Z. Wahrscheinlichkeitstheorie verw. Gebiete
61, 59–73 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On the Longest Run of Coincidences

Tibor Nemetz[1] and Norbert Kusolitsch[2]

[1] Mathematical Institute of the Hungarian Academy of Sciences, Realtanoda u. 13–15, Budapest, Hungary
[2] Institut für Statistik und Wahrscheinlichkeitstheorie, Technische Universität Wien, Argentinierstrasse 8/7, Wien, Österreich

Summary. Consider the rectangles of the first $k(n)$ lines of length n in the right-upper integer-lattice, and suppose that its points are labelled randomly by the numbers $1, 2, \ldots, m$. The time i is called coincidence if the points $(i, 1)$, $(i, 2)$, ..., $(i, k(n))$ are labelled identically. Asymptotic properties of the longest run of coincidences are discussed under different conditions on $k(n)$.

The results are related to a problem of P. Révész: If the points of an $n \times n$ integer-lattice are coloured red and white randomly, what is the largest area of rectangles with red points only.

A conjecture is formulated, indicating some peculiar number-theoretic characteristics of some limit-relations in this area.

1. Introduction

In this paper we are concerned with a problem of P. Révész [3]. Our formulation differs, however, slightly from his one.

Let all lattice points (i,j), $i,j \in \mathcal{N}$ be labelled randomly with one of the numbers $1, 2, \ldots, m$. Let $\xi_{i,j}$ denote the label of the point (i,j). We will suppose that the random variables $\xi_{i,j}$ are independent, identically distributed (i.i.d.) with the probability distribution

$$P(\xi_{i,j} = s) = p_s, \qquad s = 1, \ldots, m. \tag{1}$$

Let us furthermore consider a deterministic, monotone non-decreasing sequence of integers $k(n)$ with the property $k(n) \leqq n$. For fixed n we can interpret the rectangle $\{(i,j): 1 \leqq i \leqq k(n), 1 \leqq j \leqq n\}$ as $k(n)$ lines with n points on each. We can also ask if the labels in the j-th column coincide or not and if they do, what is the common label. To be more formal let

$$\eta_j = \eta_j(r, k(n), n) = \begin{cases} 1 & \text{if } 1 \leqq \xi_{1,j} = \ldots = \xi_{k,j} \leqq r \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

0044-3719/82/0061/0059/$03.00

In this paper we are concerned with the asymptotic behaviour of the longest run of 1's in $\eta_1, \eta_2, \ldots, \eta_n$. We recall that a run of 1's is a sequence of consecutive 1's preceeded and followed by a 0, except at the two end-times, where the definition is the natural one. The length of a run is the number of its elements.

Properties of coincidences were first utilized by Shannon [4]. He used the total number of coincidences rather than runs of them as a statistic, which provides an effective way to check the identity of two cryptosystems. In this case the sequence of $k(n)$ is a constant sequence $k(n) \equiv k$ with small values of k and $r = m = 26$ or 27.

For the case of $k(n) \equiv k$ one can easily deduce the asymptotic behaviour of the longest run from known results. Indead, the probabilities

$$P_{k(n)} = P(\eta_1(r, k(n), n) = 1) = \sum_{s=1}^{r} p_s^k \tag{3}$$

do not depend on n in this case. Let us denote their joint value by p, i.e. $p = \sum_{s=1}^{r} p_s^k$. Then we can apply the following theorem.

Theorem I. *Let $\eta_1, \eta_2, \ldots$ be i.i.d. random variables with*

$$P(\eta_i = 1) = p, \qquad P(\eta_i = 0) = 1 - p.$$

Let $v(n)$ be the length of the longest run of 1's among $\eta_1, \ldots, \eta_n$. Then

$$P\left(\lim_n \frac{v(n)}{\log n} = \frac{1}{\log \frac{1}{p}}\right) = 1.$$

(For a proof of this theorem see, e.g. Petrov [2].)

This means, that it suffices to consider the case of non-bounded $k(n)$, which is most interesting when $\dfrac{k(n) \log \frac{1}{p}}{\log n}$ has a positive limit value. If this limit value is not the reciprocal of an integer then the length of the longest run converges to a constant. Otherwise it is still concentrated on 2 values, but surprisingly it does not converge even in distribution.

In part 2 we will prove limit theorems, in part 3 questions on expected values will be discussed. The last part utilizes these results for investigating the area of the largest rectangle of identical elements. It is shown that this area is essentially the same as that of the largest square of identical elements.

This latter was found by Révész in [3].

2. Convergence of the Length of the Longest Coincidence-Run: Weak and Strong Type Results

First we prove two inequalities on the longest run.

Lemma 1. *Let $\zeta_1, \zeta_2, \ldots, \zeta_n, \ldots$ be i.i.d. random variables with*

$$P(\zeta_1 = 1) = p \quad \textit{and} \quad P(\zeta_1 = 0) = 1 - p$$

and let us denote the length of the longest run of 1's *among the first n variables by* $\nu(n)$, *then we have*

$$P(\nu(n)<N)\leqq(1-p^N)^{-1}\exp(-np^N/N) \tag{4}$$

and

$$P(\nu(n)\geqq N)\leqq np^N \tag{5}$$

for all integers N.

Proof. Let A_i denote the event, that at least one of the random variables $\zeta_i,\dots,\zeta_{i+N-1}$ is zero i.e. that no run of the length N or more start with the index i. Then

$$P(\nu(n)<N)\leqq P(A_1 A_{1+N}\cdots A_{1+hN})$$

with $h=\left[\frac{n-N}{N}\right]$. $[x]$ is the largest integer $\leqq x$.

A_i and A_j are independent for $|i-j|\geqq N$, we get

$$P(\nu(n)<N)\leqq P(A_1)^{h+1}=(1-p^N)^{\left[\frac{n}{N}\right]}\leqq(1-p^N)^{-1}\exp(-np^N/N).$$

On the other hand, $\nu(n)\geqq N$ means that at least one of the complementary events $\bar{A}_i$, $i=1,\dots,n-N+1$ occurs. Therefore

$$P(\nu(n)\geqq N)\leqq\sum_{i=1}^{n-N+1}P(\bar{A}_i)=(n-N+1)p^N\leqq np^N. \quad \square$$

Let us turn back to our original problem. Let the random variables $\xi_{i,j}$ and a deterministic monotone nondecreasing sequence of integers $k(n)$ be defined as in the introduction and suppose that $k(n)$ goes to infinity. Then the distribution of the random variables η_j given by (2) depends on n. The probability

$$P_{k(n)}=P(\eta_1(r,k(n),n)=1)=\sum_{s=1}^{r}p_s^{k(n)}$$

is essentially determined by $p_0=\max_{1\leqq s\leqq r}p_s$: Indeed,

$$p_0^{k(n)}\leqq P_{k(n)}\leqq r\,p_0^{k(n)}. \tag{6}$$

Applying this inequality, Lemma 1 implies the following lemma.

Lemma 2. *Let* $\alpha(k,n)=\alpha(k(n),n)$ *be the length of the longest run of coincidences in the first r labels that is the length of the longest run of* 1's *in* $\eta_1(r,k(n),n),\dots,\eta_n(r,k(n),n)$. *Then for arbitrary* $\varepsilon>0$, *we have*

$$P\left(\frac{\alpha(k,n)\,k(n)\log\frac{1}{p_0}}{\log n}>1+\varepsilon\right)\leqq n^{-\varepsilon/2}.$$

Proof. Applying (5) and the second inequality of (6) we get

$$P\left(\frac{\alpha(k,n)\,k(n)\log\frac{1}{p_0}}{\log n}>1+\varepsilon\right)=P\left(\alpha(k,n)\geqq\left[\frac{(1+\varepsilon)\log n}{k(n)\log\frac{1}{p_0}}\right]+1\right)$$

$$\leqq n(r\,p_0^{k(n)})^{\left[\frac{(1+\varepsilon)\log n}{k(n)\log\frac{1}{p_0}}\right]+1}$$

$$\leqq n^{-\varepsilon}\,r^{\frac{(1+\varepsilon)\log n}{k(n)\log\frac{1}{p_0}}+1}\leqq n^{-\varepsilon/2}$$

if $k(n)$ is large enough. □

This lemma leads us to the following result:

Theorem 1. *We have*

$$P\left(\limsup_n \frac{\alpha(k,n)\,k(n)\log\frac{1}{p_0}}{\log n}\leqq 1\right)=1.$$

Proof. The above lemma implies

$$P\left(\frac{\alpha(k,2^m)\,k(2^m)\log\frac{1}{p_0}}{\log 2^m}>1+\varepsilon\right)\leqq 2^{-\frac{m\varepsilon}{2}}$$

for every $\varepsilon>0$ and m large enough.

Since $\sum_{m=1}^{\infty} 2^{-m\varepsilon/2}<\infty$, this inequality and the Borel-Cantelli-lemma yield

$$\limsup \frac{\alpha(k,2^m)\,k(2^m)\log\frac{1}{p_0}}{\log 2^m}\leqq 1 \quad \text{a.s.} \tag{7}$$

On the other hand, for every integer n with $2^m\leqq n\leqq 2^{m+1}-1$ we get

$$\frac{\alpha(k,n)\,k(n)}{\log n}\leqq\frac{\alpha(k(n),2^{m+1})\,k(n)}{\log 2^m}.$$

Therefore

$$\limsup_n \frac{\alpha(k(n),n)\,k(n)}{\log n}\leqq\limsup\frac{\alpha(k(n),2^{m+1})\,k(n)}{\log 2^m}$$

$$=\limsup_m \frac{\alpha(k(n),2^{m+1})\,k(n)}{\log 2^{m+1}}.$$

The same reasoning that leads to (7) implies

$$\limsup \frac{\alpha(k(n), 2^{m+1})\, k(n) \log \frac{1}{p_0}}{\log 2^{m+1}} \leqq 1 \quad \text{a.s.}$$

This completes our proof. □

If $k(n)$ tends to infinity fast enough, then $\alpha(k, n)$ becomes bounded:

Corollary 1. *If* $\liminf_n \frac{k(n) \log \frac{1}{p_0}}{\log n} > x$ $(x > 0)$, *then*

$$P\left(\limsup_n \alpha(k, n) < \frac{1}{x}\right) = 1.$$

Proof. The assumption $\liminf_n \frac{k(n) \log \frac{1}{p_0}}{\log n} > x$ and Theorem 1 imply

$$\limsup \alpha(k, n) < \frac{1}{x} \quad \text{a.s.} \quad \square$$

Remark. Corollary 1 shows that, if $\liminf \frac{k(n) \log \frac{1}{p_0}}{\log n} > 1$, then $\alpha(k, n)$ becomes 0 a.s. Accordingly we may restrict ourselves to the case, where $x < 1$ in the above corollary.

Theorem 2. *If* $\limsup_n \frac{k(n) \log \frac{1}{p_0}}{\log n} < x$ $(x > 0)$, and $k(n) \to +\infty$, then

$$P\left(\liminf_n \alpha(k, n) \geqq \left[\frac{1}{x}\right]\right) = 1.$$

Proof. From our assumption follows the existence of some $\varepsilon > 0$ such that for all large enough integers n

$$k(n) \leqq x(1-\varepsilon) \log n \Big/ \log \frac{1}{p_0} \tag{8}$$

holds true.

The probability $P\left(\alpha(k, n) < \left[\frac{1}{x}\right]\right)$ can be estimated from above by using (4) with $N = \left[\frac{1}{x}\right]$ and $p = P_{k(n)}$:

$$P\left(\alpha(k,n)<\left[\frac{1}{x}\right]\right)\leqq\left(1-P_{k(n)}^{\left[\frac{1}{x}\right]}\right)^{-1}\exp\left(-nP_{k(n)}^{\left[\frac{1}{x}\right]}\Big/\left[\frac{1}{x}\right]\right)$$

$$\leqq 2\exp\left(-nx\;p_0^{k(n)\left(\frac{1}{x}-1\right)}\right)$$

if $k(n)$ is large enough.

Applying (8) we get

$$P\left(\alpha(k,n)<\left[\frac{1}{x}\right]\right)\leqq 2\exp\left(-\frac{nxp_0}{n^{(1-x)(1-\varepsilon)}}\right)\leqq 2\exp(-xn^{\varepsilon}).$$

Since the sum $\sum_{n=1}^{\infty}\exp(-xn^{\varepsilon})$ is finite, an application of the Borel-Cantelli-lemma proves the theorem. □

Summing up the results of Corollary 1 and Theorem 2, we get

Corollary 2. *If* $\lim_n \frac{k(n)\log\frac{1}{p_0}}{\log n}=x$ *and* $\frac{1}{x}$ *is not an integer, then* $\alpha(k,n)$ *converges to* $\left[\frac{1}{x}\right]$ *almost surely.*

Remark. If $\frac{1}{x}$ in Corollary (2.7) is an integer, then $\alpha(k,n)$ does not necessarily converge, as is shown in the next example.

Example 1. Let $r=1$ and suppose $p_1<2^{-x}$. Assume, furthermore, that the sequence $k(n)$ satisfies

$$\left|k(n)-\frac{x\log n}{\log\frac{1}{p_1}}\right|<C \quad \text{for some constant } C. \tag{9}$$

Then neither $\lim_n P\left(\alpha(k,n)=\frac{1}{x}\right)$ nor $\lim P\left(\alpha(k,n)=\frac{1}{x}-1\right)$ exist. Note that $\lim_n P(\alpha(k,n)=y)=0$ for $y\neq\frac{1}{x}-1,\ \frac{1}{x}$.

Proof. We need a result of Erdös-Révész (see [1] Theorem 5), so first we recall their result in our terms:

For all integers n and M the following inequalities hold true

$$1-\left(1-p_1^{k(n)M}(1+M)\right)^{\left[\frac{1}{2}\left[\frac{n-2M}{M}\right]\right]+1}$$

$$\leqq P(\alpha(k,n)\geqq M)\leqq 1-\left(1-p_1^{k(n)M}(1+M)\right)^{\left[\frac{n-2M}{M}\right]+1} \tag{10}$$

For the sequence $k(n)$ is an integer sequence and the differences $\left|k(n)-\frac{x\log n}{\log\frac{1}{p_1}}\right|$ are bounded, there must exist two subsequences $(k(n_i^1))$ and $(k(n_i^2))$ of $(k(n))$ and a real constant T such that

$$\lim_i \left(k(n_i^1) - \frac{x \log n_i^1}{\log \frac{1}{p_1}} \right) \leqq T$$

and

$$\lim_i \left(k(n_i^2) - \frac{x \log n_i^2}{\log \frac{1}{p_1}} \right) \geqq T+1.$$

Because both bounds of (10), considered as functions of $k(n)$, are monotone decreasing, it follows that

$$\lim_i \left(1 - \left(1 - \frac{p_1^{T/x}}{n_i^1} \left(1 + \frac{1}{x} \right) \right)^{\left[\frac{1}{2}\left[\frac{n_i^1 - 2/x}{1/x}\right]\right]+1} \right) \leqq \liminf_i P\left(\alpha(k, n_i^1) \geqq \frac{1}{x} \right) \tag{11}$$

and

$$\limsup_i P\left(\alpha(k, n_i^2) \geqq \frac{1}{x} \right) \leqq \lim_i \left(1 - \left(1 - \frac{p_1^{T/x} p_1^{1/x}}{n_i^2} (1+M) \right)^{\left[\frac{n_i^2 - 2/x}{1/x}\right]+1} \right) \tag{12}$$

The limit expressions in the inequalities (11) and (12) are equal to

$$1 - \exp\left(-\frac{1}{2} x p_1^{T/x} \left(1 + \frac{1}{x} \right) \right) \quad \text{resp.} \quad 1 - \exp\left(-p_1^{1/x} x p_1^{T/x} \left(1 + \frac{1}{x} \right) \right).$$

By the monotony of the function $f(y) = \exp\left(-y x p_1^{T/x} \left(1 + \frac{1}{x} \right) \right)$ and our assumption $p_1 < 2^{-x}$ it follows immediately that

$$\limsup_i P\left(\alpha(k, n_i^2) \geqq \frac{1}{x} \right) < \liminf_i P\left(\alpha(k, n_i^1) \geqq \frac{1}{x} \right). \tag{13}$$

So $P\left(\alpha(k, n) \geqq \frac{1}{x} \right)$ cannot converge to a limit value. This proves our assertion. □

Conjecture. *We are fairly sure that the situation described in Example 1 holds generally true under condition* (9), *i.e. there is no need for the additional assumptions* $r=1$ *and* $p<2^{-x}$. *However we can prove the general assertion just for* $\frac{1}{x}=1$ *and* $\frac{1}{x}=2$.

The following theorem presents two cases where $\alpha(k, n)$ converges in weak sense even when $\frac{1}{x}$ is an integer.

Theorem 3. *Let* $\frac{1}{x}$ $(0<x\leqq 1)$ *be an integer, say* $\frac{1}{x}=M$, *and*

$$\lim_n \frac{k(n) \log \frac{1}{p_0}}{\log n} = x.$$

a) *If for some positive function $f(n)$, which tends to infinity,*

$$\liminf_n \left(k(n) - \frac{x \log n}{\log \frac{1}{p_0}} - f(n) \right) \geqq 0$$

holds true, then

$$\lim_n P(\alpha(k, n) = M-1) = 1.$$

b) *Similarly, if* $\limsup_n \left(k(n) - \frac{x \log n}{\log \frac{1}{p_0}} + f(n) \right) \leqq 0$, *then*

$$\lim_n P(\alpha(k, n) = M) = 1.$$

Proof. a) (5) and (6) induce

$$P(\alpha(k, n) = M) \leqq P(\alpha(k, n) \geqq M) \leqq n r^M p_0^{M k(n)} \leqq r^M p_0^{M f(n)}.$$

Therefore a) follows immediately.

b) (4) and (6) imply

$$P(\alpha(k, n)) \leqq (1 - P_{k(n)}^M)^{-1} \exp\left(-\frac{n p_0 k(n) M}{M} \right) \leqq 2 \exp(-p_0^{-f(n)}).$$

This proves b).

Theorem 4. *If* $\lim_n \frac{k(n)}{\log n} = 0$, *then*

$$P\left(\lim_n \frac{\alpha(k, n)\, k(n) \log \frac{1}{p_0}}{\log n} = 1 \right) = 1.$$

Proof. For every $\varepsilon > 0$ we get by (4) and (6)

$$P\left(\alpha(k, n) < \frac{(1-\varepsilon) \log n}{k(n) \log \frac{1}{p_0}} \right)$$

$$\leqq P\left(\alpha(k, n) < \left[\frac{(1-\varepsilon) \log n}{k(n) \log \frac{1}{p_0}} \right] + 1 \right)$$

$$\leqq \left(1 - P_{k(n)}^{\left[\frac{(1-\varepsilon) \log n}{k(n) \log \frac{1}{p_0}} \right]} \right)^{-1} \exp\left(- \frac{n p_0^{\frac{(1-\varepsilon) \log n}{k(n) \log \frac{1}{p_0}} + 1}}{\frac{(1-\varepsilon) \log n}{k(n) \log \frac{1}{p_0}} + 1} \right)$$

$$\leqq 2 \exp\left(- \frac{n^\varepsilon p_0 \log \frac{1}{p_0}}{2 \log n} \right).$$

The proof is completed by an application of the Borel-Cantelli-lemma, as in the proof of Theorem 2. □

3. Convergence of the Expected Value of $\alpha(k, n)$

In this chapter we will show that the asymptotic behaviour of $E\alpha(k, n)$ is exactly the same as that of $\alpha(k, n)$. First we will prove a lemma. In this chapter $E(\cdot)$ denotes the expectation.

Lemma 3. For every integer M ($M \leqq n$) and $k(n)$ large enough, we get

$$\sum_{i=M}^{n} i P(\alpha(k, n)=i) \leqq 2 n r^M p_0^{Mk(n)}.$$

Proof. By applying the inequalities (5) and (6) one gets

$$\sum_{i=M}^{n} i P(\alpha(k, n)=i) = \sum_{i=M}^{n} P(\alpha(k, n) \geqq i) \leqq \sum_{i=M}^{n} n P_{k(n)}^{i}$$
$$\leqq n P_{k(n)}^{M} (1-P_{k(n)})^{-1} \leqq 2 n r^M p_0^{Mk(n)}. \quad \square$$

The next theorem is an analogon to Theorem 1.

Theorem 5. *We have*

$$\limsup_n E \frac{\alpha(k, n)\, k(n) \log \dfrac{1}{p_0}}{\log n} \leqq 1.$$

Proof. For every $\varepsilon > 0$ the following holds true:

$$E \frac{\alpha(k, n)\, k(n) \log \dfrac{1}{p_0}}{\log n}$$
$$\leqq (1+\varepsilon) P\left(\alpha(k, n) \leqq \frac{(1+\varepsilon)\log n}{k(n)\log \dfrac{1}{p_0}}\right) + \sum_{i=\left[\frac{(1+\varepsilon)\log n}{k(n)\log \frac{1}{p_0}}\right]+1}^{n} i P(\alpha(k, n)=i).$$

Therefore (3.1) induces

$$E \frac{\alpha(k, n)\, k(n) \log \dfrac{1}{p_0}}{\log n} \leqq 1+\varepsilon+2 n r^{\left[\frac{(1+\varepsilon)\log n}{k(n)\log \frac{1}{p_0}}\right]+1} n^{-1-\varepsilon} \leqq 1+\varepsilon+2 n^{-\varepsilon/2}$$

if $k(n)$ is large enough. □

The following corollary is an immediate consequence of Theorem 5.

Corollary 3. *If* $\liminf_n \frac{k(n)\log\frac{1}{p_0}}{\log n} > x$ $(x>0)$, *then*

$$\limsup_n E\,\alpha(k,n) < \frac{1}{x}.$$

For $x=1$ *we get a slightly stronger result, namely:*

Theorem 6. *If* $\liminf_n \frac{k(n)\log\frac{1}{p_0}}{\log n} > 1$, *then we have*

$$\lim E\,\alpha(k,n)\,k(n) = 0.$$

Proof. By means of Lemma 3, inequalities (5) and (6), we get for every large enough integer n:

$$\begin{aligned} E\,k(n)\,\alpha(k,n) &= k(n)\,P(\alpha(k,n)=1) + \sum_{i=2}^{n} i\,k(n)\,P(\alpha(k,n)=i) \\ &\leqq r\,n\,k(n)\,p_0^{k(n)} + 2\,n\,r^2\,p_0^{2k(n)}. \end{aligned} \tag{14}$$

Because of our assumption there exists an $\varepsilon>0$, such that $(1+\varepsilon)\frac{\log n}{\log\frac{1}{p_0}} < k(n)$ for all large enough n. Now we can distinguish two cases:

i) $(1+\varepsilon)\frac{\log n}{\log\frac{1}{p_0}} < k(n) \leqq \frac{3\log n}{\log\frac{1}{p_0}}$

ii) $\frac{3\log n}{\log\frac{1}{p_0}} < k(n) \leqq n.$

In case i) (14) implies

$$E\,k(n)\,\alpha(k,n) \leqq \frac{3\,r\,n\log n}{n^{1+\varepsilon}\log\frac{1}{p_0}} + \frac{2r^2 n}{n^2}.$$

In case ii) we get again from (14)

$$E\,k(n)\,\alpha(k,n) \leqq \frac{r\,n^2}{n^3} + \frac{2r^2 n}{n^6}.$$

Thus in both cases $E\,K(n)\,\alpha(k,n)$ converges to zero.

Now we turn to a lower bound for the limit value of $E\,\alpha(k,n)$

Theorem 7. *If* $\limsup_n \frac{k(n)\log\frac{1}{p_0}}{\log n} < x$ $(x>0)$, *then*

$$\liminf_n E\,\alpha(k,n) \geqq \left[\frac{1}{x}\right].$$

Proof. Because $\left[\frac{1}{x}\right] P\left(\alpha(k,n) \geqq \left[\frac{1}{x}\right]\right) \leqq E\,\alpha(k,n)$ for every n, Theorem 2 yields

$$\left[\frac{1}{x}\right] \leqq \liminf_n E\,\alpha(k,n). \quad \square$$

In a similar manner the following theorem can be proved by an application of Theorem 4:

Theorem 8. *If $k(n)/\log n$ converges to zero, then* $E\dfrac{\alpha(k,n)\,k(n)\log\frac{1}{p_0}}{\log n}$ *converges to one.*

Proof. Because of Theorem 5 it suffices to show that

$$1 \leqq \lim_n E\,\frac{\alpha(k,n)\,k(n)\log\frac{1}{p_0}}{\log n}.$$

But this follows immediately from Theorem 4 and the fact, that

$$P\left(\alpha(k,n) \geqq \frac{\log n}{k(n)\log\frac{1}{p_0}}\right) \leqq E\,\frac{\alpha(k,n)\,k(n)\log\frac{1}{p_0}}{\log n}. \quad \square$$

Remark. The following statements are immediate consequences of Corollary 3 and Theorem 7.

a) If $\dfrac{k(n)\log\frac{1}{p_0}}{\log n}$ converges to some positive x, whose reciprocal $\frac{1}{x}$ is not an integer, then $E\,\alpha(k,n)$ converges to $\left[\frac{1}{x}\right]$.

b) If $\lim_n \dfrac{k(n)\log\frac{1}{p_0}}{\log n} = x$, where x is the reciprocal of an integer, then

$$\frac{1}{x} - 1 \leqq \liminf_n E\,\alpha(k,n) \leqq \limsup_n E\,\alpha(k,n) \leqq \frac{1}{x}.$$

The following theorem presents two cases, where $E\,\alpha(k,n)$ converges even when $\frac{1}{x}$ is an integer (see Theorem 3).

Theorem 9. *Let $\frac{1}{x}$ $(0 < x \leqq 1)$ be an integer, say $\frac{1}{x} = M$, and*

$$\lim_n \frac{k(n)\log\frac{1}{p_0}}{\log n} = x.$$

i) *If for some positive function $f(n)$, which tends to infinity, $\liminf_n \left(k(n) - \frac{x \log n}{\log \frac{1}{p_0}} - f(n)\right) \geqq 0$ holds true, then*

$$\lim_n E\alpha(k, n) = M - 1.$$

ii) *Similarly, if* $\limsup_n \left(k(n) - \frac{x \log n}{\log \frac{1}{p_0}} + f(n)\right) \leqq 0$, *then*

$$\lim E\alpha(k, n) = M.$$

Proof. i) From Theorem 7 it follows

$$M-1 \leqq E\alpha(k, n) \leqq (M-1) P(\alpha(k, n) \leqq M-1) + \sum_{i=M}^{n} i P(\alpha(k, n) = i).$$

With this and Lemma 3 we get for large enough n

$$E\alpha(k, n) \leqq M - 1 + 2nr^M p_0^{Mk(n)} \leqq M - 1 + 2r^M p_0^{Mf(n)}.$$

This proves the first part.

ii) The assertion of part ii) follows from Theorem 3/b, Corollary 3 and from the fact that

$$M P(\alpha(k, n) = M) \leqq E\,\alpha(k, n) \text{ for every integer } n. \quad \square$$

Remark. Under the assumptions of Example 1, $E\alpha(k, n)$ is not converging.

Proof. Let us define $M = \frac{1}{x}$, as above. Then we have

$$\begin{aligned}(M-1) &P(\alpha(k, n) = M-1) + MP(\alpha(k, n) = M) \\ &\leqq E\alpha(k, n) \leqq (M-1) P(\alpha(k, n) \leqq M-1) + MP(\alpha(k, n) = M) + \frac{2nr^{M+1}}{n^{1+x} p_1^{(M+1)}}.\end{aligned}$$

Therefrom follows

$$M - 1 + \liminf_n P(\alpha(k, n) = M) \leqq \liminf_n E\alpha(k, n)$$

and

$$\limsup_n E\alpha(k, n) \leqq M - 1 + \limsup_n P(\alpha(k, n) = M). \quad \square$$

Now we know that inequality (13) is fullfilled by the two subsequences $(k(n_i^1))$ and $(k(n_i^2))$ of Example 1. This, together with the above inequalities, proves the remark.

4. The Area of the Largest Rectangle of Identical Labels

We will assume throughout this chapter, that the random variables $\xi_{i,j}$ can take on just two labels 0 and 1 with probabilities $1-p$ and p, i.e.

$$P(\xi_{i,j}=0)=1-p, \qquad P(\xi_{i,j}=1)=p.$$

Let us denote

$$\alpha(k, n, i)=\max\left\{N: \max_{1\leq j\leq n-N+1} \sum_{u=i}^{i+k-1} \sum_{v=j}^{j+N-1} \xi_{uv}=Nk\right\},$$

$$\gamma(n)=\max_{1\leq k\leq n} \max_{1\leq i\leq n-k+1} k\alpha(k, n, i)$$

and

$$\kappa(n)=\max\left\{N^2: \max_{1\leq u\leq n-N+1} \max_{1\leq v\leq n-N+1} \sum_{i=u}^{u+N-1} \sum_{j=v}^{v+N-1} \xi_{i,j}=\mathrm{N}^2\right\}.$$

Thus $\kappa(n)$ can be interpreted as the "area" of the largest square, consisting just of 1's, in the $n\times n$ matrix $(\xi_{i,j})_{j=1,\ldots,n}^{i=1,\ldots,n}$, whereas $\gamma(n)$ is the "area" of the largest rectangle of 1's.

Révész proved the following theorem (see e.g. [3]):

Theorem R. *We have*

$$P\left(\lim_n \frac{\kappa(n)\log\frac{1}{p}}{2\log n}=1\right)=1.$$

Our aim is to show that the same holds true for $\gamma(n)$.

Theorem 10. *We have*

$$P\left(\lim \frac{\gamma(n)\log\frac{1}{p}}{2\log n}=1\right)=1.$$

Proof. Because of Theorem R it suffices to prove

$$P\left(\limsup_n \frac{\gamma(n)\log\frac{1}{p}}{2\log n}\leq 1\right)=1.$$

Let us introduce the following notations, where $\varepsilon>0$ is some arbitrary constant:

$$B_n=\left\{\gamma(n)>\frac{2(1+\varepsilon)\log n}{\log\frac{1}{p}}\right\},$$

$$B_{n,k,i}=\left\{\alpha(k, n, i)>\frac{2(1+\varepsilon)k\log n}{\log\frac{1}{p}}\right\}.$$

We have

$$B_n = \bigcup_{1 \leq k \leq \frac{4 \log n}{\log \frac{1}{p}}} \; \bigcup_{1 \leq i \leq n-k+1} B_{n,k,i} \cup \bigcup_{\frac{4 \log n}{\log \frac{1}{p}} < k \leq n} \; \bigcup_{1 \leq i \leq n-k+1} B_{n,k,i}. \tag{15}$$

Following the proof of Lemma 2, step by step, it can be seen that for all integers k

$$P(B_{n,k,i}) \leq n^{-1-\varepsilon}$$

and

$$P(B_{n,k,i}) \leq n^{-3}, \quad \text{if } k > \frac{4 \log n}{\log \frac{1}{p}}.$$

This together with (15) yields

$$P(B_n) \leq \frac{4 \log n}{n^{\varepsilon} \log \frac{1}{p}} + \frac{1}{n}.$$

From this inequality we can deduce by means of the Borel-Cantelli-lemma, that

$$P\left(\limsup_m \frac{\gamma(2^m) \log \frac{1}{p}}{2 \log 2^m} \leq 1 \right) = 1. \tag{16}$$

If n is an integer with $2^{m-1} < n \leq 2^m$, then

$$\frac{\gamma(n) \log \frac{1}{p}}{2 \log n} \leq \frac{\gamma(2^m) \log \frac{1}{p}}{2 \log 2^{m-1}} = \frac{\gamma(2^m) \log \frac{1}{p}}{2 \log 2^m - 2 \log 2}. \tag{17}$$

From (16) and (17) follows the assertion of our theorem immediately. □

Not only the largest square of 1's and the largest rectangle are of the same order of magnitude, but also the longest run of 1's in one of n lines, each with n places.

Theorem 11. *Let* $\beta(n) = \max_{1 \leq i \leq n} \alpha(1, n, i)$, *then*

$$P\left(\lim_n \frac{\beta(n) \log \frac{1}{p}}{2 \log n} = 1 \right) = 1.$$

Proof. Because of Theorem 10 we have just to prove, that

$$P\left(\liminf_n \frac{\beta(n) \log \frac{1}{p}}{2 \log n} \geq 1 \right) = 1.$$

Let us introduce the following notations ($\varepsilon > 0$ some arbitrary, but fixed constant):

$$D_n = \left\{ \beta(n) < \frac{2(1-\varepsilon)\log n}{\log \frac{1}{p}} \right\}$$

and

$$D_{n,i} = \left\{ \alpha(1, n, i) < \frac{2(1-\varepsilon)\log n}{\log \frac{1}{p}} \right\}.$$

From (4) it follows:

$$P(D_n) = P\Big(\bigcap_{1 \leqq i \leqq n} D_{n,i} \Big) \leqq \left(1 - \frac{1}{n^{2-2\varepsilon}}\right)^{-n} \left(\exp\left(- \frac{n \cdot n^{-2+2\varepsilon}}{\left[\frac{2(1+\varepsilon)\log n}{\log \frac{1}{p}} \right] + 1} \right) \right)^n$$

$$\leqq \exp\left(- \frac{1}{n^{1-2\varepsilon}} - \frac{n^2 \varepsilon \log \frac{1}{p}}{2(1+\varepsilon)\log n} \right).$$

Therefore

$$\sum_{n=1}^{\infty} P(Dn) < \infty.$$

An application of the Borel-Cantelli-lemma completes the proof. □

References

1. Erdös, P., Révész, P.: On the length of the longest head-run. In: "Topics in Information Theory", ed.: Csíszár, Colloquia Mathematica Societatis Janos Bolyai, Budapest, **1975**, pp. 219–228
2. Petrov, V.V.: On the probabilities of large deviations for sums of independent random variables. Teor. Verojatnost i Primenen **10** (1965) pp. 310–322, Theor. Probab. Appl. **10**, 287–298 (1965)
3. Révész, R.: A Characterization of the Asymptotic Properties of a Stochastics Process by four Classes of Deterministic Curves. [To be published]
4. Shannon, C.: Communication theory of secrecy systems. Bell System Techn. Journal **28**, 656–715 (1949)

Received February 10, 1981