Large deviation behavior for the longest head run in IID Bernoulli sequence

Xian-Yuan Wu

School of mathematical science. CNU.

(7-17-2012. BNU).

1.

$Z_1, Z_2, \cdots$ i.i.d Bernoulli sequence.

$p \in (0,1)$, $p+q=1$.

$\mathbb{P}(Z_1=1)=p$, $\mathbb{P}(Z_1=0)=q$.

S_N : the length of the longest run of 1's with the first N tosses.

Examples: 1)

$Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7, Z_8, Z_9, Z_{10}$

1 0 0 1 1 1 0 1 1 0

$N=10, \quad S_N = 3$

2)

$Z_1, Z_2, Z_3, Z_4, Z_5, Z_6$

1 1 1 1 0 0

$N=6, \quad S_N = 4.$

3)

$Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7, Z_8, Z_9, Z_{10}$

0 1 0 1 0 1 0 0 1 0

$N=10, \quad S_N = 1.$

2. Erdös-Rényi Law

$$\lim_{N\to\infty} \frac{S_N}{\ln N} = \xi(p) := \left[\log \frac{1}{p}\right]^{-1}.$$

almost surely.

A new law of large number!

3. Asymptotic distribution:

- $S_N - \xi(p)\ln N$ possesses NO limit distribution!

But:

$$S_{N_j} - \lfloor \xi(p) \ln N_j \rfloor \xrightarrow{d} Y_r \quad , \; j \to \infty .$$

where: 1) $r \in [0,1]$, $\{N_j\} \subset \{N\}$, s.t.

$$\lim_{j\to\infty} \{ \xi(p)\ln N_j - \lfloor \xi(p) \ln N_j \rfloor \} = r.$$

2) $Y_r = \lfloor Y + r \rfloor$, and

$$\mathbb{P}(Y \le x) = e^{-qp^x}$$

Y — extreme valued random variable.

4. Other results.

$p = q = \frac{1}{2}$.

$$\mathbb{E} S_N = (\ln N + \gamma)/\ln 2 + \rho_0(N) + o(1).$$

$\gamma = 0.577\cdots$ Euler-Mascheroni Constant

$|\rho_0(N)| < 1.6 \times 10^{-6}$. with no limit!

$\mathrm{Var}\, S_N$ is bounded!

... ...

Problem:

Asymptotic behaviors of

$$\mathbb{P}\left(\frac{S_N}{\ln N} > \xi(p) + x \right) \quad \text{and}$$

$$\mathbb{P}\left(\frac{S_N}{\ln N} \leq \xi(p) - x \right). \qquad ?$$

Large deviation ?

$r \in [0, 1]$

$N_j \in \{N\}$. Such that

$$\lim_{j \to \infty} \left\{ \xi(p) \ln N_j - \lfloor \xi(p) \ln N_j \rfloor \right\} = r.$$

Theorem 1 (Main results).

i) $\forall\, x>0$, $$\lim_{j\to\infty} \frac{\mathbb{P}\left(S_{N_j}/\ln N_j > \xi(p)+x\right)}{q p^{1-r} N_j^{-x/\xi(p)}} = 1.$$

ii). $\forall\, 0<x<\xi(p)$. $\forall\, \varepsilon>0$.

$$e^{-(q+\varepsilon)p^{1-r}N_j^{x/\xi(p)}} \le \mathbb{P}\left(\frac{S_{N_j}}{\ln N_j} \le \xi(p)-x\right) \le e^{-(q-\varepsilon)p^{1-r}N_j^{x/\xi(p)}}$$

For large N_j. In particular, for $0<x<\frac{\xi(p)}{2}$.

$$\lim_{j\to\infty} \frac{\mathbb{P}\left(\frac{S_{N_j}}{\ln N_j} \le \xi(p)-x\right)}{e^{-q p^{1-r} N_j^{x/\xi(p)}}} = 1.$$

Theorem 2. (Large deviation).

For each $x>0$,

$$\lim_{N\to\infty} \frac{-1}{\ln N} \ln \mathbb{P}\left(\frac{S_N}{\ln N} \geq \xi(p)+x\right) = \frac{x}{\xi(p)}.$$

For every $0<x<\xi(p)$,

$$\lim_{N\to\infty} \frac{1}{\ln N} \ln\left\{-\ln \mathbb{P}\left[\frac{S_N}{\ln N} \leq \xi(p)-x\right]\right\} = \frac{x}{\xi(p)}$$

On the proofs:

Strategy: give good bounds (lower and upper)

to probability:

$$\mathbb{P}(S_N < k).$$

A known result in reliability theory:

$$(1-p^k)^{N-k+1} \leq \mathbb{P}(S_N < k) \leq (1-qp^k)^{N-k+1}$$

for $k = 1, 2, \cdots, N$ and $0 < p < 1$.

Our new lower bound for $\mathbb{P}(S_N < k)$.

Theorem 3 : For any $p \in (0,1)$ and $0 < \psi < \xi(p)^{-1}$, there exists $k_1 = k_1(p, \psi)$ such that

$$\left(1 - e^{-\psi k}\right)\left(1 - (1 + e^{-\psi k})\, q p^k\right)^{N-k+1}$$

$$\leq \mathbb{P}(S_N < k) \leq \left(1 - q p^k\right)^{N-k+1}$$

for all $N \geq 4k$ and $k \geq k_1$.

The bounds given in theorem 3 are optimal! (for large k).

How to prove Th3?

Let $\{X_n : n \geq 0\}$ be Markov Chain ($\mathbb{Z}_+$ - valued)

$$X_0 = 0$$

$$P_{ij} = \begin{cases} p & j = i+1 \\ q & j = 0 \end{cases}$$

For $k \geq 0$, Let

$$T_k := \inf\{n : X_n = k\}$$

Dual relation between S_N and T_k :

$$\mathbb{P}(S_N \geq k) = \mathbb{P}(T_k \leq N)$$

$$0 \leq k \leq N$$

Theorem 4: $\forall\ p \in (0,1),\ \exists\ C_1(p,k),\ C_2(p,k)$ and $f(p,k)$ such that

$$\boxed{\begin{aligned} & C_1(p,k)\, q p^k \left[1 - f(p,k)\, q p^k\right]^n \\ & \quad \leq \mathbb{P}\left(T_k = k+n\right) \\ & \quad \leq C_2(p,k)\, q p^k \left[1 - q p^k\right]^n \end{aligned}}$$

for $n > 3k \geq 1$. $\lim\limits_{k\to\infty} C_1(p,k) = \lim\limits_{k\to\infty} C_2(p,k)$

$= \lim\limits_{k\to\infty} f(p,k) = 1.$

Three Lemmas: Let $P_n = \mathbb{P}(T_k = k+n)$, $n = 0, 1, 2, \cdots$

Lemma 1: $\{P_n : n \geq 0\}$ forms the following Fibonacci sequence

$$P_n = \begin{cases} p^k; & n = 0 \\ a_1 P_{n-1} + a_2 P_{n-2} + \cdots + a_n P_0; & 1 \leq n \leq k \\ a_1 P_{n-1} + a_2 P_{n-2} + \cdots + a_k P_{n-k}. & n \geq k+1. \end{cases}$$

where ~~a_k~~ $a_i = qp^{i-1}$, $i = 1, 2, \cdots, k$.

In particular, $P_k = P_{k-1} = \cdots = P_2 = P_1 = qp^k$.

Lemma 2: Let $\alpha_k = 1 - (q \vee p)p^{k-1}$ and $\beta_k = 1 - P_1$. then

$$\alpha_k \leq \frac{P_{n+1}}{P_n} \leq \beta_k, \quad \text{for all } n \geq k$$

$q \vee p = \max\{q, p\}$.

Lemma 3. Let $\gamma_K = 1 - f(p,k)P_1$, where $f(p,k) = \dfrac{1}{1-g(p,k)}$

and

$$g(p,k) = \left[p^k \sum_{i=0}^{k-1} \beta_K^i\right] \Big/ \left[\sum_{i=0}^{K-1} p^{k-1-i} \alpha_K^i\right]$$

then

$$\frac{P_{n+1}}{P_n} \geqslant \gamma_K$$

for all $n \geqslant 3k$.

$$\boxed{1 - f(p,k)P_1 \leq \frac{P_{n+1}}{P_n} \leq 1 - P_1} \quad \forall\, n \geq 3k$$

$$P_1 = qp^k$$

$f(p,k) > 1$ and $f(p,k) \longrightarrow 1$ as $k \to \infty$.

In fact, $\boxed{f(p,k) \leq 1 + e^{-\psi k}}$ $0 < \psi < \xi^{-1}(p)$.

for $k \geq k_1(p,\psi)$.

Thank you !