DIMACS Series in Discrete Mathematics
and Theoretical Computer Science
Volume 1, 1990

Random Volumes in the n-Cube

M. E. DYER, Z. FÜREDI, AND C. McDIARMID

ABSTRACT. Consider the n-cube $[0, 1]^n$ in $\mathbf{R}^n$. This has 2^n vertices and volume 1. Pick $N = N(n)$ vertices independently at random, form their convex hull, and let V_n be its expected volume. How large should $N(n)$ be to pick up significant volume?

Let $\kappa = 2/\sqrt{e} \approx 1.213$, and let $\epsilon > 0$. We have recently shown that, as $n \to \infty$, $V_n \to 0$ if $N(n) \leq (\kappa - \epsilon)^n$, and $V_n \to 1$ if $N(n) \geq (\kappa + \epsilon)^n$. We discuss this and related results.

1. Introduction

We are interested in the n-cube $Q_n = [0, 1]^n$ in n-dimensional real space $\mathbf{R}^n$. This polytope has the set $\{0, 1\}^n$ of 2^n vertices and has volume 1. Let $N = N(n)$, and let $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_N$ be independent random variables, each uniformly distributed over $\{0, 1\}^n$. Form the convex hull S_n of these random points and let V_n be its expected volume, that is, $V_n = \mathbf{E}[\mathrm{vol}(S_n)]$. How large should $N(n)$ be to pick up significant volume? The answer is surprisingly (?) small. The following theorem is given in **[2]**. We shall sketch the outlines of the proof here.

THEOREM 1.1. *Let $\kappa = 2/\sqrt{e} \approx 1.213$ and let $\epsilon > 0$. Then, as $n \to \infty$,*

$$V_n \to \begin{cases} 0 & \text{if } N(n) \leq (\kappa - \epsilon)^n, \\ 1 & \text{if } N(n) \geq (\kappa + \epsilon)^n. \end{cases}$$

□

What happens if we pick points *within* the n-cube? Suppose now that we sample N times uniformly from $[0, 1]^n$ and let V_n be the expected volume of the convex hull of the points picked. The next theorem is also from **[2]**; it is proved along exactly the same lines as Theorem 1.1.

THEOREM 1.2. *Let $\lambda = \int_0^\infty (1 - \coth t + 1/t)^2\, dt \approx 2.13969$, and let $\epsilon > 0$.*

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 52A25, 60D05.

©1990 American Mathematical Society
1052-1798/90 \$1.00 + \$.25 per page

Then, as $n \to \infty$,

$$V_n \to \begin{cases} 0 & \text{if } N(n) \le (\lambda - \epsilon)^n, \\ 1 & \text{if } N(n) \ge (\lambda + \epsilon)^n. \end{cases}$$

□

These theorems concerning the n-cube are of course tight, but an even tighter result holds for the unit ball B_n in $\mathbf{R}^n$ in a sense that we are about to explain. Denote the volume of this ball by γ_n. Suppose now that we sample N times uniformly from B_n, and let V_n be the expected volume of the convex hull of the points picked.

THEOREM 1.3. *If* $\omega(n) \to \infty$ *as* $n \to \infty$, *and*

$$N(n) = n^{(\frac{1}{2} + \frac{\omega(n)}{\log n})n},$$

then $V_n/\gamma_n \to 1$. □

(Natural logarithms are used throughout.)

However, now define $V_B(n)$ to be the maximum volume over all sets S which are the convex hull of $N(n)$ points in B_n. (There is no randomness here.) Bárány and Füredi [**1**] extend an idea of Elekes [**3**] to show that, as $n \to \infty$, $V_B(n)/\gamma_n \to 1$ only if the conditions of Theorem 1.3 hold. Thus, roughly speaking, as soon as N is large enough that it is possible to place N points so as to pick up most of the volume of B_n, then a random choice of N points will do.

Is there a similar phenomenon for the n-cube Q_n? We may define $V_Q(n)$ analogously to $V_B(n)$ above. Then, using Elekes' idea, we can show

THEOREM 1.4. *Let* $\eta = 1.18858$. *Then* $V_Q(n) \to 0$ *as* $n \to \infty$ *if* $N(n) = O(\eta^n)$. □

This leads us to pose the following

QUESTION 1.5. *Is it the case that, for* $\epsilon > 0$, $V_Q(n) \to 0$ *when* $N(n) = O((\frac{2}{\sqrt{\epsilon}} - \epsilon)^n)$? □

2. Sketch of the proof of Theorem 1.1

In this section we sketch the outlines of the proof in [**2**] of Theorem 1.1.

Which points $\mathbf{x} = (x_1, x_2, \dots, x_n)$ of Q_n are not likely to be included in S_n? This will happen if some half-space H contains $\mathbf{x}$ but contains few vertices of Q_n. Thus, given $\mathbf{x}$ in Q_n, let $q(\mathbf{x})$ be the infimum, over all half-spaces H containing $\mathbf{x}$, of the quantity $\mathbf{P}(\mathbf{Z} \in H)$. Here $\mathbf{Z}$ is uniformly distributed over all the vertices of Q_n. Clearly, if $\mathbf{x}$ is in H, but none of $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N$ is, then $\mathbf{x} \notin S_n$. Thus

$$\mathbf{P}(\mathbf{x} \in S_n) \le Nq(\mathbf{x}).$$

For $\alpha > 0$, let the α-*center* Q_n^α be the convex subset of Q_n defined by

$$Q_n^\alpha = \{\mathbf{x} \in Q_n : q(\mathbf{x}) \geq e^{-\alpha n}\}.$$

LEMMA 2.1 (central lemma). *Let* $\alpha > 0$.

(a) *If* $\mathrm{vol}(Q_n^\alpha) = o(1)$ *and* $N(n) = o(e^{\alpha n})$, *then* $\mathbf{E}[\mathrm{vol}(S_n)] = o(1)$.

(b) *If* $\mathrm{vol}(Q_n^\alpha) = 1 - o(1)$ *and* $N(n) \geq \beta n^2 e^{\alpha n}$ *where* $\beta > \alpha$, *then* $\mathbf{E}[\mathrm{vol}(S_n)] = 1 - o(1)$.

By this lemma it suffices to show that

$$\mathrm{vol}(Q_n^\alpha) = \begin{cases} o(1) & \text{if } \alpha < \nu, \\ 1 - o(1) & \text{if } \alpha > \nu \end{cases}$$

where $\nu = \log 2 - \frac{1}{2}$. To do this we approximate Q_n^α by a more easily handled body. We would like to find a suitable "separable penalty function"

$$F(\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} f(x_j),$$

such that if we set

$$F_n^\alpha = \{\mathbf{x} \in (0, 1)^n : F(\mathbf{x}) \leq \alpha\},$$

then F_n^α approximates Q_n^α in a suitable way.

Let us pull a rabbit out of a hat. Suppose we take

$$f(x) = x \log x + (1 - x)\log(1 - x) + \log 2,$$

for $0 < x < 1$. Then we can show that

(a) $F_n^\alpha \subseteq Q_n^\alpha$, and

(b) if $0 < \beta < \alpha$ then $Q_n^\beta \cap (0, 1)^n \subseteq F_n^\alpha$ for n sufficiently large.

To prove (a) we use the Bernstein (or Markov) inequality; to prove (b) we use "exponential centering" together with a uniform version of the central limit theorem **[4]**—the details are messy. From (a), (b), it suffices to show that

$$\mathrm{vol}(F_n^\alpha) = \begin{cases} o(1) & \text{if } \alpha < \nu, \\ 1 - o(1) & \text{if } \alpha > \nu. \end{cases}$$

But this is easy. Let $X_1, X_2, \ldots, X_n$ be independent random variables each uniformly distributed on $(0, 1)$. Then $\mathbf{E}[f(X_1)]$ turns out to be ν—this is the "explanation" of the constant. Also, by the weak law of large numbers

$$\begin{aligned} \mathrm{vol}(F_n^\alpha) &= \mathbf{P}((X_1, X_2, \ldots, X_n) \in F_n^\alpha) \\ &= \mathbf{P}\left(\frac{1}{n}\sum_{j=1}^{n} f(X_j) \leq \alpha\right) \\ &= \begin{cases} o(1) & \text{if } \alpha < \nu, \\ 1 - o(1) & \text{if } \alpha > \nu. \end{cases} \end{aligned}$$

3. Sampling from the unit ball B_n

In this section we shall prove Theorem 1.3. Let $N = N(n) = n^{(\frac{1}{2}+\frac{\omega(n)}{\log n})n}$, where $\omega = \omega(n) \to \infty$ as $n \to \infty$. Sample N times uniformly from the unit ball $B_n = B(0, 1)$ in $\mathbf{R}^n$, let S_n be the convex hull of the points picked, and let $V_n = \mathbf{E}[\mathrm{vol}(S_n)]$. Let

$$\gamma_n = \mathrm{vol}(B_n) = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

We must show that $V_n/\gamma_n \to 1$ as $n \to \infty$. We shall in fact show more, that if $\epsilon > 0$, then

$$\mathbf{P}(S_n \supseteq B(0, 1-\epsilon/n)) \to 1 \quad \text{as } n \to \infty. \tag{1}$$

Note that

$$\mathrm{vol}(B(0, 1-\epsilon/n)) = \gamma_n(1-\epsilon/n)^n \geq \gamma_n(1-\epsilon),$$

and so if (1) holds, then

$$V_n/\gamma_n \geq 1 - \epsilon + o(1),$$

and we are done.

Let $r = 1 - \epsilon/n$ and

$$V_\epsilon = \mathrm{vol}(\{\mathbf{x} \in B_n : x_1 \geq r\}).$$

By the argument used in the proof in [2] of the central lemma, it suffices for us to show that

$$\binom{N}{n}(1 - V_\epsilon/\gamma_n)^{N-n} \to 0 \quad \text{as } n \to \infty. \tag{2}$$

But

$$\begin{aligned}\frac{V_\epsilon}{\gamma_n} &= \frac{\gamma_{n-1}}{\gamma_n}\int_r^1 (1-x^2)^{n/2}\,dx \\ &\sim \frac{\gamma_{n-1}}{\gamma_n}\int_r^1 (1-x^2)^{n/2}\,dx \\ &= \frac{\gamma_{n-1}}{\gamma_n}\left[-\frac{(1-x^2)^{(n+1)/2}}{n+1}\right]_r^1 \\ &= \frac{\gamma_{n-1}}{\gamma_n}\frac{(1-r^2)^{(n+1)/2}}{n+1}.\end{aligned}$$

Now $\gamma_{n-1}/\gamma_n \sim \sqrt{2\pi/n}$ and $(1-r^2)^{(n+1)/2} \sim (2\epsilon/n)^{(n+1)/2}e^{-\epsilon/4}$. So $V_\epsilon/\gamma_n \sim (2\pi)^{-1/2}e^{-\epsilon/4}(2\epsilon)^{(n+1)/2}n^{-(n/2+1)}$. Hence

$$\begin{aligned}NV_\epsilon/\gamma_n &= \exp\left\{\left(\frac{1}{2}+\frac{\omega}{\log n}\right)n\log n + \frac{n+1}{2}\log(2\epsilon) - \left(\frac{n}{2}+1\right)\log n + O(1)\right\} \\ &= \exp\{(1+o(1))\omega n\}.\end{aligned}$$

We can now establish (2). We have

$$\begin{aligned}\binom{N}{n}(1 - V_\epsilon/\gamma_n)^{N-n} &\le \exp\{n \log N - (N-n)V_\epsilon/\gamma_n\}\\ &= \exp\left\{\left(\frac{1}{2} + \frac{\omega}{\log n}\right) n^2 \log n - \exp\{(1+o(1))\omega n\}\right\}\\ &\to 0 \quad \text{as } n \to \infty.\end{aligned}$$

4. Deterministic lower bound

In this section we shall prove Theorem 1.4. We wish to prove a lower bound to the maximum volume that can be achieved by the convex hull of N points placed *anywhere* in Q_n. This will obviously hold also when the points are restricted to be vertices. However, by Carathéodory's theorem, any internal point of Q_n is contained in a simplex whose vertices are also vertices of Q_n. Thus the maximum volume that can be achieved by the convex hull S_n of any N points of Q_n is no more than that which can be achieved by the convex hull S of $N' = (n+1)N$ of Q_n's vertices. Thus we may restrict attention to the vertices of Q_n at the cost of inflating the number of points by a factor $(n+1)$. This factor turns out to be insignificant, but the argument below can, in fact, be modified without great difficulty to avoid its introduction.

Using a theorem of Elekes [**3**] we describe a set of balls whose union is guaranteed to include S. These balls are defined by any chosen point and the vertices of S. It is natural to consider the center $(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})$ of Q_n as the chosen point. Each ball in the set is then the smallest that contains this point and a particular vertex of S. For the "typical" vertex $(0, 0, \dots, 0)$ of Q_n, the relevant ball is

$$K = \left\{\mathbf{x} : \sum_{j=1}^{n} \left(x_j - \frac{1}{4}\right)^2 \le n/16\right\}.$$

For any other vertex, the corresponding ball can be determined by symmetry. Observe that

$$\operatorname{vol}(K \cap Q_n) = \mathbf{P}\left(\sum_{j=1}^{n} \left(X_j - \frac{1}{4}\right)^2 \le n/16\right),$$

where the X_j are distributed independently, with each uniform on $[0, 1]$. For any $t > 0$, therefore, the Bernstein inequality gives

$$\begin{aligned}\operatorname{vol}(K \cap Q_n) &\le \mathbf{E}\left[\exp\left\{2t\left(n/16 - \sum_{j=1}^{n}\left(X_j - \frac{1}{4}\right)^2\right)\right\}\right]\\ &= (\mathbf{E}[\exp\{t(X - 2X^2)\}])^n\end{aligned}$$

where X is uniform on $[0, 1]$. Thus, since $t > 0$ is arbitrary,

$$\mathrm{vol}(K \cap Q_n) \leq \left\{\inf_{t>0} g(t)\right\}^n ,$$

where

$$g(t) = \int_0^1 e^{t(x-2x^2)}\,dx . \tag{3}$$

It is easy to show, by differentiating twice, that $g(t)$ is a strictly convex function of t. It is also easy to see that $g(0) = 1$, $g'(0) < 0$, and $g(t) \to \infty$ as $t \to \infty$. Thus $g(t)$ has a unique minimum in $(0, \infty)$. In the region of the minimizing value $t_{\min}$ (which turns out to be around $2\frac{1}{2}$), close numerical approximation of $g(t)$ can easily be achieved as follows. We substitute $y = (x - \frac{1}{4})$ in the integrand of (3) and then perform term-by-term integration of its expansion as a power series in y. Hence we can minimize $g(t)$ numerically to high accuracy by (say) Fibonacci search. We find $t_{\min} \approx 2.52635$ and $g(t_{\min}) < 0.841339$. Now S is the convex hull of $N' = (n+1)N$ vertices, so

$$\mathrm{vol}(S) \leq N' \,\mathrm{vol}(K \cap Q_n) = o(1),$$

if

$$N = O(1.18858^n) = O(0.841339^{-n}). \quad \square$$

Acknowledgments

Thanks to Imre Bárány and Geoff Grimmett for useful conversations and insights.

References

1. I. Bárány and Z. Füredi, *Approximation of the sphere by polytopes having few vertices*, Proc. Amer. Math. Soc. **102** (1988), 651–659.
2. M. E. Dyer, Z. Füredi, and C. McDiarmid, *Volumes spanned by random points in the hypercube*, to appear, 1990.
3. G. Elekes, *A geometric inequality and the complexity of computing the volume*, Discrete Comput. Geom. **1** (1986), 289–292.
4. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II, Wiley, New York, 1971.

University of Leeds

Hungarian Academy of Sciences

University of Oxford