Z. Wahrscheinlichkeitstheorie verw. Gebiete
70, 91–115 (1985)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On the Erdös-Rényi Theorem for Random Fields and Sequences and Its Relationships with the Theory of Runs and Spacings

Paul Deheuvels

7, Avenue du Château, F-92340 Bourg-la-Reine, France

Summary. We prove in this paper a law of Erdös-Rényi type for arrays of independent and identically distributed random variables. The relationships of our theorem with similar results obtained in the theory of runs and spacings are investigated. Applications include the evaluation of the rate of convergence of Erdös-Rényi maxima in limiting cases and a generalization of a Theorem of Erdös and Révész on runs.

1. Introduction

Let $X_{i_1,\dots,i_d}$, $i_1 \geqq 1, \dots, i_d \geqq 1$, be a d-indexed array of i.i.d. random variables. Let, for $n_1 \geqq 0, \dots, n_d \geqq 0$, $0 \leqq K \leqq N$,

$$S(n_1, \dots, n_d, K) = \sum_{j_1=n_1+1}^{n_1+K} \dots \sum_{j_d=n_d+1}^{n_d+K} X_{j_1,\dots,j_d},$$

and

$$I(N, K) = \max_{0 \leqq n_1, \dots, n_d \leqq N-K} S(n_1, \dots, n_d, K).$$

The main purpose of this paper is to establish the following Theorem, which appears as a direct extension of the Erdös-Rényi-Shepp law in the case of d-dimensional arrays.

Theorem 0. *Suppose that* $X_i = X_{i_1,\dots,i_d}$ *is such that*

(i) $E(X_i) = 0 < E(X_i^2) < \infty$.

(ii) $t_0 = \sup\{t;\ \phi(t) = E(e^{tX_i}) < \infty\} > 0$.

Then, for each $\alpha \in \left\{\frac{\phi'(t)}{\phi(t)},\ 0 < t < t_0\right\}$, *if* $c = c(\alpha)$ *is related to* α *via the equation*

$$\rho = \rho(\alpha) = \exp\left(-\frac{1}{c}\right) = \inf_t \phi(t)\, e^{-t\alpha},$$

for any $d \geqq 2$, *we have*

$$\limsup_{N \to \infty} \frac{|I(N, [\{cd \operatorname{Log} N\}^{1/d}]) - \alpha [\{cd \operatorname{Log} N\}^{1/d}]^d|}{\{cd \operatorname{Log} N\}^{(d-1)/d}} \leqq c^{-1} d$$

almost surely. Here, $\operatorname{Log} u = \log_e u$ *is the Neperian logarithm of* u, *and* $[u]$ *stands for the integer part of* u.

The second purpose of the following discussion is to point out the close relationship which exists between Erdös-Rényi type results and the asymptotic theory of longest runs in Bernoulli sequences, maximal uniform spacings, and largest gaps in Poisson processes.

As we shall see, all these probabilistic models yield very similar results which can be described in a unified way by the use of strong approximation techniques. Aside from the resulting simplification of the description of apparently different problems, this has the advantage of giving a geometrical interpretation of the Erdös-Rényi theorem in higher dimensional spaces.

In itself, the proof of Theorem 0 makes use of methods based on the techniques of Csörgő and Steinebach (1981) and of Deheuvels and Devroye (1983). This explains why we have postponed it to the end (§ 8), prefering to discuss from the start problems related to runs and spacings (§ 2–7).

Related results on the limiting behavior of $I(N, K)$ have been given in Steinebach (1983). However, Steinebach's statements have to be formulated more precisely with $[K_N]$ instead of K_N, where $\{K_N, N \geqq 1\}$ is a noninteger sequence satisfying conditions (I1, I2, I3) on page 62 of his paper. Otherwise, there is no possible ultimately non constant integer sequence $\{K_N, N \geqq 1\}$ fulfilling the forementioned assumptions.

In the sequel, $\operatorname{Log} u = \log_e u$ will denote the Neperian logarithm of u, while $\operatorname{Log}_p u = \operatorname{Log}(\operatorname{Log}_{p-1} u)$, with $\operatorname{Log}_1 u = \operatorname{Log} u$, stands for the p times iterated logarithm of u.

2. The Classical Erdös-Rényi Theorem

In 1970, Erdös and Rényi proved the following theorem (see also Shepp (1964)).

Theorem A (Erdös-Rényi). *Let* $X_1, X_2, \ldots$ *be an i.i.d. sequence of random variables with partial sums* $S_0 = 0$, $S_n = X_1 + \ldots + X_n$. *Suppose that*

(i) $E(X_1) = 0 < E(X_1^2) = 1$.

(ii) $t_0 = \sup\{t;\ \phi(t) = E(e^{tX_1}) < \infty\} > 0$.

For $0 \leqq K \leqq N$, *put* $I(N, K) = \max\limits_{0 \leqq n \leqq N-K} \{S_{n+K} - S_n\}$.

Then, for each $\alpha \in \left\{\frac{\phi'(t)}{\phi(t)},\ 0 < t < t_0\right\}$, *if* $c = c(\alpha)$ *is related to* α *via the equation*

$$\rho = \rho(\alpha) = \exp\left(-\frac{1}{c}\right) = \inf_t \phi(t)\, e^{-t\alpha}$$

we have

$$\lim_{N\to\infty} \frac{I(N,[c\operatorname{Log}N])}{[c\operatorname{Log}N]}=\alpha \quad a.s.$$

In 1981, Csörgő and Steinebach precised Theorem A by proving:

Theorem B (Csörgő-Steinebach). *Under the assumptions of Theorem A, we have*

$$\lim_{N\to\infty}\left\{\frac{I(N,[c\operatorname{Log}N])}{[c\operatorname{Log}N]^{1/2}}-\alpha[c\operatorname{Log}N]^{1/2}\right\}=0 \quad a.s.$$

The best possible rate of convergence of the Erdös-Rényi Theorem was obtained by Deheuvels and Devroye (1983) (see Deheuvels, Devroye and Lynch (1985)) as follows:

Theorem C (Deheuvels-Devroye). *Under the assumptions of Theorem A, we have*

$$\limsup_{N\to\infty}\frac{I(N,[c\operatorname{Log}N])-\alpha[c\operatorname{Log}N]}{\operatorname{LogLog}N}=\frac{1}{2t^*} \quad \text{a.s.},$$

$$\liminf_{N\to\infty}\frac{I(N,[c\operatorname{Log}N])-\alpha[c\operatorname{Log}N]}{\operatorname{LogLog}N}=\frac{-1}{2t^*} \quad \text{a.s.},$$

where $t^*=t^*(\alpha)$ is the unique solution of the equation $\dfrac{\phi'(t)}{\phi(t)}=\alpha$.

Let us now consider the particular case where $X_1, X_2, \dots$ is an i.i.d. sequence of Bernoulli $B(p)$ random variables such that

$$P(X_1=1)=p, \qquad P(X_1=0)=1-p.$$

We get then easily from Theorem C (see Deheuvels and Devroye, 1983, Theorem 9):

Corollary 1. *Let $X_1, X_2, \dots$ be an i.i.d. sequence of Bernoulli $B(p)$ random variables with $0<p<1$. Then, for any $\alpha\in(p,1)$, or equivalently for any $c>\dfrac{1}{-\operatorname{Log}p}$, if*

$$c=c(\alpha)=\left\{\alpha\operatorname{Log}\frac{\alpha}{p}+(1-\alpha)\operatorname{Log}\frac{1-\alpha}{1-p}\right\}^{-1},$$

we have

$$\limsup_{N\to\infty}\frac{I(N,[c\operatorname{Log}N])-\alpha[c\operatorname{Log}N]}{\operatorname{LogLog}N}=\beta(\alpha,p),$$

$$\liminf_{N\to\infty}\frac{I(N,[c\operatorname{Log}N])-\alpha[c\operatorname{Log}N]}{\operatorname{LogLog}N}=-\beta(\alpha,p),$$

where

$$\beta(\alpha,p)=\left\{2\operatorname{Log}\frac{\alpha(1-p)}{p(1-\alpha)}\right\}^{-1}.$$

In the limiting cases $\alpha=p$ and $\alpha=1$, for which Theorem C is not valid, one gets respectively $c(p)=\infty$ (corresponding to the law of large numbers), and $c(1)$

$=\lim_{\alpha\uparrow 1} c(\alpha)=\frac{1}{-\operatorname{Log} p}$. It seems reasonable therefore to expect that the results of Corollary 1 extend themselves to the case $\alpha=1$ and $c=\frac{1}{-\operatorname{Log} p}$.

For this, it can be noted that, for $0\leqq L\leqq K\leqq N$, we have

$$0\leqq I(N,L)\leqq I(N,K)\leqq K.$$

If we take $p<\alpha=1-\varepsilon<1$ in Corollary 1 and let $\varepsilon\downarrow 0$, we get easily, since then $c(1-\varepsilon)\to\frac{1}{-\operatorname{Log} p}$:

$$\lim_{N\to\infty}\frac{I\left(N,\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right]\right)}{\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right]}=1 \quad \text{a.s.}$$

Unfortunately, this does not give any information concerning the speed of convergence of the preceding expression to its limit. We shall prove here namely that:

Proposition 1. *Under the assumptions of Corollary* 1, *we have, almost surely*

$$\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right]-1\leqq I\left(N,\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right]\right)\leqq\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right] \quad \textit{as } N\to\infty.$$

This result can be deduced from Theorem 1* in Erdös-Révész (1975). Its main interest comes from the fact that its proof is based on the asymptotic theory of runs which is discussed in the next paragraph. An interesting extension of Proposition 1 completes Theorem C for the rate of convergence of Erdös-Rényi maxima. It is given as follows:

Theorem 1. *Let $X_1, X_2, \ldots$ be an i.i.d. sequence of random variables. Assume that the hypotheses of Theorem A are satisfied and further that*

$$A=\operatorname{ess\,sup}(X_1)<\infty, \quad \textit{and} \quad P(X_1=A)>0.$$

Then, for $c=\frac{1}{-\operatorname{Log} P(X_1=A)}$, we have almost surely as $N\to\infty$:

$$[c\operatorname{Log} N]-0(1)\leqq A^{-1}I(N,[c\operatorname{Log} N])\leqq[c\operatorname{Log} N],$$

the upper bound in the above inequality being reached infinitely often with probability one.

For $c<\frac{1}{-\operatorname{Log} P(X_1=A)}$, $c>0$, we have almost surely as $N\to\infty$:

$$A^{-1}I(N,[c\operatorname{Log} N])=[c\operatorname{Log} N].$$

3. A Generalization and a New Proof of a Theorem of Erdös and Révész

Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernoulli $B(p)$ random variables such that

$$P(X_1=1)=p, \quad P(X_1=0)=1-p, \quad 0<p<1.$$

Let $S_0=0$, $S_n=X_1+\ldots+X_n$, and put, for $0\leqq K\leqq N$,

$$I(N,K)=\max_{0\leqq n\leqq N-K}\{S_{n+K}-S_n\}.$$

Let Z_n be the largest integer for which $I(N,Z_N)=Z_N$. In other words, Z_N is the length of the *longest run* of 1's in $X_1,\ldots,X_N$.

In 1975, Erdös and Révész gave the following characterization of the upper and lower almost sure classes of Z_N as $N\to\infty$:

Theorem D (Erdös-Révész). *Let* $p=\frac{1}{2}$. *Then, for any* $\varepsilon>0$, *we have*

$$P\left(Z_N\leqq\left[\frac{\operatorname{Log}N-\operatorname{Log}_3 N}{\operatorname{Log}2}-\varepsilon\right]-2 \text{ i.o.}\right)=0,$$

and

$$P\left(Z_N\leqq\left[\frac{\operatorname{Log}N-\operatorname{Log}_3 N}{\operatorname{Log}2}+\varepsilon\right]-1 \text{ i.o.}\right)=1.$$

Furthermore, if $\{\alpha(N), N\geqq 1\}$ *is a sequence of positive numbers, then*

$$P(Z_N\geqq\alpha(N) \text{ i.o.})=0 \text{ or } 1, \quad \text{according as } \sum_{N=1}^{\infty}2^{-\alpha(N)}<\infty \text{ or } =\infty.$$

Noting that the methods of Erdös and Révész can be used in a similar manner to obtain the almost sure upper and lower classes of Z_N for an arbitrary p, we shall give in the sequel a new proof of their results, which shall be extended to the *k-th longest run*.

A *run* of 1's in $X_1,\ldots,X_N$ will be defined as any subsequence $X_r,\ldots,X_s$ such that $1\leqq r\leqq s\leqq N$, and

- If $r>1$, then $X_{r-1}=0$;
- If $s<N$, then $X_{s+1}=0$;
- For any t: $r\leqq t\leqq s$, $X_t=1$.

The *length* of the run $X_r,\ldots,X_s$ will be defined as $s-r+1$.

For any $N\geqq 1$, let $R(N)$ denote the number of runs of 1's in $X_1,\ldots,X_N$. It can be verified that $0\leqq R(N)\leqq\frac{N+1}{2}$, and that $R(N)\to\infty$ a.s. as $N\to\infty$. Hence, for any fixed $k\geqq 1$, there exists almost surely an N_k such that, for any $N\geqq N_k$, we have $R(N)\geqq k$.

We shall denote by $Z_N^{(k)}$ the *k-th longest run* of 1's in $X_1,\ldots,X_N$, defined as the k-th largest length of the successive runs of 1's in $X_1,\ldots,X_N$ if $R(N)\geqq k$, and by 0 if $R(N)<k$.

Evidently $Z_N=Z_N^{(1)}$. We shall prove now the following result.

Theorem 2. *Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernoulli $B(p)$ random variables with $0<p<1$. Let $k \geqq 1$ be fixed, and let $Z_N^{(k)}$ be the k-th longest run of* 1's *observed in $X_1, \ldots, X_N$. Then, for any $r \geqq 3$, we have*

$$P\left(Z_N^{(k)} \geqq \frac{\operatorname{Log} N + \frac{1}{k}(\operatorname{Log}_2 N + \ldots + \operatorname{Log}_{r-1} N + (1+\varepsilon)\operatorname{Log}_r N)}{-\operatorname{Log} p} \text{ i.o.}\right) = \begin{vmatrix} 0 & \text{if } \varepsilon>0, \\ 1 & \text{if } \varepsilon \leqq 0. \end{vmatrix}$$

Furthermore, for any $\varepsilon>0$,

$$P\left(Z_N^{(k)} \leqq \left[\frac{\operatorname{Log} N - \operatorname{Log}_3 N + \operatorname{Log}(1-p)}{-\operatorname{Log} p} - \varepsilon\right] - 1 \text{ i.o.}\right) = 0,$$

and

$$P\left(Z_N^{(k)} \leqq \left[\frac{\operatorname{Log} N - \operatorname{Log}_3 N + \operatorname{Log}(1-p)}{-\operatorname{Log} p}\right] \text{ i.o.}\right) = 1.$$

Remark. This last result makes precise the corresponding statement of Theorem D, which is extended to $\varepsilon=0$.

Proof. Let $G_1 = \min\{n \geqq 1;\, X_n = 0\}$, and, for $m=2,3,\ldots$, define recursively G_m by

$$G_m = \min\{n > G_1 + \ldots + G_{m-1};\, X_n = 0\} - (G_1 + \ldots + G_{m-1}).$$

It can be seen that the length of the successive runs of 1's in $X_1, X_2, \ldots$ coincide with the values greater or equal to one in the sequence $\{G_m - 1, m \geqq 1\}$, and that $G_1, G_2, \ldots$ form an i.i.d. sequence with the following geometric distribution:

$$P(G_1 \geqq r) = p^{r-1}, \quad r = 1, 2, \ldots.$$

We shall now make use of a strong approximation argument consisting in considering the geometrically distributed random variables $G_1, G_2, \ldots$ as the integer parts of exponentially distributed random variables. This can be done by an application of Lemma 2 in Deheuvels (1982), giving:

Lemma 1. *Without loss of generality, there exists on the probability space on which $X_1, X_2, \ldots$ is defined an i.i.d. sequence $\{\omega_m, m \geqq 1\}$ of exponentially distributed random variables such that*

$$P(\omega_1 > t) = e^{-t}, \quad t > 0,$$

and

$$G_m = \left[\frac{\omega_m}{-\operatorname{Log} p}\right] + 1, \quad m = 1, 2, \ldots.$$

Let, for $m=1,2,\ldots$, $U_m = e^{-\omega_m}$. The sequence $\{U_m, m \geqq 1\}$ is an i.i.d. sequence of random variables uniformly distributed on (0, 1). If we denote by

$$U_{1,n} < U_{2,n} < \ldots < U_{n-1,n} < U_{n,n}$$

the order statistic of $U_1, \ldots, U_n$, and

$$\omega_n^{(n)} < \omega_n^{(n-1)} < \ldots < \omega_n^{(2)} < \omega_n^{(1)}$$

the order statistic of $\omega_1, \ldots, \omega_n$, then evidently we have:

$$\omega_n^{(k)} = -\operatorname{Log} U_{k,n}, \quad 1 \leqq k \leqq n.$$

It follows that if

$$G_n^{(k)} = \left[\frac{\omega_n^{(k)}}{-\operatorname{Log} p}\right] + 1 = \left[\frac{-\operatorname{Log} U_{k,n}}{-\operatorname{Log} p}\right] + 1,$$

then $G_n^{(k)}$ is the k-th largest value among $G_1, \ldots, G_n$, since

$$G_n^{(n)} \leqq G_n^{(n-1)} \leqq \ldots \leqq G_n^{(2)} \leqq G_n^{(1)}.$$

Next, we use the following lemma.

Lemma 2. *For any fixed* $k \geqq 1$ and $r \geqq 5$,

$$\begin{aligned} &P(n U_{k,n} \geqq \operatorname{Log}_2 n + (k+1) \operatorname{Log}_3 n + \operatorname{Log}_4 n + \ldots + \operatorname{Log}_{r-1} n + (1+\varepsilon) \operatorname{Log}_r n \text{ i.o.}) \\ &\quad = P(n U_{k,n} \leqq \{(\operatorname{Log} n)(\operatorname{Log}_2 n) \ldots (\operatorname{Log}_{r-1} n)(\operatorname{Log}_r n)^{1+\varepsilon}\}^{-1/k} \text{ i.o.}) = 0 \quad \text{or} \quad 1, \end{aligned}$$

according as $\varepsilon > 0$ *or* $\varepsilon \leqq 0$.

This last result is due to Geffroy (1958) for the lower class of $U_{k,n}$ when $k = 1$, Barndorff-Nielsen (1961) for the upper class of $U_{k,n}$ when $k=1$, Kiefer (1973) for the lower class of $U_{k,n}$ when $k \geqq 1$, and Deheuvels (1974) for the upper class of $U_{k,n}$ when $k \geqq 1$.

Note that Kiefer (1973, Theorem 2), proved that for a fixed $k \geqq 1$, we have

$$\limsup_{n \to \infty} \frac{n U_{k,n}}{\operatorname{Log}_2 n} = 1 \quad \text{a.s.}$$

Let us now put, for any $N \geqq 1$,

$$v(N) = \max\left\{m; \sum_{i=1}^{m} G_i \leqq N\right\}, \quad \text{where } \sum_{i=1}^{0} G_i = 0.$$

For any N such that $R(N) \geqq k$, we must then evidently have

$$G_{v(N)}^{(k)} - 1 = \left[\frac{-\operatorname{Log} U_{k,v(N)}}{-\operatorname{Log} p}\right] \leqq Z_N^{(k)} \leqq G_{v(N)+1}^{(k)} - 1 = \left[\frac{-\operatorname{Log} U_{k,v(N)+1}}{-\operatorname{Log} p}\right].$$

Next, by a simple renewal argument, we get, almost surely as $N \to \infty$,

$$\begin{aligned} v(N) &= N(1-p) + O(N \operatorname{Log}_2 N)^{1/2}, \\ \operatorname{Log} v(N) &= \operatorname{Log} N + \operatorname{Log}(1-p) + O(N^{-1/2}(\operatorname{Log}_2 N)^{1/2}), \\ \operatorname{Log}_r v(N) &= \operatorname{Log}_r N + O(1/\operatorname{Log} N), \quad r \geqq 2. \end{aligned}$$

It follows from Lemma 1 and Lemma 2 that

$$P\left(G_n^{(k)} - 1 \leqq \left[\frac{\operatorname{Log} n - \operatorname{Log}_3 n - \operatorname{Log}\left\{1 + \frac{(k+1)\operatorname{Log}_3 n}{\operatorname{Log}_2 n}\right\}}{-\operatorname{Log} p}\right] \text{ i.o.}\right) = 1,$$

and that, for any $\varepsilon>0$,

$$P\left(G_n^{(k)}-1\leqq\left[\frac{\operatorname{Log} n-\operatorname{Log}_3 n-\operatorname{Log}\left\{1+\frac{(k+1+\varepsilon)\operatorname{Log}_3 n}{\operatorname{Log}_2 n}\right\}}{-\operatorname{Log} p}\right]-1 \text{ i.o.}\right)=0.$$

Hence, we have

$$P\left(G_{\nu(N)+1}^{(k)}-1\leqq\left[\frac{\operatorname{Log} n-\operatorname{Log}_3 n+\operatorname{Log}(1-p)-\frac{(k+1)\operatorname{Log}_3 n}{2\operatorname{Log}_2 n}}{-\operatorname{Log} p}\right] \text{ i.o.}\right)=1,$$

and

$$P\left(G_{\nu(N)}^{(k)}-1\leqq\left[\frac{\operatorname{Log} n-\operatorname{Log}_3 n+\operatorname{Log}(1-p)-\frac{2(k+1)\operatorname{Log}_3 n}{\operatorname{Log}_2 n}}{-\operatorname{Log} p}\right]-1 \text{ i.o.}\right)=0.$$

This suffices to prove the second part of Theorem 2.

Likewise, it follows from Lemma 1 and Lemma 2 that, for any $\varepsilon>0$,

$$P\left(G_n^{(k)}-1\geqq\left[\frac{\operatorname{Log} n+\frac{1}{k}(\operatorname{Log}_2 n+\ldots+\operatorname{Log}_{r-1} n+(1+\varepsilon)\operatorname{Log}_r n)}{-\operatorname{Log} p}\right]+1 \text{ i.o.}\right)=0,$$

while

$$P\left(G_n^{(k)}-1\geqq\left[\frac{\operatorname{Log} n+\frac{1}{k}(\operatorname{Log}_2 n+\ldots+\operatorname{Log}_{r-1} n+\operatorname{Log}_r n)}{-\operatorname{Log} p}\right] \text{ i.o.}\right)=1.$$

The first half of Theorem 2 follows from these results, when applied for different values of r.

Remarks. 1°) The upper class given in Theorem 2 is less precise as the one given in Theorem C. Sharper results could here be obtained by using the complete characterization of the lower class of $U_{k,n}$ due to Kiefer (1973).

2°) Let $R(r,N)$ denote the number of runs of 1's in $X_1,\ldots,X_N$ which are greater or equal to r. Since $R\left(r,\sum_{i=1}^{n} G_i\right)=\#\{G_i\geqq r+1, 1\leqq i\leqq n\}$, we get, by Glivenko-Cantelli and the law of large numbers,

$$R\left(m,\left[\frac{n}{1-p}\right]\right)\sim np^r \quad \text{a.s. as } n\to\infty.$$

This can be restated in:

Proposition 2. *For any* $r=1,2,\ldots$, *if* $R(r,N)$ *denotes the number of runs of* 1's *in* $X_1,\ldots,X_N$, *which are greater or equal to* r, *then, almost surely,*

$$\lim_{N\to\infty}\frac{R(r,N)}{N}=p^r(1-p), \qquad r=1,2,\ldots.$$

It is well known (see Fisz (1963) for further references) that $R(1,N) = R(N) \sim Np(1-p)$ as $N \to \infty$. This enables to obtain likewise upper and lower strong bounds for $G_N^{(k_N)}$ when $k_N \geqq 1$ is such that $\limsup_{N\to\infty} N^{-1} k_N < p(1-p)$.

As an example, we give a solution for the following problem, due to P. Révész. What is the number of runs of 1's in $X_1, \ldots, X_N$, larger than $\{\operatorname{Log} N - \operatorname{Log}_3 N - C\}/\{-\operatorname{Log} p\}$?

For this, we need a result due to Kiefer (1973):

Lemma 3 (Kiefer). *Let* $v \in (0, +\infty)$ *be given, and let* $c'_v > 0$, $c''_v > 0$, $\beta'_v > 1$, $\beta''_v < 1$, *be solutions of the equations*

$$\beta'_v(\operatorname{Log}\beta'_v - 1) = (1 - c'_v)/c'_v, \qquad c'_v \beta'_v = v,$$
$$\beta''_v(\operatorname{Log}\beta''_v - 1) = (1 - c''_v)/c''_v, \qquad c''_v \beta''_v = v.$$

Then, if $k = k_n \sim v \operatorname{Log}_2 n$ *as* $n \to \infty$, *we have*

$$c'_v = \liminf \frac{n U_{k,n}}{\operatorname{Log}_2 n} < \limsup \frac{n U_{k,n}}{\operatorname{Log}_2 n} = c''_v \quad \text{a.s.}$$

It follows from Lemma 2 that, almost surely,

$$\operatorname{Log}\left(\frac{1-p}{c''_v}\right) = \liminf_{N\to\infty}(-\operatorname{Log} p)^{-1}\{G^{(k)}_{v(N)} - \operatorname{Log} N + \operatorname{Log}_3 N\}$$
$$< \limsup_{N\infty}(-\operatorname{Log} p)^{-1}\{G^{(k)}_{v(N)} - \operatorname{Log} N + \operatorname{Log}_3 N\} = \operatorname{Log}\left(\frac{1-p}{c'_v}\right).$$

We have just proved:

Theorem 3. *Let* $v \in (0, +\infty)$ *be given, and let* $c'_v > 0$, $c''_v > 0$, $\beta'_v > 1$, $\beta''_v < 1$, *be solutions of the equations*

$$\beta'_v(\operatorname{Log}\beta'_v - 1) = (1 - c'_v)/c'_v, \qquad c'_v \beta'_v = v,$$
$$\beta''(\operatorname{Log}\beta''_v - 1) = (1 - c''_v)/c''_v, \qquad c''_v \beta''_v = v,$$

Then, for any $\varepsilon > 0$, *we have*

$$P\left(Z_N^{([v \operatorname{Log}_2 N])} < \left[\frac{\operatorname{Log} N - \operatorname{Log}_3 N + \operatorname{Log}(1-p) - \operatorname{Log} c'_v}{-\operatorname{Log} p} - 1 - \varepsilon\right] \text{ i.o.}\right) = 0,$$

$$P\left(Z_N^{([v \operatorname{Log}_2 N])} \leqq \left[\frac{\operatorname{Log} N - \operatorname{Log}_3 N + \operatorname{Log}(1-p) - \operatorname{Log} c'_v}{-\operatorname{Log} p} + \varepsilon\right] \text{ i.o.}\right) = 1.$$

$$P\left(Z_N^{([v \operatorname{Log}_2 N])} \geqq \left[\frac{\operatorname{Log} N - \operatorname{Log}_3 N + \operatorname{Log}(1-p) - \operatorname{Log} c''_v}{-\operatorname{Log} p} - 1 - \varepsilon\right] \text{ i.o.}\right) = 1,$$

and

$$P\left(Z_N^{([v \operatorname{Log}_2 N])} > \left[\frac{\operatorname{Log} N - \operatorname{Log}_3 N + \operatorname{Log}(1-p) - \operatorname{Log} c''_v}{-\operatorname{Log} p} + \varepsilon\right] \text{ i.o.}\right) = 0.$$

Remarks. In 1980, Guibas and Odlyzko extended by a different technique the results of Erdös and Révész (1975) to a complete characterization of the lower a.s. class of $Z_N^{(1)}$ as $N \to \infty$.

4. Runs and the Erdös-Rényi Theorem

A direct consequence of Theorem 2 taken with $k=1$ (or equivalently of Theorem D of Erdös and Révész) is that, almost surely as $N\to\infty$,

$$I\left(N,\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right]\right)=\left[\frac{\operatorname{Log} N}{-\operatorname{Log} p}\right]+O(\operatorname{Log}_3 N).$$

We shall make this result more precise by proving Proposition 1 in the slightly stronger version (which could also be deduced from Theorem 1* in Erdös-Révész (1975)):

Theorem 4. *Let $\{d(N), N\geqq 1\}$ be a sequence such that*

$$-1<\liminf_{n\to\infty}\frac{d(n)}{\operatorname{Log}_3 n}\quad \textit{and}\quad \limsup_{n\to\infty}\frac{d(n)}{\operatorname{Log}_2 n}<1.$$

Then, under the assumptions of Corollary 1, we have, almost surely as $N\to\infty$:

$$\left[\frac{\operatorname{Log} N+d(N)}{-\operatorname{Log} p}\right]-1\leqq I\left(N,\left[\frac{\operatorname{Log} N+d(N)}{-\operatorname{Log} p}\right]\right)\leqq\left[\frac{\operatorname{Log} N+d(N)}{-\operatorname{Log} p}\right].$$

Furthermore, in the preceding inequality, the upper and lower bounds are reached infinitely often with probability one.

Proof. With the notations of § 3, let $\theta'_n=\max\left\{\omega_{2i-1}+\omega_{2i};\ i=1,2,\ldots,\left[\frac{n}{2}\right]\right\}$ and $\theta''_n=\max\left\{\omega_{2i}+\omega_{2i+1};\ i=1,2,\ldots,\left[\frac{n-1}{2}\right]\right\}$. It is straightforward that θ'_n (resp. θ''_n) is the partial maximum $\max\{\xi_1,\ldots,\xi_M\}$, with $M=\left[\frac{n}{2}\right]$ $\left(\text{resp. } M=\left[\frac{n-1}{2}\right]\right)$, of an i.i.d. sequence $\{\xi_m, m\geqq 1\}$ of random variables with a Gamma $\Gamma(2)$ distribution:

$$L(t)=P(\xi_m>t)=P(\omega_1+\omega_2>t)=(t+1)e^{-t},\qquad t>0,\ m=1,2,\ldots.$$

Define, for $0<u<1$, $H(u)$ by $L(H(u))=u$. We have evidently

$$H(u)-\operatorname{Log}(1+H(u))=-\operatorname{Log} u.$$

When $u\to 0$, it follows that

$$H(u)=\{-\operatorname{Log} u\}\{1+o(1)\},$$

which in turn implies that, as $u\to 0$,

$$H(u)=-\operatorname{Log} u+\operatorname{Log} H(u)+O\left(\frac{1}{H(u)}\right)=-\operatorname{Log} u+\operatorname{Log}(-\operatorname{Log} u)+o(1).$$

Next, we use the fact that $\{U_m=L(\xi_m), m\geqq 1\}$ is an i.i.d. sequence of random variables, uniformly distributed on (0, 1). It follows that

$$\max\{\xi_1,\ldots,\xi_M\}=H(\min\{U_1,\ldots,U_M\}).$$

By Lemma 2, for any $\varepsilon>0$, there exists almost surely an M_ε such that $M \geqq M_\varepsilon$ implies that

$$\min\{U_1, \dots, U_M\} \leqq \frac{1+\varepsilon}{M} \operatorname{Log}_2 M.$$

It follows that, almost surely as $M \to \infty$, we have

$$\begin{aligned}\max\{\xi_1, \dots, \xi_M\} &\geqq -\operatorname{Log}\left(\frac{1+\varepsilon}{M}\operatorname{Log}_2 M\right) + \operatorname{Log}\left(-\operatorname{Log}\left(\frac{1+\varepsilon}{M}\operatorname{Log}_2 M\right)\right) + o(1)\\ &\geqq \operatorname{Log} M + \operatorname{Log}_2 M - \operatorname{Log}_3 M - \operatorname{Log}(1+\varepsilon) + o(1).\end{aligned}$$

Hence, we obtain $\left(\text{by putting successively } M = \left[\frac{n}{2}\right] \text{ and } M = \left[\frac{n-1}{2}\right]\right)$ that, almost surely as $n \to \infty$, we have

$$\theta_n' \geqq \operatorname{Log} n + \operatorname{Log}_2 n - \operatorname{Log}_3 n + O(1) \quad \text{and} \quad \theta_n'' \geqq \operatorname{Log} n + \operatorname{Log}_2 n - \operatorname{Log}_3 n + O(1).$$

By the notations of § 3, we have, for $m = 1, 2, \dots$,

$$G_m = \left[\frac{\omega_m}{-\operatorname{Log} p}\right] + 1.$$

It follows that, almost surely as $n \to \infty$, we have

$$\max\{G_m + G_{m+1}, 1 \leqq m \leqq n-1\} \geqq \frac{\operatorname{Log} n + \operatorname{Log}_2 n - \operatorname{Log}_3 n + O(1)}{-\operatorname{Log} p},$$

which, in turn, implies that, almost surely as $n \to \infty$,

$$\max\{G_m + G_{m+1}, 1 \leqq m \leqq \nu(n) - 1\} \geqq \frac{\operatorname{Log} n + \operatorname{Log}_2 n - \operatorname{Log}_3 n + O(1)}{-\operatorname{Log} p}.$$

We have also used in the proof of Theorem 2 the fact that, almost surely as $n \to \infty$,

$$\max\{G_m, 1 \leqq m \leqq \nu(n)\} \geqq \frac{\operatorname{Log} n - \operatorname{Log}_3 n + O(1)}{-\operatorname{Log} p}.$$

The meaning of these two results is the following. There exists almost surely a constant C and an integer n_0, such that any sequence $X_1, \dots, X_n$ with $n \geqq n_0$ contains a run of 1's of length larger than

$$\frac{\operatorname{Log} n - \operatorname{Log}_3 n + C}{-\operatorname{Log} p},$$

and two successive runs of 1's (say 1 ... 101 ... 1) whose added lengths are larger than

$$\frac{\operatorname{Log} n + \operatorname{Log}_2 n - \operatorname{Log}_3 n + C}{-\operatorname{Log} p}.$$

If we now consider any moving average $X_{m+1}+\ldots+X_{m+K}$, $0\leqq m\leqq m+K\leqq n$, where

$$K\leqq\frac{\operatorname{Log} n+\operatorname{Log}_2 n-\operatorname{Log}_3 n+C}{-\operatorname{Log} p},$$

then we have always $X_{m+1}+\ldots+X_{m+K}\leqq K$, and there exists a choice of m, $0\leqq m\leqq n-K$, such that $X_{m+1}+\ldots+X_{m+K}\geqq K-1$.

This proves the first half of Theorem 4. The second half will follow from the fact that, if

$$\max\{G_m, 1\leqq m\leqq v(n)\}=\frac{\operatorname{Log} n+\rho_n \operatorname{Log}_2 n}{-\operatorname{Log} p}=\frac{\operatorname{Log} n+\delta_n \operatorname{Log}_3 n}{-\operatorname{Log} p},$$

then, by Lemma 2, we have

$$\liminf_{n\to\infty}\delta_n=-1 \quad \text{and} \quad \limsup_{n\to\infty}\rho_n=1 \qquad \text{a.s.}$$

The proof of Theorem 4 is now complete.

The following result completes Corollary 1 and Proposition 1. It follows directly from the preceding arguments.

Proposition 2. *Under the assumptions of Corollary* 1, *for any* $c\in\left]0,\frac{1}{-\operatorname{Log} p}\right[$, *we have, almost surely as* $N\to\infty$:

$$I(N,[c\operatorname{Log} N])=[c\operatorname{Log} N].$$

It may be remarked that $1/c(\alpha)=\operatorname{Log}\frac{\alpha}{p}+(1-\alpha)\operatorname{Log}\frac{1-\alpha}{1-p}$ decreases from $-\operatorname{Log}(1-p)$ to 0 as α increases from 0 to p and increases from 0 up to $-\operatorname{Log} p$ as α increases from p to 1. It follows evidently that:

Corollary 3. *Under the assumptions of Corollary* 1, *if*

$$I(N,K)=\max_{0\leqq n\leqq N-K}\{S_{n+K}-S_n\}, \quad \text{and} \quad J(N,K)=\min_{0\leqq n\leqq N-K}\{S_{n+K}-S_n\},$$

then, almost surely,

$$\lim_{N\to\infty}\frac{I(N,[c\operatorname{Log} N])}{[c\operatorname{Log} N]}=\begin{cases}1 & \text{for } 0<c\leqq\frac{1}{-\operatorname{Log} p},\\ \alpha & \text{for } c>\frac{1}{-\operatorname{Log} p},\end{cases}$$

$$\lim_{N\to\infty}\frac{J(N,[c\operatorname{Log} N])}{[c\operatorname{Log} N]}=\begin{cases}0 & \text{for } 0<c\leqq\frac{1}{-\operatorname{Log}(1-p)},\\ \beta & \text{for } c>\frac{1}{-\operatorname{Log}(1-p)},\end{cases}$$

where α (resp. β) *is defined as the unique solution of* $c=c(\alpha)$ (resp. $c=c(\beta)$) *such that* $p<\alpha<1$ (resp. $0<\beta<p$).

Corollary 3 can be considered as a special case of the extension of the Erdös-Rényi theorem given in Csörgő (1979), p. 785, which covers case (ii) of Deheuvels and Devroye (1983), Theorem 10.

The arguments above show that when $X_1, X_2, \dots$ is a Bernoulli sequence, the asymptotic limiting behavior of $I(N, [c \operatorname{Log} N])$, $N \to \infty$, is closely related to the behavior of the longest run of 1's in $X_1, \dots, X_N$. This can be used to cover the case of Theorem 1 in a straightforward extension.

We shall now discuss some further applications of the methods used in § 2 for the study of runs. We begin with the lemma:

Lemma 4. *Let* $\{G_n, n \geqq 1\}$ *be an i.i.d. sequence of geometrically distributed random variables such that*

$$P(G_1 \geqq r) = p^{r-1}, \quad r = 1, 2, \dots.$$

Put

$$T_0 = 0, \quad T_n = G_1 + \dots + G_n.$$

Let, for $0 \leqq K \leqq N$,

$$H(N, K) = \max_{0 \leqq n \leqq N-K} \{T_{n+K} - T_n\}.$$

Then, for any $a > \dfrac{1}{1-p}$, *we have*

$$\limsup_{N\to\infty} \frac{H(N, [\gamma \operatorname{Log} N]) - a[\gamma \operatorname{Log} N]}{\operatorname{Log}_2 N} = \frac{1}{2}\left\{\operatorname{Log}\left(\frac{a-1}{ap}\right)\right\}^{-1},$$

and

$$\liminf_{N\to\infty} \frac{H(N, [\gamma \operatorname{Log} N]) - a[\gamma \operatorname{Log} N]}{\operatorname{Log}_2 N} = -\frac{1}{2}\left\{\operatorname{Log}\left(\frac{a-1}{ap}\right)\right\}^{-1},$$

where $\gamma = \gamma(a) = \left\{-\operatorname{Log}(1-p) + (a-1)\operatorname{Log}\left(\frac{a-1}{p}\right) - a \operatorname{Log} a\right\}^{-1}$.

Proof. It follows from Theorem C. If we remark that k represents the total number of zeroes in $\{X_j, 1 \leqq j \leqq T_k\}$, we can deduce from Lemma 4 the following result.

Theorem 5. *Let* $X_1, X_2, \dots$ *be an i.i.d. sequence of Bernoulli* $B(p)$ *random variables with* $0 < p < 1$. *Let* $S_0 = 0$ *and* $S_n = X_1 + \dots + X_n$. *For any* $a > \dfrac{1}{1-p}$, *put*

$$\gamma = \gamma(a) = \left\{-\operatorname{Log}(1-p) + (a-1)\operatorname{Log}\left(\frac{a-1}{p}\right) - a \operatorname{Log} a\right\}^{-1},$$

and for $n = 0, 1, \dots$ *and* $N > e^{1/\gamma}$, *define* $K(\gamma, N, n)$ *as the minimum value of* $m \geqq 1$ *such that* $S_{n+m} - S_n = m - [\gamma \operatorname{Log} N]$, *or equivalently, such that there are* $[\gamma \operatorname{Log} N]$ *zeroes in* $\{X_{n+i}, 1 \leqq i \leqq m\}$.

Then we have

$$\limsup_{N\to\infty} \max_{0 \leqq n + K(\gamma, N, n) \leqq N} \left\{\frac{K(\gamma, N, n) - a[\gamma \operatorname{Log} N]}{\operatorname{Log}_2 n}\right\} = \frac{1}{2}\left\{\operatorname{Log}\left(\frac{a-1}{ap}\right)\right\}^{-1},$$

and

$$\liminf_{N\to\infty} \max_{0\leqq n+K(\gamma,N,n)\leqq N} \left\{\frac{K(\gamma,N,n)-a[\gamma \operatorname{Log} N]}{\operatorname{Log}_2 N}\right\} = -\frac{1}{2}\left\{\operatorname{Log}\left(\frac{a-1}{ap}\right)\right\}^{-1}.$$

Remarks. 1°) The result of Theorem 5 cannot be extended to the case $a=1$.

2°) By Theorem 5, it can be seen that if $K=K_N$ is such that, for some $\varepsilon>0$,

$$K \geqq a[\gamma \operatorname{Log} N] + \frac{(1+\varepsilon)\operatorname{Log}_2 N}{2}\left\{\operatorname{Log}\left(\frac{a-1}{ap}\right)\right\}^{-1}$$

then, as $N\to\infty$, there exists almost surely in the sequence $\{X_n, 1\leqq n\leqq N+K\}$ a subsequence $\{X_k, n\leqq k\leqq n+K\}$ with at least $[\gamma \operatorname{Log} N]$ zeros. If we assume further that $K_N \sim a[\gamma \operatorname{Log} N]$, then we have

$$\liminf_{n\to\infty} \max_{0\leqq n\leqq N+K} \frac{S_{n+K}-S_n}{[\gamma \operatorname{Log} N]} \geqq \lim_{N\to\infty} \frac{K-[\gamma \operatorname{Log} N]}{[\gamma \operatorname{Log} N]} = a-1, \quad \text{a.s.}$$

Likewise, if $K=K_N\sim a[\gamma \operatorname{Log} N]$ is such that, for some $\varepsilon>0$,

$$K \leqq a[\gamma \operatorname{Log} N] - \frac{(1+\varepsilon)\operatorname{Log}_2 N}{2}\left\{\operatorname{Log}\left(\frac{a-1}{ap}\right)\right\}^{-1},$$

then, there exists a subsequence $\{X_k, n\leqq k\leqq n+K\}$ as above with at most $[\gamma \operatorname{Log} N]$ zeroes. This result, jointly with the preceding, implies that

$$\lim_{n\to\infty} \max_{0\leqq n\leqq N-a[\gamma \operatorname{Log} N]} \frac{S_{n+a[\gamma \operatorname{Log} N]}-S_n}{a[\gamma \operatorname{Log} N]} = \frac{a-1}{a} \quad \text{a.s.}$$

A comparison with Theorem A and Corollary 3 shows that we must have $a\gamma(a)=c\left(\frac{a-1}{a}\right)$, which can be verified directly.

A direct consequence of the preceding arguments is expressed in the following:

Theorem 6. *Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernoulli $B(p)$ random variables with $0<p<1$. For any $\alpha\in]p,1[$, or equivalently for any $c>\frac{1}{-\operatorname{Log} p}$, if c and α are related by*

$$c=c(\alpha)=\left\{\operatorname{Log}\frac{\alpha}{p}+(1-\alpha)\operatorname{Log}\frac{1-\alpha}{1-p}\right\}^{-1}.$$

Then, for any $\varepsilon>0$, there exists almost surely an N_ε such that, for any $N>N_\varepsilon$, there exists a $K=K(N)$ such that

$$\max_{0\leqq n\leqq N-K} \{S_{n+K}-S_n\} = [\alpha c \operatorname{Log} N],$$

and that

$$|K(N)-[c \operatorname{Log} N]| \leqq \frac{(1+\varepsilon)\operatorname{Log}_2 N}{2\operatorname{Log}\frac{\alpha}{p}}.$$

Proof. It suffices to put in Theorem 5, $c=c(\alpha)=a\gamma(a)$, and $\alpha=\frac{a-1}{a}$, noting that $\gamma(a)=\frac{1}{a}c\left(\frac{a-1}{a}\right)$.

It must be noted that, even though Theorem 6 is closely related to Theorem C in the specific case of Bernoulli random variables, both results are not, strictly speaking, equivalent.

5. Runs, Poisson Processes and Spacings

We shall assume throughout this paragraph that $X_1, X_2, \ldots$ is an i.i.d. sequence of Bernoulli $B(p)$ random variables with $0<p<1$. Consider, on the same probability space (eventually enlarged) an i.i.d. sequence $\xi_1, \xi_2, \ldots$ of random variables independent of $X_1, X_2, \ldots$ and such that

$$P(\xi_1=r)=\frac{\lambda^r e^{-\lambda}}{r!(1-e^{-\lambda})}, \quad r=1,2,\ldots, \text{ where } \lambda=-\operatorname{Log} p.$$

It is then easily seen (see Serfling, 1978) that the sequence $\{Y_n=(1-X_n)\xi_n, n\geqq 1\}$ is an i.i.d. sequence of Poisson $P(\lambda)$ distributed random variables, such that

$$P(Y_1=r)=\frac{\lambda^r}{r!}e^{-\lambda}, \quad r=0,1,\ldots,$$

and such that $Y_n=0$ iff $X_n=1$.

Without loss of generality, it is also possible to define a standard Poisson process $\{N(t), t\geqq 0\}$, such that, for any $n\geqq 1$,

$$N(n\lambda)-N((n-1)\lambda)=Y_n.$$

Let $0<z_1<z_2<\ldots$ be the times of arrivals of $\{N(t), t\geqq 0\}$, and define, for an arbitrary $T>0$, the *largest gap* in $N(t)$, $0\leqq t\leqq T$ as

$$G(T)=T \quad \text{if } N(T)=0,$$

$$G(T)=\operatorname{Max}\{z_1, z_2-z_1, \ldots, z_{N(T)}-z_{N(T)-1}, T-z_{N(T)}\}, \quad \text{if } N(T)\geqq 1.$$

Let also (as in § 3) Z_N stand for the longest run of 1's in the sequence $X_1, \ldots, X_N$. The relationship between Z_N and $G(T)$ follows from the following straightforward Lemma:

Lemma 5. *For any $N\geqq 1$, we have*

$$\frac{G(N\lambda)}{\lambda}-2\leqq Z_N\leqq\frac{G(N\lambda)}{\lambda}.$$

By using exactly the same arguments as in the proof of Theorem 2, it can be seen that:

Lemma 6. *For any $r \geqq 4$ and $\varepsilon > 0$, there exists almost surely a T_0, such that, for any $T \geqq T_0$, we have*

$$\operatorname{Log} T - \operatorname{Log}_3 T - \frac{2 \operatorname{Log}_3 T + \operatorname{Log}_4 T + \ldots + (1+\varepsilon) \operatorname{Log}_r T}{\operatorname{Log}_2 T} \leqq G(T)$$
$$\leqq \operatorname{Log} T + \operatorname{Log}_2 T + \ldots + \operatorname{Log}_{r-1} T + (1+\varepsilon) \operatorname{Log}_2 T,$$

while for $\varepsilon \leqq 0$, for any $T_1 > 0$, there exists almost surely $T_2 > T_1$ (resp. $T_3 > T_2$) *such that the first (resp. the second) inequality is not satisfied.*

This result enables to obtain again upper and lower classes for maximal runs, and, by Lemma 5, a limiting Erdös-Rényi type theorem for Bernoulli sequences. The interest of this approach is that it may be extended without modification in the multivariate case. This will be precised in § 6.

If one considers an i.i.d. sequence $U_1, U_2, \ldots$ of uniformly distributed random variables on (0, 1), whose order statistics will be denoted by

$$0 = U_{0,n} < U_{1,n} < \ldots < U_{n,n} < U_{n+1,n} = 1,$$

the (uniform) *spacings* of order n are defined by $S_i^{(n)} = U_{i,n} - U_{i-1,n}$, $n = 1, 2, \ldots$; the *maximal spacing* will be denoted by

$$\Delta_n = \max_{1 \leqq i \leqq n+1} S_i^{(n)}.$$

The upper and lower almost sure classes of Δ_n as $n \to \infty$ have been described in Devroye, 1981, 1982, and Deheuvels, 1982. There is a close relationship between Δ_n and $G(n)$. In fact, if one takes $\Delta_{\Pi(n)}$, where $\Pi(n)$ is a Poisson $P(n)$ random variable independent of $X_1, X_2, \ldots$, then $\Delta_{\Pi(n)}$ is *identical in distribution* with $G(n)/n$. However, the limiting almost sure behavior of $\Delta_{\Pi(n)}$ and of $G(n)/n$ differ slightly as shown in the following.

Theorem E (Devroye-Deheuvels). *For any $r \geqq 5$,*

$$P(n\Delta_n \geqq \operatorname{Log} n + 2 \operatorname{Log}_2 n + \operatorname{Log}_3 n + \ldots + \operatorname{Log}_{r-1} n + (1+\varepsilon) \operatorname{Log}_r n \text{ i.o.})$$
$$= P(n\Delta_n \geqq \operatorname{Log} n - \operatorname{Log}_3 n - \operatorname{Log} 2 - \varepsilon \text{ i.o.}) = 0 \text{ or } 1,$$

according as $\varepsilon > 0$ or $\varepsilon \leqq 0$.

It is easily verified that the limiting almost sure behavior of $\Delta_{N(n)}$ and of Δ_n as $n \to \infty$ are identical. It follows that the corresponding upper and lower bounds differ from those of $G(n)/n$ after the second (resp. the third) term.

The conclusion which can be made from the discussions of § 2, 3, 4 and 5 is that *the limiting behavior of runs, gaps in a Poisson process, and spacings, are closely related together* and to *the Erdös-Rényi theorem.* This general idea has driven us toward the multivariate extensions which will be described in the next paragraphs.

We shall make use of a general inequality concerning Poisson processes. This inequality which has interest in itself will be described in the next paragraph.

6. An Inequality for the Poisson Process

Theorem 7. *Let $(S, \mathscr{S}, \mu)$ be a measure space, on which is defined a Poisson process $N(.)$ such that, for each $A \in \mathscr{S}$,*

$$P(N(A)=k)=\frac{\mu(A)^k}{k!}e^{-\mu(A)}, \qquad k=0,1,\ldots.$$

Let $A_1, \ldots, A_N$ be arbitrary μ-integrable subsets of $\mathscr{S}$. Then we have

$$P\left(\bigcap_{i=1}^{N}\{N(A_i)\geqq 1\}\right)\geqq \prod_{i=1}^{N} P(N(A_i)\geqq 1).$$

Proof. Let 1_A stand for the indicator function of A. We have

$$\begin{aligned}1_{\bigcap_{i=1}^{N}\{N(A_i)\geqq 1\}}&=1_{\bigcap_{i=1}^{N-1}\{N(A_i)\geqq 1\}}-1_{\bigcap_{i=1}^{N-1}\{N(A_i)\geqq 1\}\cap\{N(A_N)=0\}}\\&=1_{\bigcap_{i=1}^{N-1}\{N(A_i)\geqq 1\}}-1_{\bigcap_{i=1}^{N-1}\{N(A_i-A_N)\geqq 1\}\cap\{N(A_N)=0\}}.\end{aligned}$$

By taking expectations, it follows that

$$P\left(\bigcap_{i=1}^{N}\{N(A_i)\geqq 1\}\right)=P\left(\bigcap_{i=1}^{N-1}\{N(A_i)\geqq 1\}\right)-P\left(\bigcap_{i=1}^{N-1}\{N(A_i-A_N)\geqq 1\}\right)P(N(A_N)=0).$$

Since evidently

$$P\left(\bigcap_{i=1}^{N-1}\{N(A_i-A_N)\geqq 1\}\right)\leqq P\left(\bigcap_{i=1}^{N-1}\{N(A_i)\geqq 1\}\right),$$

it follows that

$$P\left(\bigcap_{i=1}^{N}\{N(A_i)\geqq 1\}\right)\geqq P\left(\bigcap_{i=1}^{N-1} N(A_i)\geqq 1\right)P(N(A_N)\geqq 1).$$

The result follows by induction on N.

Remarks. 1°) The result of Theorem 7 is trivial if the A_i are disjoint.

2°) The result is not true if we replace the Poisson process $N(.)$ by a multinomial point process. For instance, if we throw a point randomly in S, and if $A_1 \cap A_2 = \emptyset$, then $\{N(A_1)\geqq 1\}$ and $\{N(A_2)\geqq 1\}$ are exclusive events, and $P(\{N(A_1)\geqq 1\}\cap\{N(A_2)\geqq 1\})=0$.

The following theorem gives an extension of Theorem 7 in the case where $A_1, \ldots, A_N$ are random.

Theorem 8. *Let S be a metrizable locally compact topological space, $\mathscr{S}$ the algebra of Borel subsets of S, μ a positive Radon measure on S, and $N(.)$ the Poisson process with expectancy $\mu(.)$ on S. Let $N\geqq 1$ be a fixed integer, and* *assume that* $A_1, \ldots, A_N$ *are random elements of* $\mathscr{S}$ *such that, for any* $k = 2, 3, \ldots, n$,

$$P\left(\bigcap_{i=1}^{k-1} \{N(A_i - A_k) \geqq 1\} \cap \{N(A_k) = 0\}\right) \leqq P\left(\bigcap_{i=1}^{k} N(A_i) \geqq 1\right) P(N(A_k) = 0),$$

then, we have

$$P\left(\bigcap_{i=1}^{N} \{N(A_i) \geqq 1\}\right) \geqq \prod_{i=1}^{N} P(N(A_i) \geqq 1).$$

Proof. It follows by the same arguments as in the proof of Theorem 6.

The bound given in Theorem 8 is reached in the following simple example. Let $\{N(t), t > 0\}$ be a standard Poisson process with times of arrivals $0 < z_1 < z_2 < \ldots$. For a given $\delta > 0$, put $A_1 = (0, \delta)$, $A_2 = (z_1, z_1 + \delta), \ldots, A_N = (z_{N-1}, z_{N-1} + \delta)$, and $z_0 = 0$. We get then

$$P\left(\bigcap_{i=1}^{N} \{N(A_i) \geqq 1\}\right) = P\left(\bigcap_{i=1}^{N} \{z_i - z_{i-1} < \delta\}\right) = \prod_{i=1}^{N} (1 - e^{-\delta})$$
$$= \prod_{i=1}^{N} P(N(A_i) \geqq 1).$$

7. The Area of the Largest Head Square

Let $\{X_{ij}, i \geqq 1, j \geqq 1\}$ be a double array of i.i.d. random variables with $P(X_{11} = 1) = p$ and $P(X_{11} = 0) = 1 - p$, $0 < p < 1$.

Let

$$S(n, m, K) = \sum_{j=m+1}^{m+K} \sum_{i=n+1}^{n+K} X_{ij},$$
$$I(N, K) = \operatorname*{Max}_{0 \leqq m, n \leqq N-K} S(m, n, K).$$

Révész (1981) has given the following theorem concerning the asymptotic almost sure behavior of the *largest head square*, defined as the largest integer Z_N such that $I(N, Z_N) = Z_N$.

Theorem F (Révész). *Let* $p = 1/2$, *and put*

$$A(N) = \left\{\frac{2 \operatorname{Log} N}{-\operatorname{Log} p}\right\}^{1/2}, \quad B(N) = \left\{\frac{2 \operatorname{Log} N}{-\operatorname{Log} p}\right\}^{1/2} - \left[\left\{\frac{2 \operatorname{Log} N}{-\operatorname{Log} p}\right\}^{1/2}\right].$$

Then, for any $C > \frac{1}{2}$, *we have, almost surely as* $N \to \infty$,

$$Z_N = \begin{cases} [A(N)] - 1 \quad \text{or} \quad [A(N)] & \text{if } B(N) \leqq \varepsilon_N, \\ [A(N)] & \text{if } \varepsilon_N < B(N) < 1 - \varepsilon_N, \\ [A(N)] \quad \text{or} \quad [A(N)] + 1 & \text{if } B(N) \geqq 1 - \varepsilon_N, \end{cases}$$

where

$$\varepsilon_N = \frac{c \operatorname{Log}_2 N}{\left\{2 (\operatorname{Log} N) \left(\operatorname{Log} \frac{1}{p}\right)\right\}^{1/2}}.$$

It can be seen without great difficulty that Révész's proof enables to handle the general case where $p \in]0, 1[$ is arbitrary. The corresponding result is stated in the Theorem.

The rather complicate form of Theorem F does not enable to get a clear view of the meaning of the constant $A(N)$ and of the definition of the upper and lower class. In fact, a close look to Theorem F enables to give the following equivalent version:

Theorem F* (Révész). *For any* $p \in]0, 1[$ *and* $\varepsilon > 0$, *almost surely as* $N \to \infty$,

$$Y_N^2 \leqq Z_N^2 \leqq \frac{\operatorname{Log} N^2 + (1+\varepsilon) \operatorname{Log}_2 N^2}{-\operatorname{Log} p},$$

where Y_N *is the largest integer such that*

$$Y_N^2 \leqq \frac{\operatorname{Log} N^2 - (1+\varepsilon) \operatorname{Log}_2 N^2}{-\operatorname{Log} p}.$$

From there, the extension in higher dimensions is straightforward.

Let $\{X_{i_1, \ldots, i_d}, i_1 \geqq 1, \ldots, i_d \geqq 1\}$ be a d-indexed array of i.i.d. random variables with a Bernoulli $B(p)$ distribution. Let

$$S(n_1, \ldots, n_d, K) = \sum_{j_1 = n_1 + 1}^{n_1 + K} \cdots \sum_{j_d = n_d + 1}^{n_d + K} X_{j_1, \ldots, j_d},$$

$$I(N, K) = \operatorname*{Max}_{0 \leqq n_1, \ldots, n_d \leqq N - K} S(n_1, \ldots, n_d, K).$$

Let Z_N be defined as the largest integer for which

$$I(N, Z_N) = Z_N^d.$$

It follows, under these assumptions, that:

Theorem 9. *For any* $d \geqq 1$, $p \in]0, 1[$, *and* $\varepsilon > 0$, *we have, almost surely as* $N \to \infty$,

$$Y_N^d \leqq Z_N^d \leqq \frac{\operatorname{Log} N^d + (1+\varepsilon) \operatorname{Log}_2 N^d}{-\operatorname{Log} p},$$

where Y_N *is the largest integer such that*

$$Y_N^d \leqq \frac{\operatorname{Log} N^d - (1+\varepsilon) \operatorname{Log}_2 N^d}{-\operatorname{Log} p}.$$

Proof. As it is, the proof of Theorem 9 can be made by the same statements as given in Révész (1981) for $d = 2$, $p = 1/2$. Hence, details will be omitted.

It is of some interest to know if the similarity between the asymptotic behavior of Z_n (runs), $G(T)$ (gaps in a Poisson process), and Δ_n (spacings) which has been described for $d=1$ in §5 holds again in higher dimensions ($d \geqq 2$). It is indeed the case as we shall now see.

In the first place, consider an i.i.d. sequence $U_1, U_2, \ldots$ of random vectors, uniformly distributed in $(0,1)^d$. Then, the *maximal spacing* Δ_n will be defined as the maximal value of δ such that there exists a hypercube of side δ in $(0,1)^d$ which has no interior point among $U_1, \ldots, U_n$.

It has been proved in Deheuvels (1983), that:

Theorem G (Deheuvels). *For any* $d \geqq 1$, $r \geqq 5$, *and* $\varepsilon > 0$, *almost surely as* $n \to \infty$, *we have*

$$\begin{aligned}\operatorname{Log} n - \operatorname{Log}_3 n - \operatorname{Log} 2 - \varepsilon \leqq n \Delta_n^d \leqq & \operatorname{Log} n + (d+1) \operatorname{Log}_2 n \\ & + \operatorname{Log}_3 n + \ldots + (1+\varepsilon) \operatorname{Log}_r n.\end{aligned}$$

A similar (but weaker) result is given for the Poisson process in:

Theorem 10. *Let* $N(.)$ *be a standard Poisson process in* $(0, +\infty)^d$. *For any* $T > 0$, *let* $G(T)$ *be the maximal value of* δ *such that there exists a hypercube of side* δ *in* $(0, T)^d$ *which has no interior points among the points of the process* $N(.)$.

Then, we have, almost surely as $T \to \infty$,

$$G(T)^d = \operatorname{Log} T^d + O(\operatorname{Log} T)^{(d-1)/d}.$$

Proof. Note that one could obtain in the above evaluation $O(\operatorname{Log}_2 T)$ instead of $O(\operatorname{Log} T)^{(d-1)/d}$, by using similar methods as in Deheuvels, 1983. Since the corresponding proofs are somewhat lengthy, they will be given elsewhere (see the sequel for an upper bound) and we shall limit ourselves to this weaker result.

We shall deduce Theorem 10 from Theorem 9, by using the Poisson approximation technique of §5. For this, define $X_{i_1, \ldots, i_d}$ as

$$X_{i_1, \ldots, i_d} = 1_{\left\{N\left(\prod_{j=1}^{d} (\alpha i_j, \alpha i_j + \alpha)\right) = 0\right\}}, \qquad \text{where } \alpha = \{-\operatorname{Log} p\}^{1/d}.$$

It follows that $\{X_{i_1, \ldots, i_d}, i_1 \geqq 1, \ldots, i_d \geqq 1\}$ is an array of i.i.d. Bernoulli $B(p)$ random variables. Further, we have

$$\frac{G(N\alpha)}{\alpha} - 2 \leqq Y_N \leqq \frac{G(N\alpha)}{\alpha}.$$

It follows evidently that, as $N \to \infty$, $G(N\alpha)^d = \alpha^d Y_N^d + O(Y_N^{d-1})$. This, jointly with Theorem 9, proves Theorem 10.

In order to make precise the result of Theorem 10, we can use the following argument. Let $\delta > 0$ be given, and consider $T > \delta$, and an arbitrary integer $M > 2T/\delta$. For any d-uple $I = \{i_1, \ldots, i_d\}$ of integers such that $0 \leqq i_j \leqq M - \left[\frac{\delta M}{T}\right]$

$-1, j=1, \ldots, d$, put

$$L_j^I = \left]\frac{T i_j}{M}, \delta + \frac{T(i_j-2)}{M}\right[, \quad 0 \leqq i_j \leqq M - \left[\frac{\delta M}{T}\right] - 2,$$

$$L_j^I = \left]T - \delta + \frac{2T}{M}, T\right[, \quad i_j = M - \left[\frac{\delta M}{T}\right] - 1.$$

Next, if $A_I(\delta, M)$ stands for the event that there is no point of arrival of $N(.)$ inside the hypercube $\prod_{j=1}^{d} L_j^I$, we have evidently

$$\{G(T) \geqq \delta\} \subset \bigcup_I A_I(\delta, M), \quad \text{or equivalently,} \quad \{G(T) \leqq \delta\} \supset \bigcap_I \bar{A}_I(\delta, M).$$

By Theorem 7, it follows that

$$\begin{aligned} P(G(T) \geqq \delta) &\leqq 1 - \left(1 - \exp\left(-\left\{\delta - \frac{2T}{M}\right\}^d\right)\right)^{(M - [\frac{\delta M}{T}] - 1)^d} \\ &\leqq 1 - \left(1 - \exp\left(-\left\{\delta - \frac{2T}{M}\right\}^d\right)\right)^{M^d} \\ &\leqq 1 - \left(1 - \exp\left(-\delta^d + \frac{2Td\delta^{d-1}}{M}\right)\right)^{M^d}. \end{aligned}$$

Put now

$$\delta^d = d \operatorname{Log} T + (d-1) \operatorname{Log}_2 T + x,$$

and

$$M = [\lambda T (\operatorname{Log} T)^{\frac{d-1}{d}}],$$

where λ is a constant. We get then

$$P(G(T) \geqq \delta) \leqq 1 - \exp\left\{-\exp\left(-x + d \operatorname{Log} \lambda + \frac{2d^{2-\frac{1}{d}}}{\lambda} + o(1)\right)\right\}, \quad \text{as } T \to \infty.$$

We now specify $\lambda = 2d^{\frac{d-1}{d}}$. We obtain the following limiting results:

Proposition 3. *For any x, we have*

$$\liminf_{T \to \infty} P(G^d(T) < d \operatorname{Log} T + (d-1) \operatorname{Log}_2 T + x) \geqq \exp(-e^{-x+C}),$$

where $C = d \operatorname{Log}(2ed^{\frac{d-1}{d}})$.

Proposition 4. *For any $\varepsilon > 0$ and $r \geqq 5$, we have, almost surely as $T \to \infty$*

$$G^d(T) \leqq \operatorname{Log} T^d + (d+1) \operatorname{Log}_2 T^d + \operatorname{Log}_3 T^d + \ldots + (1+\varepsilon) \operatorname{Log}_r T^d.$$

Proof. It suffices to prove the result for $T = e^k$, $k = 1, 2, \ldots$, and for an arbitrary $\varepsilon > 0$. This follows from Borel-Cantelli (see Deheuvels (1983)).

Proposition 4 makes precise the upper bound of Theorem 10. The derivation of lower bounds will not be treated here.

Following our discussion of § 5, it can be seen that the *similarities between runs, spacings, and gaps in a Poisson process, exist in the d-dimensional space for any* $d \geqq 1$.

From there, one is tempted to extend this to the Erdös-Rényi theorem. Unfortunately, there is no such result up to now for $d \geqq 2$ (not to mention Steinebach's (1983) paper, which cannot be used to cover the case we discuss here).

This remark gives the origin of Theorem 0, whose detailed proof is to be found in the next paragraph.

8. Multidimensional Erdös-Rényi Theorems

Let $X_{i_1,\ldots,i_d}$, $i_1 \geqq 1, \ldots, i_d \geqq 1$ be a d-indexed array of i.i.d. random variables. Assume that the hypotheses of Theorem 0 are satisfied, with the notations of § 1. We shall prove in the sequel Theorem 0, by using the same methods as Csörgő and Steinebach, 1981.

Lemma 7. *Put, for* $n_1 \geqq 1, \ldots, n_d \geqq 1$, $S_{n_1,\ldots,n_d} = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} X_{i_1,\ldots,i_d}$. *Then, there there exists a* $\delta > 0$ *such that, for any* $K \geqq 1$ *and u, we have*

$$P_1 = P\Big(\max_{0 \leqq n_j \leqq K,\, 1 \leqq j \leqq d} S_{n_1,\ldots,n_d} \geqq K^d \alpha + u\Big) \leqq K^d \rho^{K^d}(\alpha) e^{-\delta u}.$$

Proof. We follow Csörgő and Steinebach, 1981, Lemma 1. Let $t = t^*(\alpha) > 0$ be such that $\phi'(t)/\phi(t) = \alpha$. Then

$$\begin{aligned} P_1 &\leqq \sum_{0 \leqq n_j \leqq K,\, 1 \leqq j \leqq d} P(S_{n_1,\ldots,n_d} \geqq K^d \alpha + u) \\ &\leqq \sum_{0 \leqq n_j \leqq K,\, 1 \leqq j \leqq d} E(\exp(t(S_{n_1,\ldots,n_d} - K^d \alpha - u))) \\ &\leqq K^d \phi^{K^d}(t) e^{-K^d t\alpha - ut} = K^d \rho^{K^d}(\alpha) e^{-ut}, \end{aligned}$$

hence result.

Proof of Theorem 0 (i). We chose now $K = [\{cd \operatorname{Log} N\}^{1/d}]$, and remark that

$$\begin{aligned} P_2(N) &= P(I(N, K) \geqq K^d \alpha + u) \leqq N^d K^d \exp(-K^d/c) e^{-u\delta} \\ &\leqq N^d \{cd \operatorname{Log} N\} \exp(-\{d \operatorname{Log} N\}\{1 - d(cd \operatorname{Log} N)^{-1/d}\} - u\delta) \\ &= cd \exp(\operatorname{Log}_2 N + c^{-1} d(cd \operatorname{Log} N)^{(d-1)/d} - u\delta). \end{aligned}$$

Let us take now take $u = c^{-1} d(cd \operatorname{Log} N)^{(d-1)/d}(1+\varepsilon)$ if $d \geqq 2$. If we take $N = N_j$ to be the greatest integer such that $[\operatorname{Log} N_j] = j$. We get then that, for any $\varepsilon > 0$, $\sum_j P_2(N_j) < \infty$, proving:

Proposition 5. *Under the assumptions of Theorem* 0, *we have, for any* $d \geqq 2$,

$$\operatorname*{Lim\,Sup}_{N \to \infty} \frac{I(N, [\{cd \operatorname{Log} N\}^{1/d}]) - \alpha[\{cd \operatorname{Log} N\}^{1/d}]^d}{(cd \operatorname{Log} N)^{(d-1)/d}} \leqq c^{-1} d \quad \text{a.s.}$$

For the lower bound, we shall need the following result of Csörgö and Steinebach, 1981 (Lemma 1, (16)), which we adapt in d dimensions:

Lemma 8. *For any* $\varepsilon>0$, *there exists constants* $\delta=\delta(\varepsilon)$ *and* $A=A(\varepsilon)>0$ *such that*

$$P(S_{K,\ldots,K}>K^d\alpha-K^{d/2}\varepsilon)\geqq A\rho^{K^d}(\alpha)e^{\delta K^{d/2}},$$

for K sufficiently large.

Proof. Let $t=t^*(\alpha)>0$ be such that $\phi'(t)/\phi(t)=\alpha$. Consider the probability measure $P_{t,K}$ defined by $P_{t,L}(E)=\int_E e^{tS_K}/\phi^K(t)\,dP$, where S_K is distributed as the sum of K i.i.d. random variables with the distribution of $X_{1,\ldots,1}$.

We have

$$\begin{aligned}P(S_K>K\alpha-\varepsilon K^{1/2})&=\phi^K(t)\int_{\{S_K-K\alpha>-\varepsilon K^{1/2}\}} e^{-tS_K}\,dP_{t,K}\\&\geqq\phi^K(t)e^{-Kt\alpha+t\frac{\varepsilon}{2}K^{1/2}}P_{t,K}\left(-\varepsilon<\frac{S_K-K\alpha}{K^{1/2}}\leqq-\frac{\varepsilon}{2}\right).\end{aligned}$$

The result follows from the central limit theorem applied to the probability above. The lemma is obtained by replacing K by K^d in the preceding expressions, and by letting $K\to\infty$.

Proof of Theorem 0 (ii). We have, as in Csörgö and Steinebach, 1981, p. 993,

$$\begin{aligned}P_3(N)&=P(I(N,K)\leqq K^d\alpha-\varepsilon K^{d/2})\leqq(1-A\rho^{K^d}(\alpha)e^{\delta K^{d/2}})^{[N/K]^d}\\&\leqq\exp\left(-A[N/K]^d\exp\left(-\frac{K^d}{c}+\delta K^{d/2}\right)\right).\end{aligned}$$

Put now $K=[\{cd\,\mathrm{Log}\,N\}^{1/d}]$. We have then

$$P_3(N)\leqq\exp(-A\{cd\,\mathrm{Log}\,N\}^{-1}\{1+o(1)\}\exp(\delta[\{cd\,\mathrm{Log}\,N\}^{1/d}]^{d/2})).$$

From there, it follows that $\sum_N P_3(N)<\infty$. We have proved the following result:

Proposition 6. *Under the assumptions of Theorem* 0, *we have, for any* $d\geqq1$,

$$\operatorname*{Lim\,Inf}_{N\to\infty}\frac{I(N,[\{cd\,\mathrm{Log}\,N\}^{1/d}])-\alpha[\{cd\,\mathrm{Log}\,N\}^{1/d}]^d}{(cd\,\mathrm{Log}\,N)^{1/2}}\geqq0\quad\text{a.s.}$$

The proof of Theorem 0 is now complete, since, for $d\geqq2$, $1/2\leqq\frac{d-1}{d}$.

Remarks. 1°) The upper bound in Proposition 5 follows from the inequalities

$$u^d-du^{d-1}\leqq(u-1)^d\leqq[u]^d\leqq u^d,\qquad u>0,$$

applied with $u=\{cd\,\mathrm{Log}\,N\}^{1/d}$. It appears in fact that the total number of random variables taken in the "moving average" $S(n_1,\ldots,n_d,[u])$ may differ from $[u^d]$ by a number as high as $d(1-\varepsilon)u^{d-1}$, with $0<\varepsilon<1$.

This may be improved by changing the definition of $I(N,[u])$. If, for instance, we define

$$S^*(n_1,\ldots,n_d,u)=\sum_{j_1=n_1+1}^{n_1+K}\ldots\sum_{j_d=n_d+1}^{n_d+K_d} X_{j_1,\ldots,j_d} \quad \text{(see also Steinebach (1983))},$$

where $K_1=K_1(u),\ldots,K_d=K_d(u)$ are chosen in such a way that $K_1\ldots K_d$ is closer of $[u^d]$ than $[u]^d$, it seems then possible to improve the upper bound corresponding to the associated maximum $I^*(N,u)$ in the corresponding version of Proposition 5.

2°) It seems that the most general form of Erdös-Rényi problem in the d-dimensional space can be stated in the following terms. Assume that $X_{j_1,\ldots j_d}=R(j_1,\ldots,j_d)$, and consider the random function $R(x_1,\ldots,x_d)=R([x_1],\ldots,[x_d])$. Next, consider a kernel $K(.)$ and the associated moving average

$$A(x,\lambda)=\int_{R^d} R(u)K(\lambda^{-1}(x-u))du.$$

Finally, consider a sequence of increasing sets $S_1\subset S_2\subset\ldots$.

The general Erdös-Rényi problem is then to find the asymptotic behavior as $n\to\infty$ of

$$\operatorname*{Max}_{x\in S_n} A(x,\lambda_n),$$

when $\lambda_n^d\sim(\text{Constant})\times\text{Log}(\text{Volume of } S_n)$.

We have discussed here the particular case where S_n is an increasing sequence of hypercubes, and where $K(.)$ is the indicator function of a hypercube.

This problem is very much related to the question of finding the asymptotic behavior as $n\to\infty$ of

$$\operatorname*{Min}_{x} f_n(x) \quad \text{and} \quad \operatorname*{Max}_{x} f_n(x),$$

where

$$f_n(x)=(n\delta_n^d)^{-1}\sum_{i=1}^{n} K(\delta_n^{-1}(X_i-x))$$

is a Parzen-Rosenblatt density estimator in R^d, and where $n\delta_n^d\sim C\,\text{Log}\,n$.

3°) The estimations given in Theorem 0 and Proposition 5–6 are probably not the best available, due to the crude method of evaluation which has proved their validity.

4°) There are many other Erdös-Rényi type theorems which could be given on the same lines as Theorem 0. Most of them appear as direct generalizations of the corresponding results in one dimension, with the appropriate definition of maximum given above. Details on this will be given elsewhere.

5°) Theorem 0 has obvious applications in strong approximation theory.

Acknowledgement. I would like to thank the referee for his comments and remarks.

References

Barndorff-Nielsen, O.: On the rate of growth of the partial maxima of a sequence of independent indentically distributed random variables. Math. Scand. **9**, 383–394 (1961)

Csörgő, S.: Erdös-Rényi laws. Ann. Stat. **7**, 772–787 (1979)

Csörgő, M., Steinebach, J.: Improved Erdös-Rényi and strong approximation laws for increments of partial sums. Ann. Probab. **9**, 988–996 (1981)

Deheuvels, P.: Majoration et minoration presque sûre optimale des éléments de la statistique ordonnée d'un échantillon croissant de variables aléatoires indépendantes. Rendi Conti della Academia Nazionale dei Lincei, Ser. 8, **61**, 707–719 (1974)

Deheuvels, P.: Strong limiting bounds for maximal uniform spacings. Ann. Probab. **10**, 1058–1065 (1982)

Deheuvels, P.: Strong bounds for multidimensional spacings. Z. Wahrscheinlichkeitstheor. verw. Gebiete **64**, 411–424 (1983)

Deheuvels, P., Devroye, L.: Limit laws related to the Erdös-Rényi theorem. Tech. Report n°83-6, L.S.T.A., Université Paris VI (1983)

Deheuvels, P. Devroye, L., Lynch, J.: Exact convergence rate in the limit theorems of Erdös-Rényi and Shepp. to appear in Ann. Probab. (1985)

Devroye, L.: Laws of the iterated logarithm for order statistics of uniform spacings. Ann. Probab. **9**, 860–867 (1981)

Devroye, L.: A LogLog law for maximal uniform spacings. Ann. Probab. **10**, 863–868 (1982)

Erdös, P., Rényi, A.: On a new law of large numbers. J. Anal. Math. **23**, 103–111 (1970)

Erdös, P., Révész, P.: On the length of the longest head run. Colloq. Math. Soc. J. Bolyai **16**, pp. 219–228. In: Topics in information theory, Csiszár, I., Elias, P. (eds.). Amsterdam: North Holland 1975

Fisz, M.: Probability theory and mathematical statistics (3d edition). New York: Wiley 1963

Geffroy, J.: Contributions à la théorie des valeurs extrêmes. Publ. Inst. Stat. Univ. Paris **7/8**, 37–185 (1958/59)

Guibas, L.J., Odlyzko, A.M.: Long repetitive patterns in random sequences. Z. Wahrscheinlichkeitstheor. verw. Gebiete **53**, 241–262 (1980)

Kiefer, J.: Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. Proc. Sixth Berkeley Sympos. Univ. Calif. (1973). Math. Stat. Probab. **3**, 227–244 (1973)

Révész, P.: How random is random?. 14th Meeting of the European Statisticians (Closing address) Wroclaw, Poland (1981)

Serfling, R.J.: Some elementary results on Poisson approximation in a sequence of Bernoulli trials. SIAM Rev. **20**, 3, 567–579 (1978)

Shepp, L.A.: A limit theorem concerning moving averages. Ann. Math. Stat. **35**, 424–428 (1964)

Steinebach, J.: On the increments of partial sum processes with multidimensional indices. Z. Wahrscheinlichkeitstheor. verw. Gebiete **63**, 59–70 (1983)

Received December 10, 1983; in revised form January 16, 1984; in final form October 25, 1984