SIAM J. APPL. MATH.
Vol. 46, No. 1, February 1986

010

EXTREME VALUE DISTRIBUTION FOR THE LARGEST CUBE IN A RANDOM LATTICE*

R. W. R. DARLING† AND MICHAEL S. WATERMAN†‡

Abstract. Suppose that the sites of a finite d-dimensional lattice ($d \geqq 2$) of side n are occupied by independent, identically distributed random variables with value 0 or 1. The length of the side of the largest cube of 1's is found to have (approximately) an integerized extreme value distribution. The distribution becomes increasingly concentrated on three consecutive integers, as n increases. Applications to clustering are discussed.

Introduction. Long success runs in coin tossing have long been studied in probability theory. The applications range from gambling to quality control and pattern recognition. In this paper we consider higher dimensional analogues of "long success runs". That is, we derive the asymptotic probability distribution of the largest cube of 1's in a d-dimensional random lattice of 0's and 1's. We also derive the asymptotic distribution along a subsequence when up to b zeros are allowed in the cube of 1's. These distributions are fundamental for pattern recognition and have application to a number of areas such as vision (Glatz [8]), uranium prospecting (Conover et al. [3]), ecology, radar astronomy (Schwager [13]) and cosmology (Abell [1]). Our result verifies a conjecture of Diaconis and Freedman [19, p. 120].

The organization of the paper is as follows. Section 1 contains the statement of the theorem and two corollaries. In § 2 we explain precisely how to use the theorem to test for existence of clustering in spatial data. Sections 3, 4 and 5 are concerned with the technicalities of the proof, which uses only elementary combinatorial methods, without recourse to any theorems of probability. In § 6 we prove that the distribution of the side of the largest cube of 1's (except for at most b zeros) becomes increasingly concentrated on three consecutive integers (depending on n), as the lattice size n goes to infinity. Section 7 presents the results of a Monte Carlo simulation, showing that the asymptotic probability distribution is a good approximation to the actual distribution even for a 30×30 lattice.

The almost sure behaviour of the length of the longest head run in one dimension has been studied by Erdos and Renyi [5], Naus [11], and Erdos and Revesz [6]. The probability distribution of the length of the longest head run was obtained only recently by Gordon, Schilling and Waterman [9], using methods of extreme value theory. Almost sure behaviour of the area of the largest square and the largest rectangle of 1's in an $n \times n$ lattice of 0's and 1's was studied by Nemetz and Kusolitsch [12].

The methods of the present paper were inspired by those of Watson [14], who was studying extreme values of a stationary stochastic process. It may be appropriate to embed the result of the present paper into the theory of maxima for discrete random fields. For example in dimension 2, for each site $\mathbf{i}$ of the $n \times n$ lattice, let $U(\mathbf{i})$ denote the length of side of the largest square of 1's whose lowest vertex is $\mathbf{i}$. The subject of our present study is $\max\{U(\mathbf{i})\}$ over all lattice sites $\mathbf{i}$. For a detailed survey of extreme value theory for stochastic processes, see Leadbetter et al. [10].

1. Description of the results. Let $(X(\mathbf{i}): \mathbf{i} = (i_1, \cdots, i_d) \in \mathbf{Z}^d)$ be a collection of independent, identically distributed Bernoulli random variables, with

$$P(X(\mathbf{i})=1)=p=1-P(X(\mathbf{i})=0), \quad \text{each } \mathbf{i}.$$

* Received by the editors July 3, 1984, and in revised form April 25, 1985.
† Department of Mathematics, University of Southern California, Los Angeles, California 90089-1113.
‡ This author's work was supported by the System Development Foundation.

For $a = 1, 2, \cdots$ and each index $\mathbf{i}$, define

$$C_a(\mathbf{i}) = \{\mathbf{j} = (j_1, \cdots, j_d) \in \mathbf{Z}^d\colon i_m \leqq j_m \leqq i_m + a - 1, \text{ all } m\}.$$

Think of $C_a(\mathbf{i})$ as a cube of side a in the integer lattice $\mathbf{Z}^d$, with lowest vertex $\mathbf{i}$. Let $\mathbf{1}$ denote the lowest vertex of all, namely $(1, 1, \cdots, 1)$ in $\mathbf{Z}^d$. For each nonnegative integer b, let $Z_b(n)$ denote the length of the side of the largest cube of sites in $C_n(\mathbf{1})$ consisting entirely of 1's, except for at most b zeros. More formally,

$$Z_b(n) = \max_{\mathbf{i} \in C_n(\mathbf{1})} \{a\colon X(\mathbf{j}) = 1 \text{ for all but at most } b \text{ sites } \mathbf{j} \text{ in } C_a(\mathbf{i})\}.$$

For example, when $b = 0$ and $d = 2$, $Z_0(n)$ is the length of the side of the largest square consisting entirely of 1's inside the set of sites $\{(i_1, i_2)\colon 1 \leqq i_1 \leqq n,\ 1 \leqq i_2 \leqq n\}$. For the following example, $d = 2$, $n = 5$, $Z_0(5) = 2$, and $Z_2(5) = 3$.

$$\begin{matrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{matrix}$$

THEOREM. *Suppose $d \geqq 2$, and fix an integer $b \geqq 0$ and a real number t. For each positive integer a, define*

$$n = n(a) = [e^{(u-t)/d}] + a - 1, \tag{1.1}$$

where

$$u = -\ln\left(\sum_{j=0}^{b} \binom{a^d}{j} (1-p)^j p^{a^d - j}\right) \qquad (= a^d \ln(1/p) \text{ if } b = 0), \tag{1.2}$$

and

$$[x] = \text{greatest integer less than or equal to } x.$$

The asymptotic distribution of the side of the largest cube of 1*'s, except for at most b zeros, along the subsequence $(n(a),\ a = 1, 2, \cdots)$ is*

$$\lim_{a \to \infty} P(Z_b(n(a)) \leqq a - 1) = \exp(-e^{-t}) \tag{1.3}$$

and the convergence is uniform in t.

COROLLARY 1 (rectangular lattice, dimension 2). *Take $d = 2$, and for integers $m \geqq 1$, $n \geqq 1$, and $b \geqq 0$, define*

$$Z_b(m, n) = \max_{\substack{1 \leqq i_1 \leqq m \\ 1 \leqq i_2 \leqq n}} \{a\colon X(\mathbf{j}) = 1 \textit{ for all but at most } b \textit{ sites } \mathbf{j} \textit{ in } C_a(i_1, i_2)\}.$$

Fix real numbers t and γ, with $\gamma > 1$. For each positive integer a, define

$$\begin{aligned} n(a) &= \left[\left(\gamma^{-1} e^{u-t} + \left(\frac{\gamma - 1}{2\gamma}\right)^2 (a-1)^2\right)^{1/2} + \left(\frac{\gamma+1}{2\gamma}\right)(a-1)\right], \\ m(a) &= [n(a)\gamma] \end{aligned} \tag{1.4}$$

where u is the number defined in (1.2) *and $[\cdot]$ is the greatest integer function, as before. Then*

$$\lim_{a \to \infty} P(Z_b(m(a), n(a)) \leqq a - 1) = \exp(-e^{-t})$$

and the convergence is uniform in t.

Remark. The derivation of formula (1.4) depends on the equations

$$(m-a+1)(n-a+1)e^{-u}=e^{-t}+\text{truncation error},$$
$$m=n\gamma+\text{truncation error}.$$

In practical applications, the integers m and n will be given. Assume $m>n$, and take $\gamma=m/n$. The number u is known, so select a pair (a, t) such that (1.4) holds. Then $P(Z(m, n)\leqq a-1)$ is approximately $\exp(-e^{-t})$.

No separate proof of Corollary 1 will be given, since the proof of the Theorem can easily be modified to deal with this case.

COROLLARY 2. *Fix $b\geqq 0$, and for each positive integer n, define*

$$m(n)=[(d\log_{1/p}n)^{1/d}].$$

For each $\varepsilon>0$, there exists an integer $K(\varepsilon)$ such that for all $n\geqq K(\varepsilon)$,

$$P(m(n)-1\leqq Z_b(n)\leqq m(n)+1)>1-\varepsilon.$$

Remarks. (1) Notice that the three integers on which $Z_b(n)$ becomes concentrated do not depend on b; this is because the b zeros make a negligible contribution to a cube with volume $(Z_b(n))^d$ as n tends to infinity.

(2) It would be desirable to have an estimate of $K(\varepsilon)$. Unfortunately $K(\varepsilon)$ is related to the rate of convergence in equation (1.3), whose calculation seems intractable at the present time.

(3) The proof uses (1.3) to estimate the probability that $Z_b(n)$ lies in the given range. An even sharper result in the case $d=2$, $b=0$ and $p=0.5$ has been obtained recently by P. Revesz [17], who is able to restrict to two integers. (Revesz's result was not known to us when we wrote this article.) The method in [17] is entirely different.

2. Application to clustering. The previous theorem can be used to construct a test for clustering. For the sake of concreteness, we shall give an example from astronomy. The book of Diggle [15] gives an excellent survey of methods for treating this class of problems.

Earlier in this century Shane and Wirtanen produced a series of photographic plates, designed to show all the galaxies above a certain magnitude (brightness). A typical plate might show 500 galaxies. More recent studies using red shifts make it possible to add a third dimension to the location of galaxies in the sky. Many astronomers have analyzed such data to discover evidence of superclustering and other structure in the positions of clusters of galaxies (see Peebles [16]).

Abell [1] divided the photographic plate into equal rectangular regions, and counted the number of clusters in each region. He then performed a chi-squared goodness-of-fit test to compare the counts with the best-fitting Poisson distribution in the plane. However this method does not, strictly speaking, test for the existence of "super-clusters", but only for goodness of fit to the Poisson distribution.

We now propose an alternate way to analyze data of this kind. Suppose K clusters appear on the plate. Choose n to be the greatest integer less than or equal to $\sqrt{2K}$. Define

$$p=1-\left(1-\frac{1}{n^2}\right)^K$$

which is approximately $\frac{1}{2}$. Put an $n\times n$ grid of equally spaced lines over the plate, dividing the plate into n^2 equal rectangular sites. For each site (i,j), $i, j\leqq n$, define

$$X(i,j)=\begin{cases}0 & \text{if no galaxy appears in site } (i,j),\\ 1 & \text{if 1 or more galaxies appear in site } (i,j).\end{cases}$$

Let H_0 be the null hypothesis that these K clusters have positions which are distributed like K independent, identically distributed random variables, each with the uniform distribution over the area of the plate. Then under H_0,

$$P(X(i,j)=0) = P(0 \text{ out of } K \text{ uniform random variables takes values in rectangle } (i,j))$$

$$= \left(\frac{n^2-1}{n^2}\right)^K = \left(1-\frac{1}{n^2}\right)^K = 1-p$$

which implies

$$P(X(i,j)=1)=p.$$

For various integers a (for example $2, 3, 4, 5, \cdots$) calculate corresponding values of

$$e^{-t} = p^{a^2}(n-a+1)^2.$$

(This is obtained from (1.1).) The theorem states that under H_0, $P(Z_0(n) \leqq a-1)$ is approximately $\exp(-e^{-t})$. Select the probability α of a Type I error, say $\alpha = 10^{-3}$, and find the smallest integer a_1 so that

$$\exp(-e^{-t}) \equiv \exp(-p^{a_1^2}(n-a_1+1)^2) > 1-\alpha.$$

Then under H_0,

$$P(Z_0(n) \geqq a_1) < \alpha,$$

neglecting the errors involved in the limiting operation described in the theorem. Finally, we examine the $(X(i,j)\colon 1 \leqq i, j \leqq n)$ derived from the photographic plate, and obtain a value for $Z_0(n)$, the side of the biggest square of 1's in the $n \times n$ array of 0's and 1's. If $Z_0(n)$ is greater than or equal to a_1, we reject H_0 with approximate significance level α.

The procedure could easily be modified to allow for the processing of m separate plates: simply perform a goodness of fit test against the predicted distribution of $Z_0(n)$. To detect large-scale clustering, it may be more appropriate to use $Z_b(n)$, for some $b>0$. The same test can easily be performed in dimension $d>2$. If the researcher were interested in the existence of large "gaps" containing no galaxies, then he could count the side of the largest square of zeros instead.

An even more natural context for such a method is when a grid of lattice sites is defined a priori, and the result of the experiment is to label each site either occupied (1) or unoccupied (0). In that case we would estimate p by the proportion of occupied sites.

3. Technical preliminaries—intersection numbers. Consider two squares in $\mathbb{Z}^d$ of side a, one of which has lowest vertex $(1,1)$, while the other has lowest vertex $(3,5)$. The *intersection number* for these two squares simply means the number of sites in common, which in this case will be $(a-2)(a-4)$. We shall formalize this notion in d dimensions. For $\mathbf{i}$ in $\mathbb{Z}^d$ and $a \geqq 1$, recall that

$$C_a(\mathbf{i}) = \{\mathbf{j} \in \mathbb{Z}^d : i_m \leqq j_m \leqq i_m + a - 1, m = 1, 2, \cdots, d\}.$$

Thus $C_a(\mathbf{i})$ is a cube of side a, with lowest vertex $\mathbf{i}$. For pairs of sites $\mathbf{i}$ and $\mathbf{k}$ in $\mathbb{Z}^d$, define

$$N_a(\mathbf{i}, \mathbf{k}) = |C_a(\mathbf{i}) \cap C_a(\mathbf{k})|,$$

$$I_a(\mathbf{i}) = \{\mathbf{k} \neq \mathbf{i} : N_a(\mathbf{i}, \mathbf{k}) > 0\}.$$

We refer to $N_a(\mathbf{i},\mathbf{k})$ as the *intersection number* of cubes $C_a(\mathbf{i})$ and $C_a(\mathbf{k})$, and $I_a(\mathbf{i})$ may be called *the set of sites which a-interact with* $\mathbf{i}$. In the sequel we shall usually abbreviate $N_a(\mathbf{i},\mathbf{k})$ to $N(\mathbf{i},\mathbf{k})$ and $I_a(\mathbf{i})$ to $I(\mathbf{i})$, whenever a is fixed. This section is devoted to proving the following algebraic inequality:

PROPOSITION 1. *Let* $0<p<1$, *and let* $a\geqq 1$ *be fixed. Then*

$$\sum_{\mathbf{k}\in I(\mathbf{i})} p^{-N(\mathbf{i},\mathbf{k})}\leqq c_1(d)h^d p^{-a^d+a^{d-1}} \tag{3.1}$$

where $h=(1-p)^{-1}$ *and* $c_1(d)=2^d+d$.

Let us begin with some algebraic preliminaries. For any integer $a\geqq 1$ and a real variable w, define a sequence of polynomials $(J_m(a,w),\ m=0,1,2,\cdots)$ according to the following recursive scheme:

$$\begin{aligned} &J_0(a,w)=w,\\ &J_m(a,w)=\sum_{j=1}^{a-1} J_{m-1}(a,w^j)+\sum_{j=1}^{a} J_{m-1}(a,w^j), \qquad m=1,2,\cdots. \end{aligned} \tag{3.2}$$

For integers $1\leqq r\leqq d$, define

$$A(d,r)=\{\mathbf{i}=(i_1,\cdots,i_d)\in\mathbb{Z}^d: 1\leqq i_m\leqq a-1 \text{ if } 1\leqq m\leqq r, \text{ and } 1\leqq i_m\leqq a \text{ if } (r+1)\leqq m\leqq d\}.$$

In other words, $A(d,r)$ is the product of r copies of $\{1,2,\cdots,a-1\}$ and $(d-r)$ copies of $\{1,2,\cdots,a\}$.

A routine induction argument proves the following:

LEMMA 1.

$$J_d(a,w)=\sum_{r=0}^{d}\binom{d}{r}\sum_{(i_1,\cdots,i_d)\in A(d,r)} w^{i_1\cdots i_d} \quad \textit{for } d\geqq 1. \tag{3.3}$$

The connection between the polynomials $J_d(a,w)$ and intersection numbers is as follows:

LEMMA 2. *The following identity holds for all* w *and all* $a\geqq 1; d\geqq 1$:

$$\sum_{\mathbf{k}\in I(\mathbf{i})} w^{N(\mathbf{i},\mathbf{k})}=J_d(a,w)-w^{(a^d)}. \tag{3.4}$$

Proof. Observe that for all $\mathbf{k}$ in $I_a(\mathbf{i})=I(\mathbf{i})$,

$$N(\mathbf{i},\mathbf{k})=(a-|i_1-k_1|)(a-|i_2-k_2|)\cdots(a-|i_d-k_d|).$$

Consequently

$$\sum_{\mathbf{k}\in I(\mathbf{i})} w^{N(\mathbf{i},\mathbf{k})}=\sum_{j_1=-a+1}^{a-1}\cdots\sum_{j_d=-a+1}^{a-1} w^{(a-|j_1|)\cdots(a-|j_d|)}-w^{a^d}. \tag{3.5}$$

(The reason for subtracting w^{a^d} from the sum on the right is that the site $\mathbf{i}$ itself is not a member of $I(\mathbf{i})$.) When $d=1$, this expression becomes

$$\sum_{k=1}^{a-1} w^k+\sum_{k=1}^{a} w^k-w^a$$

which is precisely $J_1(a,w)-w^a$. Therefore (3.4) holds when $d=1$. Suppose (3.4) holds when $d = m$. When $d = (m+1)$, the left side of (3.4) may be calculated by (3.5); it equals

$$\sum_{k=-a+1}^{a-1}\left(\sum_{j_1=-a+1}^{a-1}\cdots\sum_{j_m=-a+1}^{a-1} w^{a-|j_1|)\cdots(a-|j_m|)(a-|k|)}\right)-w^{a^{m+1}}$$

$$=\sum_{k=-a+1}^{a-1} J_m(a, w^{a-|k|})-w^{a^{m+1}}$$

by the assumption that (3.4) holds when $d = m$. This equals

$$=\sum_{j=1}^{a-1} J_m(a, w^j)+\sum_{j=1}^{a} J_m(a, w^j)-w^{a^{m+1}},$$

$$=J_{m+1}(a, w)-w^{a^{m+1}}$$

by the recursive definition (3.2) of $J_{m+1}(a, w)$. This completes the induction and the lemma is proved.

Proof of Proposition 1. Let $w=1/p>1$, and let $h=1/(1-p)$. Then for $m=1,2,3,\cdots$

$$\frac{w^m}{w^m-1}=(1-1/w^m)^{-1}\leqq(1-1/w)^{-1}=(1-p)^{-1}=h$$

and so

$$\sum_{j=1}^{a} w^{mj}=w^m\left(\frac{w^{am}-1}{w^m-1}\right)\leqq hw^{am}. \tag{3.6}$$

Combining Lemmas 1 and 2 gives

$$\sum_{\mathbf{k}\in J(\mathbf{i})} w^{N(\mathbf{i},\mathbf{k})}=\sum_{r=0}^{d}\binom{d}{r}\sum_{\mathbf{i}\in A(d,r)} w^{i_1\cdots i_d}-w^{a^d}$$

$$=\sum_{r=1}^{d}\binom{d}{r}\sum_{\mathbf{i}\in A(d,r)} w^{i_1\cdots i_d}+\sum_{i_1=1}^{a}\cdots\sum_{i_d=1}^{a} w^{i_1\cdots i_d}-w^{a^d}.$$

An elementary but lengthy calculation shows that this is

$$=\sum_{r=1}^{d}\binom{d}{r}\sum_{\mathbf{i}\in A(d,r)} w^{i_1\cdots i_d}+\sum_{s=1}^{d}\left(\sum_{i_s=1}^{a-1}\sum_{i_{s+1}=1}^{a}\cdots\sum_{i_d=1}^{a} w^{a^{(s-1)}i_s\cdots i_d}\right).$$

Applying (3.6) repeatedly to this expression, we deduce that it is

$$\leqq\sum_{r=1}^{d}\binom{d}{r}h^d w^{(a-1)^r a^{d-r}}+w^{(a-1)a^{d-1}}\sum_{s=1}^{d} h^{d-s+1}$$

$$<h^d w^{(a-1)a^{d-1}}\left\{\sum_{r=1}^{d}\binom{d}{r}+\sum_{s=1}^{d} h^{-s+1}\right\}<(2^d+d)h^d w^{(a-1)a^{d-1}}$$

which completes the proof.

4. Intersecting cubes containing zeros. Let b be a fixed nonnegative integer. For each integer $a\geqq 1$, and each lattice site $\mathbf{i}$, let $S_a(\mathbf{i})$ denote the event that the cube of side a, with lowest vertex $\mathbf{i}$, consists entirely of ones except for at most b zeros. More formally,

$$S_a(\mathbf{i})=\{X(\mathbf{j})=1 \text{ for all but at most } b \text{ sites } \mathbf{j} \text{ in } C_a(\mathbf{i})\}.$$

When a is fixed, we may abbreviate $S_a(\mathbf{i})$ to $S(\mathbf{i})$. The purpose of this section is to prove the following:

PROPOSITION 2. *For each integer $a \geqq 1$,*

$$\sum_{\mathbf{k}\in I_a(\mathbf{i})} P(S_a(\mathbf{i}) \cap S_a(\mathbf{k})) \leqq c_2(d,p)\, e^{-u} a^{bd} p^{a^d-1} \tag{4.1}$$

where, as in (1.2)

$$e^{-u} = \sum_{j=0}^{b} \binom{a^d}{j} (1-p)^j p^{a^d-j},$$

and $c_2(d,p)$ is a constant, given by

$$c_2(d,p) = (2^d + d) h^d e \max\left(1, \left(\frac{1-p}{p}\right)^b\right), \qquad h = (1-p)^{-1}.$$

Proof. Suppose $\mathbf{k}$ belongs to $I_a(\mathbf{i})$, so that the intersection of the cubes $C_a(\mathbf{i})$ and $C_a(\mathbf{k})$ is nonempty. Define random variables Z_1, Z_2 and Z_3 as follows:

$$Z_1 = \text{number of zeros in } C_a(\mathbf{i}) \cap C_a(\mathbf{k}),$$

$$Z_2 = \text{number of zeros in } C_a(\mathbf{i}) \setminus C_a(\mathbf{k}),$$

$$Z_3 = \text{number of zeros in } C_a(\mathbf{k}) \setminus C_a(\mathbf{i}).$$

Then

$$P(S_a(i) \cap S_a(k)) = \sum_{s=0}^{b} \sum_{q=0}^{b-s} \sum_{r=0}^{b-s} P(Z_1 = s, Z_2 = q, Z_3 = r).$$

Since Z_1, Z_2 and Z_3 are independent, this equals

$$\begin{aligned}
&= \sum_{s=0}^{b} \sum_{q=0}^{b-s} P(Z_1 = s, Z_2 = q) \sum_{r=0}^{b-s} P(Z_3 = r) \\
&= \sum_{s=0}^{b} P(Z_1 + Z_2 \leqq b, Z_1 = s) \sum_{r=0}^{b-s} P(Z_3 = r) \\
&\leqq \max_{0 \leqq r \leqq b} P(Z_3 \leqq b - r) \sum_{s=0}^{b} P(Z_1 + Z_2 \leqq b, Z_1 = s) \\
&\leqq P(Z_3 \leqq b) P(Z_1 + Z_2 \leqq b).
\end{aligned} \tag{4.2}$$

Notice that $\binom{n}{k} \leqq n^k/k!$. Hence for $b \leqq n$, we have

$$\sum_{k=0}^{b} \binom{n}{k} \leqq \sum_{k=0}^{b} \frac{n^k}{k!} \leqq n^b \sum_{k=0}^{b} \frac{1}{k!} < n^b e. \tag{4.3}$$

Define

$$L = \max\left(1, \left(\frac{1-p}{p}\right)^b\right). \tag{4.4}$$

Now

$$P(Z_3 \leqq b) = \sum_{r=0}^{b \wedge (a^d - N)} \binom{a^d - N}{r} p^{a^d - N - r} (1-p)^r$$

where $N=N(\mathbf{i},\mathbf{k})$; notice that (a^d-N) is the number of sites in $C_a(\mathbf{k})\backslash C_a(\mathbf{i})$, and $b\wedge c$ denotes min (b,c). By (4.3) and (4.4)

$$P(Z_3\leqq b)p^{a^d-N}(a^d-N)^b Le<eLa^{bd}p^{a^d-N}.$$

On the other hand,

$$P(Z_1+Z_2\leqq b)=P(S_a(\mathbf{i}))=\sum_{j=0}^{b}\binom{a^d}{j}(1-p)^j p^{a^d-j}=e^{-u}$$

where u is the constant defined in (1.2).

Therefore from (4.2)

$$P(S_a(\mathbf{i})\cap S_a(\mathbf{k}))<Le^{1-u}a^{bd}p^{a^d-N(\mathbf{i},\mathbf{k})}.$$

Now apply Proposition 1, inequality (3.1):

$$\sum_{\mathbf{k}\in I_a(\mathbf{i})}P(S_a(\mathbf{i})\cap S_a(\mathbf{k}))<Le^{1-u}c_1(d)h^d a^{bd}p^{a^{d-1}}$$

as desired.

5. Proof of the theorem.

Step I. As in the previous section, define the event $S_a(\mathbf{i})$ by

$$S_a(\mathbf{i})=\{X(\mathbf{j})=1 \text{ for all but at most } b \text{ sites } \mathbf{j} \text{ in } C_a(\mathbf{i})\}$$

where the integer b is fixed throughout. Then for any integer $a\geqq 1$;

$$P(Z_b(n)\geqq a)=P\left(\bigcup_{\mathbf{i}\in C_n(\mathbf{1})}S_a(\mathbf{i})\right). \tag{5.1}$$

By the inclusion-exclusion formula (see Feller [7, p. 99]), for each even integer r,

$$Q_r\leqq P\left(\bigcup_{\mathbf{i}}S_a(\mathbf{i})\right)\leqq Q_{r+1} \tag{5.2}$$

where Q_r is an alternating sum of r terms as follows:

$$\begin{aligned}Q_r=&\sum_{\mathbf{i}}P(S(\mathbf{i}))-\sum_{\mathbf{i}}\sum_{\mathbf{j}\neq\mathbf{i}}P(S(\mathbf{i})\cap S(\mathbf{j}))+\cdots\\&+(-1)^{r+1}\sum_{\substack{\mathbf{i}(1)\\ \text{distinct}}}\cdots\sum_{\mathbf{i}(r)}P(S(\mathbf{i}(1))\cap\cdots\cap S(\mathbf{i}(r))).\end{aligned} \tag{5.3}$$

Recall the definition (1.2) of u and the relationship (1.1) of $n(a)$ and a. The first term on the right side of (5.3) is equal to

$$\sum_{\mathbf{i}\in C_n(\mathbf{1})}P(S(\mathbf{i}))=(n-a+1)^d e^{-u}\equiv\lambda(a)\equiv\lambda. \tag{5.4}$$

For all values of a and n(a), λ is approximately e^{-t}.

Step II. We proceed to estimate the general term in (5.3) Let us introduce the following notation, for $r=2,3,4,\cdots$:

$V_r=\{$collections of r distinct lattice sites in $C_{n-a+1}(\mathbf{1})\}$,
$A_r=\{\{\mathbf{i}(1),\cdots,\mathbf{i}(r)\}$ in V_r such that $N_a(\mathbf{i}(j),\mathbf{i}(k))=0$ for $j\neq k$ (see § 3)$\}$,
$B_r=V_r-A_r$.

Let us fix an integer $q \geqq 2$. Then by independence

$$\begin{aligned}
&\sum_{\{\mathbf{i}(1), \cdots, \mathbf{i}(q)\} \text{ in } V_q} P(S(\mathbf{i}(1)) \cap \cdots \cap S(\mathbf{i}(q))) \\
&= \sum_{A_q} P(S(\mathbf{i}(1))) \cdots P(s(\mathbf{i}(q))) + \sum_{B_q} P(S(\mathbf{i}(1)) \cap \cdots \cap S(\mathbf{i}(q))) \\
&= ((n-a+1)^d e^{-u})^q / q! + R(q,1) + R(q,2) \\
&= \frac{\lambda^q}{q!} + R(q,1) + R(q,2)
\end{aligned}$$

where

$$R(q,1) = e^{-qu}(|A_q| - (n-a+1)^{qd}/q!), \tag{5.5}$$

$$R(q,2) = \sum_{B_q} P(S(\mathbf{i}(1)) \cap \cdots \cap S(\mathbf{i}(q))). \tag{5.6}$$

Now

$$|A_q| \geqq (1/q!)(n-a+1)^d((n-a+1)^d - (2a-1)^d) \cdots ((n-a+1)^d - (q-1)(2a-1)^d).$$

It is easily verified by induction that for $0 < \delta < 1$ and $m = 1, 2, \cdots, [1/\delta]$, $(1-\delta)(1-2\delta) \cdots (1-m\delta) \geqq 1 - m^2\delta$. Thus

$$|A_q| \geqq (1/q!)(n-a+1)^{qd}(1-(q-1)^2[(2a-1)/(n-a+1)]^d).$$

Hence

$$0 > R(q,1) \geqq -(e^{-qu}/q!)(n-a+1)^{qd}(q-1)^2[(2a-1)/(n-a+1)]^d,$$
$$0 > R(q,1) \geqq -(\lambda^q/q!)(q-1)^2[(2a-1)/(n+1)]^d.$$

From (1.1) and (1.2)

$$e^{-t}(n-a+1)^{-d} \leqq e^{-u} = \sum_{j=0}^{b} \binom{a^d}{j} (1-p)^j p^{a^d-j}.$$

Using (4.3) and (4.4), we see that

$$e^{-u} \leqq p^{a^d} \max(1, ((1-p)/p)^b) \sum_{j=0}^{b} \binom{a^d}{j} < Lea^{db}p^{a^d}. \tag{5.7}$$

Therefore

$$0 > R(q,1) \geqq -(\lambda^q/q!)(q-1)^2 Le^{t+1} p^{a^d}(2a-1)^d a^{db}. \tag{5.8}$$

Step III. We now proceed to estimate the second remainder term $R(q,2)$ defined in (5.6). First we introduce some new terminology. Any collection of r sites $\{\mathbf{i}(1), \cdots, \mathbf{i}(r)\}$ in $\mathbb{Z}^d$ may be considered as the vertices of a graph G, where distinct sites $\mathbf{i}(j)$ and $\mathbf{i}(k)$ are joined by an edge whenever $N_a(\mathbf{i}(j), \mathbf{i}(k)) > 0$. If G is connected, we shall refer to $\{\mathbf{i}(1), \cdots, \mathbf{i}(r)\}$ as a *cluster of size r*.

Using this construction, each collection of sites $\{\mathbf{i}(1), \cdots, \mathbf{i}(r)\}$ in B_r has a unique decomposition into m_1 clusters of size 1, m_2 clusters of size 2, and so on, where the integers $m_1, m_2, \cdots$ satisfy:

$$m_j \geqq 0, \quad m_1 \leqq r-2, \quad m_1 + 2m_2 + 3m_3 + \cdots + rm_r = r. \tag{5.9}$$

We shall call the vector $\mathbf{m} = (m_1, \cdots, m_r)$ the *vector of cluster-numbers* for $\{\mathbf{i}(1), \cdots, \mathbf{i}(r)\}$. Let $M(r)$ denote the set of all vectors $\mathbf{m} = (m_1, \cdots, m_r)$ in $\mathbb{Z}^r$ satisfying

(5.9). An asymptotic formula for $|M(r)|$ may be found in Hall [18, p. 40]; in any case

$$|M(r)| \leqq (r+1)^r/r!. \tag{5.10}$$

For each $\mathbf{m}$ in $M(r)$, let $H(r, \mathbf{m})$ denote the set of collections of sites $\{\mathbf{i}(1), \cdots, \mathbf{i}(r)\}$ in B_r whose vector of cluster-numbers is $\mathbf{m}$; thus

$$B_r = \bigcup_{\mathbf{m} \in M(r)} H(r, \mathbf{m}). \tag{5.11}$$

Of special interest is $H(r, (0, \cdots, 0, 1)) = G(r)$, the set of collections $\{\mathbf{i}(1), \cdots, \mathbf{i}(r)\}$ which form a single cluster of size r. Our immediate goal is to estimate

$$g(a, r) = \sum_{G(r)} P(S(\mathbf{i}(1)) \cap \cdots \cap S(\mathbf{i}(r))).$$

By (5.1), $g(a, 1) = (n-a+1)^d e^{-u} = \lambda$. By (4.1),

$$\begin{aligned} g(a, 2) &= \sum_{\mathbf{j}} \sum_{\mathbf{k} \in I_a(\mathbf{j})} P(S(\mathbf{j}) \cap S(\mathbf{k}))/2! \\ &\leqq (n-a+1)^d c_2(d, p)\, e^{-u} a^{bd} p^{a^{d-1}}/2! = \lambda c_2(d, p) a^{bd} p^{a^{d-1}}/2!. \end{aligned}$$

Estimate $g(a, 3)$ as follows: choose $\mathbf{i}(1)$ from $(n-a+1)^d$ possible sites, choose $\mathbf{i}(2)$ in $I_a(\mathbf{i}(1))$, and $\mathbf{i}(3)$ in $I_a(\mathbf{i}(1)) \cup I_a(\mathbf{i}(2))$ (at most $2(2a-1)^d$ choices). Then by inclusion $P(S(\mathbf{i}(1)) \cap S(\mathbf{i}(2)) \cap S(\mathbf{i}(3))) \leqq P(S(\mathbf{i}(1)) \cap S(\mathbf{i}(2)))$, and

$$g(a, 3) \leqq \left\{ \sum_{\mathbf{i}(1)} \sum_{\mathbf{i}(2) \in I_a(\mathbf{i}(1))} P(S(\mathbf{i}(1)) \cap S(\mathbf{i}(2))) \right\} 2(2a-1)^d/3!$$

where we divide by $3!$ because permuting the labels gives the same element of $G(3)$. For $r \geqq 4$, we use the same trick, restricting $\mathbf{i}(k+1)$ to lie in $\bigcup_{1 \leqq j \leqq k} I_a(\mathbf{i}(j))$ for $k = 2, \cdots, r-1$; we obtain

$$g(a, r) \leqq \{(n-a+1)^d c_2(d, p)\, e^{-u} a^{bd} p^{a^{d-1}}\}(r-1)!(2a-1)^{d(r-2)}/r!,$$

which gives

$$g(a, r) \leqq (c_2(d, p) \lambda 2^{d(r-2)}/r) a^{d(r+b-2)} p^{a^{d-1}}, \qquad r \geqq 2. \tag{5.12}$$

Next we wish to estimate

$$h(a, \mathbf{m}) = \sum_{H(q, \mathbf{m})} P(S(\mathbf{i}(1)) \cap \cdots \cap S(\mathbf{i}(q)))$$

for an arbitrary $\mathbf{m}$ in $M(1)$, for $q \geqq 2$. Suppose first that the element $\{\mathbf{i}(1), \cdots, \mathbf{i}(q)\}$ of $H(q, \mathbf{m})$ is labelled so that members of the same cluster appear in consecutive order, with the clusters in increasing order of size. An obvious property of independence between clusters shows that, for example, if $q = 5$ and $\mathbf{m} = (1, 2, 0, 0, 0)$, then

$$P(S(\mathbf{i}(1)) \cap \cdots \cap S(\mathbf{i}(5))) = P(S(\mathbf{i}(1)))P(S(\mathbf{i}(2)) \cap S(\mathbf{i}(3)))P(S(\mathbf{i}(4)) \cap S(\mathbf{i}(5))).$$

The obvious generalization shows that

$$h(a, \mathbf{m}) = \prod_{1 \leqq r \leqq q} g(a, r)^{m_r}/m_r!$$

where the factors $m_r!$ appear because switching clusters of the same size yields the same element of $H(q, \mathbf{m})$. Hence by (5.12)

$$h(a, \mathbf{m}) \leqq \lambda^{m_1} \prod_{2 \leqq r \leqq q} (c_2(d, p) \lambda 2^{dr} 2^{-2d} a^{dr} a^{d(b-2)} p^{a^{d-1}}/r)^{m_r}/m_r!.$$

Let $v=\sum_{j=2}^{q} jm_j$ and $z=\sum_{j=2}^{q} m_j$; then

$$h(a,\mathbf{m}) \leqq \lambda^{m_1}(c_2(d,p)\lambda 2^{-2d} a^{d(b-2)} p^{a^{d-1}})^z (2^d a^d)^v \Big/ \prod_{2\leqq r\leqq q} r^{m_r} m_r!. \tag{5.13}$$

However (5.10) and (5.11) imply

$$R(q,2)=\sum_{\mathbf{m}\in M(q)} h(a,\mathbf{m}) \leqq |M(q)| \max\{h(a,\mathbf{m}) : \mathbf{m}\in M(q)\}.$$

The inequality for $h(a,\mathbf{m})$ shows that as $a\to\infty$,

$$h(a,\mathbf{m})=O(a^{d(v+z(b-2))} p^{za^{d-1}}).$$

Since $v\leqq q$ and $1\leqq z\leqq q/2$,

$$\max\{h(a,\mathbf{m}) : \mathbf{m}\in M(q)\}=O(a^{dq(1+0.5(b-2))} p^{a^{d-1}}).$$

Combining all this with (5.8) gives

$$\begin{aligned}&\sum_{V_q} P(S(i(1))\cap\cdots\cap S(i(q)))=\lambda^q/q!+R(q,1)+R(q,2),\\ &R(q,1)=O(a^{d(b+1)}p^{a^d}), \qquad R(q,2)=O(a^{dq(1+0.5(b-2))}p^{a^{d-1}}).\end{aligned} \tag{5.14}$$

(The constants in the $O(\cdot)$ estimates are easily deduced from the above.) Notice, by the way, that $R(q,2)$ goes to zero as $a\to\infty$ only when $d\geqq 2$.

Step IV. Let us now return to (5.3) in Step I. Equation (5.14) shows that for a fixed even integer r, and for $m=r, r+1$,

$$Q_m=\lambda-\frac{\lambda^2}{2!}+\cdots+(-1)^{m+1}\frac{\lambda^m}{m!}+O(a^{dm\beta}\, e^{-a^{d-1}\ln(1/p)}),$$

where

$$\lambda=(n-a+1)^d e^{-u}, \qquad \beta=1+0.5(b-2).$$

From (5.1) and (5.2),

$$1-Q_{r+1}\leqq P(Z_b(n)\leqq a-1)\leqq 1-Q_r,$$

which may be written

$$\begin{aligned}&1-\lambda+\frac{\lambda^2}{2!}-\cdots-\frac{\lambda^{r+1}}{(r+1)!}+O(f(a))\\ &\leqq P(Z_b(n)\leqq a-1)\leqq 1-\lambda+\frac{\lambda^2}{2!}-\cdots+\frac{\lambda^r}{r!}+O(f(a))\end{aligned} \tag{5.15}$$

where

$$f(a)=a^{d(r+1)\beta}\, e^{-a^{d-1}\ln(1/p)}. \tag{5.16}$$

We now show how to replace λ by e^{-t} in (5.15). First, (1.1) gives

$$(n-a+1)^d\leqq e^{u-t}<(n-a+2)^d,$$

so that

$$(n-a+1)^d e^{-u}\equiv\lambda\leqq e^{-t}<\lambda\left(1+\frac{1}{n-a+1}\right)^d. \tag{5.17}$$

However

$$\frac{1}{n-a+2} < e^{(t-u)/d} \leqq e^{t/d}(Lea^{db}p^{a^d})^{1/d}$$

by (5.7). This proves that if $\delta = \delta(a) = (n-a+1)^{-1}$, then

$$\delta(a) = O(a^b p^{(a^d/d)}). \tag{5.18}$$

Now apply the inequality

$$(1+\delta)^m \leqq 1 + 2^{m-1} m\delta \quad \text{for } 0 \leqq \delta < 1, \quad m = 2, 3, \cdots.$$

This together with (5.17) shows that

$$\frac{\lambda^k}{k!} \leqq \frac{e^{-kt}}{k!} < \frac{\lambda^k}{k!}(1+\delta)^{dk} \leqq \frac{\lambda^k}{k!}(1+2^{dk-1}\, dk\delta).$$

Thus

$$\left| \sum_{k=0}^{m} \frac{(-\lambda)^k}{k!} - \sum_{k=0}^{m} \frac{(-e^{-t})^k}{k!} \right|$$

$$\leqq \sum_{k=0}^{m} \left| \frac{\lambda^k}{k!} - \frac{e^{-kt}}{k!} \right| \leqq \sum_{k=1}^{m} \frac{\lambda^k}{k!} 2^{dk-1}\, dk\,\delta < (\lambda\, d 2^{d-1}\, e^{\lambda 2^d})\delta = c_4(\lambda, d)\delta.$$

Hence (5.15) can be rewritten as

$$\sum_{k=0}^{r+1} \frac{(-e^{-t})^k}{k!} + O(a^b p^{a^d/d}) + O(f(a))$$

$$\leqq P(Z_b(n) \leqq a-1) \leqq \sum_{k=0}^{r} \frac{(-e^{-t})^k}{k!} + O(a^b p^{a^d/d}) + O(f(a)).$$

Keeping r fixed, let a tend to infinity. Then

$$\sum_{k=0}^{r+1} \frac{(-e^{-t})^k}{k!} \leqq \lim_{a\to\infty} P(Z_b(n(a)) \leqq a-1) \leqq \sum_{k=0}^{r} \frac{(-e^{-t})^k}{k!}.$$

Finally let r tend to infinity, giving

$$\lim_{a\to\infty} P(Z_b(n(a)) \leqq a-1) = \exp(-e^{-t})$$

as desired.

6. Length of the side becomes concentrated on three integers. Let $m(n)$ be the integer function of n defined in (6.1) below. It was proved by Nemetz and Kusolitsch [12] that when $d=2$, $Z_0(n)^2/2\log_{1/p} n$ converges almost surely to 1. The theorem above allows us to prove a much stronger result, namely that for any $d \geqq 2$ and any $b \geqq 0$, as n increases $Z_b(n)$ becomes concentrated on the three integers $\{m(n)-1, m(n), m(n+1)\}$. This is reminiscent of the tightness result in one dimension proved by Gordon, Schilling and Waterman [9], and also of a general property of extreme value distributions proved by Anderson [2]. For the reader's convenience, we restate:

COROLLARY 2. *Fix $b \geqq 0$, and for each positive integer n, define*

$$m(n) = [(d \log_{1/p} n)^{1/d}]. \tag{6.1}$$

For each $\varepsilon > 0$, there exists an integer $K(\varepsilon)$ such that for all $n \geqq K(\varepsilon)$,

$$P(m(n)-1 \leqq Z_b(n) \leqq m(n)+1) > 1-\varepsilon. \tag{6.2}$$

Proof. Abbreviate $m(n)$ to m. It follows from (6.1) that

$$m^d \leqq \log_{1/p} n^d < (m+1)^d, \qquad p^{-m^d} \leqq n^d < p^{-(m+1)^d},$$

and

$$p^{(m+1)^d} < \frac{1}{n^d} \leqq p^{m^d}. \tag{6.3}$$

Define real numbers $u = u(n)$ and $v = v(n)$ by the formulas

$$u = -\ln\left\{\sum_{j=0}^{b}\binom{(m-1)^d}{j}(1-p)^j p^{(m-1)^d-j}\right\}, \tag{6.4}$$

$$v = -\ln\left\{\sum_{j=0}^{b}\binom{(m+2)^d}{j}(1-p)^j p^{(m+2)^d-j}\right\}. \tag{6.5}$$

Define another pair of real numbers $s = s(n)$ and $t = t(n)$ by the formulas

$$e^{-s} = e^{-u}(n-m+2)^d, \tag{6.6}$$

$$e^{-t} = e^{-v}(n-m-1)^d. \tag{6.7}$$

Then (6.3) shows that

$$\begin{aligned} e^{-s} &= p^{(m-1)^d}(n-m+2)^d \sum_{j=0}^{b}\binom{(m-1)^d}{j}\left(\frac{1-p}{p}\right)^j \\ &\geqq p^{(m-1)^d-m^d}\left(1-\frac{m-2}{n}\right)^d b \min\left(1, \left(\frac{1-p}{p}\right)^b\right) \\ &\geqq (1/p)^{d(m-1)^{d-1}}\left(1-\frac{m-2}{n}\right)^d c_5(p,b). \end{aligned}$$

Since $0<p<1$, and since m and $(m-2)/n$ tend to infinity and to zero respectively as n tends to infinity, by (6.1), it follows that

$$\lim_{n\to\infty} e^{-s(n)} = \infty. \tag{6.8}$$

Likewise by (6.3) and (4.3),

$$\begin{aligned} e^{-t} &= p^{(m+2)^d}(n-m-1)^d \sum_{j=0}^{b}\binom{(m+2)^d}{j}\left(\frac{1-p}{p}\right)^j \\ &< p^{(m+2)^d-(m+1)^d}\left(1-\frac{m+1}{n}\right)^d e(m+2)^{db} \max\left(1, \left(\frac{1-p}{p}\right)^b\right) \\ &< p^{d(m+1)^{d-1}}(m+2)^{db}\left(1-\frac{m+1}{n}\right)^d c_6(p,b). \end{aligned}$$

Since $0<p<1$, and since m and $(m+1)/n$ tend to infinity and to zero respectively as n tends to infinity, it follows that

$$\lim_{n\to\infty} e^{-t(n)} = 0. \tag{6.9}$$

Given $\varepsilon > 0$, (6.8) and (6.9) imply that there exists $L(\varepsilon)$ such that for $n \geqq L(\varepsilon)$,

$$\max(\exp(-e^{-s(n)}), 1-\exp(-e^{-t(n)})) < \varepsilon/4. \tag{6.10}$$

We now proceed to derive formula (6.2). First, write

$$\begin{aligned} P\Big(m(n)-1 \leqq Z_b(n) \leqq m(n)+1\Big) &= P\Big(Z_b(n) \leqq (m(n)+2)-1\Big) \\ &\quad - P\Big(Z_b(n) \leqq (m(n)-1)-1\Big). \end{aligned} \tag{6.11}$$

For the sake of comparison with definitions (1.1) and (1.2) in the theorem of § 1, we shall express n as if it were a function first of $(m-1)$, then of $(m+2)$; using (6.6) and (6.7), we may write

$$n = [e^{(u-s)/d}] + (m-1) - 1, \tag{6.12}$$

$$n = [e^{(v-t)/d}] + (m+2) - 1. \tag{6.13}$$

Apply the Theorem first with $(m-1)$ in place of a, then with $(m+2)$ in place of a: there exists an integer $M(\varepsilon)$ such that for $n \geqq M(\varepsilon)$,

$$P(Z_b(n) \leqq m(n)-2) < \exp(-e^{-s(n)}) + \varepsilon/4$$

and

$$P(Z_b(n) \leqq m(n)+1) > \exp(-e^{-t(n)}) - \varepsilon/4.$$

Equations (6.10) and (6.11) show that when $n \geqq K(\varepsilon) = \max(L(\varepsilon), M(\varepsilon))$, we have

$$\begin{aligned} P(m(n)-1 \leqq Z_b(n) \leqq m(n)+1) &> \exp(-e^{-t(n)}) - \exp(-e^{-s(n)}) - \varepsilon/2 \\ &> 1 - \varepsilon/4 - \varepsilon/4 - \varepsilon/2 = 1 - \varepsilon, \end{aligned}$$

as desired.

7. Using simulations to estimate the speed of convergence. The main theorem of this paper gives a limit for the probability that the largest cube of 1's, except for at most b zeros, in a d-dimensional lattice of side $n(a)$, has side less than a, as a tends to infinity. How far is the actual probability $P(Z_b(n(a)) \leqq a-1)$ from the limit for moderate values of a, such as $a=4$, $n=50$?

A Monte Carlo simulation was performed for the case $p=0.5$, $d=2$, $b=0$. For $n=30$ and $n=50$, pairs a and t were chosen to satisfy (1.1) and (1.2) and the relative frequency of the event $\{Z_0(n) \leqq a-1\}$ was compared with the theoretical limit $\exp(-e^{-t})$. To find the size of the largest square of 1's in each simulated lattice, we used the algorithm described in Darling and Waterman [4]; this algorithm operates in $O(n^2)$ time.

The results are presented in the table below. The main conclusion from the simulation is that the theorem provides a good estimate for $P(Z_0(n) \leqq a-1)$ when a is greater than $[(d \log_{1/p} n)^{1/d}]$, even for moderate values of n such as $n=30$. (The approximation improves as a increases, because the remainder terms diminish when $\lambda \equiv (n-a+1)^d p^{a^d}$ becomes small.) This is encouraging from the point of view of applications, because researchers will usually be interested in estimating probabilities of "large" squares consisting entirely, or almost entirely, of 1's.

TABLE

	Theoretical frequency/Monte Carlo frequency	
Largest square	$n=30$, Sample size $=10^5$	$n=50$, Sample size $=10^4$
2	0.216/0.308	0.011/0.029
3	0.733/0.682	0.956/0.939
4	0.011/0.010	0.033/0.031

Explanation. Each site in an $n\times n$ lattice takes values 0 or 1, each with probability 0.5, and all sites are independent. The "largest square" means the side of the largest square of sites consisting entirely of 1's. The "theoretical frequency" of a square of side a is simply $e^{-\lambda(a+1)}-e^{-\lambda(a)}$, where $\lambda(a)=(n-a+1)^2 2^{-a^2}$. The "Monte Carlo frequency" is the proportion of pseudorandom $n\times n$ lattices for which the largest square was as shown, out of the sample of size 10^4 or 10^5.

Acknowledgments. The authors are grateful to Louis Gordon and Richard Arratia for useful discussions. Astronomer Gibson Reaves provided much helpful information about the data on clustering of galaxies. Mark Eggert is responsible for the computer program which performed the Monte Carlo simulation. We thank the referee for several improvements.

REFERENCES

[1] G. O. ABELL, *The distribution of rich clusters of galaxies*, Astrophysical J. Supplement, III (1958), pp. 211-288.
[2] C. W. ANDERSON, *Extreme value theory for a class of discrete distributions with applications to some stochastic processes*, J. Appl. Prob., 7 (1970), pp. 99-113.
[3] W. J. CONOVER, T. R. BEMENT AND R. L. IMAN, *On a method for detecting clusters of possible uranium deposits*, Technometrics, 21 (1979), pp. 277-282.
[4] R. W. R. DARLING AND M. S. WATERMAN, *Matching rectangles in d dimensions: algorithms and laws of large numbers*, Advances in Math., 55 (1985), pp. 1-12.
[5] P. ERDOS AND A. RENYI, *On a new law of large numbers*, J. Analyse. Math., 22 (1970), pp. 103-111.
[6] P. ERDOS AND P. REVESZ, *On the length of the longest head-run*, Topics in Information Theory, in Collo. Math. Soc. J. Bolyai 16, Keszthely (Hungary), 1975, pp. 219-228.
[7] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. I, Third edition, John Wiley, New York, 1968.
[8] J. GLATZ, *Expected waiting time for the visual response*, Biol. Cyber., 35 (1979), pp. 39-41.
[9] L. GORDON, M. SCHILLING AND M. S. WATERMAN, *An extreme value theory for long head runs*, Z. Wahrsch (in press, 1985).
[10] M. R. LEADBETTER, G. LINDGREN AND H. ROOTZEN, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1983.
[11] J. I. NAUS, *Probabilities for a generalized birthday problem*, J. Amer. Stat. Assoc., 69 (1974), pp. 810-815.
[12] T. NEMETZ AND N. KUSOLITSCH, *On the longest run of coincidences*, Z. Wahrsch., 61 (1982), pp. 59-73.
[13] S. J. SCHWAGER, *Run probabilities in sequences of Markov-dependent trials*, J. Amer. Stat. Assoc., 78 (1983), pp. 168-175.
[14] G. S. WATSON, *Extreme values in samples from m-dependent stationary stochastic processes*, Ann. Math. Stat., 25 (1954), pp. 798-800.
[15] P. DIGGLE, *Statistical Analysis of Spatial Point Patterns*, Academic Press, New York, 1983.
[16] P. J. E. PEEBLES, *The origin of galaxies and clusters of galaxies*, Science, 224 (1984), pp. 1385-1391.
[17] P. REVESZ, *How random is random?* Probability and Mathematical Statistics, 4 (1984), pp. 109-116.
[18] M. HALL, *Combinatorial Theory*, Academic Press, New York, 1967.
[19] P. DIACONIS AND D. FREEDMAN, *On the statistics of vision*, J. Math. Psych., 24 (1981), pp. 112-138.