J. Appl. Prob. **53,** 747–764 (2016)
*doi:*10.1017/jpr.2016.38

LAPLACE TRANSFORM ASYMPTOTICS AND LARGE DEVIATION PRINCIPLES FOR LONGEST SUCCESS RUNS IN BERNOULLI TRIALS

TAKIS KONSTANTOPOULOS,* *Uppsala University*
ZHENXIA LIU ** *** AND
XIANGFENG YANG,** **** *Linköping University*

Abstract

The longest stretch $L(n)$ of consecutive heads in n independent and identically distributed coin tosses is seen from the prism of large deviations. We first establish precise asymptotics for the moment generating function of $L(n)$ and then show that there are precisely two large deviation principles, one concerning the behavior of the distribution of $L(n)$ near its nominal value $\log_{1/p} n$ and one away from it. We discuss applications to inference and to logarithmic asymptotics of functionals of $L(n)$.

Keywords: Large deviation principle; rate function; Fenchel–Legendre transform; Laplace transform; moment generating function; run; longest run; Bernoulli trial; confidence interval

2010 Mathematics Subject Classification: Primary 60F10
Secondary 44A10; 60G50; 60G70

1. Introduction

The earliest reference to the longest stretch of consecutive successes in 'random' trials is (as we learn in the 1981 English translation [16, p. 138] of the 1928 book of von Mises) in a 1916 paper of the German philosopher Karl Marbe and concerns the longest stretch of consecutive births of children of the same sex as appearing in the birth register of a Bavarian town. (This was actually used by parents to 'predict' the sex of their child.) The longest stretch of same-sex births in two hundred thousand birth registrations was actually $17 \approx \log_2(200 \times 10^3)$. Von Mises [15] was apparently the first to study the problem rigorously and his result can be seen in Feller [6, Section XIII.12].

If $X_1, X_2, \dots$ are independent and identically distributed (i.i.d.) Bernoulli trials, $\mathbb{P}(X_i = 1) = p$, $\mathbb{P}(X_i = 0) = q := 1 - p$, and if $L(n)$ is the largest ℓ such that $X_{i+1} + \cdots + X_{i+\ell} = \ell$ for some $0 \le i \le n - \ell$, then we call the base-$1/p$ logarithm $\log_{1/p} n$ of n the *nominal value* of $L(n)$ because, as Erdős and Rényi [4] showed (in a more general setup in fact; see also [5] and [12]),

$$\lim_{n\to\infty} \frac{L(n)}{\log_{1/p} n} = 1 \quad \text{almost surely.} \tag{1}$$

Received 16 July 2015; revision received 22 October 2015.
* Postal address: Department of Mathematics, Uppsala University, SE-751 06 Uppsala, Sweden.
Email address: takiskonst@gmail.com
** Postal address: Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden.
*** Email address: zhenxia.liu@liu.se
**** Email address: xiangfeng.yang@liu.se

The distribution of $L(n)$ is not explicit. Yet, there are many estimates. The literature is littered with them and one of us recently contributed to it in [8] (where other quantities, such as the number of times that longest or shortest runs occur, are also explored).

Our principal interest in this paper is to see to what extent large deviations theory can be applied to the problem of squeezing something useful about the distribution of $L(n)$. We first establish logarithmic asymptotics for the moment generating function $\mathbb{E}\exp(\lambda L(n))$ as $n \to \infty$. The asymptotics split into three parts: the subcritical regime, $\lambda < \ln(1/p)$, the supercritical regime, $\lambda > \ln(1/p)$, and the critical regime when $\lambda = \ln(1/p)$. These asymptotics can be used in combination with the Gärtner–Ellis theorem (but see Remark 2 below) to derive a full large deviations principle (LDP). There are precisely two LDPs. One concerning the behavior of the distribution of $L(n)$ near its nominal value $\log_{1/p} n$ and another far away from it.

We outline the results below. Our starting point is asymptotics for the moment generating function and this is what we do right away. Note that we use ln for natural logarithm and $\log_b$ for logarithm with base b. The symbol $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$. Note also that we use the term 'Laplace transform' interchangeably with the term 'moment generating function'. (The variable λ ranges over the whole real line.)

Theorem 1. *The moment generating function of $L(n)$ has the following asymptotics.*

(i) *Subcritical regime: for $\lambda < \ln(1/p)$,*

$$\ln \mathbb{E}\exp\{\lambda L(n)\} \sim \lambda \log_{1/p} n.$$

(ii) *Critical regime: for $\lambda = \ln(1/p)$,*

$$\ln \mathbb{E}\exp\{\lambda L(n)\} \sim 2\lambda \log_{1/p} n.$$

(iii) *Supercritical regime: for $\lambda > \ln(1/p)$,*

$$\ln \mathbb{E}\exp\{\lambda L(n)\} \sim \left(\lambda - \ln\left(\frac{1}{p}\right)\right)n.$$

To the best of the authors' knowledge, the asymptotics on the moment generating function in Theorem 1 have not explicitly appeared in the literature. To show Theorem 1 there are several options. One option is the use of the recursion formula

$$\mathbb{E}\exp\{\lambda L(n)\} = q\sum_{j=0}^{n-1} p^j \mathbb{E}\exp(\lambda \max\{L(n-j-1), j\}) + p^n \exp(\lambda n),$$

appearing in [8]. Another possible option is to use Fibonacci-type polynomials, as appearing in the combinatorially-derived expressions for the moment generating function in [11]. But the simplest method is a good estimate for the distribution of $L(n)$; see Lemma 2. Why this lemma works to establish the asymptotics in the subcritical and critical regimes is the subject of Section 2 (Lemmas 3 and 4).

One implication of Theorem 1 is that it immediately suggests the form of large deviations of $L(n)$. In [7], a large deviations type probability was established in the following form:

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(L(n) < k) = -\beta \tag{2}$$

for a fixed k, where β is positive constant. Since, however, $\log_{1/p} n$ is the nominal value of $L(n)$, in the sense that (1) holds, the limit (2) is not strictly speaking a result in the theory of large deviations since it is not about the deviation from the most probable point $\log_{1/p} n$ of the random variables $L(n)$. A partial answer was recently included in [9] where it was proved that

$$\lim_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \geq 1+x\bigg) = -x \ln\bigg(\frac{1}{p}\bigg), \qquad x > 0. \tag{3}$$

Despite the fact that the research on head runs is a classical topic with many applications (see, for instance, [1]), no explicit general LDPs can be found in the literature.

The subcritical asymptotics of Theorem 1 correspond to the convergence $L(n)/\log_{1/p} n \to 1$ almost surely as $n \to \infty$. Therefore, we can study the large deviations on $L(n)/\log_{1/p} n$. Let us first define the function $\Lambda^*(x)$ as

$$\Lambda^*(x) = \begin{cases} +\infty, & x < 1, \\ (x-1)\ln\bigg(\dfrac{1}{p}\bigg), & x \geq 1. \end{cases} \tag{4}$$

Note that Λ^* is lower semicontinuous with $\{x \in \mathbb{R} \colon \Lambda^*(x) \leq c\}$ compact for all $c \geq 0$. This means that Λ^* is a *good* rate function (in the terminology of [3]). Our references to large deviations theory are Dembo and Zeitouni [3] and Wentzell [17]. The following full LDP is obtained as a corollary to Theorem 1.

Corollary 1. (LDP near the nominal value.) *The normalized longest head run $L(n)/\log_{1/p} n$ satisfies a large deviation principle with a good rate function $\Lambda^*(x)$ given by (4) and speed $\log_{1/p} n$. Namely,*

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

$$\varliminf_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \in O\bigg) \geq -\inf_{x\in O} \Lambda^*(x); \tag{5}$$

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

$$\varlimsup_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \in F\bigg) \leq -\inf_{x\in F} \Lambda^*(x). \tag{6}$$

Remark 1. Evidently, the large deviation principle presented in Corollary 1 generalizes the result (3) due to [9], which comes from choosing the open set $O = (1+x, \infty)$ and the closed set $F = [1+x, \infty)$.

Remark 2. (*Connections with the Gärtner–Ellis theorem.*) The proof of the large deviation upper bound (6) comes directly from the Gärtner–Ellis theorem (cf. [3]). We note that the rate function Λ^* is the Fenchel–Legendre transform of the following function:

$$\Lambda(\lambda) = \begin{cases} +\infty, & \lambda > \ln(1/p), \\ 2\lambda, & \lambda = \ln(1/p), \\ \lambda, & \lambda < \ln(1/p), \end{cases}$$

that is, $\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}}[\lambda x - \Lambda(\lambda)]$. There is a slight catch here. In order to establish the lower bound, the Gärtner–Ellis theorem requires that the function Λ be essentially smooth; namely, that $\lim_{k\to\infty} |\Lambda'(\lambda_k)| = \infty$ as $\lambda_k \to \ln(1/p)$. But this does not hold here. Therefore, the Gärtner–Ellis theorem does not cover our case. If, instead, we look at the lower bound proposed in the Gärtner–Ellis theorem, then we have for any open set O,

$$\underline{\lim}_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{P}\left(\frac{L(n)}{\log_{1/p} n} \in O\right) \geq -\inf_{x\in O\cap H} \Lambda^*(x),$$

where H is the so called *set of* exposed points [3, p. 44] of Λ^*. In our case, it is easy to see that the set H consists of only one point $H = \{1\}$. So the proposed lower bound from the Gärtner–Ellis theorem becomes trivial since

$$\inf_{x\in O\cap H} \Lambda^*(x) = \Lambda^*(1) = 0.$$

In summary, our large deviation principle in Theorem 1 gives a nontrivial example which the Gärtner–Ellis theorem does not cover.

The supercritical regime of Theorem 1 gives another large deviation result with a good rate function $\widetilde{\Lambda}^*(x)$ defined by

$$\widetilde{\Lambda}^*(x) = \begin{cases} +\infty, & x < 0, \\ x \ln\left(\dfrac{1}{p}\right), & 0 \leq x \leq 1, \\ +\infty, & x > 1. \end{cases} \tag{7}$$

Corollary 2. (LDP away from the nominal value.) *The normalized longest head run $L(n)/n$ satisfies a LDP with a good rate function $\widetilde{\Lambda}^*(x)$ given by (7) and speed n. Namely,*

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

$$\underline{\lim}_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \in O\right) \geq -\inf_{x\in O} \widetilde{\Lambda}^*(x);$$

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

$$\overline{\lim}_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \in F\right) \leq -\inf_{x\in F} \widetilde{\Lambda}^*(x).$$

Another implication of Theorem 1 and its Corollaries 1 and 2 is in obtaining asymptotics for other functionals of $L(n)$. We summarize the results as follows.

Corollary 3. (i) *If $f : \mathbb{R}_+ \to \mathbb{R}$ is continuous and satisfies one of the two conditions*

$$\lim_{m\to\infty} \overline{\lim}_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{E}\left[\exp\left\{\log_{1/p} n f\left(\frac{L(n)}{\log_{1/p} n}\right)\right\} \mathbf{1}_{\{f(L(n)/\log_{1/p} n)\geq m\}}\right] = -\infty, \tag{8}$$

$$\overline{\lim}_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{E} \exp\left\{\log_{1/p} n \gamma f\left(\frac{L(n)}{\log_{1/p} n}\right)\right\} < \infty \quad \textit{for some } \gamma > 1, \tag{9}$$

then it holds that

$$\lim_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{E} \exp\left\{\log_{1/p} n f\left(\frac{L(n)}{\log_{1/p} n}\right)\right\} = \max_{x\in\mathbb{R}} [f(x) - \Lambda^*(x)].$$

(ii) *If* $g\colon \mathbb{R}_+ \to \mathbb{R}$ *is continuous and satisfies one of the two conditions*

$$\lim_{m\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\ln\mathbb{E}\Big[\exp\Big\{ng\Big(\frac{L(n)}{n}\Big)\Big\}\mathbf{1}_{\{g(L(n)/n)\ge m\}}\Big] = -\infty, \tag{10}$$

$$\overline{\lim_{n\to\infty}}\frac{1}{n}\ln\mathbb{E}\exp\Big\{n\gamma g\Big(\frac{L(n)}{n}\Big)\Big\} < \infty \quad \text{for some } \gamma > 1, \tag{11}$$

then it holds that

$$\lim_{n\to\infty}\frac{1}{n}\ln\mathbb{E}\exp\Big\{ng\Big(\frac{L(n)}{n}\Big)\Big\} = \max_{x\in\mathbb{R}}[g(x) - \widetilde{\Lambda}^*(x)].$$

Here we list several functions f and g for which the conclusions of Corollary 3 hold. The verification is included in Section 4.

- $f(x)$ and $g(x)$ are continuous and bounded. In this case, (8)–(11) hold.
- $f(x) = cx^\alpha$, $x \in \mathbb{R}_+$, where $c > 0$ and $0 < \alpha < 1$. It is proved in Section 4 that (8) holds.
- $g(x)$ satisfies the condition: there is $m > 0$ such that if $|g(x)| \ge m$ then $x > 1$. For instance, with $c_1, c_2, c_3, c_4, \alpha$ positive constants, the functions

  $$c_1x^\alpha, \qquad c_2e^{c_3x^\alpha}, \qquad c_4\ln(x+\alpha)$$

  satisfy this condition. Condition (10) is fulfilled for this type of functions since

  $$\mathbf{1}_{\{g(L(n)/n)\ge m\}} \le \mathbf{1}_{\{L(n)/n>1\}} = 0.$$

Some easy conclusions of Theorem 1 concern well-known asymptotics for the moments of $L(n)$. Formally taking a derivative at $\lambda = 0$ of the expression in the subcritical regime gives

$$\mathbb{E}L(n)^k \sim (\log_{1/p} n)^k, \qquad k \in \mathbb{N}.$$

The asymptotic expressions of the first two moments can be found in [14], and the higher-order moments are discussed in [13, p. 63]. For convenience, we include the asymptotic mean as follows:

$$\mathbb{E}L(n) = \log_{1/p} n + \log_{1/p}(1-p) + \log_{1/p}(\exp(\gamma)) - \tfrac{1}{2} + \varepsilon(n), \tag{12}$$

where $\gamma = 0.5772\ldots$ is Euler's constant, and $\varepsilon(n)$ is 'small'.

The rest of the paper is organized as follows. In Section 2 we prove Theorem 1, along with some auxiliary results. In Section 3 we prove the large deviation principles, stated in Corollaries 1 and 2. Some other asymptotics related to Corollary 3 are given in 4. We discuss an application to inference in Section 5, and some open problems in Section 6. In order to simplify notation we use the abbreviation

$$\ell(n) := \log_{1/p} n$$

whenever convenient. As usual, we let $\lfloor x\rfloor$ to be the largest integer n such that $n \le x$ and $\lceil x\rceil$ to be the smallest integer n such that $n \ge x$.

2. Laplace transform asymptotics

We obtain logarithmic asymptotics for $\mathbb{E}\exp\{\lambda L(n)\}$, for all $\lambda \in \mathbb{R}$, in several steps. First, we obtain a lower bound valid for all $\lambda \in \mathbb{R}$. Then we obtain an upper bound for the subcritical case ($\lambda < \ln(1/p)$). These two bounds combined give the exact logarithmic asymptotics for the subcritical case. The limit in the critical case ($\lambda = \ln(1/p)$) requires special care and is treated separately. Finally, we obtain asymptotics for the supercritical case ($\lambda > \ln(1/p)$).

Lemma 1. *It holds that*

$$\varliminf_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{E} \exp\{\lambda L(n)\} \geq \lambda \quad \textit{for all } \lambda \in \mathbb{R}.$$

Proof. The $\lambda = 0$ case is trivial. Assume that $\lambda > 0$. Then, for $0 < \epsilon < 1$,

$$\begin{aligned}\mathbb{E} \exp\{\lambda L(n)\} &\geq \mathbb{E}[\exp\{\lambda L(n)\}; L(n) \geq (1-\epsilon) \log_{1/p} n] \\ &\geq \exp\{\lambda(1-\epsilon)\log_{1/p} n\} \mathbb{P}(L(n) \geq (1-\epsilon)\log_{1/p} n).\end{aligned}$$

Hence,

$$\frac{1}{\log_{1/p} n} \ln \mathbb{E} \exp\{\lambda L(n)\} \geq \lambda(1-\epsilon) + \frac{1}{\log_{1/p} n} \ln \mathbb{P}(L(n) \geq (1-\epsilon)\log_{1/p} n).$$

Since $\mathbb{P}(L(n) \geq (1-\epsilon)\log_{1/p} n) \to 1$,

$$\varliminf_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{E} \exp\{\lambda L(n)\} \geq \lambda(1-\epsilon),$$

and letting $\epsilon \downarrow 0$ we obtain the result. When $\lambda < 0$, we use

$$\mathbb{E} \exp\{\lambda L(n)\} \geq \mathbb{E}[\exp\{\lambda L(n)\}; L(n) \geq (1+\epsilon)\log_{1/p} n]$$

and proceed similarly. □

The following bound for the distribution of $L(n)$ is known in the literature, but we give a simple proof below for completeness.

Lemma 2. *For all* $k, n \in \mathbb{N}$, $1 \leq k \leq n$,

$$(1-p^k)^{n-k+1} \leq \mathbb{P}(L(n) < k) \leq (1-qp^k)^{n-k+1}.$$

Proof. Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{P}(X_1 = 1) = p$, $\mathbb{P}(X_1 = 0) = q$. Let $S_i = X_1 + \cdots + X_i$, $i \geq 1$. Note that $L(n) < k$ if and only if $S_m - S_{m-k} < k$ for all $k \leq m \leq n$. By a standard correlation inequality,

$$\mathbb{P}\bigg(\bigcap_{m=k}^{n} \{S_m - S_{m-k} < k\}\bigg) \geq \prod_{m=k}^{n} \mathbb{P}(S_m - S_{m-k} < k) = \prod_{m=k}^{n} (1-p^k) = (1-p^k)^{n-k+1},$$

and this is the lower bound. For the upper bound, since, trivially, $L(k-1) < k$, we have

$$\mathbb{P}(L(n) < k) = \prod_{m=k}^{n} \frac{\mathbb{P}(L(m) < k)}{\mathbb{P}(L(m-1) < k)}.$$

But, since, trivially again, $L(m) \geq L(m-1)$ for all m,

$$\mathbb{P}(L(m-1) < k) = \mathbb{P}(L(m) < k) + \mathbb{P}(L(m-1) < k \leq L(m)),$$

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

and observe that

$$\begin{aligned}\mathbb{P}(L(m-1)<k\le L(m)) &= \mathbb{P}(L(m-k-1)<k,\, X_{m-k}=0,\, X_{m-k+1}=\cdots=X_m=1)\\ &= \mathbb{P}(L(m-k-1)<k)qp^k\\ &\ge \mathbb{P}(L(m-1)<k)qp^k.\end{aligned}$$

Substituting this into the previous display gives $\mathbb{P}(L(m)<k)\le(1-qp^k)\mathbb{P}(L(m-1)<k)$ which implies that $\mathbb{P}(L(n)<k)\le\prod_{m=k}^{n}(1-qp^k)=(1-qp^k)^{n-k+1}$, as claimed. □

We next obtain an upper bound in the subcritical regime. Remember that $\ell(n)=:\log_{1/p}n$.

Lemma 3. *It holds that*

$$\varlimsup_{n\to\infty}\frac{1}{\log_{1/p}n}\ln\mathbb{E}\exp\{\lambda L(n)\}\le\lambda\quad \textit{for } -\infty<\lambda<\ln(1/p).$$

Proof. Suppose first that $0<\lambda<\ln(1/p)$, pick $\epsilon>0$, and write

$$\begin{aligned}\mathbb{E}\exp(\lambda L(n)) &= \mathbb{E}\bigg(\exp(\lambda L(n));\,\frac{L(n)}{\ell(n)}-1\le\epsilon\bigg)+\mathbb{E}\bigg(\exp(\lambda L(n));\,\frac{L(n)}{\ell(n)}-1>\epsilon\bigg)\\ &=: \mathsf{A}_+(n)+\mathsf{B}_+(n).\end{aligned}\tag{13}$$

The first term is estimated as

$$\mathsf{A}_+(n)\le\exp(\lambda(1+\epsilon)\ell(n))\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}-1\le\epsilon\bigg),\tag{14}$$

and so

$$\frac{\ln\mathsf{A}_+(n)}{\ell(n)}\le\lambda(1+\epsilon)+o(1),$$

implying that

$$\varlimsup_{n\to\infty}\frac{\ln\mathsf{A}_+(n)}{\ell(n)}\le\lambda.$$

For the second term, we write

$$\begin{aligned}\mathsf{B}_+(n) &:= \sum_{k=1}^{\infty}\mathbb{E}\bigg(\exp(\lambda L(n));\,1+k\epsilon<\frac{L(n)}{\ell(n)}\le1+(k+1)\epsilon\bigg)\\ &\le \sum_{k=1}^{\infty}\exp(\lambda(1+(k+1)\epsilon)\ell(n))\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}>1+k\epsilon\bigg).\end{aligned}$$

From Lemma 2, observe that

$$\mathbb{P}(L(n)\ge k)=1-\mathbb{P}(L(n)<k)\le1-(1-p^k)^{n-k+1}\le(n-k+1)p^k\le np^k$$

for all $0\le k\le n$, and, trivially, for all $k>n$ also. This implies that

$$\mathbb{P}(L(n)>t)\le np^t,\qquad t\ge0,$$

and so

$$P\bigg(\frac{L(n)}{\ell(n)}>1+k\epsilon\bigg)\le np^{(1+k\epsilon)\ell(n)}=nn^{-(1+k\epsilon)}=n^{-k\epsilon}.$$

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

Therefore,

$$\begin{aligned}\mathsf{B}_+(n) &\le \exp(\lambda(1+\epsilon)\ell(n))\sum_{k=1}^{\infty}\exp(\lambda k\epsilon\ell(n))n^{-k\epsilon}\\ &= \exp(\lambda(1+\epsilon)\ell(n))\sum_{k=1}^{\infty} n^{-(1-\lambda/\ln(1/p))k\epsilon}\\ &= \exp(\lambda(1+\epsilon)\ell(n))(n^{(1-(\lambda/\ln(1/p)))\epsilon}-1)^{-1},\end{aligned}$$

whence

$$\varlimsup_{n\to\infty}\frac{\ln \mathsf{B}_+(n)}{\ell(n)} \le \lambda.$$

Since

$$\varlimsup_{n\to\infty}\frac{\ln \mathbb{E}\exp(\lambda L(n))}{\ell(n)} = \max\bigg\{\varlimsup_{n\to\infty}\frac{\ln \mathsf{A}_+(n)}{\ell(n)},\ \varlimsup_{n\to\infty}\frac{\ln \mathsf{B}_+(n)}{\ell(n)}\bigg\},$$

the result follows.

Next suppose that $\lambda < 0$. For $0<\epsilon<1$, write

$$\begin{aligned}\mathbb{E}\exp(\lambda L(n)) &= \mathbb{E}\bigg(\exp(\lambda L(n));\ \frac{L(n)}{\ell(n)}-1 > -\epsilon\bigg) + \mathbb{E}\bigg(\exp(\lambda L(n));\ \frac{L(n)}{\ell(n)}-1 \le -\epsilon\bigg)\\ &=: \mathsf{A}_-(n)+\mathsf{B}_-(n).\end{aligned}$$

For the first term, we have

$$\mathsf{A}_-(n) \le \exp(\lambda(1-\epsilon)\ell(n))\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}-1>-\epsilon\bigg)$$

implying that

$$\varlimsup_{n\to\infty}\frac{\ln \mathsf{A}_-(n)}{\ell(n)} \le \lambda.$$

As for the second term,

$$\begin{aligned}\mathsf{B}_-(n) &= \sum_{k=1}^{\lfloor 1/\epsilon\rfloor-1}\mathbb{E}\bigg(\exp(\lambda L(n));\ 1-(k+1)\epsilon \le \frac{L(n)}{\ell(n)} < 1-k\epsilon\bigg)\\ &\le \sum_{k=1}^{\lfloor 1/\epsilon\rfloor-1}\exp(\lambda(1-(k+1)\epsilon)\ell(n))\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}<1-k\epsilon\bigg).\end{aligned}$$

Since there are only finitely many terms in the sum, we can simply write

$$\begin{aligned}\varlimsup_{n\to\infty}\frac{\ln \mathsf{B}_-(n)}{\ell(n)} &\le \max_{1\le k\le\lfloor 1/\epsilon\rfloor-1}\bigg\{\lambda(1-(k+1)\epsilon)+\varlimsup_{n\to\infty}\frac{1}{\ell(n)}\ln\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}<1-k\epsilon\bigg)\bigg\}\\ &\le \max_{1\le k\le\lfloor 1/\epsilon\rfloor-1}\{\lambda(1-(k+1)\epsilon)-\infty\}\\ &= -\infty,\end{aligned}$$

where $-\infty$ appears because of Lemma 7 below. We again conclude that

$$\varlimsup \ell(n)^{-1}\ln\mathbb{E}\exp(\lambda L(n)) \le \lambda. \qquad \square$$

The critical case is treated next.

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

Lemma 4. *When* $\lambda = \ln(1/p)$, *it holds that*

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

Proof. Fix sufficiently small $\epsilon > 0$. Using the probability estimates of Lemma 2, it follows that there exist positive constants c_1 and c_2 such that

$$\begin{aligned} c_2\, n^{-(1+k\epsilon)}(n+1-(1+k\epsilon)\ell(n)) &\le \mathbb{P}\bigg(\frac{L(n)}{\ell(n)} > 1+k\epsilon\bigg) \\ &\le c_1 n^{-(1+k\epsilon)}(n+1-(1+k\epsilon)\ell(n)) \end{aligned}$$

uniformly over all k such that

$$1 \le k \le \bigg\lfloor \frac{1}{\epsilon}\bigg(\frac{n}{\ell(n)} - 1\bigg)\bigg\rfloor =: N_n. \tag{15}$$

We first obtain a lower bound. From the estimate above,

$$\begin{aligned} \mathbb{E}\exp(\lambda L(n)) &\ge \mathbb{E}\bigg(\exp(\lambda L(n));\ \frac{L(n)}{\ell(n)} > 1+\epsilon\bigg) \\ &\ge \sum_{k=1}^{N_n} \mathbb{E}\bigg(\exp(\lambda L(n));\ 1+k\epsilon < \frac{L(n)}{\ell(n)} \le 1+(k+1)\epsilon\bigg) \\ &\ge \sum_{k=1}^{N_n} \mathbb{E}\bigg(\exp(\lambda \ell(n)(1+k\epsilon));\ 1+k\epsilon < \frac{L(n)}{\ell(n)} \le 1+(k+1)\epsilon\bigg). \end{aligned}$$

Since $\lambda = \ln(1/p)$ and $\ell(n) = (\ln n)/\ln(1/p)$, we have $\exp(\lambda\ell(n)) = \exp(\ln n) = n$. Hence,

$$\begin{aligned} \mathbb{E}\exp(\lambda L(n)) &\ge n\sum_{k=1}^{N_n} n^{k\epsilon}\bigg[\mathbb{P}\bigg(\frac{L(n)}{\ell(n)} > 1+k\epsilon\bigg) - \mathbb{P}\bigg(\frac{L(n)}{\ell(n)} > 1+(k+1)\epsilon\bigg)\bigg] \\ &\ge n\sum_{k=1}^{N_n} n^{k\epsilon}[c_2 n^{-(1+k\epsilon)}(n+1-(1+k\epsilon)\ell(n)) \\ &\qquad\qquad - c_1 n^{-(1+(k+1)\epsilon)}(n+1-(1+(k+1)\epsilon)\ell(n))] \\ &= n\sum_{k=1}^{N_n} [c_2 n^{-1}(n+1-(1+k\epsilon)\ell(n)) \\ &\qquad\qquad - c_1 n^{-(1+\epsilon)}(n+1-(1+(k+1)\epsilon)\ell(n))] \\ &=: n\mathsf{S}(n). \end{aligned}$$

Hence,

$$\frac{\ln \mathbb{E}\exp(\lambda L(n))}{\ell(n)} \ge \frac{\ln n}{\ell(n)} + \frac{\ln \mathsf{S}(n)}{\ell(n)} = \ln\bigg(\frac{1}{p}\bigg) + \ln\bigg(\frac{1}{p}\bigg)\frac{\ln \mathsf{S}(n)}{\ln n}.$$

We now claim that the last ratio converges to 1. This follows by direct computation:

$$\ln \mathsf{S}(n) \sim \ln\bigg[\frac{c_2 n}{2\epsilon \log_{1/p} n} - \frac{c_1 n^{1-\epsilon}}{2\epsilon \log_{1/p} n}\bigg] \sim \ln\bigg[\frac{c_2 n}{2\epsilon \log_{1/p} n}\bigg] = \ln n + o(\ln n).$$

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

Hence, we have proved a lower bound:

$$\varliminf_{n\to\infty} \frac{\ln \mathbb{E}\exp(\lambda L(n))}{\ell(n)} \geq 2\ln\left(\frac{1}{p}\right).$$

To obtain an upper bound, we use the decomposition (13) as in the proof of Lemma 3, but with $\lambda = \ln(1/p)$. The first term is estimated in precisely the same manner; see (14). Hence,

$$\varlimsup_{n\to\infty} \frac{\ln \mathsf{A}_+(n)}{\ell(n)} \leq \lambda = \ln\left(\frac{1}{p}\right). \tag{16}$$

For the second term, we write

$$\begin{aligned}\mathsf{B}_+(n) &= \mathbb{E}\left(\exp(\lambda L(n));\ \frac{L(n)}{\ell(n)} - 1 > \epsilon\right)\\ &= \sum_{k=1}^{N_n} \mathbb{E}\left(\exp(\lambda L(n));\ 1 + k\epsilon < \frac{L(n)}{\ell(n)} \leq 1 + (k+1)\epsilon\right),\end{aligned}$$

where N_n is as in (15), giving

$$\begin{aligned}\mathsf{B}_+(n) &\leq \sum_{k=1}^{N_n} \exp(\lambda\ell(n)[1 + (k+1)\epsilon])\mathbb{P}\left(\frac{L(n)}{\ell(n)} > 1 + k\epsilon\right)\\ &= n^{1+\epsilon}\sum_{k=1}^{N_n} n^{k\epsilon}\mathbb{P}\left(\frac{L(n)}{\ell(n)} > 1 + k\epsilon\right)\\ &\leq n^{1+\epsilon}\sum_{k=1}^{N_n} c_1 n^{-1}(n + 1 - (1 + k\epsilon)\ell(n)),\end{aligned}$$

from which

$$\frac{\ln \mathsf{B}_+(n)}{\ell(n)} \leq (1+\epsilon)\ln\left(\frac{1}{p}\right) + \ln\left(\frac{1}{p}\right)\frac{1}{\ln n}\ln\sum_{k=1}^{N_n} c_1 n^{-1}(n + 1 - (1 + k\epsilon)\ell(n)).$$

By direct computation,

$$\ln\sum_{k=1}^{N_n} c_1 n^{-1}(n + 1 - (1 + k\epsilon)\ell(n)) \sim \ln\left[\frac{c_1 n}{\epsilon\log_{1/p} n}\right] = \ln n + o(\ln n).$$

Combining the last two displays and letting $\epsilon \downarrow 0$, we obtain

$$\varlimsup_{n\to\infty} \frac{\ln \mathsf{B}_+(n)}{\ell(n)} \leq 2\ln\left(\frac{1}{p}\right). \tag{17}$$

From the decomposition (13), with the estimates (16) and (17), we conclude that

$$\begin{aligned}\varlimsup_{n\to\infty} \frac{\ln \mathbb{E}\exp(\lambda L(n))}{\ell(n)} &= \max\left\{\varlimsup_{n\to\infty} \frac{\ln \mathsf{A}_+(n)}{\ell(n)}, \varlimsup_{n\to\infty} \frac{\ln \mathsf{B}_+(n)}{\ell(n)}\right\}\\ &\leq \max\left\{\ln\left(\frac{1}{p}\right), 2\ln\left(\frac{1}{p}\right)\right\}\\ &= 2\ln\left(\frac{1}{p}\right).\end{aligned}$$

□

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

In order to study the asymptotic behavior of $\mathbb{E}\exp\{\lambda L(n)\}$ when $\lambda > \ln(1/p)$, we use the following result.

Lemma 5. *For fixed* $0 \le x \le 1$, *it holds that*

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{L(n)}{n} \ge x\right) = -x \ln\left(\frac{1}{p}\right).$$

Proof. The $x = 0$ case is trivial. For $x > 0$, we apply the inequalities of Lemma 2 with $k = \lceil nx \rceil$ and obtain

$$1 - (1 - qp^{\lceil nx \rceil})^{n-\lceil nx \rceil+1} \le \mathbb{P}\left(\frac{L(n)}{n} \ge x\right) \le 1 - (1 - p^{\lceil nx \rceil})^{n-\lceil nx \rceil+1}.$$

Since $1 - (1 - a)^N \le Na$ for all $0 \le a \le 1$, and since $1 - (1 - a)^N \ge (N - 1)a$ for all sufficiently small $a \ge 0$, we have

$$(n - \lceil nx \rceil)qp^{\lceil nx \rceil} \le \mathbb{P}\left(\frac{L(n)}{n} \ge x\right) \le (n - \lceil nx \rceil + 1)p^{\lceil nx \rceil} \quad \text{for all sufficiently large } n.$$

Taking logarithms, dividing by n, and sending n to ∞ finishes the proof. □

Lemma 6. *It holds that*

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\exp\{\lambda L(n)\} = \lambda - \ln\left(\frac{1}{p}\right) \quad \textit{for } \lambda > \ln(1/p).$$

Proof. For the lower bound, fix $0 < x < 1$, write

$$\mathbb{E}\exp(\lambda L(n)) \ge \mathbb{E}\left(\exp(\lambda L(n));\ \frac{L(n)}{n} > x\right) \ge \exp(\lambda x n)\mathbb{P}\left(\frac{L(n)}{n} > x\right),$$

and use Lemma 5 to obtain

$$\varliminf_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\exp(\lambda L(n)) \ge \lambda x - x \ln\left(\frac{1}{p}\right) \to \lambda - \ln\left(\frac{1}{p}\right) \quad \text{as } x \to 1.$$

For the upper bound, pick $\epsilon > 0$ and write

$$\begin{aligned}
\mathbb{E}\exp(\lambda L(n)) &= \mathbb{E}\left(\exp(\lambda L(n));\ \frac{L(n)}{n} \le \epsilon\right) + \mathbb{E}\left(\exp(\lambda L(n));\ \frac{L(n)}{n} > \epsilon\right) \\
&\le \exp(\lambda \epsilon n) + \sum_{k=1}^{[1/\epsilon]-1} \mathbb{E}\left(\exp(\lambda L(n));\ k\epsilon < \frac{L(n)}{n} \le (k+1)\epsilon\right) \\
&\le \exp(\lambda \epsilon n) + \sum_{k=1}^{[1/\epsilon]-1} \exp(\lambda(k+1)\epsilon n)\mathbb{P}\left(\frac{L(n)}{n} > k\epsilon\right).
\end{aligned}$$

Hence (with $a \vee b := \max(a, b)$),

$$\begin{aligned}
\varlimsup_{n\to\infty} \frac{1}{n}\mathbb{E}\exp(\lambda L(n)) &\le (\lambda\epsilon) \vee \max_{1\le k\le[1/\epsilon]-1} \left\{\lambda(k+1)\epsilon - k\epsilon \ln\left(\frac{1}{p}\right)\right\} \\
&\le \lambda\epsilon + \lambda - \ln\left(\frac{1}{p}\right) \\
&\to \lambda - \ln\left(\frac{1}{p}\right) \quad \text{as } \epsilon \to 0,
\end{aligned}$$

where we used Lemma 5 again and the assumption that $\lambda - \ln(1/p) > 0$. □

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

Lemma 7. (See [9, Theorem 1.1].) *For each* $x > 0$, *we have*

$$\lim_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \geq 1 + x\bigg) = -x \ln\bigg(\frac{1}{p}\bigg).$$

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

$$\lim_{n\to\infty} \frac{1}{\log_{1/p} n} \ln\bigg[-\ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \leq 1 - x\bigg)\bigg] = x \ln\bigg(\frac{1}{p}\bigg).$$

Note that this lemma can be simply derived based on Lemma 2, but what has actually been proved in [9] is precise asymptotics without the logarithm.

3. LDPs

We study the LDPs announced in Corollaries 1 and 2. Consider the logarithmic moment generating function of $L(n)/\log_{1/p} n$, defined by

$$\Lambda_n(\lambda) = \ln \mathbb{E} \exp\bigg\{\frac{\lambda L(n)}{\log_{1/p} n}\bigg\}, \qquad \lambda \in \mathbb{R}.$$

The proof of Corollary 1 is based on the *cumulant*; namely,

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{\log_{1/p} n} \Lambda_n(\lambda \log_{1/p} n). \tag{18}$$

That this limit exists is a direct consequence of Theorem 1.

Proposition 1. *The limit in (18) exists and is given by*

$$\Lambda(\lambda) = \begin{cases} +\infty, & \lambda > \ln(1/p), \\ 2\lambda, & \lambda = \ln(1/p), \\ \lambda, & \lambda < \ln(1/p). \end{cases}$$

The Fenchel–Legendre transform of Λ is the function $x \mapsto \sup_{\lambda\in\mathbb{R}}[\lambda x - \Lambda(\lambda)]$ which (as an easy calculation shows) is given by the function Λ^* defined in (4):

$$\sup_{\lambda\in\mathbb{R}}[\lambda x - \Lambda(\lambda)] = \Lambda^*(x) = \begin{cases} +\infty, & x < 1, \\ (x-1)\ln\Big(\dfrac{1}{p}\Big), & x \geq 1. \end{cases}$$

Proof of Corollary 1. To prove the upper bound (6) we apply the Gärtner–Ellis theorem (cf. [3, Section 2.3]). For the lower bound (5), we must provide a separate argument. It suffices to prove that for a fixed point $y > 1$,

$$\lim_{\delta\to 0} \varliminf_{n\to\infty} \frac{1}{\log_{1/p} n} \ln \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \in B_{y,\delta}\bigg) \geq -(y-1)\ln\bigg(\frac{1}{p}\bigg), \tag{19}$$

where $B_{y,\delta}$ is the open ball centered at y with a radius δ. To achieve (19), we write

$$\mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \in B_{y,\delta}\bigg) = \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} > y - \delta\bigg) - \mathbb{P}\bigg(\frac{L(n)}{\log_{1/p} n} \geq y + \delta\bigg).$$

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

In order to analyze the logarithm, we apply an inequality in the form

$$\ln(a-b) \geq \ln(a) - \frac{b}{a-b} \quad \text{for } a > b > 0.$$

Therefore,

$$\begin{aligned}
&\lim_{\delta\to 0}\varliminf_{n\to\infty}\frac{1}{\ell(n)}\ln\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}\in B_{y,\delta}\bigg)\\
&\quad\geq \lim_{\delta\to 0}\varliminf_{n\to\infty}\frac{1}{\ell(n)}\bigg(\ln\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}>y-\delta\bigg)\\
&\qquad\qquad - \frac{\mathbb{P}(L(n)/\ell(n)\geq y+\delta)}{\mathbb{P}(L(n)/\ell(n)>y-\delta)-\mathbb{P}(L(n)/\ell(n)\geq y+\delta)}\bigg). \qquad (20)
\end{aligned}$$

We can apply Lemma 7 to handle the first limit as follows:

$$\lim_{\delta\to 0}\varliminf_{n\to\infty}\frac{1}{\ell(n)}\ln\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}>y-\delta\bigg) = \lim_{\delta\to 0} -(y-1-\delta)\ln\bigg(\frac{1}{p}\bigg) = -(y-1)\ln\bigg(\frac{1}{p}\bigg). \quad (21)$$

For the last ratio term in (20), it follows from applying Lemma 7 twice that

$$\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}\geq y+\delta\bigg) \leq \exp\bigg\{\bigg[-(y-1+\delta)\ln\bigg(\frac{1}{p}\bigg)+\varepsilon_1\bigg]\ell(n)\bigg\}$$

and

$$\mathbb{P}\bigg(\frac{L(n)}{\ell(n)}> y-\delta\bigg) \geq \exp\bigg\{\bigg[-(y-1-\delta)\ln\bigg(\frac{1}{p}\bigg)-\varepsilon_2\bigg]\ell(n)\bigg\}$$

for sufficiently small $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$. Thus, assuming $2\delta\ln(1/p) - \varepsilon_1 - \varepsilon_2 > 0$,

$$\begin{aligned}
&\frac{\mathbb{P}(L(n)/\ell(n)\geq y+\delta)}{\mathbb{P}(L(n)/\ell(n)>y-\delta)-\mathbb{P}(L(n)/\ell(n)\geq y+\delta)}\\
&\quad = \frac{1}{\mathbb{P}(L(n)/\ell(n)>y-\delta)/\mathbb{P}(L(n)/\ell(n)\geq y+\delta)-1}\\
&\quad \leq \frac{1}{\exp((2\delta\ln(1/p)-\varepsilon_1-\varepsilon_2)\ell(n))-1}\\
&\quad \to 0 \quad \text{as } n\to\infty. \qquad (22)
\end{aligned}$$

Then (19) follows by substituting (21) and (22) back into (20). □

We now pass to the second LDP. Consider the logarithmic moment generating function of $L(n)/n$,

$$\widetilde{\Lambda}_n(\lambda) = \ln\mathbb{E}\exp\bigg\{\frac{\lambda L(n)}{n}\bigg\}, \qquad \lambda\in\mathbb{R},$$

and define its cumulant by

$$\widetilde{\Lambda}(\lambda) := \lim_{n\to\infty}\frac{1}{n}\Lambda_n(\lambda n). \quad (23)$$

It is again Theorem 1 that is responsible for the existence of the cumulant.

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

Proposition 2. *The limit in (23) exists and is given by*

$$\widetilde{\Lambda}(\lambda) = \begin{cases} \lambda - \ln\Big(\dfrac{1}{p}\Big), & \lambda \geq \ln(1/p), \\ 0, & \lambda < \ln(1/p). \end{cases}$$

An easy calculation shows that

$$\sup_{\lambda\in\mathbb{R}}[\lambda x - \widetilde{\Lambda}(\lambda)] = \widetilde{\Lambda}^*(x) = \begin{cases} +\infty, & x < 0, \\ x\ln\Big(\dfrac{1}{p}\Big), & 0 \leq x \leq 1, \\ +\infty, & x > 1, \end{cases}$$

which is the function announced in (7). The proof of Corollary 2 now proceeds along the same lines as that of Corollary 1 and is therefore omitted.

4. Exponential functionals

Proof of Corollary 3. The proof is straightforward and follows from Varadhan's integral lemma (cf. [3, Section 4.3]). □

We next verify that the function $f(x) = cx^{\alpha}, 0 < \alpha < 1$, satisfies the condition (8) in Corollary 3. Without loss of generality, we assume that $c > 0$ and obtain

$$\begin{aligned} &\frac{1}{\ell(n)} \ln \mathbb{E}\bigg[\exp\bigg(\ell(n) f\bigg(\frac{L(n)}{\ell(n)}\bigg)\bigg); f\bigg(\frac{L(n)}{\ell(n)}\bigg) \geq m\bigg] \\ &\quad = \frac{1}{\ell(n)} \ln \mathbb{E}\bigg[\exp\bigg(c\ell(n)\bigg(\frac{L(n)}{\ell(n)}\bigg)^{\alpha}\bigg); \bigg(\frac{L(n)}{\ell(n)}\bigg)^{\alpha} > m\bigg] \\ &\quad = \frac{1}{\ell(n)} \ln \sum_{k=0}^{\infty} \mathbb{E}\bigg[\exp\bigg(c\ell(n)\bigg(\frac{L(n)}{\ell(n)}\bigg)^{\alpha}\bigg); m+k < \bigg(\frac{L(n)}{\ell(n)}\bigg)^{\alpha} \leq m + (k+1)\bigg] \\ &\quad \leq \frac{1}{\ell(n)} \ln \sum_{k=0}^{\infty} \exp(c\ell(n)(m+k+1))\mathbb{P}\bigg(m+k < \bigg(\frac{L(n)}{\ell(n)}\bigg)^{\alpha}\bigg) \\ &\quad = c(m+1) + \frac{1}{\ell(n)} \ln \sum_{k=0}^{\infty} \exp(ck\ell(n))\mathbb{P}\bigg((m+k)^{1/\alpha} < \frac{L(n)}{\ell(n)}\bigg). \end{aligned}$$

We now apply Lemma 2 with $k = \lceil (m+k)^{1/\alpha}\ell(n)\rceil + 1$ and obtain

$$\begin{aligned} \mathbb{P}\bigg(\frac{L(n)}{\ell(n)} > (m+k)^{1/\alpha}\bigg) &= \mathbb{P}(L(n) > \lceil \ell(n)(m+k)^{1/\alpha}\rceil) \\ &= 1 - \mathbb{P}(L(n) < \lceil \ell(n)(m+k)^{1/\alpha}\rceil + 1) \\ &\leq 1 - (1 - p^{\lceil \ell(n)(m+k)^{1/\alpha}\rceil + 1})^{n - \lceil \ell(n)(m+k)^{1/\alpha}\rceil} \\ &\leq (n - \lceil \ell(n)(m+k)^{1/\alpha}\rceil) p^{\lceil \ell(n)(m+k)^{1/\alpha}\rceil + 1} \\ &\leq np^{\ell(n)(m+k)^{1/\alpha}} \\ &= n^{1-(m+k)^{1/\alpha}}. \end{aligned}$$

Combining the previous two estimates, we have

$$\begin{aligned}
&\frac{1}{\ell(n)}\ln\mathbb{E}\bigg[\exp\bigg(\ell(n)f\bigg(\frac{L(n)}{\ell(n)}\bigg)\bigg);\ f\bigg(\frac{L(n)}{\ell(n)}\bigg)\geq m\bigg]\\
&\quad\leq c(m+1)+\frac{1}{\ell(n)}\ln\bigg(\sum_{k=0}^{\infty}\exp(ck\ell(n))n^{-(m+k)^{1/\alpha}+1}\bigg)\\
&\quad= c(m+1)+\frac{1}{\ell(n)}\ln\bigg(\sum_{k=0}^{\infty}n^{ck/\ln(1/p)}n^{-(m+k)^{1/\alpha}+1}\bigg)\\
&\quad\leq c(m+1)+\frac{1}{\ell(n)}\ln\bigg(\sum_{k=0}^{\infty}n^{ck/\ln(1/p)}n^{-(m^{1/\alpha}+k^{1/\alpha})/2+1}\bigg)\\
&\quad= c(m+1)-\frac{(m^{1/\alpha}-1)\ln(1/p)}{2}+\frac{1}{\ell(n)}\ln\bigg(\sum_{k=0}^{\infty}n^{ck/\ln(1/p)}n^{-k^{1/\alpha}/2}\bigg)\\
&\quad\to c(m+1)-\frac{(m^{1/\alpha}-1)\ln(1/p)}{2}\qquad\text{as } n\to\infty \text{ (since } \alpha<1).
\end{aligned}$$

Therefore, (8) follows by taking $m\to\infty$.

With $f(x)=tx^{\alpha}$, $t>0$, $0<\alpha<1$, we have

$$\max_{x\in\mathbb{R}}\{f(x)-\Lambda^*(x)\}=\max_{x\geq 1}\{tx^{\alpha}-\lambda_p(x-1)\},$$

where $\lambda_p:=\ln(1/p)$ for brevity. There are two cases.

Case 1: $t>\ln(1/p)/\alpha$. Then the maximum above is achieved at $x^*=(\alpha t/\lambda_p)^{1/(1-\alpha)}$ and is equal to

$$t^{1/(1-\alpha)}\lambda_p^{-\alpha/(1-\alpha)}C_\alpha+\lambda_p,$$

where C_α is the positive quantity

$$C_\alpha=\alpha^{\alpha/(1-\alpha)}-\alpha^{1/(1-\alpha)}.$$

Since $\ell(n)f(L(n)/\ell(n)) = t\ell(n)^{1-\alpha}L(n)^{\alpha} = t\lambda_p^{\alpha-1}(\ln n)^{1-\alpha}L(n)^{\alpha}$, from Corollary 3, we have

$$\begin{aligned}
\ln\mathbb{E}[\exp(t\lambda_p^{\alpha-1}(\ln n)^{1-\alpha}L(n)^{\alpha})]&\sim\frac{\ln n}{\lambda_p}(t^{1/(1-\alpha)}\lambda_p^{-\alpha/(1-\alpha)}C_\alpha+\lambda_p)\\
&=(\ln n)(t^{1/(1-\alpha)}\lambda_p^{-1/(1-\alpha)}C_\alpha+1).
\end{aligned}$$

Case 2: $t\leq\ln(1/p)/\alpha$. Then the maximum is achieved at $x^*=1$ and is equal to t. Hence,

$$\ln\mathbb{E}[\exp(t\lambda_p^{\alpha-1}(\ln n)^{1-\alpha}L(n)^{\alpha})]\sim\frac{t}{\lambda_p}\ln n.$$

The expressions become neater upon a change of variables and are summarized below.

Corollary 4. *For all $t>0$, for all $0<\alpha<1$, as $n\to\infty$,*

$$\ln\mathbb{E}[\exp(t(\ln n)^{1-\alpha}L(n)^{\alpha})]
\sim\begin{cases}\dfrac{t}{\ln^{\alpha}(1/p)}\ln n & \textit{if } t\leq\ln^{\alpha}(1/p)/\alpha,\\[2ex]
\bigg[\bigg(\dfrac{t}{\ln^{\alpha}(1/p)}\bigg)^{1/(1-\alpha)}(\alpha^{\alpha/(1-\alpha)}-\alpha^{1/(1-\alpha)})+1\bigg]\ln n & \textit{otherwise.}\end{cases}$$

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

5. An application to inference

Let us consider a classical problem in confidence intervals. Let $\{X_k\}_{1\le k\le n}$ be an i.i.d. random sample from a Bernoulli population X with $\mathbb{P}(X=1)=p$ and $\mathbb{P}(X=0)=1-p$, $0<p<1$. Our aim in this section is to construct a $100(1-\alpha)\%$ confidence interval for p with a given significance level α, when p is close to 1 (or 0) and n is not very large.

The normal approximation to the binomial random variable $K:=\sum_{i=1}^{n} X_i$ does not work well when p is close to 1 (or 0). Nevertheless, there are several alternatives in this case: Wilson's score interval [18], the Clopper–Pearson interval [2], and others (such as Jeffreys' interval, Agresti–Coull interval etc.). In this section we propose another confidence interval based on the longest head run $L(n)$ with the help of Corollary 1. It turns out that this type of confidence intervals works much better than others.

To construct such confidence intervals, on the one hand, from Corollary 1 it follows that, for each $x>0$,

$$\lim_{n\to\infty}\frac{1}{\log_{1/p} n}\ln\mathbb{P}\biggl(\frac{L(n)}{\log_{1/p} n}\ge 1+x\biggr)=-x\ln\biggl(\frac{1}{p}\biggr).$$

On the other hand, Lemma 7 below states that, for every $0<x<1$,

$$\lim_{n\to\infty}\frac{1}{\log_{1/p} n}\ln\biggl[-\ln\mathbb{P}\biggl(\frac{L(n)}{\log_{1/p} n}\le 1-x\biggr)\biggr]=x\ln\biggl(\frac{1}{p}\biggr).$$

Combining these two asymptotics gives a $100(1-\alpha)\%$ confidence interval of p as follows:

$$I_p=\biggl(\exp\biggl\{-\frac{\ln(n)-\ln(\alpha/2)}{\widehat{L}(n)}\biggr\},\exp\biggl\{-\frac{\ln(n)-\ln(-\ln(\alpha/2))}{\widehat{L}(n)}\biggr\}\biggr),\tag{24}$$

where $\widehat{L}(n)$ is a point estimate of $L(n)$. A reasonable point estimate of $L(n)$ is

$$\widehat{L}(n)=L_{\mathrm{obs}}(n)-\bigl[\log_{1/\widehat{p}}(1-\widehat{p})+\log_{1/\widehat{p}}(\exp(\gamma))-\tfrac{1}{2}\bigr]$$

with $L_{\mathrm{obs}}(n)$ being the observed longest head run in n trials, and $\widehat{p}:=k/n$ being the sample proportion. To see this, firstly we know that in the long run $L(n)/\log_{1/p} n\to 1$, therefore we want an estimate that satisfies $\mathbb{E}\widehat{L}(n)\to\log_{1/p} n$. Secondly, it follows from the mean (12) that

$$\mathbb{E}\widehat{L}(n)=\log_{1/p} n+[\log_{1/p}(1-p)+\log_{1/p}(\exp(\gamma))]-[\log_{1/\widehat{p}}(1-\widehat{p})+\log_{1/\widehat{p}}(\exp(\gamma))]+\varepsilon(n),$$

which is quite close to $\log_{1/p} n$. This explains that (24) is an appropriate confidence interval for p.

In Table 1 we show the results of simulations for the derived confidence interval I_p in (24) when p is close to 1 (the case when p is close to 0 can be similarly handled), and we make several comparisons with Wilson score intervals and Clopper–Pearson intervals. Based on the simulations, it is evident that our confidence interval (24) works much better than others when p is close to 1 and n is not very large. In Table 2, for larger p and n we apply the normal approximation to the Binomial random variable. In this case it turns out that the lower bound of the normal approximation intervals works better than Wilson score intervals and Clopper–Pearson intervals, but the upper bound does not. In any case, our confidence interval (24) still works the best among them.

https://doi.org/10.1017/jpr.2016.38 Published online by Cambridge University Press

TABLE 1: Values for the confidence interval I_p when p is close to 1 for the Wilson score interval (WS), Clopper–Pearson interval (CP), and the longest run interval (LR).

	$p = 0.9500, n = 200, \alpha = 0.05$				
	$\widehat{p} = 0.9650$	$\widehat{p} = 0.9450$	$\widehat{p} = 0.9600$	$\widehat{p} = 0.9500$	$\widehat{p} = 0.9700$
WS	(0.9295, 0.9829)	(0.9042, 0.9690)	(0.9231, 0.9796)	(0.9104, 0.9726)	(0.9361, 0.9862)
CP	(0.9292, 0.9858)	(0.9037, 0.9722)	(0.9227, 0.9826)	(0.9100, 0.9758)	(0.9358, 0.9889)
LR	(0.9329, 0.9696)	(0.9145, 0.9611)	(0.9243, 0.9656)	(0.9325, 0.9694)	(0.9484, 0.9767)
	$p = 0.98, n = 200, \alpha = 0.05$				
	$\widehat{p} = 0.9800$	$\widehat{p} = 0.9850$	$\widehat{p} = 0.9700$	$\widehat{p} = 0.9800$	$\widehat{p} = 0.9750$
WS	(0.9497, 0.9922)	(0.9568, 0.9949)	(0.9361, 0.9862)	(0.9497, 0.9922)	(0.9428, 0.9893)
CP	(0.9496, 0.9945)	(0.9568, 0.9969)	(0.9358, 0.9889)	(0.9496, 0.9945)	(0.9426, 0.9918)
LR	(0.9657, 0.9846)	(0.9751, 0.9889)	(0.9578, 0.9810)	(0.9703, 0.9867)	(0.9606, 0.9821)

TABLE 2: Values for the confidence interval I_p for larger p and n than in Table 1 for the Wilson score interval (WS), Clopper–Pearson interval (CP), longest run interval (LR), and the normal approximation (N).

	$p = 0.9950, n = 1000, \alpha = 0.05$				
	$\widehat{p} = 0.9950$	$\widehat{p} = 0.9940$	$\widehat{p} = 0.9950$	$\widehat{p} = 0.9960$	$\widehat{p} = 0.9960$
N	(0.9906, 0.9994)	(0.9892, 0.9988)	(0.9906, 0.9994)	(0.9921, 0.9999)	(0.9921, 0.9999)
WS	(0.9883, 0.9979)	(0.9870, 0.9972)	(0.9883, 0.9979)	(0.9898, 0.9984)	(0.9898, 0.9984)
CP	(0.9884, 0.9984)	(0.9870, 0.9978)	(0.9884, 0.9984)	(0.9898, 0.9989)	(0.9898, 0.9989)
LR	(0.9915, 0.9955)	(0.9909, 0.9952)	(0.9919, 0.9957)	(0.9941, 0.9969)	(0.9938, 0.9967)

6. Open problems

A problem for future research would be the study of a large deviation principle for the random-dimensional random vector $R(n) = (R_1(n), R_2(n), \ldots, R_{L(n)}(n))$ of counts of successive runs of all lengths. That is, let $R_\ell(n)$ be the number of head runs of length ℓ up to the nth coin toss. Distributional relations for $R(n)$ were studied in [8].

Furthermore, it would be interesting to obtain large deviation principles for longest runs in a Markov chain. In other words, assume that (X_n) is a Markov chain with finite (or countable) state space S and let $L(x, n)$ be the longest sojourn time at a state $x \in S$ before time n. Although there are Stein–Chen-type estimates [10], [19] for the distribution of such quantities, the errors in these estimates are too big for the study of a large deviation principle. We would like to obtain an LDP for $L(x, n)$ or for the vector $(L(x, n),\ x \in S)$ which would, by contraction principle, give us an LDP for $L(n) := \sup_{x \in S} L(x, n)$.

Acknowledgements

Takis Konstantopoulos was supported by the Swedish Research Council (grant number 2013-4688). The authors thank the anonymous referee for several suggestions and comments.

References

[1] BALAKRISHNAN, N. AND KOUTRAS, M. V. (2002). *Runs and Scans with Applications.* John Wiley, New York.
[2] CLOPPER, C. J. AND PEARSON, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26,** 404–413.

[3] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications,* 2nd edn. Springer, Berlin.
[4] Erdős, P. and Rényi, A. (1970). On a new law of large numbers. *J. Analyse Math.* **23,** 103–111.
[5] Erdős, P. and Révész, P. (1977). On the length of the longest head-run. In *Topics in Information Theory* (Colloq. Math. Soc. Janos Bolyai **16**), North-Holland, Amstersdam, pp. 219–228.
[6] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Vol. I,* 3rd edn. John Wiley, New York.
[7] Fu, J. C., Wang, L. and Lou, W. Y.W. (2003). On exact and large deviation approximations for the distribution of the longest run in a sequence of two-state Markov dependent trials. *J. Appl. Prob.* **40,** 346–360.
[8] Holst, L. and Konstantopoulos, T. (2015). Runs in coin tossing: a general approach for deriving distributions for functionals. *J. Appl. Prob.* **52,** 752–770.
[9] Mao, Y.-H., Wang, F. and Wu, X.-Y. (2015). Large deviation behavior for the longest head run in an IID Bernoulli sequence. *J. Theoret. Prob.* **28,** 259–268.
[10] Novak, S. Y. (1992). Longest runs in a sequence of m-dependent random variables. *Prob. Theory Relat. Fields* **91,** 269–281.
[11] Philippou, A. N. and Makri, F. S. (1985). Longest success runs and Fibonacci-type polynomials. *Fibonacci Quart.* **23,** 338–346.
[12] Rényi, A. (1970). *Foundations of Probability.* Holden-Day, San Francisco, CA.
[13] Sandmann, K. and Schönbucher, P. (eds) (2002). *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann*. Springer, Berlin.
[14] Schilling, M. F. (1990). The longest run of heads. *College Math. J.* **21,** 196–207.
[15] Von Mises, R. (1921). Zur Theorie der Iterationen. *Z. Angew. Math. Mech.* **1,** 298–307.
[16] Von Mises, R. (1981). *Probability, Statistics, and Truth*, 2nd edn. Dover, New York.
[17] Wentzell, A. D. (1990). *Limit Theorems on Large Deviations for Markov Stochastic Processes.* Kluwer, Dordrecht.
[18] Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *J. Amer. Statist. Assoc.* **22,** 209–212.
[19] Zhang, Y. Z. and Wu, X. Y. (2013). Some results associated with the longest run in a strongly ergodic Markov chain. *Acta Math. Sin. (Engl. Ser.)* **29,** 1939–1948.