Can. J. Math. Vol. **45** (6), 1993 pp. 1245–1262

THE A.S. LIMIT DISTRIBUTION OF THE LONGEST HEAD RUN

TAMÁS F. MÓRI

ABSTRACT It is well known that the length Z_n of the longest head run observed in n tosses with a fair coin is approximately equal to $\log_2 n$ with a stochastically bounded remainder term Though $Z_n - \log_2 n$ does not converge in law, in the present paper it is shown to have *almost sure* limit distribution in the sense of the a s central limit theorem having been studied recently The results are formulated and proved in a general setup covering other interesting problems connected with patterns and runs such as the longest monotone block or the longest tube of a random walk

1. **Introduction.** Consider an infinite sequence of independent coin tossings. Let T_m denote the number of tosses needed until m consecutive heads occur. A strongly related quantity is the length Z_n of the longest pure head run in the first n tosses. These random variables have been in the mainstream of research on the nature of randomness for a long time. They appeared as early as in 1738, in de Moivre's *Doctrine of Chances*. Exact distributional results such as generating functions were obtained mainly by combinatorial arguments. In addition, the weak and a. s. asymptotic behaviour of Z_n has also been characterized completely, see [ER] or [GSW]. It turned out that T_m has exponential limit distribution as $m \longrightarrow \infty$; a general property shared by a wide class of first visit type stopping times in Markov processes. On the other hand, for large n the quantity Z_n is approximately equal to $\log_2 n$ with a stochastically bounded remainder term. However, $Z_n - \log_2 n$ does not converge in distribution.

Recently much attention has been given to the so-called a. s. extensions of classical weak limit theorems. Although similar results in particular cases have been known for a longer time, the a. s. central limit theorem, considered as the starting point of these studies, was first proposed independently by Brosamler and Schatte in 1988. Their results have been extended and generalized by several authors. An excellent representative of these investigations is [BD] which provides a systematic study of analogues of classical limit theorems in terms of logarithmic average and logarithmic density.

Typical results of this kind follow a common pattern. They start from a certain sequence of random variables (or even stochastic processes) $\xi_1, \xi_2, \ldots$ converging in distribution to a limit law Q, then prove that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I(\xi_i \in A) = Q(A) \text{ a. s.}$$

Research supported by the Hungarian National Foundation for Scientific Research, Grant No 1905
Received by the editors June 30, 1992
AMS subject classification 60C05, 60F15
Key words and phrases runs, patterns, waiting times, almost sure central limit theorem

for all Q-contınuıty Borel sets A Such a result has been consıdered as a strengthened versıon of the dıstrıbutıonal theorem, but ıt turned out not to be the case Counter examples have been found even for normalızed sums of ı ı d random varıables obeyıng an a s lımıt theorem of the above type but not beıng convergent ın dıstrıbutıon, whıle the opposıte dırectıon of ımplıcatıon ıs often true

In the present paper another class of non-convergent random varıables wıth a s logarıthmıc lımıts ıs ınvestıgated a class contaınıng the sequence $(Z_n, n \geq 1)$ ıntroduced above

Let $(\mathcal{X}, \mathcal{F})$ be a measurable space and $X_1, X_2,$ ı ı d $\mathcal{X}$-valued random varıables Let $X_{n\,k}$ denote the block $(X_n, X_{n+1}, \quad, X_{n+k-1})$ Suppose for every posıtıve ınteger m we are gıven a measurable set $B_m \in \mathcal{F}^m$ such that $B_m \subset B_{m\;1} \times \mathcal{X}$ Let

$$T_m = \min\{n \geq m \quad X_{n-m+1\,m} \in B_m\},$$

then $T_1 < T_2 <$ Let $p(m) = \mathbf{P}(X_{1\,m} \in B_m) = \mathbf{P}(T_m = m)$, thıs ıs decreasıng ın m Assume that $p(m) > 0$ for every m, then T_m ıs fınıte wıth probabılıty 1, furthermore ıt has fınıte moments of arbıtrary order We shall also suppose a very mıld condıtıon on the rate of decrease of $p(m)$, namely

$$\sum_m mp(m) < \infty \tag{1 1}$$

Because $p(m)$ ıs decreasıng, (1 1) ımplıes that $\lim_{m\to\infty} m^2 p(m) = 0$

Let us ıntroduce $Z_n = \max\{m \quad T_m \leq n\}$ We are goıng to deal wıth the a s logarıthmıc lımıt behavıour of the sequence Z_n, $n \geq 1$

2 **Auxiliary results for waiting times.** In thıs sectıon three lemmas wıll be proved on the asymptotıc mean, dıstrıbutıon, and joınt dıstrıbutıon of the waıtıng tımes T_m Sımılar results were obtaıned ın [M85] or [CsFK] Although they could certaınly be adapted to the present model, we also gıve here (sımple) proofs for the sake of the reader's convenıence

LEMMA 2 1 *Let us abbrevıate* $\mathbf{E}(T_m)$ *by* $E(m)$ *Then*

$$1 \leq p(m)E(m) \leq m, \tag{2 1}$$

$$\liminf_{m\to\infty} p(m)E(m) = \lim_{k\to\infty} \liminf_{m\to\infty} \frac{p(m)}{\mathbf{P}(T_m = m+k)}, \tag{2 2}$$

$$\limsup_{m\to\infty} p(m)E(m) = \lim_{k\to\infty} \limsup_{m\to\infty} \frac{p(m)}{\mathbf{P}(T_m = m+k)}, \tag{2 3}$$

$$E(m) \sim \frac{1}{\mathbf{P}(T_m = 2m)} \tag{2 4}$$

Note that ın (2 2–3) the lımıts as $k \to \infty$ do exıst, sınce, for fıxed m, $\mathbf{P}(T_m = m+k)$ ıs a decreasıng functıon of k

PROOF. It is quite easy to see that

$$\mathbf{P}(T_m > n)\mathbf{P}(T_m = m+k) \geq \mathbf{P}(T_m = n+m+k). \tag{2.5}$$

On the other hand,

$$\mathbf{P}(T_m > n)\mathbf{P}(T_m = m+k) - \mathbf{P}(T_m = n+m+k) \\ \leq \mathbf{P}(T_m > n-m) \sum_{i=1}^{m-1} \mathbf{P}(X_{n-m+i+1,m} \in B_m,\ X_{n+k+1,m} \in B_m).$$

For $i \leq k$ clearly

$$\mathbf{P}(X_{n-m+i+1,m} \in B_m,\ X_{n+k+1,m} \in B_m) = p(m)^2,$$

and for $i > k$

$$\mathbf{P}(X_{n-m+i+1,m} \in B_m,\ X_{n+k+1,m} \in B_m) \leq \mathbf{P}(X_{n-m+i+1,m+k-i} \in B_{m+k-i},\ X_{n+k+1,m} \in B_m) \\ = p(m+k-i)\,p(m).$$

Hence

$$\mathbf{P}(T_m > n)\mathbf{P}(T_m = m+k) \leq \mathbf{P}(T_m = n+m+k) + \mathbf{P}(T_m > n-m)\,c(k,m)\,p(m), \tag{2.6}$$

where

$$c(k,m) = \min(m,k)\,p(m) + \sum_{i=k+1}^{m-1} p(i). \tag{2.7}$$

In particular, $c(k,m) = mp(m)$ if $k \geq m$.

Summing (2.5) and (2.6) for $n = 0, 1, \ldots$ we obtain

$$\mathbf{P}(T_m \geq m+k) \leq E(m)\mathbf{P}(T_m = m+k) \leq \mathbf{P}(T_m \leq m+k) + 2E(m)\,p(m)\,c(k,m). \tag{2.8}$$

Here we have used that

$$\sum_{n=0}^{\infty} \mathbf{P}(T_m > n-m) = \mathbf{E}(m+T_m) \leq 2E(m).$$

After rearrangement we have, using the trivial inequality $1-kp(m) \leq \mathbf{P}(T_m \geq m+k) \leq 1$, that

$$\frac{1-kp(m)}{E(m)\,p(m)} \leq \frac{\mathbf{P}(T_m = m+k)}{p(m)} \leq \frac{1}{E(m)\,p(m)} + 2c(k,m).$$

Since $\lim_{k\to\infty} \lim_{m\to\infty} c(k,m) = 0$, (2.2) and (2.3) follow immediately. For $k = 0$ the left-hand inequality in (2.8) gives $1 \leq E(m)\,p(m)$. On the other hand, the upper bound $E(m)\,p(m) \leq m$ is obtained when T_m is majorized by the smallest multiple of m, say km, for which $X_{(k-1)m+1,m} \in B_m$. Hence (2.1). Finally, substituting $k = m$ into (2.8) we obtain

$$1 - mp(m) \leq E(m)\mathbf{P}(T_m = 2m) \leq 1 + 2E(m)\,p(m)^2 m \leq 1 + 2m^2 p(m)$$

by (2.1). This implies (2.4). ∎

https://doi.org/10.4153/CJM-1993-070-x Published online by Cambridge University Press

LEMMA 2.2.

(2.9)
$$\left(1-\frac{1}{E(m)}\right)^k\big(1-mp(m)\big)-2m^2p(m) \le \mathbf{P}(T_m \ge m+k) \le \left(1-\frac{1}{E(m)}\right)^k, \quad k=0,1,\dots .$$

Consequently,

(2.10)
$$\lim_{m\to\infty} \mathbf{P}\big(T_m/E(m) > t\big) = e^{-t}$$

uniformly in $t \ge 0$.

PROOF. From (2.8) it follows that

(2.11)
$$\begin{aligned}\mathbf{P}(T_m \ge m+k)\left(1-\frac{1}{E(m)}\right) &\ge \mathbf{P}(T_m \ge m+k+1)\\ &\ge \mathbf{P}(T_m \ge m+k)\left(1-\frac{1}{E(m)}\right) - 2p(m)\,c(k,m).\end{aligned}$$

Iterating (2.11) we obtain

$$\mathbf{P}(T_m \ge m+k) \le \left(1-\frac{1}{E(m)}\right)^k,$$

and for $k \ge m$

$$\begin{aligned}\mathbf{P}(T_m \ge m+k) &\ge \left(1-\frac{1}{E(m)}\right)^{k-m}\mathbf{P}(T_m \ge 2m) - 2mp(m)^2\sum_{i=m}^{k-1}\left(1-\frac{1}{E(m)}\right)^{k-i}\\ &\ge \left(1-\frac{1}{E(m)}\right)^k\big(1-mp(m)\big) - 2mE(m)\,p(m)^2.\end{aligned}$$

This formula trivially holds true for $k < m$. Applying (2.1) we arrive at (2.9) which, in turn, implies (2.10). ∎

LEMMA 2.3. *Let* $m_1 \le m_2$ *and* n_1, n_2 *be arbitrary positive integers. Then*

$$\begin{aligned}\big|\mathbf{P}(T_{m_1} > n_1, T_{m_2} > n_2) &- \mathbf{P}(T_{m_1} > n_1)\mathbf{P}(T_{m_2} > n_2)\big|\\ &\le \frac{E(m_1)}{E(m_2)}\big(1+2m_2^2p(m_2)\big) + O\big(p(m_2)\big)\end{aligned}$$

uniformly in n_1, n_2, m_1.

PROOF. On the one hand,

$$\mathbf{P}(T_{m_1} > n_1, T_{m_2} > n_2) \le \mathbf{P}(T_{m_1} > n_1)\mathbf{P}(T_{m_2} > n_2 - n_1).$$

Thus

$$\begin{aligned}\Delta &= \mathbf{P}(T_{m_1} > n_1, T_{m_2} > n_2) - \mathbf{P}(T_{m_1} > n_1)\mathbf{P}(T_{m_2} > n_2)\\ &\le \mathbf{P}(T_{m_1} > n_1)\mathbf{P}(n_2 - n_1 < T_{m_2} \le n_2).\end{aligned}$$

From (2.8) it follows that

$$\mathbf{P}(T_m = m+k) \le \frac{1}{E(m)} + 2p(m)\,c(k,m)$$

($c(k,m)$ is defined in (2.7). Hence for arbitrary $a < b$

$$\mathbf{P}(a < T_m \le b) \le \frac{b-a}{E(m)} + 2p(m)\Big(\sum_{k=0}^{m-1} c(k,m) + (b-a)mp(m)\Big). \tag{2.12}$$

Here

$$\sum_{k=0}^{m-1} c(k,m) = \binom{m}{2} p(m) + \sum_{k=1}^{m-1} kp(k) = o(1) + \sum_{k=1}^{\infty} kp(k).$$

Substitute $a = n_2 - n_1$, $b = n_2$, $m = m_2$ into (2.12) and use that $n_1\mathbf{P}(T_{m_1} > n_1) \le E(m_1)$ to obtain

$$\Delta \le \frac{E(m_1)}{E(m_2)}\big(1 + 2m_2^2 p(m_2)\big) + O\big(p(m_2)\big).$$

On the other hand, if $n_1 \ge n_2$, then $T_{m_1} > n_1$ implies $T_{m_2} > n_2$; thus $\Delta \ge 0$. For $n_1 < n_2$ we have

$$\begin{aligned}\mathbf{P}(T_{m_1} > n_1, T_{m_2} > n_2) \ge \mathbf{P}(T_{m_1} > n_1)\mathbf{P}(T_{m_2} > n_2 - n_1)\\ - \mathbf{P}(n_1 + m_2 - m_1 < T_{m_2} < n_1 + m_2).\end{aligned}$$

To the last term we can apply (2.12) with $a = n_1 + m_2 - m_1$, $b = n_1 + m_2$, $m = m_2$. We obtain that

$$\Delta \ge -\frac{m_1}{E(m_1)}\big(1 + 2m_2^2 p(m_2)\big) - O\big(p(m_2)\big),$$

and, since $m_1 \le E(m_1)$, this completes the proof. ∎

3. **The** a. s. **limit distribution of** Z_n.

THEOREM 3.1. *Suppose f is a positive, increasing, differentiable function such that $E(m) \sim f(m)$ and the limit*

$$c = \lim_{t\to\infty}\big(\log f(t)\big)', \quad 0 \le c \le \infty \tag{3.1}$$

exists. Denote $g = f^{-1}$.

CASE (i) $c = 0$. *Then for every $t \in \mathbb{R}$*

$$\lim_{n\to\infty} \frac{1}{\log n}\sum_{i=1}^{n} \frac{1}{i} I\big(Z_i - g(i) < t\big) = \frac{1}{e} \text{ a. s.} \tag{3.2}$$

CASE (ii) $0 < c < \infty$. *Then for every $t \in \mathbb{R}$*

$$\lim_{n\to\infty} \frac{1}{\log n}\sum_{i=1}^{n} \frac{1}{i} I\big(Z_i - g(i) < t\big) = \int_0^1 F\big(c(t+z)\big)\,dz \text{ a. s.}, \tag{3.3}$$

where $F(z) = \exp\big(-\exp(-z)\big)$.

CASE (iii) $c = \infty$. *Suppose, in addition, that*

$$(\log\log f(t))' \leq \beta(t), \tag{3.4}$$

where β *is a positive nonincreasing function,* $\int_0^\infty \beta^2(t)\,dt < \infty$. *Then*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I\big(Z_i - g(i) < t\big) = \begin{cases} 0 & \text{if } t \leq -1, \\ 1+t & \text{if } -1 < t < 0, \\ 1 & \text{if } 0 \leq t. \end{cases} \tag{3.5}$$

DISCUSSION. 1. While $p(m)$ is quite easy to compute in many cases, $E(m)$ is often not. Lemma 2.1 sometimes helps to find a suitable f.

2. In Case (i) the following generalization can also be proved.

THEOREM 3.2. *Suppose* $\varphi(z)$, $z > 0$ *is a positive, nondecreasing function such that for every t belonging to a finite or infinite interval I the function* $z + t\varphi(z)$ *is eventually increasing as* $z \longrightarrow \infty$, *and*

$$\lim_{z\to\infty} \frac{f\big(z + t\varphi(z)\big)}{f(z)} = \lambda(t) \tag{3.6}$$

exists. Then

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I\left(\frac{Z_i - g(i)}{\varphi\big(g(i)\big)} < t\right) = \exp\{-1/\lambda(t)\} \text{ a.s.} \tag{3.7}$$

at every λ*-continuity point* $t \in I$. *(The right-hand side is meant* 0 *and* 1 *when* $\lambda(t)$ *is* 0 *and* ∞, *resp.)*

This theorem contains Theorem 3.1, Case (i), when $\varphi(z) \equiv 1$. Indeed, $c = 0$ implies that $\frac{f(z+t)}{f(z)} = \exp\{t(\log f)'(z + t\theta)\} \longrightarrow 1$ as $z \longrightarrow \infty$ $(0 \leq \theta \leq 1)$; thus $\lambda(t) = 1$ for every t.

What types of limit distributions can be obtained? From the classical theory of extremes and de Haan's results (see Chapter 3 of [BGT]) it follows that apart from trivial cases such as $\lambda(t) \equiv 0, 1$ or ∞, *etc.*, $\lambda(t)$ is either a power or an exponential function. More precisely,

if f is regularly varying at infinity with index ϱ *(*$\varrho \geq 2$ *by (1.1)), then*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I\left(\frac{Z_i}{g(i)} < t\right) = \exp\{-t^{-\varrho}\} \text{ a.s.} \tag{3.8}$$

for every positive t, and

if f is Γ*-varying with auxiliary function* φ *(hence f is rapidly varying), then*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I\left(\frac{Z_i - g(i)}{\varphi\big(g(i)\big)} < t\right) = F(t) \text{ a.s.} \tag{3.9}$$

for every real t, and φ *can be chosen as* $\varphi(z) = \frac{1}{f(z)} \int_0^1 f(x)\,dx$.

Indeed, one can easily see that in (3.6) it suffices to consider the sequence $z = g(n)$, $n \geq 1$. Now let $G(x) = \exp\{-1/f(x)\}$, then G is a probability distribution function and (3.6) can be rephrased as

$$\lim_{n\to\infty} G^n(a_n + tb_n) = \exp\{-1/\lambda(t)\},$$

that is, we arrived at the classical problem of describing all possible limit distributions of maxima of i. i. d. random variables.

3. If $B_m \subset \mathcal{X} \times B_{m-1}$ holds for every $m > 1$, then clearly $p(m) \leq p(i)p(m-i)$, $1 \leq i \leq m-1$; hence the limit $\alpha = \lim_{m\to\infty} p(m)^{1/m}$ exists and is equal to $\inf_m p(m)^{1/m}$; thus $\alpha < 1$. This means that f grows at least at an exponential rate: $\lim_{t\to\infty} \frac{\log f(t)}{t} = \log\frac{1}{\alpha}$, which implies $c = \log(1/\alpha) > 0$.

4. Suppose $(\log\log f)' = \frac{(\log f)'}{\log f}$ vanishes at infinity. Then $\log\log f(z+1) - \log\log f(z) = (\log\log f)'(z+\theta)$, where $\theta = \theta(z) \in [0,1]$, hence $\frac{\log f(z+1)}{\log f(z)} \longrightarrow 1$ as $z \longrightarrow \infty$. Conversely, from the asymptotic equality $\log f(z+1) \sim \log f(z)$, it follows that $(\log\log f)' \longrightarrow 0$ unless $(\log f)'$ exhibits some oscillatory behaviour. Thus our condition (3.4) is only a little more restrictive than the assumption $\log f(z+1) \sim \log f(z)$. On the other hand, we show that the latter is already necessary if we want

$$S_n = \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I\big(Z_i - g(i) < t\big)$$

to converge with positive probability to something different from 0 and 1, even for a single value of t.

Since $\sum_m \mathbf{P}(T_m < m+k) \leq k\sum_m p(m) < \infty$, it follows that the first k experiments can be forgotten when dealing with T_m with a large m. Hence for large m the quantity Z_m does not depend on $X_1, X_2, \ldots, X_k$, k fixed. The zero-one law then gives that $(S_n, n \geq 1)$ converges with probability 1 and the limit is constant a. s.

Now let $f(m-1-t) < i < T_m$, then $Z_i - g(i) < (m-1) - (m-1-t) = t$. Thus

$$S_{T_m} = S_{f(m-1-t)} \frac{\log f(m-1-t)}{\log T_m} + \left(1 - \frac{\log f(m-1-t)}{\log T_m}\right) + o(1)$$

on the event $\{f(m-1-t) < T_m\}$. Consequently,

$$\liminf_{m\to\infty} \frac{\log f(m-1-t)}{\log T_m} \geq 1 \quad \text{a. s.} \tag{3.10}$$

Similary, for $T_m \leq i \leq f(m-t)$ we can write $Z_i - g(i) \geq m - (m-t) = t$; hence on the event $\{T_m < f(m-t)\}$

$$S_{f(m-t)} = S_{T_m} \frac{\log T_m}{\log f(m-t)},$$

implying

$$\liminf_{m\to\infty} \frac{\log T_m}{\log f(m-t)} \geq 1 \quad \text{a. s.}$$

Combining this with (3.10) we get that $\log f(m-t-1) \sim \log f(m-t)$ as $m \longrightarrow \infty$.

5. Cases (ii) and (iii) are more interesting because there $\big(Z_n - g(n), n \geq 1\big)$ is stochastically bounded but does not have a limit distribution, whatever additive normalization should be applied in place of $g(n)$.

Indeed, with the notation $m(n) = \lceil g(n) + t \rceil$ where $\lceil \cdot \rceil$ stands for the ceiling function (upper integer part), we have

$$\mathbf{P}\big(Z_n - g(n) < t\big) = \mathbf{P}\big(Z_n < m(n)\big) = \mathbf{P}(T_{m(n)} > n) = \exp\left\{-\frac{n}{f\big(m(n)\big)}\right\} + o(1)$$
$$= F\Big(\log f\big(m(n)\big) - \log f\big(g(n)\big)\Big) + o(1).$$

Here

$$\log f\big(m(n)\big) - \log f\big(g(n)\big) \approx \big(m(n) - g(n)\big)c$$

for large n, and this can be beyond arbitrary positive or negative bounds for appropriate choices of t. Thus $\big(Z_n - g(n), n \geq 1\big)$ is stochastically bounded.

On the other hand, let us forget the definition of $g(n)$ for a moment and let $g(n)$ be redefined as an arbitrary centralizing sequence tending to infinity. Then for every fixed t there is an infinite sequence of indices n with $m(n) \neq m(n+1)$. As above, we have again

$$\mathbf{P}\big(Z_n - g(n) < t\big) = \exp\left\{-\frac{n}{f\big(m(n)\big)}\right\} + o(1)$$

and

$$\mathbf{P}\big(Z_{n+1} - g(n+1) < t\big) = \exp\left\{-\frac{n+1}{f\big(m(n+1)\big)}\right\} + o(1).$$

The exponentials on the right-hand sides cannot be close unless they approach 0 or 1 because the ratio of the exponents is bounded away from 1 (clearly, $f(m) \sim f(m+1)$ only holds in Case (i)). Thus $Z_n - g(n)$ cannot have a proper limit distribution.

The almost sure central limit theorem was once thought to be a stronger result than its classical counterpart. This was not the case, however. Almost sure convergence is a consequence of the high effectivity of logarithmic weighting and is often implied by the corresponding weak limit theorem. For normalized sums of independent random variables it was shown by Berkes and Dehling [BD]. Furthermore, this implication cannot always be reversed, as shown in [BDM], that is, the *almost sure* limit distribution appears to be even weaker a notion than the ordinary one. This surprising fact is given further support by the above properties of Z_n.

4. **Proofs of the theorems.** In the proof we shall use the following lemma which is a version of Serfling's strong law of large numbers (see [S]).

Define $\ell_1(x) = \log x$ for $x \geq e$ and $\ell_1(x) = 1$ for $x < e$. For $k \geq 2$ let $\ell_k(x) = \ell_1\big(\ell_{k-1}(x)\big)$.

LEMMA 4.1 [M92]. *Let $\xi_1, \xi_2, \ldots$ be arbitrary random variables with finite variances. Suppose there exist a positive non-increasing function $h(\cdot)$ on the positive numbers and a positive integer m such that*

$$\int_1^\infty h(z)\frac{\ell_m(z)}{z\ell_1(z)}\,dz < \infty \tag{4.1}$$

and

$$|\mathbf{E}(\xi_i\xi_j)| \le h(j/i) \quad \text{for all } 1 \le i \le j. \tag{4.2}$$

Then

$$\lim_{n\to\infty} \frac{1}{\log n}\sum_{i=1}^n \frac{1}{i}\xi_i = 0 \quad \text{a.s.} \tag{4.3}$$

If, in addition, the random variables $\{\xi_n, n \ge 1\}$ are uniformly bounded, (1) can be weakened to require that

$$\int_1^\infty \frac{h(z)}{z\ell_1(z)}\,dz < \infty. \tag{4.1$'$}$$

PROOF OF THEOREM 3.1. First we show that in all cases it suffices to deal with the (non-random) sequence

$$\frac{1}{\log f(n-t)}\sum_{i=1}^{f(n-t)} \frac{1}{i}\mathbf{P}\big(Z_i - g(i) < t\big), \quad n \ge 1. \tag{4.4}$$

In Cases (i) and (ii) we can apply the above lemma to the random variables

$$\xi_i = I\big(Z_i - g(i) < t\big) - \mathbf{P}\big(Z_i - g(i) < t\big).$$

By Lemma 2.3 we have for every $i \le j$

$$\begin{aligned} |\mathbf{E}(\xi_i\xi_j)| &= |\mathbf{P}(T_{m(i)} > i, T_{m(j)} > j) - \mathbf{P}(T_{m(i)} > i)\mathbf{P}(T_{m(j)} > j)| \\ &\le \text{const}\cdot\left(\frac{f\big(m(i)\big)}{f\big(m(j)\big)} + p\big(m(j)\big)\right), \end{aligned}$$

where $m(z) = \lceil g(z) + t\rceil$. Firstly, let $h(z) = p\big(m(z)\big)$, then

$$\begin{aligned} \int_{f(m-t-1)}^\infty \frac{h(z)}{z\ell_1(z)}\,dz &= \sum_{i=m}^\infty p(i)\int_{f(i-t-1)}^{f(i-t)} \frac{dz}{z\ell_1(z)} \\ &= \sum_{i=m}^\infty p(i)\big(\log\log f(i-t) - \log\log f(i-t-1)\big). \end{aligned} \tag{4.5}$$

Since $\lim_{i\to\infty}\big(\log\log f(i-t) - \log\log f(i-t-1)\big) = 0$, by (1.1) we obtain that $\frac{h(z)}{z\ell_1(z)}$ is integrable at infinity.

On the other hand,

$$(4.6)\quad \frac{f\big(m(i)\big)}{f\big(m(j)\big)} \le \frac{f\big(g(i)+t+1\big)}{f\big(g(j)+t\big)} = \frac{i}{j}\frac{f\big(g(i)+t+1\big)}{f\big(g(i)\big)}\frac{f\big(g(j)\big)}{f\big(g(j)+t\big)} \sim \frac{i}{j}e^{c(t+1)}e^{-ct} = e^{c}\frac{i}{j}.$$

Thus $|\mathbf{E}(\xi_i\xi_j)| \le \text{const}\cdot\left(\frac{i}{j} + h(\frac{j}{i})\right)$, and hence by Lemma 4.1 $\lim_{n\to\infty}\frac{1}{\log n}\sum_{i=1}^{n}\frac{1}{i}\xi_i =$ 0 holds a. s., showing that attention can be turned from the original sequence of logarithmically weighted sums of indicators to the corresponding sequence of expectations. Besides, it is clearly sufficient to consider a subsequence of the form (4.4), since $\log f(n-t) \sim \log f(n+1-t)$ and the terms of the weighted sum, being probabilities, are bounded.

In Case (iii) the above proof breaks down. In fact, it is no longer possible to apply Lemma 4.1 directly, because $f\big(m(i)\big) = f\big(m(j)\big)$ whenever $f(m-t-1) < i \le j \le f(m-t)$ for some m, and this still allows the ratio j/i to be arbitrary large.

Instead, we can copy the proof of Lemma 4.1 with straightforward modifications.

Let $\varepsilon > 0$ be fixed and $N_n = \max\{m : \log f(m-t) \le (1+\varepsilon)^n\}$. Since $\log f(m+1) \sim \log f(m)$, we have $\log f(N_n - t) \sim (1+\varepsilon)^n$. It is sufficient to show that

$$(4.7)\qquad S = \sum_n \mathbf{E}\left(\frac{1}{\log f(N_n-t)}\sum_{i=1}^{f(N_n-t)}\frac{1}{i}\xi_i\right)^2 < \infty.$$

This will imply that

$$Q_n = \frac{1}{\log f(N_n-t)}\sum_{i=1}^{f(N_n-t)}\frac{1}{i}\xi_i \longrightarrow 0 \quad \text{a. s.}$$

Since for $f(N_{n-1}-t) < m \le f(N_n-t)$ clearly $\frac{\log f(N_n-t)}{\log m} \le 1$ and

$$\left|\frac{1}{\log m}\sum_{i=1}^{m}\frac{1}{i}\xi_i - \frac{\log f(N_n-t)}{\log m}Q_n\right| \le \frac{\log m - \log f(N_n-t)}{\log m} + o(1) \le 1 - \frac{1}{1+\varepsilon} + o(1) < \varepsilon$$

eventually, we obtain that

$$\limsup_{m\to\infty}\left|\frac{1}{\log m}\sum_{i=1}^{m}\frac{1}{i}\xi_i\right| \le \varepsilon \quad \text{a. s.}$$

however small ε be. As in Cases (i) and (ii), here too we can deal with the subsequence (4.4).

In order to prove (4.7) let us first expand the squares in the sum then apply Lemma 2.3 to the terms.

$$\sum_n \mathbf{E}\left(\frac{1}{\log f(N_n-t)}\sum_{i=1}^{f(N_n-t)}\frac{1}{i}\xi_i\right)^2$$
$$= \sum_n \frac{1}{\log^2 f(N_n-t)}\sum_{i=1}^{f(N_n-t)}\sum_{j=1}^{f(N_n-t)}\frac{1}{ij}\mathbf{E}(\xi_i\xi_j)$$
$$\le \text{const}\cdot\sum_n \frac{1}{\log^2 f(N_n-t)}\sum_{1\le i\le j\le f(N_n-t)}\frac{1}{ij}\left(\frac{f\big(m(i)\big)}{f\big(m(j)\big)} + p\big(m(j)\big)\right).$$

Let us group the terms of the inner sum according to the values of $m(i)$ and $m(j)$ and then change the order of summation. We get

$$
\begin{aligned}
\mathcal{S} &\le \text{const} \cdot \sum_n \frac{1}{\log^2 f(N_n - t)} \sum_{v=1}^{N_n} \left(\frac{a(v)}{f(v)} \sum_{u=1}^{v} f(u)\, a(u) + p(v)\, a(v) \log f(v) \right) \\
&= \text{const} \cdot \sum_{v=1}^{\infty} \left(\frac{a(v)}{f(v)} \sum_{u=1}^{v} f(u)\, a(u) + p(v)\, a(v) \log f(v) \right) \sum_{N_n \ge v} \frac{1}{\log^2 f(N_n - t)},
\end{aligned} \tag{4.8}
$$

where

$$a(u) = \sum_{f(u-t-1) < \imath \le f(u-t)} \frac{1}{i} \sim \log f(u-t) - \log f(u-t-1) \le \log f(u-t-\theta)\, \beta(u-t-\theta)$$

with $\theta \in [0,1]$. We can and will suppose that $\beta(z) \sim \beta(z+1)$ as $z \longrightarrow \infty$. Otherwise we can always pass to $\beta_1(z) = \sup\{\varepsilon\beta(\varepsilon z) : 0 \le \varepsilon \le 1\}$; it is easy to see that $\beta_1(z)$ is decreasing but $z\beta_1(z)$ is increasing; thus $\beta_1(z) \sim \beta_1(z+1)$; furthermore, $\beta(z) \le \beta_1(z)$ and $\beta_1^2(2z) \le \frac{1}{4}\beta_1^2(z) + \beta^2(z)$, from which $\int_0^\infty \beta_1^2(z)\, dz \le 4 \int_0^\infty \beta^2(z)\, dz < \infty$. Consequently,

$$a(u) \le \text{const} \cdot \log f(u)\, \beta(u).$$

Since $f(u)$ grows faster than exponentially, it follows that

$$\sum_{u=1}^{v} f(u) \log f(u)\, \beta(u) \le \text{const} \cdot f(v) \log f(v)\, \beta(v).$$

Similarly, by the definition of N_n we have that

$$\sum_{n\, N_n \ge v} \frac{1}{\log^2 f(N_n - t)} \le \text{const} \cdot \frac{1}{\log^2 f(v)}.$$

From all these we obtain

$$\mathcal{S} \le \text{const} \cdot \sum_{v=1}^{\infty} \left(\beta^2(v) + p(v)\, \beta(v) \right).$$

Since $\beta(v) \longrightarrow 0$ as $v \longrightarrow \infty$, we have $\sum_v p(v)\, \beta(v) \log v < \infty$. In addition, $\sum_v \beta^2(v) < \infty$. Thus $\mathcal{S}$ is finite.

Let us turn our attention to (4.4). Since

$$\frac{1}{\log f(n-t)} \sum_{\imath=1}^{f(n-t)} \frac{1}{i} \mathbf{P}\big(Z_\imath - g(i) < t\big) = \frac{1}{\log f(n-t)} \sum_{m=1}^{n} \sum_{\imath\, m(\imath)=m} \frac{1}{i} \mathbf{P}(T_m > i),$$

it is sufficient to study the asymptotics of

$$\sigma_m = \frac{1}{a(m)} \sum_{\imath\, m(\imath)=m} \frac{1}{i} \mathbf{P}(T_m > i) = \frac{1}{a(m)} \sum_{f(m-t-1) < \imath \le f(m-t)} \frac{1}{i} \exp\left\{ -\frac{i}{f(m)} \right\} + o(1)$$

as $m \longrightarrow \infty$.

https://doi.org/10.4153/CJM-1993-070-x Published online by Cambridge University Press

In Case (i), since $f(m+\delta) \sim f(m)$ for every fixed real δ, the exponential terms converge to e^{-1} and thus $\sigma_m \longrightarrow e^{-1}$.

In Cases (ii) and (iii) let us introduce z by

$$i = f(m-1-t)^{1-z} f(m-t)^z.$$

Then

$$\frac{\Delta i}{i} \sim \big(\log f(m-t) - \log f(m-1-t)\big)\Delta z \sim a(m)\Delta z$$

and

$$\begin{aligned}\frac{i}{f(m)} &= \exp\Big\{-(1-z)\big(\log f(m) - \log f(m-t-1)\big) - z\big(\log f(m) - \log f(m-t)\big)\Big\}\\ &= \exp\Big\{-\big((1-z)(t+1) + zt\big)(\log f)'(m)\big(1+o(1)\big)\Big\}.\end{aligned}$$

In Case (ii) this is approximately $\exp\{-c(t+1-z)\}$, and thus we obtain that

$$\sigma_m \sim \int_0^1 F\big(c(t+1-z)\big)\,dz = \int_0^1 F\big(c(t+z)\big)\,dz.$$

In Case (iii) the limit of the exponent is $-\infty$ when $i \longrightarrow \infty$ through values with $z < t+1-\varepsilon$, and $+\infty$ when $z > t+1+\varepsilon$ ($\varepsilon > 0$ arbitrary). Hence for $-1 < t < 0$

$$\sigma_m \sim (t+1)F(+\infty) - tF(-\infty) = t+1. \qquad \blacksquare$$

PROOF OF THEOREM 3.2. Let us see what has to be changed in the above proof when applied to (3.7) with $\varphi(z) \longrightarrow \infty$.

Now (4.4) reads

$$\text{(4.9)} \qquad \frac{1}{\log f\big(n - t\varphi(n)\big)} \sum_{i=1}^{f(n-t\varphi(n))} \frac{1}{i}\mathbf{P}\Big(Z_i - g(i) < t\varphi\big(g(i)\big)\Big), \quad n \ge 1,$$

and the lines below that are to be corrected correspondingly. Again,

$$\log\log f\big(i - t\varphi(i)\big) - \log\log f\big(i-1-t\varphi(i-1)\big) \longrightarrow 0,$$

because

$$\log f\big(i - t\varphi(i)\big) = \log f(i) + \log\lambda(-t) + o(1) \sim \log f\big(i-1-t\varphi(i-1)\big).$$

Hence (4.5) can be adapted.

This time in (4.6) $m(i) = \big\lceil g(i) + t\varphi\big(g(i)\big)\big\rceil$, and hence for every λ-continuity point $t \in I$ we have

$$\frac{f\big(m(i)\big)}{f\big(m(j)\big)} \le \frac{f\Big(g(i) + t\varphi\big(g(i)\big) + 1\Big)}{f\big(g(j) + t\varphi(j)\big)} \sim \frac{i\lambda(t)}{j\lambda(t)} = \frac{i}{j}.$$

The rest is even simpler, namely, the limit relation

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z_i - g(i)}{\varphi\big(g(i)\big)} < t\right) = \exp\{-1/\lambda(t)\}$$

can be shown directly, without any summation procedure. Indeed, we can suppose that $0 < \lambda(t) < \infty$; then

$$\mathbf{P}\left(\frac{Z_\iota - g(i)}{\varphi\big(g(i)\big)} < t\right) = P(T_{m(\iota)} > i) = \exp\left\{-\frac{i}{f\big(m(i)\big)}\right\} + o(1).$$

Since $c = 0, f\big(m(i)\big) \sim f\Big(g(i)+t\varphi\big(g(i)\big)\Big) \sim f\big(g(i)\big)\lambda(t) = i\lambda(i)$, completing the proof. ■

5. **Examples.** Following [CsFK] we are going to specialize our results to obtain interesting corollaries in three important particular cases.

5.1 *The longest (k-interrupted) head run.* In this example $X_1, X_2, \ldots$ are i. i. d. Bernoulli random variables with $\mathbf{P}(X_1 = 1) = 1 - \mathbf{P}(X_1 = 0) = p$ $(0 < p < 1)$. Interpreting the values 1 and 0 as heads and tails, we can think of the sequence $(X_n, n \geq 1)$ as successive coin tosses with a possibly biased coin. For a fixed non-negative integer k let

$$B_m = \{(x_1, \ldots, x_m) \in \{0, 1\}^m : x_1 + \cdots + x_m \geq m - k\}.$$

Then T_m is the number of tosses needed for a k-interrupted head run of length m to appear, *i.e.*, T_m is the first time when the number of tails among the last m outcomes is at most k. The corresponding Z_n is the length of the longest k-interrupted head run observed in n experiments. Particularly, when $k = 0$, T_m is the waiting time for a pure head run of length m and Z_n is the longest head run in n tosses.

Clearly, $p(m) = \sum_{\iota=0}^{k} \binom{m}{\iota} p^{m-\iota} q^\iota \sim (\frac{qm}{p})^k \frac{1}{k!} p^m$ $(q = 1 - p)$ satisfies (1.1). The asymptotics of the expectation $E(m)$ can be found by using Lemma 2.1, but one can also turn to Theorem 3.A of [F], where the following limit theorem is found.

$$\lim_{n\to\infty} \mathbf{P}\left(T_m\left(\frac{qm}{p}\right)^k \frac{1}{k!} p^m q > t\right) = e^{-t}, \quad t > 0$$

(in fact, only for a fair coin, but the proof can easily be extended to the non-symmetric case). Hence

$$f(m) = k! \left(\frac{p}{qm}\right)^k \frac{1}{p^m q}$$

will do: this is Case (ii) of our Theorem 3.1 with $c = \log \frac{1}{p}$. Let Log denote the logarithm to the base $1/p$, then

$$g(n) = \operatorname{Log} n + k \operatorname{Log}\operatorname{Log} n + \operatorname{Log} \frac{q}{k!}\left(\frac{q}{p}\right)^k + o(1).$$

Thus (3.3) gives the following theorem.

COROLLARY 5.1. *Let Z_n denote the length of the longest k-interrupted head run in n tosses. Then*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{\iota=1}^{n} \frac{1}{i} I(Z_\iota - \operatorname{Log} i - k \operatorname{Log}\operatorname{Log} i < t) = \int_t^{t+1} \exp\left\{-\frac{q}{k!}\left(\frac{q}{p}\right)^k p^z\right\} dz \quad \text{a. s.}$$ ■

https://doi.org/10.4153/CJM-1993-070-x Published online by Cambridge University Press

5.2 *The longest tube.* Let $X_1, X_2, \ldots$ be i. i. d. integer valued random variables with common distribution $\mathbf{P}(X_1 = k) = p_k$, $k \in \mathbb{Z}$. Note that neither recurrence nor finite moments are required. Assume that the random walk $S_n = X_1 + \cdots + X_n$, $n \geq 1$, is aperiodic and every integer point can be reached with positive probability, *i.e.*, for every $k \in \mathbb{Z}$ there exists an n_0 such that $\mathbf{P}(S_n = k) > 0$ for $n \geq n_0$. This particularly implies the irreducibility of the random walk.

Let d be a positive integer and

$$B_m = \Big\{(x_1, \ldots, x_m) \in \mathbb{Z}^m : \Big|\sum_{t=i}^{j} x_t\Big| < d, 1 \leq i \leq j \leq m\Big\}. \tag{5.1}$$

Then T_m is the waiting time for the random walk to stay, during m consecutive steps, in any of d consecutive integer points. Plotting the position of the random walk against time we find that the graph proceeds in a "tube" of length m and width $d + 1$. In other words, the oscillation of the random walk during m steps remains small.

The problem of the narrowest tube or small increments has been investigated in several papers, see [CsR, Section 3.3] or [CsF]. In fact, they defined the tube in a slightly different way, by

$$B_m = \Big\{(x_1, \ldots, x_m) \in \mathbb{Z}^m : \Big|\sum_{t=1}^{j} x_t\Big| < \alpha, 1 \leq j \leq m\Big\}. \tag{5.2}$$

This corresponds to $d = 2\alpha - 1$ above, but not exactly, because here the random walk is forced to start from the centre of the tube when the awaited block begins. The difference between (5.1) and (5.2) causes a constant multiplier in $f(m)$, that is, an additive constant in $g(m)$.

Let Q_d be the $d \times d$ matrix with entries $q_d(i,j) = p_{i-j}$, $1 \leq i \leq d$, $1 \leq j \leq d$. Denote the entries of the power Q_d^m by $q_d^m(i,j)$. If d is large enough, say $d \geq d_0$, there exists an m such that $q_d^m(i,j) > 0$, $1 \leq i \leq d$, $1 \leq j \leq d$. By the Perron-Frobenius theory of positive matrices (see [B]) it follows that Q_d has a unique characteristic number ϱ_d with maximal modulus, ϱ_d is positive, simple and has associated positive right and left eigenvectors u_d and v_d, resp. In terms of these we have

$$q_d^m(i,j) \sim \varrho_d^m u_d(i) v_d(j) \Big/ \sum_{t=1}^{d} u_d(t) v_d(t) \tag{5.3}$$

as $m \longrightarrow \infty$. Further, ϱ_d is strictly increasing for $d \geq d_0$.

For every pair of positive integers (r, s) let

$$B_{r,s} = \Big\{(x_1, \ldots, x_m) \in \mathbb{Z}^m : -r < \sum_{t=1}^{j} x_t < s, 1 \leq j \leq m\Big\}.$$

Then $B_m = \bigcup_{r=1}^{d} B_{r,d+1-r}$ and

$$p(m) = p_d(m) = \mathbf{P}(X_{1,m} \in B_m) = \sum_{r=1}^{d} \mathbf{P}(X_{1,m} \in B_{r,d+1-r}) - \sum_{r=1}^{d-1} \mathbf{P}(X_{1,m} \in B_{r,d-r}).$$

Clearly,

$$\mathbf{P}(X_{1,m} \in B_{r,s}) = \sum_{j=1}^{r+s-1} q_{r+s-1}^{m}(r,j);$$

hence for $d > d_0$

$$\begin{aligned} p_d(m) &= \sum_{i=1}^{d}\sum_{j=1}^{d} q_d^m(i,j) - \sum_{i=1}^{d-1}\sum_{j=1}^{d-1} q_{d-1}^m(i,j) \\ &= \varrho_d^m \sum_{i=1}^{d} u_d(i) \sum_{j=1}^{d} v_d(j) \Big/ \sum_{t=1}^{d} u_d(t) v_d(t)\big(1 + o(1)\big) - O(\varrho_{d-1}^m) \\ &\sim \varrho_d^m \sum_{i=1}^{d} u_d(i) \sum_{j=1}^{d} v_d(j) \Big/ \sum_{t=1}^{d} u_d(t) v_d(t) - \kappa_d \varrho_d^m \end{aligned}$$

as $m \longrightarrow \infty$.

To obtain asymptotics for $E(m)$ we can apply Lemma 2.1. Let us compute the probability $\mathbf{P}(T_m = m + k)$. Obviously,

$$\begin{aligned} \mathbf{P}(T_m = m+k) &= \mathbf{P}(X_{k+1,m} \in B_m \text{ but } X_{i,m} \notin B_m, 1 \le i \le k) \\ &= p_d(m) - \mathbf{P}(X_{k+1,m} \in B_m \text{ and } X_{i,m} \in B_m \text{ for some } i, 1 \le i \le k). \end{aligned}$$

Suppose $X_{i,m} \in B_m$ and $X_{k+1,m} \in B_m$ hold simultaneously, that is, we have two blocks of length m, each in a tube. If they are in the same tube, then they together make a longer block still in a tube: $X_{k,m+1} \in B_{m+1}$. If they are not in the same tube, then one can find a block of length $m - k$, namely $X_{k+1,m-k}$, in a narrower tube. Thus

$$\mathbf{P}(T_m = m+k) \ge p_d(m) - p_d(m+1) - p_{d-1}(m-k) \sim (1 - \varrho_d) p_d(m).$$

On the other hand,

$$\mathbf{P}(T_m = m+k) \le p_d(m) - \mathbf{P}(X_{k,m+1} \in B_{m+1}) = p_d(m) - p_d(m+1) \sim (1 - \varrho_d) p_d(m).$$

By Lemma 2.1 we have

$$E(m) \sim f(m) = 1/\kappa_d(1-\varrho_d)\varrho_d^m, \quad c = \log\frac{1}{\varrho_d}, \quad g(m) = \frac{\log m + \log\big(\kappa_d(1-\varrho_d)\big)}{\log(1/\varrho_d)}.$$

As a particular case suppose the step size of the random walk is bounded by 1, *i.e.*, $p_1 = p, p_0 = q, p_{-1} = r$, where $p + q + r = 1$ and let us choose p, q, r strictly positive so as to avoid periodicity. Then, after a little algebra one obtains that

$$\begin{aligned} d_0 &= 1, \\ \varrho_d &= q + \sqrt{pr}\cos\frac{\pi}{d+1}, \\ u_d(i) &= p^{\frac{i-1}{2}} r^{\frac{d-i}{2}} \sin\frac{i\pi}{d+1}, \quad 1 \le i \le d, \\ v_d(i) &= p^{\frac{d-i}{2}} r^{\frac{i-1}{2}} \sin\frac{i\pi}{d+1} = u_d(d+1-i), \quad 1 \le i \le d, \end{aligned}$$

hence

$$\kappa_d = \frac{(\sum_{\iota=1}^{d} p^{\frac{\iota-1}{2}} r^{\frac{d-\iota}{2}} \sin\frac{\iota\pi}{d+1})^2}{(pr)^{\frac{d-1}{2}} \sum_{\iota=1}^{d} \sin^2\frac{\iota\pi}{d+1}} = \frac{2}{d+1}(pr)^{\frac{1-d}{2}} \frac{(p^{\frac{d+1}{2}} + r^{\frac{d+1}{2}})^2 \sin^2\frac{\pi}{d+1}}{(1-\varrho_d)^2}.$$

In particular, in the symmetric case ($p = r$)

$$\kappa_d = \frac{2}{d+1}\left(\operatorname{cotan}\frac{\pi}{2(d+1)}\right)^2.$$

From (3.3) we obtain the following theorem.

COROLLARY 5.2. *Let Z_n denote the length of the longest d-tube of the random walk S_ι, $1 \le i \le n$. Then for $d \ge d_0$ we have*

$$\lim_{n\to\infty}\frac{1}{\log n}\sum_{\iota=1}^{n}\frac{1}{i} I\left(Z_\iota - \frac{\log i}{\log(1/\varrho_d)} < t\right) = \int_t^{t+1} \exp\{-\kappa_d(1-\varrho_d)\varrho_d^z\}\,dz \quad \text{a. s.} \quad \blacksquare$$

5.3 *Long blocks with few monotone segments.* Let $X_1, X_2, \ldots$ be i. i. d. real valued random variables with continuous common distribution. Define

$$B_m = \{(x_1,\ldots,x_m) \in \mathbb{R}^m : x_1 < x_2 < \cdots < x_m\}.$$

Then Z_m is the length of the longest increasing block in the first n experiments. This random variable was studied by Révész [R].

Clearly, in this case $p(m) = \frac{1}{m!}$ and $\mathbf{P}(T_m = m+k) = \frac{m}{(m+1)!}$, $k < m$, hence $E(m) \sim m!$. Thus $f(m) = \Gamma(m+1)$, $g(m) = \Gamma^{-1}(m) - 1$, $c = +\infty$, and f obviously satisfies condition (3.4) of Theorem 2.1. Consequently, we obtain the following theorem.

COROLLARY 5.3. *Let Z_n denote the length of the longest increasing block in the sequence $X_1, X_2, \ldots, X_n$ of* i. i. d. *continuous random variables. Then*

$$\lim_{n\to\infty}\frac{1}{\log n}\sum_{\iota=1}^{n}\frac{1}{i} I\big(Z_\iota - \Gamma^{-1}(i) < t\big) = \begin{cases} 0 & \text{if } t \le -2 \\ 2+t & \text{if } -2 < t < 1, \\ 1 & \text{if } -1 \le t. \end{cases}$$

with probability 1. $\blacksquare$

A natural generalization of the longest monotone block is the longest block that can be split into d or less monotone segments. In this case

$$B_m = \left\{(x_1,\ldots,x_m) \in \mathbb{R}^m : \sum_{\iota=2}^{m-1} I\big((x_{\iota+1}-x_\iota)(x_\iota - x_{\iota-1}) < 0\big) < d\right\}.$$

Let d be fixed and $m \longrightarrow \infty$. We are going to find the asymptotics of $p(m)$ and $E(m)$. Since each order of the random variables $X_1, \ldots, X_m$ is equally probable, we only have to count all permutations of the numbers $\{1, \ldots, m\}$ consisting of d or less monotone blocks (such permutations will be referred to as good ones). Let us distribute the numbers among the blocks one by one in increasing order Each number can be assigned to any of d blocks, and thus we have d^m possibilities Fixing the first block ascending or descending we determine the order of elements in all blocks, and this will double the number of displacements In this way every good permutation is counted at least once In almost all of these $2d^m$ cases each block has $\frac{m}{d}\big(1 + o(1)\big)$ elements Of course, good permutations with less than d monotone blocks have also been considered, besides, they were counted several times, but in a typical good permutation there are exactly d monotone blocks and they are approximately equal in length, so all the other good permutations can be left out of consideration The block where a local maximum or minimum belongs is not unique delimiters between monotone blocks can be put into any of the two neighboring blocks Thus each typical good permutation has been counted 2^{d-1} times From all these we obtain

$$p(m) \sim \frac{1}{m!} d^m 2^{2-d}$$

For computing $E(m)$ by the help of Lemma 2 1 let us estimate $\mathbf{P}(T_m = m + k)$ Suppose $T_m = m + k$ and $X_{k+1,m}$ corresponds to a typical good permutation Then $X_{k,m+1}$ can be obtained from a typical good permutation of length $m + 1$ by corrupting the order in the first block at the first place This can be done in ℓ different ways where ℓ is the length of the first monotone segment, that is, $\ell \sim \frac{m}{d}$ The last monotone segment is typically longer than k, and thus $X_{i,m}$ cannot be good for $i \leq k$ if X_k does not match $X_{k+1,m}$ Hence

$$\mathbf{P}(T_m = m + k) \sim \frac{m}{d} \frac{1}{(m+1)!} d^{m+1} 2^{2-d} \sim p(m),$$

consequently $E(m) \sim 1/p(m)$, $c = \infty$, and a possible choice of $g(m)$ is

$$g(m) = d\,\Gamma^{-1}(n^{1/d}) - \frac{d+1}{2}$$

COROLLARY 5 4 *Let Z_n denote the length of the longest block in the sequence $X_1, X_2, \ldots, X_n$ of* i i d *continuous random variables that consists of d or less monotone segments Then*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} I\big(Z_i - d\,\Gamma^{-1}(i^{1/d}) < t\big) = \begin{cases} 0 & \text{if } t \leq -\frac{d+3}{2}, \\ \frac{d+3}{2} + t & \text{if } -\frac{d+3}{2} \leq t < -\frac{d+1}{2}, \\ 1 & \text{if } -\frac{d+1}{2} \leq t \end{cases}$$

with probability 1 ∎

REFERENCES

[B] R Bellman *Introduction to Matrix Analysis* McGraw Hill New York 1960

[BD] I Berkes and H Dehling *Some limit theorems in log density* Ann Probabl to appear

[BDM] I Berkes H Dehling and T G Mori *Counterexamples related to the* a s *central limit theorem* Studia Sci Math Hungar **26**(1991) 153–164

[BGT] N H Bingham C M Goldie and J L Teugels *Regular Variation* Encyclopedia of Mathematics **27** Cambridge University Press 1987

[CsF] E Csakı and A Foldes, *The narrowest tube of a recurrent random walk*, Z Wahrsch Verw Geb **66**(1984), 387–403

[CsFK] E Csákı, A Foldes and J Komlos, *Lımıt theorems for Erdos Renyı type problems*, Studıa Scı Math Hungar **22**(1987), 321–332

[CsR] M Csorgő and P Révesz, *Strong Approxımatıon ın Probabılıty and Statıstıcs*, Akademıaı Kıado, Budapest, 1981

[ER] P Erdős and P Révész, *On the length of the longest head run* In Topıcs ın Informatıon Theory, Colloquıa Math Soc J Bolyaı **16**, Keszthely, Hungary, (1975), 219–228

[F] A Foldes, *The lımıt dıstrıbutıon of the length of the longest head-run*, Perıod Math Hungar **10**(1979), 301–310

[GSW] L Gordon, M F Schıllıng and M S Waterman, *An extreme value theory for long head runs*, Probab Th Rel Fıelds **72**(1986), 279–287

[M85] T F Mórı, *Large devıatıon results for waıtıng tımes ın repeated experıments*, Acta Math Hungar **45**(1985), 213–221

[M92] ———, *On the strong law of large numbers for logarıthmıcally weıghted sums*, Annales Unıv Scı Budapest, R Eotvos Nom , Sec Math , (1992), to appear

[R] P Révesz, *Three problems on the length of ıncreasıng runs*, Stoch Proc Appl **15**(1983), 169–179

[S] W F Stout, *Almost Sure Convergence*, Academıc Press, New York, 1974

Department of Probabılıty Theory and Statıstıcs
Eotvos Loránd Unıversıty
H-1088 Budapest
Hungary