

LAWS OF THE ITERATED LOGARITHM FOR ORDER STATISTICS OF UNIFORM SPACINGS

BY LUC DEVROYE¹

McGill University

Let X_1, X_2, \dots be a sequence of independent uniformly distributed random variables on $[0, 1]$, and let K_n be the k th largest spacing induced by the order statistics of X_1, \dots, X_{n-1} . We show that

$$\limsup(nK_n - \log n)/2 \log_2 n = 1/k \quad \text{almost surely,}$$

and

$$\liminf(nK_n - \log n + \log_3 n) = c \quad \text{almost surely,}$$

where $-\log 2 \leq c \leq 0$, and \log_j is the j times iterated logarithm.

1. Introduction. Consider a sequence X_1, X_2, \dots of independent identically distributed random variables with a uniform distribution on $[0, 1]$. If $X_{(1)} < X_{(2)} < \dots < X_{(n-1)}$ are the order statistics corresponding to X_1, \dots, X_{n-1} , then the *maximal uniform spacing* (or, the *maximal gap*) M_n is defined by

$$M_n = \max_{1 \leq i \leq n} S_i$$

where $S_1 = X_{(1)}$, $S_i = X_{(i)} - X_{(i-1)}$ for $1 < i < n$, and $S_n = 1 - X_{(n-1)}$. The S_i 's are called the *spacings*; see Pyke (1965).

Slud (1978) showed that $nM_n - \log n = O(\log_2 n)$ a.s.; we will refine Slud's result and show that

$$(1.1) \quad \limsup(nM_n - \log n)/2 \log_2 n = 1 \quad \text{a.s.}$$

and that

$$(1.2) \quad \liminf nM_n - \log n + \log_3 n = c \quad \text{a.s.}$$

where $-\log 2 \leq c \leq 0$. Along the way, we will obtain a few large deviation results for M_n . In Section 2, we state without proof a few known results about the distribution and the weak convergence of M_n . In Sections 4 and 5, we will establish (1.1) and (1.2) for K_n , the k th largest spacing among S_1, \dots, S_n , when the constant "1" in (1.1) is replaced by $1/k$.

2. Auxiliary results. It is well-known that (S_1, \dots, S_n) is uniformly distributed on the simplex $\{(x_1, \dots, x_n) \mid x_i \geq 0; \sum x_i = 1\}$, and that, therefore

$$P(S_1 > a_1; \dots; S_n > a_n) = (1 - \sum_{i=1}^n a_i)^{n-1}, \quad \sum_{i=1}^n a_i < 1$$

$$= 0, \quad \text{otherwise,}$$

where a_1, \dots, a_n are nonnegative numbers. From this, one can get Whitworth's formula (Whitworth (1897); see also Kendall and Moran (1963)):

$$P(M_n > x) = P(\cup_{i=1}^n [S_i > x]) = \sum_i P(S_i > x) - \sum_{i < j} P(S_i > x; S_j > x) + \dots$$

$$= \sum_{k \geq 1; kx < 1} (-1)^{k+1} (1 - kx)^{n-1} \binom{n}{k}, \quad \text{all } x > 0.$$

Received August 27, 1979; revised August 5, 1980.

¹ This research was sponsored in part by National Research Council of Canada Grant No. A3456. AMS 1980 subject classifications. Primary 60F15.

Key words and phrases. Law of the iterated logarithm, order statistics, spacings, strong laws, almost sure convergence.

A very useful property of uniform spacings is the following.

LEMMA 2.1. *If Y_1, \dots, Y_n are independent identically distributed exponential random variables, and if $T_n = \sum Y_i$, then (S_1, \dots, S_n) is distributed as $(Y_1/T_n, \dots, Y_n/T_n)$. In particular, M_n is distributed as L_n/T_n where $L_n = \max(Y_i)$.*

For a proof of Lemma 2.1, see Pyke (1965).

LEMMA 2.2. (Sukhatme, 1937). *If Y_1, \dots, Y_n are independent identically distributed exponential random variables with corresponding order statistics $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$, then the following random variables are also independent and exponentially distributed:*

$$nY_{(1)}, (n - 1)(Y_{(2)} - Y_{(1)}), \dots, 2(Y_{(n-1)} - Y_{(n-2)}), Y_{(n)} - Y_{(n-1)}.$$

An immediate consequence of Lemma 2.2 is the following.

LEMMA 2.3. *M_n is distributed as*

$$\sum_{i=1}^n (Y_i/i) / \sum_{i=1}^n Y_i$$

where Y_1, \dots, Y_n are independent exponentially distributed random variables.

The limit distribution of M_n was found by Levy (1939) and was rederived later by Darling (1952, 1953) and others.

LEMMA 2.4. *For all $x \in R$, $P(nM_n < \log n + x) \rightarrow \exp(-\exp(-x))$ as $n \rightarrow \infty$.*

LEMMA 2.5. *$nM_n/\log n \rightarrow 1$ in probability as $n \rightarrow \infty$.*

Note. If G_n is the distribution function of $nM_n - \log n$ and $G(x) = \exp(-\exp(-x))$, and if $a_n \log n \rightarrow \infty$ as $n \rightarrow \infty$, then

$$\begin{aligned} P(|nM_n/\log n - 1| > a_n) &= G_n(-a_n \log n) + 1 - G_n(a_n \log n) \\ (2.1) \qquad \qquad \qquad &\leq 2\sup_x |G_n(x) - G(x)| \\ &\quad + G(-a_n \log n) + 1 - G(a_n \log n) \rightarrow 0. \end{aligned}$$

The distribution function $G(x) = \exp(-\exp(-x))$ has mean $\gamma = 0.5772157\dots$ (the Euler constant) and variance $\pi^2/6$; see Gnedenko (1943), Gumbel (1958), Barndorff-Nielsen (1963) and David (1970) for a closer analysis of its properties. A careful application of Lemma 2.3 also gives

LEMMA 2.6. *$E(nM_n - \log n) \rightarrow \gamma$ as $n \rightarrow \infty$, and $\text{Var}(nM_n) \rightarrow \pi^2/6$ as $n \rightarrow \infty$.*

3. Large deviation results. We will first derive exponential estimates for the probability in the tail of the gamma density. We recall here that the sum T_n of n independent exponentially distributed random variables has the gamma density $g_n(x) = x^{n-1}e^{-x}/(n - 1)!, x \geq 0$.

LEMMA 3.1. *For all $x > 0$,*

$$P(T_n/n - 1 > x) \leq \exp(-nx^2(1 - x)/2)$$

and

$$P(T_n/n - 1 < -x) \leq \exp(-nx^2/2).$$

PROOF. Here and throughout the paper we will use these analytic inequalities, valid for all $x \geq 0$:

$$(3.1) \quad e^{x-x^2/2} \leq 1+x \leq e^x \leq 1+x+x^2e^x/2$$

$$(3.2) \quad 1-x \leq e^{-x-x^2/2-x^3/3} \leq e^{-x-x^2/2} \leq e^{-x} \leq 1-x+x^2/2.$$

Lemma 3.1 is now easily proved by Chernoff's classical technique (Chernoff, 1952). For any $0 < s < 1$, we have $P(T_n/n - 1 > x) \leq e^{-snx} E(e^{s(T_n-n)}) = e^{-sn(1+x)}(1-s)^{-n}$. This expression is minimal when $1-s = 1/(1+x)(s = x/(1+x))$, so that the said probability is not greater than $(e^{-x}(1+x))^n \leq ((1-x+x^2/2)(1+x))^n = (1-x^2/2+x^3/2)^n \leq e^{-nx^2(1-x)/2}$. Similarly, for all $s > 0$, $P(T_n/n - 1 < -x) \leq e^{-snx} E(e^{-s(T_n-n)}) = e^{sn(1-x)}(1+s)^{-n} = (e^x(1-x))^n \leq (e^{-x-x^2/2})^n = e^{-nx^2/2}$ where we let $s = x/(1-x)$ whenever $x < 1$. For $x \geq 1$, the result is trivially true.

LEMMA 3.2. *Let $k \geq 1$ be a fixed integer, and let $a_n \rightarrow 0$ and $a_n \log n \rightarrow \infty$. If K_n is the k -th largest spacing among S_1, \dots, S_n , then*

$$P(nK_n/\log n - 1 > a_n) \sim n^{-ka_n}/k!$$

and

$$P(nK_n/\log n - 1 \leq -a_n) \sim n^{(k-1)a_n} \exp(-na_n)/(k-1)!$$

PROOF. We will use the following fact about the tail of the binomial distribution. If B is a binomial random variable with parameters n and p , then $np \rightarrow 0$ implies $P(B \geq k) \sim P(B = k)$, and $np \rightarrow \infty$ implies $P(B < k) \sim P(B = k - 1)$ (Feller, 1957, page 140).

K_n is distributed as L'_n/T_n where L'_n is the k th largest of n independent identically distributed random variables with exponential density and whose sum is T_n (Lemma 2.1). For arbitrary $a, b > 0$ we have

$$(3.3) \quad \begin{aligned} P(L'_n < (1-a-b)\log n) - P(T_n < n(1-b)) &\leq P(nK_n/\log n < 1-a) \\ &\leq P(L'_n < (1-a+b)\log n) + P(T_n \geq n(1+b)) \end{aligned}$$

and

$$(3.4) \quad \begin{aligned} P(L'_n > (1+a+b)\log n) - P(T_n > n(1+b)) &\leq P(nK_n/\log n > 1+a) \\ &\leq P(L'_n > (1+a-b)\log n) + P(T_n \leq n(1-b)). \end{aligned}$$

Let us take $a = a_n$ and $b = n^{-1/4}$. Lemma 3.2 follows if we can show the following things:

- (i) $P(L'_n < (1-a)\log n) \sim \exp(-n^a)n^{(k-1)a}/(k-1)!$;
- (ii) $P(L'_n > (1+a)\log n) \sim n^{-ka}/k!$;
- (iii) $P(|T_n - n| > bn)/\min(P(L'_n < (1-a)\log n), P(L'_n > (1+a)\log n)) \rightarrow 0$;
- (iv) $P(L'_n < (1-a-b)\log n) \sim P(L'_n < (1-a+b)\log n)$;
- (v) $P(L'_n > (1+a+b)\log n) \sim P(L'_n > (1+a-b)\log n)$.

Clearly, $P(L'_n < (1-a)\log n) = P(B < k)$ where B is binomial with parameters n and $p = \exp(-(1-a)\log n) = n^a/n$. Since $np \rightarrow \infty$, we have $P(B < k) \sim P(B = k - 1) = \binom{n}{k-1} p^{k-1} (1-p)^{n-k+1} \sim (np)^{k-1} \exp(-np)/(k-1)! = n^{(k-1)a} \exp(-n^a)/(k-1)!$. Similarly, $P(L'_n > (1+a)\log n) = P(B \geq k)$ where now B is binomial with parameters n and $p = \exp(-(1+a)\log n) = 1/n^{1+a}$. Since $np \rightarrow 0$, we have $P(B \geq k) \sim P(B = k) \sim 1/n^{ka}k!$. This proves (i) and (ii). The same asymptotic results are valid if in (i) and (ii) we replace a by $(a+b)$ or $(a-b)$ on both sides. The ratio of the two terms of (v) (left divided by right) is $\sim n^{-2kb} \sim 1$. The ratio of the two terms of (iv) is $\sim n^{2(k-1)b} \exp(n^{(a-b)} - n^{(a+b)}) \sim 1$.

To prove (iii) we first use Lemma 3.1: $P(|T_n - n| > bn) \leq 2 \exp(-nb^2/4)$ for n large enough. It remains to check that $n^{ka} \exp(-nb^2/4) \rightarrow 0$ and that $n^{(k-1)a} \exp(n^a - nb^2/4) \rightarrow 0$. This follows from $a \rightarrow 0$.

4. Outer Bounds. In 1961 Barndorff-Nielsen (and independently Robbins and Siegmund (1970) and Deheuvels (1974)) established laws of the iterated logarithm for $Z_n = \min(X_1, \dots, X_n)$ where X_1, \dots, X_n is a sequence of independent uniform $[0, 1]$ random variables. These results can be summarized as follows. Let a_n be positive and nonincreasing. Then,

- (i) $Z_n < a_n$ i.o. (f.o.) when $\sum a_n = \infty$ ($\sum a_n < \infty$). See Geffroy (1958) for the first proof.
- (ii) $Z_n > a_n$ i.o. (f.o.) when $\sum a_n \exp(na_n) = \infty$ ($\sum a_n \exp(-na_n) < \infty$) under the assumption that na_n is ultimately non-decreasing (Robbins and Siegmund, 1970). Barndorff-Nielsen's result uses the series $\sum \log_2 n(1 - a_n)^n/n$ instead of $\sum a_n \exp(-na_n)$. For related work, see Frankel (1972) and Wichura (1973). For a short proof of the first order result: $Z_n > (1 + \epsilon) \log_2 n/n$ i.o. (f.o.) when $\epsilon = 0$ ($\epsilon > 0$), see Kiefer (1970). For a survey, with proofs, see Galambos (1978).

In this section we derive sufficient conditions (of the summability type) for $nK_n > (1 + a_n) \log n$ finitely often a.s. and $nK_n < (1 - a_n) \log n$ finitely often a.s.

LEMMA 4.1. *Let A_1, A_2, \dots be a sequence of events with $P(A_n) \rightarrow 0$ as $n \rightarrow \infty$. If either $\sum P(A_n^c \cap A_{n+1}) < \infty$ or $\sum P(A_n \cap A_{n+1}^c) < \infty$, then $P(A_n \text{ f.o.}) = 1$.*

PROOF. See Barndorff-Nielsen (1961).

THEOREM 4.1. *Let $a_n \rightarrow 0$ and $a_n \log n \rightarrow \infty$ as $n \rightarrow \infty$ such that $(1 + a_n) \log n/n$ is ultimately nonincreasing. Then, $P(nK_n > (1 + a_n) \log n \text{ i.o.}) = 0$ when*

$$(4.1) \quad \sum_{n=1}^{\infty} \log n/n^{1+ka_n} < \infty.$$

PROOF. Let A_n be the event $nK_n > (1 + a_n) \log n$. By (2.1), $P(A_n) \rightarrow 0$ as $n \rightarrow \infty$. Then, for n large enough,

$$\begin{aligned} P(A_n \cap A_{n+1}^c) &\leq P(nK_n > (1 + a_n) \log n) 2k(1 + a_{n+1})(\log(n + 1)/(n + 1)) \\ &= 2k(1 + o(1))n^{-ka_n}k!^{-1} \log n/n, \end{aligned}$$

from which Theorem 4.1 follows after applying Lemma 4.1.

THEOREM 4.2. *Let $a_n \rightarrow 0$ and $a_n \log n \rightarrow \infty$ as $n \rightarrow \infty$ such that $(1 - a_n) \log n/n$ is ultimately nonincreasing. Then, $P(nK_n < (1 - a_n) \log n \text{ i.o.}) = 0$ when*

$$(4.2) \quad \sum_{n=1}^{\infty} (\log n/n)n^{ka_n} \exp(-n^{a_n}) < \infty.$$

PROOF. Let A_n be the event $nK_n < (1 - a_n) \log n$. Once again, we will use Lemma 4.1. Obviously, $P(A_n) \sim n^{(k-1)a_n} \exp(-n^{a_n})/(k - 1)! \rightarrow 0$ as $n \rightarrow \infty$. Also, if K'_n is the $(k + 1)$ st largest spacing among S_1, \dots, S_n , then for n large,

$$\begin{aligned} P(A_n^c \cap A_{n+1}) &= P(A_n^c \cap A_{n+1} \cap [K'_n < (1 - a_{n+1}) \log(n + 1)/(n + 1)]) \\ &\leq P(K'_n < (1 - a_n) \log n/n) 2k \log n/n \\ &= 2k(1 + o(1))n^{ka_n} \exp(-n^{a_n})k!^{-1} \log n/n. \end{aligned}$$

REMARK 4.1. It follows trivially from Theorems 4.1 and 4.2 that $nK_n/\log n \rightarrow 1$ a.s. as $n \rightarrow \infty$. Of course, we have done too much work by invoking Lemma 3.2. For a short proof of $nM_n/\log n \rightarrow 1$ a.s., see Slud (1978) or Devroye (1979).

REMARK 4.2. Condition (4.1) is satisfied if for some $\delta > 0, J \geq 2$, we have

$$a_n = (k \log n)^{-1}(\log_2 n + \sum_{j=2}^J \log_j n + \delta \log_j n).$$

In particular, it is satisfied if we take $a_n = (2 + \delta)\log_2 n / (k \log n)$, $\delta > 0$. Hence,

$$(4.3) \quad \limsup(nK_n - \log n) / 2 \log_2 n \leq 1/k \text{ a.s.}$$

REMARK 4.3. Condition (4.2) is satisfied if for some $J \geq 3$, $\delta > 0$, we have

$$a_n = (\log n)^{-1}(\log(2 \log_2 n + k \log_3 n + \sum_{j=3}^J \log_j n + \delta \log_j n)),$$

or when for some $\delta > 0$, $a_n = \log((2 + \delta)\log_2 n) / \log n$. Hence,

$$(4.4) \quad \liminf(nK_n - \log n + \log_3 n) \geq -\log 2 \text{ a.s.,}$$

independent of k . The influence of k on the lower outer bound is only in the second order term of the sequence a_n . In other words, whenever M_n is small, it is very likely that the second and third largest spacings are very close in magnitude to M_n .

5. Inner Bounds. In this section we will prove the following theorems:

THEOREM 5.1. $\limsup(nK_n - \log n) / 2 \log_2 n = 1/k$ a.s.

THEOREM 5.2. $\liminf(nK_n - \log n + \log_3 n) = c$ a.s. for some $c \in [-\log 2, 0]$.

We will use the notation $[\cdot]$ for the integer part of a number. Furthermore, we will need two lemmas.

LEMMA 5.1. If $b_j = \exp(a\sqrt{j} \log j)$, where $a > 0$, then

$$(b_{j+1} - b_j) / b_j \sim a \log j / 2\sqrt{j} \quad \text{as } j \rightarrow \infty.$$

The same is true for $c_j = [b_j]$.

PROOF. In view of $(\sqrt{j+1} - \sqrt{j}) \sim \frac{1}{2}\sqrt{j}$ and $\log(1 + 1/j) \sim 1/j$, we have $(b_{j+1} - b_j) / b_j \sim a(\sqrt{j+1} \log(j+1) - \sqrt{j} \log j) \sim a \log j / 2\sqrt{j}$.

LEMMA 5.2. If $b_j = \exp(j \log j)$, then

$$b_j / b_{j+1} \sim 1/ej \quad \text{as } j \rightarrow \infty.$$

The same is true for $c_j = [b_j]$.

PROOF. By (3.1) and (3.2) we have $b_{j-1} / b_j = (j - 1)^{-1} \exp(j \log(1 - 1/j)) \leq 1/(e(j - 1))$, and $b_{j-1} / b_j \geq (j - 1)^{-1} \exp(-1 - 1/j) \geq (j - 1)^{-1} e^{-1} (1 - 1/j) = 1/ej$.

PROOF OF THEOREM 5.1. In view of (4.3) we need only show that $nK_n - \log n > (2/k - \delta)\log_2 n$ i.o. almost surely, for all $\delta > 0$. We define the following sequences:

$$\begin{aligned} n_j &= [\exp(\sqrt{j} \log j)], \\ t_j &= [n_j(2/k - \delta/2)\log_2 n_j / \log n_j], \\ a_j &= (2/k - \delta)\log_2 j / \log j, \\ d_j &= (1 + a_j)\log j / j, \\ d'_j &= (1 + (3/k)\log_2 n_j / \log n_j)\log n_j / n_j, \\ d''_j &= (1 - \log(3 \log_2 n_j) / \log n_j)\log n_j / n_j. \end{aligned}$$

Let us define the following events: A_N is the event that $K_n \in (d''_j, d'_j)$ for all $j \geq N$; B_N is the event that for some $j \geq N$, none of the random variables X_n, \dots, X_{n+t_j-1} belong to the set C_j , where C_j is the union of k intervals of length d'_j each, with the restriction that the leftmost point of each interval coincides with the leftmost point of one of the k largest spacings.

We will see that $t_j + n_j < n_{j+1}$ for all j large enough, and that $d_j'' > d_{n_j+t_j}$ for all j large enough. Thus, $A_N \cap B_N \subseteq [K_{n_j+t_j} > d_{n_j+t_j}$ for some $j \geq N]$. The theorem now follows if we can show that $P(A_N^c) + P(B_N^c) \rightarrow 0$ as $N \rightarrow \infty$. From Theorems 4.1 and 4.2 we deduce that $P(A_N^c) \rightarrow 0$ as $N \rightarrow \infty$. Furthermore,

$$P(B_N^c) \leq \prod_{j=N}^{\infty} (1 - (1 - kd_j')^b) \leq \exp(-\sum_{j=N}^{\infty} (1 - kd_j')^b) = 0$$

whenever

$$\sum_{j=1}^{\infty} (1 - kd_j')^b = \infty.$$

Because $(1 - kd_j')^b \geq \exp(-d_j'kt_j - k^2d_j'^2t_j/2)$ and $d_j'^2t_j \rightarrow 0$, it suffices to check whether $\sum \exp(-kd_j't_j) = \infty$. We have $\exp(-kd_j't_j) \sim \exp(-(2 - \delta k/2)\log_2 n_j \cdot (1 + (3/k)\log_2 n_j/\log n_j)) \sim \exp(-(2 - \delta k/2)\log_2 n_j) \sim (\sqrt{j} \log j)^{2-\delta k/2}$, which is not summable with respect to j .

We will now show that $n_j + t_j < n_{j+1}$ for all j large enough. Indeed, $n_{j+1} - n_j \sim n_j \log j/2\sqrt{j}$ (Lemma 5.1), while $t_j \sim (1/k - \delta/4)n_j/\sqrt{j}$.

Finally, let us establish that $d_j'' > d_{n_j+t_j}$ for all j large enough. Clearly,

$$\begin{aligned} d_{n_j+t_j} &= \log(n_j + t_j)/(n_j + t_j) + (2/k - \delta)\log_2(n_j + t_j)/(n_j + t_j) \\ &< \log n_j/(n_j + t_j) + t_j/n_j^2 + (2/k - \delta)\log_2 n_j/n_j \\ &< (\log n_j/n_j)(1 - (1 + o(1))t_j/n_j) + o(1)/n_j + (2/k - \delta)\log_2 n_j/n_j \\ &< \log n_j/n_j - ((2/k - \delta/2)(1 + o(1))\log_2 n_j - (2/k - \delta)\log_2 n_j)/n_j \\ &= \log n_j/n_j - (\delta/2)(1 + o(1))\log_2 n_j/n_j. \end{aligned}$$

Also, $d_j'' = \log n_j/n_j - \log(3 \log_2 n_j)/n_j > d_{n_j+t_j}$ for all j large enough.

PROOF OF THEOREM 5.2. We will show that for all $\delta > 0$, the inequality $nK_n < \log n - \log_3 n + \delta$ is satisfied i.o. almost surely, that is, a.s. $\liminf(nK_n - \log n + \log_3 n) \leq 0$. This result together with (4.4) imply the statement of Theorem 5.2.

For given $\delta > 0$, define $n_j = \lceil \exp(2j \log j) \rceil$, $d_j = (\log n_j - \log_3 n_j + \delta)/n_j$, $t_j = n_j - n_{j-1}$ and $a_j = (\log_3 n_j - \delta/2)/\log n_j$. Let further N_j be the k th largest gap defined by $X_{n_j-1}, \dots, X_{n_j-1}$ on $[0, 1]$. Obviously, $N_j < d_j$ i.o. implies that $K_n < d_j$ i.o. Since the N_j 's are independent, $N_j < d_j$ i.o. almost surely whenever $\sum P(N_j < d_j) = \infty$. By Lemma 3.2,

$$P(N_j < (\log t_j/t_j)(1 - a_j)) \sim t_j^{(k-1)a_j} \exp(-t_j^{a_j})/(k-1)!$$

because $a_j \log t_j \rightarrow \infty$. Also, $\exp(-t_j^{a_j}) \geq \exp(-n_j^{a_j}) = \exp(-c' \log_2 n_j) \sim (2j \log j)^{-c'}$ for some $c' < 1$. Thus, $\sum P(N_j < d_j) = \infty$ if $d_j > (\log t_j/t_j)(1 - a_j)$ for all j large enough. Now,

$$d_j t_j / \log n_j \geq (t_j/n_j)(1 - (\log_3 n_j - \delta)/\log n_j) = (1 - O(j^{-2}))(1 - (\log_3 n_j - \delta)/\log n_j)$$

which is greater than $1 - a_j = 1 - (\log_3 n_j - \delta/2)/\log n_j$ for all j large enough.

6. Applications.

EXAMPLE 6.1. Random covers. Assume that we try to cover $[0, 1]$ by intervals of length ℓ_n centered at X_1, \dots, X_{n-1} (where the X_i 's are independent and uniformly distributed on $[0, 1]$). Let A_n be the event $[[0, 1]$ is entirely covered]. Then, if $n\ell_n = \log n - \log_3 n + \delta$,

$$P(A_n \text{ i.o.}) = \begin{cases} 1, & \text{if } \delta > 0 \\ 0, & \text{if } \delta + \log 2 < 0. \end{cases}$$

If $n\ell_n = \log n + (2 + \delta)\log_2 n$, we have

$$P(A_n^c \text{ i.o.}) = \begin{cases} 1, & \text{if } \delta < 0 \\ 0, & \text{if } \delta > 0. \end{cases}$$

It is perhaps interesting to compare this result with Shepp's covering theorem (1972): let $\ell_1 \geq \ell_2 \geq \dots \geq 0$ be the lengths of arcs thrown at random on the circle with unit circumference ($\ell_1 < 1$). Then the circle is covered almost surely if and only if

$$\sum_{n=1}^{\infty} n^{-2} \exp(\ell_1 + \dots + \ell_n) = \infty.$$

If $\ell_n = (1/n)(1 - (1 + \delta)/\log n)$, then this condition is satisfied when $\delta \leq 0$ and is violated when $\delta > 0$.

EXAMPLE 6.2. *Uniform convergence of nonparametric estimates.* Assume that f is a uniformly continuous function on $[0, 1]$, and that f is estimated by

$$f_n(x) = \frac{\sum_{i=1}^n f(X_i) K\left(\frac{X_i - x}{\ell_n}\right)}{\sum_{i=1}^n K\left(\frac{X_i - x}{\ell_n}\right)}$$

where X_1, \dots, X_n are independent identically distributed uniform $[0, 1]$ random variables, and $K(u)$ is a nonincreasing nonnegative function of u when $u > 0$, and a nondecreasing nonnegative function of u when $u < 0$. Let the support of K be a compact set $[a, b]$ (clearly, $a \leq 0 \leq b$) with $a < b$.

It is clear that $\sup_x |f_n(x) - f(x)| \rightarrow 0$ a.s. for all uniformly continuous f if and only if $M_n > (b - a)\ell_n$ f.o. almost surely. Now, if we take $n(b - a)\ell_n = \log n + (2 + \delta)\log n$, then

$$\sup_x |f_n(x) - f(x)| \rightarrow 0 \text{ a.s. as } n \rightarrow \infty$$

for all uniformly continuous f if $\delta > 0$; the statement is false if $\delta < 0$.

EXAMPLE 6.3. *Estimating the minimum of a density.* Let f be a uniformly continuous density on $[0, 1]$, and let z be the unique point with the property that $f(z) = \min_x f(x)$. Assume that X_1, X_2, \dots is an independent sample from f , and that z is estimated by Z_n , the midpoint of the largest interval created by X_1, \dots, X_n . From $nM_n/\log n \rightarrow 1$ a.s. for uniform distributions, one can show that $Z_n \rightarrow z$ a.s. as $n \rightarrow \infty$. For the study of laws of the iterated logarithm of M_n in the non-uniform case, additional assumptions about the rate of increase of f near z seem necessary. Notice also that if the maximum of f were estimated by the midpoint of the smallest interval, then one would *not* obtain almost sure convergence as in the case of Z_n .

EXAMPLE 6.4. *Rate of convergence of nearest neighbor estimates.* Let f and X_1, X_2, \dots be as in Example 6.2, but consider now the nearest neighbor estimate $f_n(x) = f(X_n^N(x))$ where $X_n^N(x)$ is the nearest neighbor to x among X_1, \dots, X_n . If f is Lipschitz with constant C , then $\sup_x |f_n(x) - f(x)| \leq \max(CM_{n+1}/2; CX_{(1)}; C(1 - X_{(n)}))$ where $X_{(1)} < \dots < X_{(n)}$ are the order statistics obtained from X_1, \dots, X_n . From the properties of $X_{(1)}$ and M_n (Theorem 4.1) we have the following rate of convergence result:

$$\sup_x |f_n(x) - f(x)| (2n/C \log n) > 1 + a_n \quad \text{f.o. a.s.}$$

when $a_n \log n \rightarrow \infty$, $(1 + a_n)\log n/n$ is ultimately nonincreasing and $\sum_{n=1}^{\infty} \log n/n^{1+a_n} < \infty$. On the other hand, if $f(x) = Cx$, then the supremum is equal to the maximum of the three given terms, so that we may conclude, by Theorem 5.1, that there exists a Lipschitz function with constant C such that

$$\sup_x |f_n(x) - f(x)| (2n/(C \log n)) > 1 + (2 - \delta)\log 2n/\log n \quad \text{i.o. a.s. for all } \delta > 0.$$

In other words, in the class $\text{Lip}(C)$, we have

$$(6.1) \quad \limsup((2n/C)\sup_x |f_n(x) - f(x)| - \log n)/\log 2n \leq 2 \quad \text{a.s.}$$

but there always exists an f in $\text{Lip}(C)$ for which (6.1) is valid with equality.

REFERENCES

- [1] BARNDORFF-NIELSEN, O. (1961). On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables. *Math. Scand.* **9** 383-394.
- [2] BARNDORFF-NIELSEN, O. (1963). On the limit behaviour of extreme order statistics. *Ann. Math. Statist.* **34** 992-1002.
- [3] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493-507.
- [4] DARLING, D. A. (1952). On a test for homogeneity and extreme values. *Ann. Math. Statist.* **23** 450-456.
- [5] DARLING, D. A. (1953). On a class of problems related to the random division of an interval. *Ann. Math. Statist.* **24** 239-253.
- [6] DAVID, H. A. (1970). *Order Statistics*. Wiley, New York.
- [7] DEHEUVELS, P. (1974). Majoration et minoration presque sûre optimale des éléments de la statistique ordonnée d'un échantillon croissant de variables aléatoires indépendantes. *Rendiconti della Accademia Nazionale dei Lincei, Series 8* **61** 707-719.
- [8] DEVROYE, L. (1979). Asymptotical behavior of the maximal uniform spacing. Technical Report SOCS 79-8, McGill University.
- [9] FELLER, W. (1957). *An Introduction to Probability Theory And Its Applications I*. Wiley, New York.
- [10] FRANKEL, J. (1972). On the law of the iterated logarithm for order statistics. Ph.D. dissertation, Columbia University.
- [11] GALAMBOS, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley, New York.
- [12] GEFFROY, J. (1958). Contribution à la théorie des valeurs extrêmes. *Publ. Inst. Statist. Univ. Paris* **7** 37-121.
- [13] GNEDENKO, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44** 423-453.
- [14] GUMBEL, E. J. (1958). *Statistics of Extremes*. Columbia Univ. Press, New York.
- [15] KENDALL, M. G. and MORAN, P. A. P. (1963). *Geometric Probability*. Griffin, London.
- [16] KIEFER, J. (1970). Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 227-244. Univ. of California Press.
- [17] LEVY, P. (1939). Sur la division d'un segment par des points choisis au hasard. *Comptes Rendus Acad. Sci. Paris* **208** 147-149.
- [18] PYKE, R. (1965). Spacings. *J. Roy. Statist. Soc.* **27** 395-436.
- [19] ROBBINS, H. and SIEGMUND, D. (1970). On the law of the iterated logarithm for maxima and minima. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 51-70. Univ. of California Press.
- [20] SHEPP, L. A. (1972). Covering the circle with random arcs. *Israel J. Math.* **11** 328-345.
- [21] SLUD, E. (1978). Entropy and maximal spacings for random partitions. *Z. Wahrsch. Verw. Gebiete* **41** 341-352.
- [22] SUKHATME, P. V. (1937). Tests of significance for samples of the chi-square population with two degrees of freedom. *Ann. Eugen. London* **8** 52-56.
- [23] WHITWORTH, W. A. (1897). *Choice and Chance*. Cambridge Univ. Press.
- [24] WICHURA, M. (1973). Boundary crossing probabilities associated with Motoo's law of the iterated logarithm. *Ann. Probability* **1** 437-456.

SCHOOL OF COMPUTER SCIENCE
 MCGILL UNIVERSITY
 805 SHERBROOKE STREET WEST
 MONTREAL, CANADA H3A 2K6