)$ | $\hat{p}_{11} = 0.9450$<br>$\hat{L}(n) = 156$<br>$I_{p_{11}} = (0.9385, 1)$ | $\hat{p}_{11} = 0.9530$<br>$\hat{L}(n) = 259$<br>$I_{p_{11}} = (0.9625, 1)$ | $\hat{p}_{11} = 0.9500$<br>$\hat{L}(n) = 212$<br>$I_{p_{11}} = (0.9544, 1)$ | $\hat{p}_{11} = 0.9660$<br>$\hat{L}(n) = 319$<br>$I_{p_{11}} = (0.9694, 1)$ |
| | $T_3 = \begin{bmatrix} 0.9 & 0.1 \\ 0.02 & 0.98 \end{bmatrix}$ | $n = 1000$ | $\alpha = 0.05$ | |
| $\hat{p}_{11} = 0.8510$<br>$\hat{L}(n) = 227$<br>$I_{p_{11}} = (0.9573, 1)$ | $\hat{p}_{11} = 0.9250$<br>$\hat{L}(n) = 396$<br>$I_{p_{11}} = (0.9753, 1)$ | $\hat{p}_{11} = 0.9140$<br>$\hat{L}(n) = 203$<br>$I_{p_{11}} = (0.9524, 1)$ | $\hat{p}_{11} = 0.8200$<br>$\hat{L}(n) = 232$<br>$I_{p_{11}} = (0.9582, 1)$ | $\hat{p}_{11} = 0.7680$<br>$\hat{L}(n) = 155$<br>$I_{p_{11}} = (0.9381, 1)$ |

We remark that the lower confidence bound presented above is very conservative since Theorem 1.2 gives an equivalence up to logarithm. This can be seen from the coverage probabilities. From simulations, the coverage probabilities with the transition matrices $T_i, i = 1, 2, 3$, are all near $100\%$, which are much higher than the confidence coefficient $100(1 - 0.05)\%$.

It has been seen that Theorem 1.2 yields the lower confidence bound using $x \geqslant 1$. In the same way, Theorem 1.1 can give a two-sided confidence interval of $p_{11}$. Furthermore, hypothesis testings on $p_{11}$ can be done in a similar way.

**Acknowledgments.** The authors are grateful to Torkel Erhardsson, Takis Konstantopoulos and Xianyuan Wu for several comments, and to an anonymous referee for the insightful suggestions, especially the one on combining the proofs of Theorems 1.2 and 1.4.

**REFERENCES**


[1] N. Balakrishnan and M. V. Koutras, *Runs and Scans with Applications*, Wiley, New York 2002.

[2] E. J. Bedrick and J. Aragón, *Approximate confidence intervals for the parameters of a stationary binary Markov chain*, Technometrics 31 (4) (1989), pp. 437–448.

[3] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, corrected reprint of the second (1998) edition, Springer, Berlin 2010.

[4] P. Erdős and A. Rényi, *On a new law of large numbers*, J. Anal. Math. 23 (1970), pp. 103–111.

[5] P. Erdős and P. Révész, *On the length of the longest head-run*, in: *Topics in Information Theory*, I. Csiszár and P. Elias (Eds.), Colloq. Math. Soc. János Bolyai 16 (1975), pp. 219–228.

[6] S. Eryilmaz, *Some results associated with the longest run statistic in a sequence of Markov dependent trials*, Appl. Math. Comput. 175 (1) (2006), pp. 119–130.

[7] J. C. Fu, L. Wang, and W. Y. W. Lou, *On exact and large deviation approximation for the distribution of the longest run in a sequence of two-state Markov dependent trials*, J. Appl. Probab. 40 (2) (2003), pp. 346–360.

[8] L. Holst and T. Konstantopoulos, *Runs in coin tossing: A general approach for deriving distributions for functionals*, J. Appl. Probab. 52 (3) (2015), pp. 752–770.

[9] T. Konstantopoulos, Z. Liu, and X. Yang, *Laplace transform asymptotics and large deviation principles for longest success runs in Bernoulli trials*, J. Appl. Probab. 53 (3) (2016), pp. 747–764.

[10] Z. Liu and X. Yang, *A general large deviation principle for longest runs*, Statist. Probab. Lett. 110 (2016), pp. 128–132.

[11] Y. Mao, F. Wang, and X. Wu, *Large deviation behavior for the longest head run in an IID Bernoulli sequence*, J. Theoret. Probab. 28 (1) (2015), pp. 259–268.

[12] A. Rényi, *Foundations of Probability*, Holden-Day, San Francisco 1970.

[13] R. E. Young and A. L. Sweet, *Confidence intervals for transition probabilities in two-state Markov chains. I*, Comm. Statist. A—Theory Methods 11 (2) (1982), pp. 165–179.

[14] Y. Z. Zhang and X. Y. Wu, *Some results associated with the longest run in a strongly ergodic Markov chain*, Acta Math. Sin. (Engl. Ser.) 29 (10) (2013), pp. 1939–1948.



Zhenxia Liu
Department of Mathematics
Linköping University
581 83 Linköping, Sweden
*E-mail*: zhenxia.liu@liu.se

Xiangfeng Yang
Department of Mathematics
Linköping University
581 83 Linköping, Sweden
*E-mail*: xiangfeng.yang@liu.se



*Received on 20.8.2015;*
*revised version on 14.5.2017*