Statistics and Probability Letters 110 (2016) 128–132

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A general large deviation principle for longest runs

Zhenxia Liu, Xiangfeng Yang *

Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden

ARTICLE INFO

Article history:
Received 17 September 2015
Received in revised form 2 December 2015
Accepted 8 December 2015
Available online 18 December 2015

MSC:
60F10
60G50
60G70
60J10

Keywords:
Longest run
Large deviation principle
Bernoulli trail
Markov chain

ABSTRACT

In this note we prove a general large deviation principle (LDP) for the longest success run in a sequence of independent Bernoulli trails. This study not only recovers several recently derived LDPs, but also gives new LDPs for the longest success run. The method is based on the Bryc's inverse Varadhan lemma, which can be intuitively generalized to the longest success run in a two-state (success and failure) Markov chain.

1. Introduction

Let $\{X_k\}_{1\le k\le n}$ be n independent Bernoulli trails having $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = 0) = 1-p := q$, with '1' and '0' denoting the 'success' and 'failure' respectively. We assume that $0 < p < 1$ throughout the note. 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. Balakrishnan and Koutras, 2002). We refer to Erdös and Rényi (1970) and Erdös and Révész (1975) for the first few seminal works in 1970s, and Mao et al. (2015), Holst and Konstantopoulos (2015), Konstantopoulos et al. (in press) and Liu and Yang (in press) for the latest progress.

Although there is an explicit formula of the distribution function of $L(n)$ in the following form (for instance cf. Holst and Konstantopoulos, 2015),

$$\mathbb{P}(L(n) < k) = \sum_{r=0}^{\left[\frac{n+1}{k+1}\right]} (-1)^r p^{rk} q^{r-1} \left[\binom{n-rk}{r-1} + q\binom{n-rk}{r}\right], \tag{1.1}$$

where $[\cdot]$ denotes the integer part of a constant, limiting behaviors of the distribution of $L(n)$ as $n \to \infty$ can be hardly obtained due to the complexity of (1.1). Therefore probability estimating of $L(n)$ for large n has been an important topic. In this note, we fully study this aspect in terms of large deviations. We first recall a strong law of large numbers (cf. Erdös and Rényi, 1970, Erdös and Révész, 1975 and Rényi, 1970): as $n \to \infty$,

$$L(n)/\log_{1/p} n \to 1, \quad \text{almost surely.} \tag{1.2}$$

* Corresponding author.
E-mail addresses: zhenxia.liu@liu.se (Z. Liu), xiangfeng.yang@liu.se (X. Yang).

http://dx.doi.org/10.1016/j.spl.2015.12.015

This strong law suggests to study large deviation probabilities $\mathbb{P}(L(n)/\log_{1/p} n \in A)$ for a set A not containing the most probable point '1'. Such deviation probabilities have been fully investigated in Mao et al. (2015). The strong law more generally suggests to study LDP for the family of normalized longest runs $\{L(n)/\log_{1/p} n\}$. Indeed such a LDP has been recently established in Konstantopoulos et al. (in press) based on the moment generating function of $L(n)$. The employed (moment generating function) method in Konstantopoulos et al. (in press) naturally leads to a LDP for another family of normalized longest runs $\{L(n)/n\}$, which has also been investigated therein. The corresponding strong law of large numbers is $L(n)/n \to 0$ almost surely which is trivially from (1.2). These two LDPs were later generalized to Markov chains in Liu and Yang (in press).

It is clear that $\log_{1/p} n$ is the most common speed in LDP in view of (1.2), but other speeds also give useful information. For instance, the LDP for $\{L(n)/n\}$ in Konstantopoulos et al. (in press) not only implies the (super) large deviation probability $\ln \mathbb{P}(L(n) > c \cdot n) \sim -n \cdot c \cdot \ln(1/p)$ with $0 < c < 1$, but also yields the Laplace transform asymptotics $\ln \mathbb{E} \exp\{n \cdot \phi(L(n)/n)\} \sim n \cdot \sup_{x\in\mathbb{R}}[\phi(x) - \Lambda^*(x)]$ with Λ^* being the corresponding rate function and ϕ a continuous function satisfying appropriate assumptions. In this note, we present a LDP with a general speed $\alpha(n)$ which on one hand recovers those two LDPs obtained in Konstantopoulos et al. (in press), and on the other hand gives new forms of LDPs. What is more important, this approach includes and unifies all possible LDPs for normalized longest runs. The main result of this note is formulated as follows.

Theorem 1.1. *Suppose that* $\log_{1/p} n \le \alpha(n) \le n$ *and the limit* $\lim_{n\to\infty} \ln(n)/\alpha(n) =: \beta$ *exists. Then the family of normalized longest success runs* $\{L(n)/\alpha(n)\}$ *obeys the LDP with a speed* $\alpha(n)$ *and a good rate function* $S(x)$ *defined as*

$$S(x) = \begin{cases} x \cdot \ln(1/p) - \beta, & x \in D, \\ +\infty, & x \notin D, \end{cases} \tag{1.3}$$

where the interval D *is given by* $D = \{x \in \mathbb{R} : \beta/\ln(1/p) \le x \le \limsup_{n\to\infty} n/\alpha(n)\}$. *That is,*

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

$$\liminf_{n\to\infty} \frac{1}{\alpha(n)} \ln \mathbb{P}(L(n)/\alpha(n) \in O) \ge -\inf_{x\in O} S(x); \tag{1.4}$$

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

$$\limsup_{n\to\infty} \frac{1}{\alpha(n)} \ln \mathbb{P}(L(n)/\alpha(n) \in F) \le -\inf_{x\in F} S(x). \tag{1.5}$$

Here we discuss several special cases of Theorem 1.1, and their connections with the LDPs derived in Konstantopoulos et al. (in press).

- $\alpha(n) = \log_{1/p} n$. In this case the interval $D = \{x : x \ge 1\}$, $\beta = \ln(1/p)$ and $S(x) = (x-1)\ln(1/p)$ for $x \ge 1$. Such a LDP has been proved in Konstantopoulos et al. (in press).
- $\alpha(n) = n$. The interval $D = \{x : 0 \le x \le 1\}$, $\beta = 0$ and $S(x) = x \ln(1/p)$ for $0 \le x \le 1$. This LDP has also been derived in Konstantopoulos et al. (in press).
- $\alpha(n) = n^\alpha$ with $0 < \alpha < 1$. This is a new type of LDP and the speed is between $\log_{1/p} n$ and n. The interval $D = \{x : x \ge 0\}$, $\beta = 0$ and $S(x) = x \ln(1/p)$ for $x \ge 0$.

A direct consequence of Theorem 1.1 is Laplace transform asymptotics (which is also known as the Varadhan's integral lemma; see Dembo and Zeitouni (2009, Section 4.3)) in the following form

$$\ln \mathbb{E} \exp\{\alpha(n) \cdot \phi(L(n)/\alpha(n))\} \sim \alpha(n) \cdot \sup_{x\in\mathbb{R}}[\phi(x) - S(x)]$$

under suitable assumptions on the continuous function ϕ. The proof of Theorem 1.1 is given in Section 2 whose main ingredients are the Bryc's Inverse Varadhan Lemma and suitable estimates on the distribution function of $L(n)$ formulated in (2.1). This approach can be intuitively generalized to the longest success run in a two-state (success and failure) Markov chain.

2. Proof of Theorem 1.1

The proof is built on the Bryc's Inverse Varadhan Lemma (cf. Dembo and Zeitouni, 2009, Section 4.4) and the following important limit.

Lemma 2.1. *For every* $x > \beta/\ln(1/p)$, *it holds that*

$$\lim_{n\to\infty} \frac{1}{\alpha(n)} \ln \mathbb{P}(L(n)/\alpha(n) \ge x) = -S(x).$$