The Annals of Probability
1987, Vol. 15, No. 1, 274–280

MAXIMAL SPACINGS IN SEVERAL DIMENSIONS[1]

BY SVANTE JANSON

Uppsala University

Take n points at random in a fixed set in R^d. Define the maximal spacing, e.g., as the volume of the largest ball that is contained in the fixed set and avoids all n chosen points. The asymptotic distribution of the maximal spacing and strong bounds are given.

1. Introduction. Let n points be independently and uniformly distributed on a circle of unit length [or on $(0,1)$]. The spacings, i.e., the successive distances between these points, have been widely studied; see e.g., the review papers by Pyke [10], [11]. We denote the largest spacing by Δ_n. The exact distribution of Δ_n was first obtained by Stevens [12]. The asymptotic distribution as $n \to \infty$ was given by Lévy [9]; see also Holst [7] for a stronger theorem and further references. The result can be stated as follows:

$$n\Delta_n - \log n \to_d U \quad \text{as } n \to \infty, \tag{1.1}$$

where U has the extreme value distribution

$$P(U \le u) = e^{-e^{-u}}. \tag{1.2}$$

By adding the points one by one, we obtain a nonincreasing sequence $\{\Delta_n\}_1^\infty$ of random variables. Devroye [5] proved the following strong bounds.

$$\liminf_{n\to\infty} (n\Delta_n - \log n)/\log\log n = 0 \quad \text{a.s.}, \tag{1.3}$$

$$\limsup_{n\to\infty} (n\Delta_n - \log n)/\log\log n = 2 \quad \text{a.s.} \tag{1.4}$$

More refined results are given by Devroye [6] and Deheuvels [1], and extensions by Deheuvels [2] and Deheuvels and Devroye [4]. Note the asymmetry here and in the theorem below: $n\Delta_n$ have larger positive deviations from $\log n$ than negative ones. Similarly, the tails of the distribution of U are of different sizes.

The purpose of the present paper is to prove generalizations of the above results for higher dimensions. Deheuvels [3] defined (for points uniformly distributed in the unit cube in R^d) the maximal spacing as the size of the largest cubical gap (parallel to the unit cube). See below for a precise formulation. (He also treated kth largest spacings, but we will only be concerned with the maximal one.) Deheuvels' results include generalizations of (1.3) and (1.4) to this situation, but without exact values of lim inf and lim sup. The present paper will

Received February 1985; revised January 1986.
[1] Research partially supported by NFR.
AMS 1980 *subject classifications*. Primary 60005; secondary 60F05, 60F15.
Key words and phrases. Spacings, multi-dimensional spacings, random coverings.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability*. www.jstor.org

give these values (conjectured in [3]), and extend the result to, e.g., spherical gaps.

The proof is based on estimates derived in [8] and the equivalence between spacings and covering problems.

2. Definitions and results. Let K be a bounded set in R^d, $d \geq 1$, such that $|K| = 1$ and $|\partial K| = 0$ (where $|\cdot|$ denotes d-dimensional Lebesgue measure) and let $X_1, X_2, \ldots$ be a sequence of independent and uniformly distributed points in K. Let A be a fixed bounded convex set in R^d (with nonempty interior) and define the maximal spacings by

$$\Delta_n = \sup\{r: \exists x \text{ with } x + rA \subset K \setminus \{X_i\}_1^n\}. \tag{2.1}$$

The two main cases are A a cube (as in [3]), or a sphere.

We will formulate the results in terms of V_n defined by

$$V_n = |\Delta_n A|, \tag{2.2}$$

i.e., the volume of the maximal gap of the shape (and orientation) of A. We may without loss of generality assume that $|A| = 1$ and thus

$$V_n = \Delta_n^d. \tag{2.3}$$

REMARK. The definition involves gaps of a fixed (although rather arbitrary) shape and orientation. Other conceivable definitions such as the volume of the largest cube of any orientation in $K \setminus \{X_i\}_1^n$ (or the largest rectangular box with sides parallel to the coordinated axes, the largest convex set, etc.) are not covered by this paper.

The results do not depend on the shape of K, but the shape of A enters through the constant α defined in the following enigmatic way (see [8], Sections 5 and 9 for further details). Let ω denote the surface measure on ∂A (i.e., ω is the $d-1$-dimensional Hausdorff measure), and let, for $y \in \partial A$, $n(y)$ denote the exterior unit normal to A at y. The assumption that A is convex implies, without any further regularity assumptions, that $0 < \omega(\partial A) < \infty$ and that $n(y)$ is uniquely defined a.e. (ω). Define for $v \in R^d$ (assuming $|A| = 1$),

$$\alpha(v) = \frac{1}{d!}\int \cdots \int \left|\operatorname{Det}(n(y_i))_{i=1}^d\right| d\omega(y_1) \cdots d\omega(y_d), \tag{2.4}$$

where we integrate over all $y_1, \ldots, y_d \in \partial A$ such that v is a linear combination of $n(y_1), \ldots, n(y_d)$ with positive coefficients, and $\operatorname{Det}(n(y_i))$ is the determinant of the vectors $n(y_i)$ in an orthonormal basis. Then $\alpha(v)$ is a constant a.e. in v, and we denote this constant by α (see [8], Corollary 7.4).

THEOREM. *With notation as above*

$$nV_n - \log n - (d-1)\log\log n - \log \alpha \to U, \tag{2.5}$$

where U has the distribution given by (1.2),

$$\liminf_{n\to\infty} \frac{nV_n - \log n}{\log\log n} = d - 1 \quad a.s., \tag{2.6}$$

$$\limsup_{n\to\infty} \frac{nV_n - \log n}{\log\log n} = d + 1 \quad a.s. \tag{2.7}$$

REMARK. More information on α is given in [8], Section 9. In particular, it is shown there that:

$$\text{If } A \text{ is a cube, } \alpha = 1. \tag{2.8}$$

$$\text{If } A \text{ is a sphere, } \alpha = \frac{1}{d!}\left(\frac{\sqrt{\pi}\,\Gamma\left(\frac{d}{2}+1\right)}{\Gamma\left(\frac{d+1}{2}\right)}\right)^{d-1}. \tag{2.9}$$

If $d \geq 3$, α(sphere) $<$ α(cube) (and thus the spherical spacings tend to be somewhat smaller than the cubical ones), but for $d = 2$ there is equality. In fact, if $d = 2$, $\alpha = 1$ for every centrosymmetric set [i.e., such that $A - x = -(A - x)$ for some x].

REMARK. The theorem remains of course true if K is the torus T^d. Using the argument of [8], Section 8, we obtain the same result for spherical spacings on a sphere, and more generally for geodesic balls on any compact C^2 Riemannian manifold [with α given by (2.9)].

3. Proofs. It will be technically convenient to replace the fixed number n of points by a stochastic number. Hence, let $\{N_t\}_{t\geq 0}$ be a Poisson process with intensity one (independent of $X_1, X_2, \dots$) and put

$$\Delta(t) = \Delta_{N_t} \quad \text{and} \quad V(t) = V_{N_t} = \Delta(t)^d.$$

A routine verification, using the facts that V_n is nonincreasing, $N_t \sim \mathrm{Po}(t)$, $N_t/t \to 1$, a.s., and $(N_t/t - 1)\log t \to 0$ a.s., shows that the theorem is equivalent to

$$tV(t) - \log t - (d-1)\log\log t - \log\alpha \to U, \tag{3.1}$$

$$\liminf_{n\to\infty} \frac{tV(t) - \log t}{\log\log t} = d - 1 \quad \text{a.s.}, \tag{3.2}$$

$$\limsup_{n\to\infty} \frac{tV(t) - \log t}{\log\log t} = d + 1 \quad \text{a.s.} \tag{3.3}$$

We may without loss of generality assume that A is an open convex set. Then

$$\begin{aligned} \Delta_n \geq r &\Leftrightarrow \exists x \text{ with } x + rA \subset K \setminus \{X_i\}_1^n \\ &\Leftrightarrow \exists x \text{ with } x + rA \subset K \text{ and } x \notin \bigcup_1^n (X_i - rA). \end{aligned} \tag{3.4}$$

Thus, putting $K_r = \{x: x + rA \subset K\}$, $\Delta_n \geq r$ iff K_r not is covered by the sets $X_i - rA$, $i = 1, \dots, n$. Consequently, $\Delta(t) < r$ iff K_r is covered by the sets $X_i - rA$, $i = 1, \dots, N_t$.

Now, the random set $\{X_i\}_1^{N_t}$ may be regarded as a Poisson process with intensity t in K. Since $x - rA$ does not meet K_r unless $x \in K$, it makes no difference if we extend this Poisson process to a Poisson process with the same intensity in a larger set, or in the entire space R^d.

We may now apply the results of [8]. First, we need to introduce some further notation. Let $\mathscr{F}_s$ denote the mesh of cubes $\{\prod_1^d [n_i s, (n_i + 1)s]: (n_1, \dots, n_d) \in Z^d\}$, and let $n_s = \#\{Q \in \mathscr{F}_s: Q \subset K_r\}$ and $m_s = \#\{Q \in \mathscr{F}_s: Q \cap \partial K_r \neq \varnothing\}$, and let

$$\gamma = \gamma(r, t) = t^d |rA|^{d-1} e^{-t|rA|} = t^d r^{d(d-1)} e^{-tr^d}. \tag{3.5}$$

LEMMA 1. *Let $D = 3\sup\{|x|: x \in A\}$. Then, there exist $\alpha_+ = \alpha_+(tr^d)$ and $\alpha_- = \alpha_-(tr^d)$ such that $\alpha_+ \to \alpha$ and $\alpha_- \to \alpha$ as $tr^d \to \infty$ and such that, for all $s > \delta = rD$,*

$$e^{-\gamma\alpha_+(s+\delta)^d(n_s+m_s)} \leq P(\Delta(t) < r) = P(K_r \textit{ is covered}) \leq e^{-\gamma\alpha_-(s-\delta)^d n_s}. \tag{3.6}$$

PROOF. (3.6) follows by [8], Lemma 7.2, by replacing A and K by $-rA$ and K_r, respectively. Here (see [8], Sections 5 and 7) $\alpha_- = \alpha_-(-rA, v, t, rD) = \alpha_-(A, -v, tr^d, D)$ and $\alpha_+ = \alpha_+(-rA, v, t, rD) = \alpha_+(A, -v, tr^d, D)$, where v is some conveniently chosen fixed vector. This proves that α_+ and α_- depend only on tr^d. The fact that $\alpha_+ \to \alpha$ and $\alpha_- \to \alpha$ follows from the proof of [8], Lemma 7.3. □

LEMMA 2. *Choose $s = \sqrt{r}$. Then, as $r \to 0$, we have $m_s s^d \to 0$, $n_s s^d \to 1$, and $\delta/s \to 0$.*

PROOF. Let $\partial_a K$ denote the set of points whose distance to ∂K is at most a. It is easily seen that $\partial K_r \subset \partial_{Dr} K$. Hence, if $Q \in \mathscr{F}_s$ and $Q \cap \partial K_r \neq \varnothing$, then $Q \subset \partial_{Dr+ds} K$. It follows that $m_s s^d \leq |\partial_{Dr+ds} K|$. Likewise $|K| - |\partial_{Dr+ds} K| \leq |K - \partial_{Dr+ds} K| \leq n_s s^d \leq |K|$. Let us now choose $s = r^{1/2}$. We obtain that $|\partial_{Dr+ds} K| \to |\partial K| = 0$, as $r \to 0$. Hence $m_s s^d \to 0$ and $n_s s^d \to |K| = 1$. Finally, $\delta s = Dr/s \to 0$. □

LEMMA 3. *There exist $a_- = a_-(r, t)$ and $a_+ = a_+(r, t)$ such that $a_- \to \alpha$ and $a_+ \to \alpha$ as $r \to 0$ and $tr^d \to \infty$, and such that*

$$e^{-\gamma a_+} \leq P(\Delta(t) < r) \leq e^{-\gamma a_-}. \tag{3.7}$$

PROOF. A direct consequence of Lemmas 1 and 2. □

We substitute $w = tr^d$ in Lemma 3 and obtain, because $\Delta(t) < r \Leftrightarrow tV(t) < w$, the following reformulation.

LEMMA 4. *There exist* $a_- = a_-(w, t)$ *and* $a_+ = a_+(w, t)$ *such that* $a_- \to \alpha$ *and* $a_+ \to \alpha$ *as* $w \to \infty$ *and* $w/t \to 0$, *and such that, with* $\gamma = tw^{d-1}e^{-w}$,

$$e^{-\gamma a_+} \le P(tV(t) < w) \le e^{-\gamma a_-}. \tag{3.8}$$

Taking $w = \log t + (d-1)\log\log t + \log \alpha + u$ we obtain by Lemma 4, as $t \to \infty$ with u fixed, $\gamma \to \alpha^{-1}e^{-u}$ and $P(tV(t) < w) \to \exp(-e^{-u})$, which proves (3.1) and thus (2.5).

We turn to the strong bounds. (3.1) yields

$$(tV(t) - \log t)/\log\log t \to_P d - 1 \quad \text{as } t \to \infty,$$

whence lim inf $\le d - 1 \le$ lim sup a.s.

We will complete the proof of (3.2) and (3.3) by proving the following three inequalities, cf. [3].

LEMMA 5. $\liminf_{t\to\infty}(tV(t) - \log t)/\log\log t \ge d - 1$ *a.s.*

LEMMA 6. $\limsup_{t\to\infty}(tV(t) - \log t)/\log\log t \le d + 1$ *a.s.*

LEMMA 7. $\limsup_{t\to\infty}(tV(t) - \log t)/\log\log t \ge \liminf_{t\to\infty}(tV(t) - \log t)/\log\log t + 2$ *a.s.*

PROOF OF LEMMA 5. Choose any $c < d - 1$ and define

$$t_k = \exp(\sqrt{k}), \tag{3.9}$$

$$w_k = \log t_k + c \log\log t_k, \tag{3.10}$$

$$\gamma_k = t_k w_k^{d-1} e^{-w_k} \sim (\log t_k)^{d-1-c} = k^{(d-1-c)/2}. \tag{3.11}$$

Lemma 4 shows that for some $a > 0$ and k large enough

$$P(t_k V(t_k) < w_k) \le \exp(-ak^{(d-1-c)/2}). \tag{3.12}$$

The sum over k of the right-hand sides converges and by the Borel–Cantelli lemma

$$P(t_k V(t_k) < w_k \text{ i.o.}) = 0. \tag{3.13}$$

Now, suppose that $t_k V(t_k) \ge w_k$ and $t_{k-1} \le t \le t_k$. Then, since

$$t_{k-1}/t_k = \exp(\sqrt{k-1} - \sqrt{k}) > 1 - (\sqrt{k} - \sqrt{k-1}) > 1 - 1/2\sqrt{k-1},$$

$$\begin{aligned} tV(t) - \log t &\ge t_{k-1}V(t_k) - \log t_k \ge (t_{k-1}/t_k)w_k - \log t_k \\ &\ge c\log\log t_k - \log t_k/\sqrt{k} \ge c \log\log t - 1, \end{aligned}$$

if k is large enough. Hence a.s. $(tV(t) - \log t)/\log\log t \ge c - 1/\log\log t$, for all sufficiently large t, which proves Lemma 5. □

PROOF OF LEMMA 6. This is similar. Choose any $c > d + 1$ and let t_k, w_k, γ_k be defined by (3.9), (3.10), (3.11). By (3.8) and (3.11), for some $C > \infty$,

$$P(t_k V(t_k) \geq w_k) \leq 1 - \exp(-\gamma_k a_+) \leq \gamma_k a_+ \leq C k^{(d-1-c)/2}. \tag{3.14}$$

Since $c > d + 1$, the right-hand side is summable and it follows as above that $\limsup(tV(t) - \log t)/\log\log t \leq c$ a.s. □

PROOF OF LEMMA 7. Let $c < \liminf(tV(t) - \log t)/\log\log t$ and let $1 < b < 2$. We change the definition of t_k to

$$t_k = \exp(k^{1/b}) \tag{3.15}$$

and let w_k be defined by (3.10) as before. We define r_k by $w_k = t_k r_k^d$ and

$$t_k' = t_k(1 + b \log\log t_k/\log t_k) = t_k(1 + k^{-1/b}\log k). \tag{3.16}$$

We note that

$$t_{k+1}/t_k = \exp\left((k+1)^{1/b} - k^{1/b}\right) > 1 + b^{-1}(k+1)^{1/b-1} > t_k'/t_k, \tag{3.17}$$

if k is large enough, whence $t_{k+1} > t_k' > t_k$.

Let Ξ_t denote the random set $\{X_i\}_1^{N_t}$ and note that the increment $\Xi_{t_k'} \setminus \Xi_{t_k}$ is independent of Ξ_t, $t < t_k$. Recall that, by definition, if $\Delta(t_k) \geq r_k$, then there exist points $x \in K_{r_k}$ such that $x + r_k A \subset K \setminus \Xi_{t_k}$. Let Y_k be one of these points (e.g., the first in the lexicographic ordering) and let Y_k be any point in K_{r_k} if $\Delta(t_k) < r_k$. (We may have to ignore a few k for which $K_{r_k} = \varnothing$.) Thus Y_k is a random point in K_{r_k}, depending only on Ξ_{t_k}, such that $Y_k + r_k A \subset K \setminus \Xi_{t_k} \Leftrightarrow \Delta(t_k) \geq r_k$.

Let M_k be the number of points in $(\Xi_{t_k'} \setminus \Xi_{t_k}) \cap (Y_k + r_k A)$. Thus, M_k is Poisson distributed with parameter

$$\begin{aligned}(t_k' - t_k)|Y_k + r_k A| &= (t_k' - t_k) r_k^d = (t_k'/t_k - 1) w_k \\ &= k^{-1/b}\log k\left(k^{1/b} + cb^{-1}\log k\right) < \log k + 1,\end{aligned}$$

if k is large enough. Hence

$$P(M_k = 0) > \exp(-(\log k + 1)) = e^{-1}/k \tag{3.18}$$

and

$$\sum_k P(M_k = 0) = \infty. \tag{3.19}$$

Since the distribution of M_k is independent of Y_k, it follows that M_k is independent of Ξ_{t_k} and, since $t_k' < t_{k+1}$, that the variables M_k are independent (possibly ignoring the first few k). By the Borel–Cantelli lemma and (3.19),

$$P(M_k = 0 \text{ i.o.}) = 1. \tag{3.20}$$

By the choice of $c, \Delta(t_k) \geq r_k$ for all but a finite number of k a.s., and thus

$$P(M_k = 0 \text{ and } \Delta(t_k) \geq r_k \text{ i.o.}) = 1. \tag{3.21}$$

However, if $\Delta(t_k) > r_k$ and $M_k = 0$, $Y_k + r_k A \subset K \setminus \Xi_{t_k'}$. Hence, for such k,

$\Delta(t_k') \geq r_k$ and, provided k is large enough,

$$
\begin{aligned}
t_k' V(t_k') &\geq (1 + b \log\log t_k / \log t_k) w_k \\
&= \log t_k + c \log\log t_k + b \log\log t_k + bc(\log\log t_k)^2 / \log t_k \qquad (3.22)\\
&\geq \log t_k' + (b + c) \log\log t_k' - 1.
\end{aligned}
$$

Consequently, a.s., $\limsup (tV(t) - \log t)/\log\log t \geq b + c$, which proves Lemma 7. □

REMARK. The same method can be used to show that

$$-1 \leq \liminf (nV_n - \log n - (d-1)\log\log n)/\log\log\log n \leq 0 \quad \text{a.s.}$$

Devroye [5], [6] has shown that this lim inf equals -1 when $d = 1$. It seems more difficult to use our method to estimate

$$\limsup (nV_n - \log n - (d+1)\log\log n)/\log\log\log n.$$

Deheuvels [1], [3] has shown that this lim sup equals $+1$ when $d = 1$ and that it is ≤ 1 for $d \geq 2$ in the case of a cubical gap.

REFERENCES

[1] DEHEUVELS, P. (1982). Strong limiting bounds for maximal uniform spacings. *Ann. Probab.* **10** 1058–1065.

[2] DEHEUVELS, P. (1983). Upper bounds for k-th maximal spacings. *Z. Wahrsch. verw. Gebiete* **62** 465–474.

[3] DEHEUVELS, P. (1983). Strong bounds for multidimensional spacings. *Z. Wahrsch. verw. Gebiete* **64** 411–424.

[4] DEHEUVELS, P. and DEVROYE, L. (1984). Strong laws for the maximal k-spacing when $k \leq c \log n$. *Z. Wahrsch. verw. Gebiete* **66** 315–334.

[5] DEVROYE, L. (1981). Laws of the iterated logarithm for order statistics of uniform spacings. *Ann. Probab.* **9** 860–867.

[6] DEVROYE, L. (1982). A log log law for maximal uniform spacings. *Ann. Probab.* **10** 863–868.

[7] HOLST, L. (1981). On convergence of the coverage by random arcs on a circle and the largest spacing. *Ann. Probab.* **9** 648–655.

[8] JANSON, S. (1986). Random coverings in several dimensions. *Acta Math.* **156** 83–118.

[9] LÉVY, P. (1939). Sur la division d'un segment par des points choisis au hasard. *C. R. Acad. Sci. Paris* **208** 147–149.

[10] PYKE, R. (1965). Spacings (with discussion). *J. Roy. Statist. Soc. Ser. B* **27** 395–449.

[11] PYKE, R. (1972). Spacings revisited. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1** 417–427. Univ. California Press.

[12] STEVENS, W. L. (1939). Solution to a geometrical problem in probability. *Ann. Eugenics* **9** 315–320.

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
THUNBERGSVÄGEN 3
S-752 38 UPPSALA
SWEDEN