Volumes Spanned by Random Points in the Hypercube

M.E. Dyer
School of Computer Studies, University of Leeds, Leeds, United Kingdom

Z. Füredi
Mathematical Institute of the Hungarian Academy of Sciences, Budapest, Hungary

C. McDiarmid
Department of Statistics, University of Oxford, Oxford, United Kingdom

ABSTRACT

Consider the hypercube $[0, 1]^n$ in $\boldsymbol{R}^n$. 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 shall show that, as $n \to \infty$, $V_n \to 0$ if $N(n) \leq (\kappa - \epsilon)^n$, and $V_n \to 1$ if $N(n) \geq (\kappa + \epsilon)^n$. A similar result holds for sampling uniformly from within the hypercube, with constant $\lambda = \exp\{\int_0^\infty (1/u - 1/(e^u - 1))^2 \, du\} \approx 2.136$.

1. RESULTS

We are interested in the hypercube $Q_n = [0, 1]^n$ in n-dimensional real space $\boldsymbol{R}^n$. 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 $V_n = \mathbf{E}[\text{vol}(S_n)]$. How large should $N(n)$ be to pick up significant volume? The answer is surprisingly (?) small.

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

Random Structures and Algorithms, Vol. 3, No. 1 (1992)

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

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 points picked.

Theorem 1.2. *Let* $\lambda = \exp\{\int_0^\infty (1/u - 1/(e^u - 1))^2 \, du\} \approx 2.136$ *and let* $\epsilon > 0$. *Then as* $n \to \infty$,

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

This article is devoted to proving these two results. For some related material and discussion, see [3].

2. THE CENTRAL LEMMA

In this section we present and prove the central lemma on which the proofs of our two theorems rest.

Let Z be a nondegenerate random variable taking values in $[0, 1]$. Let Z_1, $Z_2, \ldots, Z_n$ be independent random variables, each distributed like Z, and let $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$. We shall be interested in the cases where Z is uniformly distributed either over $\{0, 1\}$ or over $[0, 1]$, so that $\mathbf{Z}$ is uniformly distributed either over the set $\{0, 1\}^n$ of vertices or over the whole of the hypercube $Q_n = [0, 1]^n$. We will call these the "vertex case" and the "solid case", respectively.

Which points x of Q_n are not likely to be included in S_n? This will happen if some halfspace H contains x but $\mathbf{P}(Z \in H)$ is small. Given x in Q_n, let $q(x)$ be the infinum, over all halfspaces H containing x, of the quantity $\mathbf{P}(Z \in H)$. Clearly, if x is in H, but none of $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_N$ is, then $x \not\in S_n$. Thus

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

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

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

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

Once we have proved this lemma, in order to prove the two theorems it will remain to show that $\mathrm{vol}(Q_n^\alpha)$ "flips from 0 to 1" at the appropriate value of α.

Proof.
(a) Suppose that $N = o(e^{\alpha n})$. *Let* $R = Q_n \backslash Q_n^\alpha$. We wish to show that $\mathbf{E}[\mathrm{vol}(S_n \cap R)] = o(1)$. Observe that

$$\mathbf{E}[\mathrm{vol}(S_n \cap R)] = \mathbf{E}\left[\int_R \mathbf{1}_{\{x \in S_n\}}\, dx\right] = \int_R \mathbf{P}(x \in S_n)\, dx .$$

But $\mathbf{P}(x \in S_n) \le Nq(x) < Ne^{-\alpha n}$ for each $x \in R$ and $\mathrm{vol}(R) \le \mathrm{vol}(Q_n) = 1$. Hence

$$\mathbf{E}[\mathrm{vol}(S_n \cap R)] < Ne^{-\alpha n} \to 0 \quad \text{as } n \to \infty .$$

(b) Suppose now that $N(n) = (1 + o(1))\beta n^2 e^{\alpha n}$, where $\beta > \alpha$. Observe that

$$\mathbf{E}[\mathrm{vol}(S_n)] \ge \mathrm{vol}(Q_n^\alpha)\mathbf{P}(Q_n^\alpha \subseteq S_n) ,$$

so it suffices to show that $\mathbf{P}(Q_n^\alpha \subseteq S_n) = 1 - o(1)$.

Let E be the event that $\dim S_n < n$. It is easy to show that E is very unlikely to occur. Let $\delta = \sup\{\mathbf{P}(Z = x) : x \in [0, 1]\}$, so that $\delta < 1$. Then, for any fixed set T of dimension less than n, $\mathbf{P}(\mathbf{Z} \in T) \le \delta$. So

$$\mathbf{P}(E) \le \binom{N}{n} \delta^{N-n} = o(1) .$$

For each set $J \subseteq \{1, 2, \ldots, N\}$ with $|J| = n$, we define a more interesting event

E_J: The points in $\{\mathbf{Z}_j : j \in J\}$ determine a hyperplane such that, for one of the two corresponding (closed) half-spaces H, both $\mathbf{P}(\mathbf{Z} \not\in H) \ge e^{-\alpha n}$ and the event $\{\mathbf{Z}_j : j \not\in J\} \subseteq H$ occurs.

Thus if E_J occurs, there is a reasonable chance that a random point $\mathbf{Z}$ "misses" H, but still all the points $\mathbf{Z}_j$ lie in H.

The key observation now is that

$$\{Q_n^\alpha \not\subseteq S_n\} \subseteq E \cup \bigcup_J E_J .$$

For, suppose that S_n is a full-dimensional polyhedron, and $x \in Q_n^\alpha \backslash S_n$. Then some set of J of n vertices of S_n determines a hyperplane such that one corresponding half-space contains S_n but excludes x. But $q(x) \ge e^{-\alpha n}$, so $\mathbf{P}(\mathbf{Z} \not\in H) \ge e^{-\alpha n}$. It follows that

$$\mathbf{P}(Q_n^\alpha \not\subseteq S_n) \le \mathbf{P}(E) + \sum_J \mathbf{P}(E_J) \le o(1) + \binom{N}{n}\mathbf{P}(E_D) ,$$

where $D = \{1, 2, \ldots, n\}$.

Next we bound $\mathbf{P}(E_D)$. Suppose that $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_n$ are affinely independent, and hence determine two half-spaces H_1, H_2. If $\mathbf{P}(\mathbf{Z} \not\in H_1) \ge e^{-\alpha n}$, then

$$\mathbf{P}(\mathbf{Z}_j \in H_1(j = n+1, \ldots, N)) \leq (1 - e^{-\alpha n})^{N-n},$$

and similarly for H_2. Hence

$$\begin{aligned}\mathbf{P}(E_D | \mathbf{Z}_1, \ldots, \mathbf{Z}_n) &\leq 2(1 - e^{-\alpha n})^{N-n} \\ &\leq 2\exp\{-(N-n)e^{-\alpha n}\}.\end{aligned}$$

Now we may remove the conditioning to obtain the same bound on $\mathbf{P}(E_D)$. Hence finally,

$$\begin{aligned}\mathbf{P}(Q_n^\alpha \not\subseteq S_n) &\leq \mathbf{P}(E) + \binom{N}{n}\mathbf{P}(E_D) \\ &\leq o(1) + 2\exp\{n \log N - (N-n)e^{-\alpha n}\} \\ &= \exp\{(1 + o(1))\alpha n^2 - (1 + o(1))\beta n^2\} \\ &= o(1).\end{aligned}$$ ∎

(All logarithms are natural.)

3. APPROXIMATING THE α-CENTER

For convenience here we shall switch from the n-cube $Q_n = [0, 1]^n$ to the n-cube $C_n = [-1, 1]^n$, which has volume 2^n. Thus Z is now assumed to have a nondegenerate distribution with values in $[-1, 1]$, and the α-center C_n^α is now defined as a subset of C_n.

By the central lemma 2.1, what we want to do (at least for the vertex case and the solid case) is to find a constant ν such that

$$2^{-n} \operatorname{vol}(C_n^\alpha) \to \begin{cases} 0 & \text{if } \alpha < \nu, \\ 1 & \text{if } \alpha > \nu. \end{cases} \tag{1}$$

To do this, we shall approximate C_n^α by a more easily handled body. We would like to find a "penalty function" $F(x)$ such that, if we set

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

then F_n^α approximates C_n^α in an appropriate manner. Indeed, we would like our penalty function F to be of the form

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

for $x = (x_1, x_2, \ldots, x_n) \in (-1, 1)^n$, where $f(x)$ is a suitable function defined on $(-1, 1)$, with say $f(0) = 0$ and f non-negative, differentiable and convex.

For such a "separable" penalty function F, the body F_n^α is easy to handle. Let $X_1, X_2, \ldots, X_n$ be independent random variables, with each uniform on $(-1, 1)$. Then, by the weak law of large numbers,

$$\begin{aligned}2^{-n}\,\mathrm{vol}(F_n^\alpha) &= \mathbf{P}((X_1, X_2, \ldots, X_n) \in F_n^\alpha)\\ &= \mathbf{P}\Big(\frac{1}{n}\sum_{j=1}^{n} f(x_j) \le \alpha\Big)\\ &\to \begin{cases} 0 & \text{if } \alpha < \mathbf{E}[f(X_1)]\,, \\ 1 & \text{if } \alpha > \mathbf{E}[f(X_1)]\,, \end{cases}\end{aligned}$$

as $n \to \infty$. Thus, if we can find a suitable function f such that F_n^α approximates C_n^α, then we are well on the way to establishing (1) and thus completing our proof.

Let us now assume that the random variable Z, taking values in $[-1, 1]$, also satisfies $\mathbf{E}[Z] = 0$, and that, for any $\delta > 0$, both $\mathbf{P}(Z > 1 - \delta)$ and $\mathbf{P}(Z < -1 + \delta)$ are both strictly positive.

It is shown in Appendix A that, for any $x \in (-1, 1)$, there is a unique $t = h(x)$, say, such that $K'(t) = x$. Here $K(t)$ is the *cumulant generating function* $K(t) = \log \mathbf{E}[e^{tZ}]$. We define the function f on $(-1, 1)$ by setting

$$f(x) = -K(h(x)) + xh(x)\,.$$

It then follows (see Appendix A) that $f(0) = 0$, $f(x) > 0$ for $x \neq 0$, F is strictly convex and $f'(x) = h(x)$. (See Appendix B for the function f corresponding to the vertex case.)

The function f does the trick for us. The key inequality which we use below is the Bernstein (or Markov) inequality $\mathbf{P}(X \ge 0) \le \mathbf{E}[e^X]$. The function f is chosen so that the inequality is tight when we use it—see also subsection 4B below.

Consider a (fixed) point $\boldsymbol{x} \in (-1, 1)^n \backslash \{\mathbf{0}\}$ with

$$F(\boldsymbol{x}) = \frac{1}{n}\sum_{j=1}^{n} f(x_j) = \beta\ (>0)\,.$$

Think of $\boldsymbol{x}$ as being on the boundary of the convex body F_n^β. We want to estimate $q(\boldsymbol{x})$, so we must find a half-space H of the form $\boldsymbol{t}.(\boldsymbol{z} - \boldsymbol{x}) \ge 0$ with $\mathbf{P}(\mathbf{Z} \in H)$ as small as possible. It is natural to consider the half-space $H(\boldsymbol{x})$ which is bounded by the tangent hyperplane at $\boldsymbol{x}$ (and which does not contain $\mathbf{0}$); so let us set $t_j = f'(x_j) = h(x_j)$ for each $j = 1, 2, \ldots, n$. Then

$$\begin{aligned}q(\boldsymbol{x}) &\ge \mathbf{P}(\mathbf{Z} \in H(\boldsymbol{x}))\\ &= \mathbf{P}\Big(\sum_{j=1}^{n} t_j(Z_j - x_j) \ge 0\Big)\\ &\le \mathbf{E}\Big[\exp\Big\{\sum_{j=1}^{n} t_j(Z_j - x_j)\Big\}\Big]\\ &= \exp\Big\{-\sum_{j=1}^{n}(x_j t_j - K(t_j))\Big\}\\ &= \exp\{-n\,F(\boldsymbol{x})\}\,.\end{aligned}$$

Hence, if $F(\boldsymbol{x}) > \alpha$, then $q(\boldsymbol{x}) < e^{-\alpha n}$ and so $C_n^\alpha \cap (-1, 1)^n \subseteq F_n^\alpha$.

Now let $\nu = \mathbf{E}[f(X_1)]$ (where X_1 is uniform on $(-1, 1)$). We saw above that if $\alpha < \nu$ then $2^{-n}\operatorname{vol}(F_n^\alpha) = o(1)$, and so since $C_n^\alpha \cap (-1,1)^n \subseteq F_n^\alpha$ also $2^{-n}\operatorname{vol}(C_n^\alpha) = o(1)$. This corresponds to half our target (1), and yields the "lower bound" half of Theorems 1.1 and 1.2, once we have evaluated the two corresponding constants ν. This is done in Appendix B.

4. UPPER BOUNDS

A. Overview

In the last section we approximated the α-center C_n^α by the "penalty function" body F_n^α. To establish the lower bounds in Theorems 1.1 and 1.2 we showed that

$$q(\boldsymbol{x}) \le \mathbf{P}(\mathbf{Z} \in H(\boldsymbol{x})) \le \exp\{-nF(\boldsymbol{x})\}$$

for each $\boldsymbol{x} \in (-1,1)^n \backslash \{\mathbf{0}\}$; and it followed that $C_n^\alpha \cap (-1,1)^n \subseteq F_n^\alpha$. To prove the upper bounds in these theorem, we must show that the approximation of C_n^α by F_n^α is sufficiently good. Let us now restrict our attention to the two special cases of interest, namely the vertex case and the solid case.

Lemma 4.1. *For $0 < \beta < \alpha$, $F_n^\alpha \subseteq C_n^\alpha$ for n sufficiently large.*

Once this lemma is proved, we are home. For then if $\nu < \beta < \alpha$, we have $2^{-n}\operatorname{vol}(F_n^\beta) = 1 - o(1)$ and so also $2^{-n}\operatorname{vol}(C_n^\beta) = 1 - o(1)$, the missing half of the target (1). To prove Lemma 4.1, it will suffice to prove the next lemma.

Lemma 4.2. *Let $\epsilon > 0$. Then for n sufficiently large,*

$$\mathbf{P}(\mathbf{Z} \in H(\boldsymbol{x})) \ge \exp\{-n(F(\boldsymbol{x}) + \epsilon)\}$$

for all $\boldsymbol{x} \in [0,1)^n$, $\boldsymbol{x} \neq 0$.

For, suppose that Lemma 4.2 holds, and let us deduce Lemma 4.1. By symmetry, we may replace the condition $\boldsymbol{x} \in [0,1)^n$ above by $\boldsymbol{x} \in (-1,1)^n$. Let $\boldsymbol{x} \in F_n^\beta$, $\boldsymbol{x} \neq \mathbf{0}$. By the convexity of F_n^β we see that

$$\begin{aligned} q(x) &\ge \inf\{\mathbf{P}(\mathbf{Z} \in H(y)) : y \in F_n^\beta, \quad y \neq 0\} \\ &\ge \exp\{-n(\beta + \epsilon)\}, \end{aligned}$$

and so $\boldsymbol{x} \in C_n^{\beta+\epsilon}$. Thus $F_n^\beta \subseteq C_n^{\beta+\epsilon}$.

Our aim now is to prove Lemma 4.2. We shall consider the vertex case first; the solid case is very similar. The main tool that we shall use is "exponential centering" [2], together with a uniform version of the central limit theorem—see, e.g., [5]. Unfortunately, we must handle "small" and "large" coefficients $t_j = h(x_j)$ (in the definition of $H(\boldsymbol{x})$) separately.

B. Exponential Centering

This subsection is the heart of the upper bound proof. We first state a lemma that follows immediately from a uniform version of the central limit theorem. See, for example, the Berry–Esséen theorem in Feller [5, p. 544].

Lemma 4.3. *Given a, $b > 0$, there exist n_0, $\eta > 0$ such that the following holds: for any $n \geq n_0$ and for any independent random variables $Y_1, Y_2, \ldots, Y_n$ with*

$$\mathbf{E}[Y_j] = 0\,, \quad \sigma_j^2 = \mathbf{E}[Y_j^2] \geq a \quad \text{and } \mathbf{E}[|Y_j|^3] \leq b\,,$$

we have

$$\mathbf{P}\Big(0 \leq \sum_{j=1}^{n} Y_j \leq \sigma\Big) \geq \eta\,, \quad \text{where } \sigma^2 = \sum_{j=1}^{n} \sigma_j^2\,.$$

Now we use exponential centering to handle the case where the t_j are uniformly bounded. Recall that we are considering the vertex case here.

Lemma 4.4. *Let $\epsilon > 0$ and $0 < a < b$. Then there exists n_0 such that for $n \geq n_0$,*

$$\mathbf{P}\Big(\sum_{j=1}^{n} t_j(Z_j - x_j) \geq 0\Big) \geq \exp\Big\{-\sum_{j=1}^{n} (f(x_j) + \epsilon\Big\}$$

for all $x \in (0,1)^n$ with $a \leq t_j \leq b$ for each j. (Here $t_j = h(x_j)$.)

Proof. Let $X_1, X_2, \ldots, X_n$ be independent discrete random variables and consider $X = \Sigma_{j=1}^n X_j$. In order to estimate $\mathbf{P}(X \geq x)$ it is useful to introduce independent random variables W_j with distribution

$$\mathbf{P}(W_j = y) = e^y \mathbf{P}(X_j = y) / \mathbf{E}[e^{X_j}]$$

(assuming that the denominator is finite). Let $W = \Sigma_{j=1}^n W_j$, and observe that

$$\begin{aligned} \mathbf{P}(W = y) &= \sum_{y_j : \Sigma y_j = y} \prod_{j=1}^{n} \{e^{y_j} \mathbf{P}(X_j = y_j) / \mathbf{E}[e^{X_j}] \\ &= \Big(\prod_{j=1}^{n} \mathbf{E}[e^{X_j}]\Big)^{-1} e^y \mathbf{P}(X = y)\,. \end{aligned}$$

Thus

$$\mathbf{P}(X = y) = \Big(\prod_{j=1}^{n} \mathbf{E}[e^{X_j}]\Big) e^{-y} \mathbf{P}(W = y)\,.$$

Now let us apply this to the case $X_j = t_j Z_j$ (where the Z_j are uniform on $\{-1, 1\}$). We find

$$\mathbf{P}\Big(\sum_{j=1}^{n} t_j Z_j = w\Big) = \exp\Big\{\sum_{j=1}^{n} K_Z(t_j)\Big\} e^{-w}\mathbf{P}(W = w) .$$

Now (see Appendix C)

$$\mathbf{E}[W_j] = t_j M_Z'(t_j)/M_Z(t_j) = t_j \tanh t_j = t_j x_j .$$

Let $Y = W - \mathbf{E}[W] = W - \Sigma_{j=1}^{n} t_j x_j$. Then

$$\begin{aligned}\mathbf{P}\Big(\sum_{j=1}^{n} t_j(Z_j - x_j) \geq 0\Big) &= \exp\Big\{\sum_{j=1}^{n} K_Z(t_j)\Big\} \sum_{w \geq \Sigma t_j x_j} e^{-w}\mathbf{P}(W = w) \\ &= \exp\Big\{\sum_{j=1}^{n}(K_Z(t_j) - t_j X_j)\Big\} \sum_{y \geq 0} e^{-y}\mathbf{P}(Y = y) \\ &= \exp\Big\{\sum_{j=1}^{n} f(x_j)\Big\} \sum_{y \geq 0} e^{-y}\mathbf{P}(Y = y)\end{aligned}$$

It remains to show that the factor $\Sigma_{y\geq 0} e^{-y}\mathbf{P}(Y = y)$ is not too small. Let $Y_j = W_j - \mathbf{E}[W_j]$, so that $Y = \Sigma_{j=1}^{n} Y_j$. Of course $\mathbf{E}[Y_j] = 0$. Also (see Appendix C)

$$\sigma_j^2 = \mathbf{E}[Y_j^2] = t_j^2/\cosh^2 t_j$$

and

$$\mathbf{E}[|Y_j|^3] = (2\cosh t_j - \operatorname{sech} t_j)\sigma_j^3 .$$

By our assumption that all t_j satisfy $a \leq t_j \leq b$, it follows that for some c_1, c_2, $c_3 > 0$

$$c_1 \leq \sigma_j^2 \leq c_2 \quad \text{and} \quad \mathbf{E}[|Y_j|^3] \leq c_3 .$$

Hence, by Lemma 4.3, there are n_0 and $\eta > 0$ such that the following holds. For $n \geq n_0$ and any $x \in (0, 1)^n$ with $a \leq t_j \leq b$,

$$\begin{aligned}\sum_{y \geq 0} e^{-y}\mathbf{P}(Y = y) &\geq e^{-\sigma} \sum_{0 \leq y \leq \sigma} \mathbf{P}(Y = y) \\ &= \eta e^{-\sqrt{c_2 n}}\end{aligned}$$

■

C. Handling Troublesome t_j

In Lemma 4.4 above, we dealt with the case when the t_j were uniformly bounded. Here we shall handle the "small" and "large" t_j for the vertex case. In the first lemma below, we consider the case when there is some slack in the inequality, and in the second (the relaxation lemma) we show that we can introduce such slack. We can then complete the proof of Lemma 4.2 and hence of Theorem 1.1.

Lemma 4.5. *Let $\epsilon > 0$. Then there exists n_1 such that for $n \geq n_1$,*

$$\mathbf{P}\Big(\frac{1}{n}\sum_{j=1}^{n} t_j(Z_j - x_j) \geq -\epsilon\Big) \geq \exp\Big\{-\sum_{j=1}^{n} (f(x_j) + \epsilon)\Big\}$$

for all $x \in [0, 1)^n$.

Proof. These exists $\delta > 0$ such that, if $0 \leq x_j < \delta$, then $f(x_j) < \epsilon$ and $t_j(1 + x_j) < \epsilon$, and if $1 - \delta < x_j < 1$, then $\log 2 - f(x_j) < \epsilon$. Let $x \in [0, 1)^n$ and let

$$J_0 = \{j : x_j < \delta\}\ , \quad J_1 = \{j : x_j > 1 - \delta\}\ , \quad J = \{1, 2, \dots, n\} \backslash (j_0 \cup J_1)\ .$$

Then

$$\frac{1}{n}\sum_{j\in J_0} t_j(Z_j - x_j) \geq -\frac{1}{n}\sum_{j\in J_0} t_j(1 + x_j) > -\epsilon\ ,$$

and so

$$\mathbf{P}\Big(\frac{1}{n}\sum_{j=1}^{n} t_j(Z_j - x_j) \geq -\epsilon\Big) \geq \mathbf{P}\Big(\sum_{j\in J_1} t_j(Z_j - x_j) \geq 0\Big) \cdot \mathbf{P}\Big(\sum_{j\in J} t_j(Z_j - x_j) \geq 0\Big)\ .$$

Now consider J_1. Clearly,

$$\begin{aligned}\mathbf{P}\Big(\sum_{j\in J_1} t_j(Z_j - x_j) \geq 0\Big) &\geq \mathbf{P}(Z_j = 1 \quad (\forall_j \in J_1)) \\ &= 2^{-|J_1|} \\ &\geq \exp\Big\{\sum_{j\in J_1} (f(x_j) + \epsilon)\Big\}\ .\end{aligned}$$

Next consider J. If $|J| < \epsilon n$, then

$$\mathbf{P}\Big(\sum_{j\in J} t_j(Z_j - x_j) \geq 0\Big) \geq 2^{-|J|} \geq \exp\{-\epsilon n \log 2\}.$$

Also if $|J| \geq \epsilon n$ and $n \geq n_1 = \lceil n_0/\epsilon \rceil$ (where n_0 is from Lemma 4.4) then

$$\mathbf{P}\Big(\sum_{j\in J} t_j(Z_j - x_j) \geq 0\Big) \geq \exp\{-\sum_{j\in J} (f(x_j) + \epsilon)\Big\}\ .$$

Hence, if $n \geq n_1$,

$$\begin{aligned}\mathbf{P}\Big(\frac{1}{n}\sum_{j=1}^{n} t_j(Z_j - x_j) \geq -\epsilon\Big) &\geq \exp\{-\sum_{j\in J_1 \cup J} (f(x_j) + \epsilon) - \epsilon n \log 2)\} \\ &\geq \exp\Big\{-\sum_{j=1}^{n} (f(x_j) + \epsilon')\Big\}\end{aligned}$$

where $\epsilon' = e(1 + \log 2)$. ∎

Lemma 4.6 (Relaxation Lemma). *Let $\mathbf{x} \in [0,1)^n$, $\mathbf{x} \neq \mathbf{0}$ and let $\epsilon > 0$. Then there exists $\hat{\mathbf{x}} \in [0,1)^n$, $\hat{\mathbf{x}} \neq 0$ such that*

$$(i)\quad \mathbf{P}\Big(\frac{1}{n}\sum_{j=1}^{n} t_j(Z_j - x_j) \geq 0\Big) = \mathbf{p}\Big(\frac{1}{n}\sum_{j=1}^{n} \hat{t}_j(Z_j - \hat{x}_j) \geq -\epsilon\Big),$$

$$(ii)\quad 0 \leq F(\hat{\mathbf{x}}) - F(\mathbf{x}) \leq \epsilon,$$

where $\mathbf{t}$, $\hat{\mathbf{t}}$ correspond to $\mathbf{x}$, $\hat{\mathbf{x}}$ as usual.

Proof. For $\lambda \geq 1$ let $\mathbf{x}^\lambda$ correspond to $\mathbf{t}^\lambda = \frac{1}{n}\lambda \mathbf{t}$. Consider the continuous function $d(\lambda) = \mathbf{t}.(\mathbf{x}^\lambda - \mathbf{x}) - \epsilon/\lambda$, defined for $\lambda \geq 1$. Clearly $d(1) = -\epsilon < 0$. Also, observe that $\mathbf{t} \geq 0$, $t_i > 0$ if $x_i > 0$ and $\mathbf{x}^\lambda \geq \mathbf{0}$, $x_i^\lambda \to 1$ as $\lambda \to \infty$ if $x_i > 0$. Hence $d(\lambda) > 0$ for λ sufficiently large. Thus there exists $\hat{\lambda} > 1$ such that $d(\hat{\lambda}) = 0$.

Now put $\hat{\mathbf{x}} = \mathbf{x}^{\hat{\lambda}}$. If we subtract $d(\hat{\lambda})$ $(=0)$ from $\frac{1}{n}\sum t_j(z_j - x_j)$ we obtain $\frac{1}{\hat{\lambda}}\left\{\frac{1}{n}\sum \hat{t}_j(z_j - \hat{x}_j) + \epsilon\right\}$. Thus (for any z) $\frac{1}{n}\sum t_j(z_j - x_j) \geq 0$ if and only if $\frac{1}{n}\sum \hat{t}_j(z_j - \hat{x}_j) \geq -\epsilon$, which establishes (i).

In order to prove (ii) write $v(\lambda) = F(\mathbf{x}^\lambda)$ and $k(\lambda) = \frac{1}{n}\mathbf{t}\cdot\mathbf{x}^\lambda$ for $\lambda \geq 1$. Thus

$$v(\lambda) = \frac{1}{n}\sum_{j=1}^{n}\{\lambda t_j K'(\lambda t_j) - K(\lambda t_j)\},$$

$$k(\lambda) = \frac{1}{n}\sum_{j=1}^{n} t_j K'(\lambda t_j),$$

and so

$$v'(\lambda) = \frac{\lambda}{n}\sum_{j=1}^{n} t_j^2 K''(\lambda t_j) = \lambda k'(\lambda) \geq 0.$$

Hence

$$\begin{aligned}\frac{\epsilon}{\hat{\lambda}} &= k(\hat{\lambda}) - k(1)\\ &= \int_1^{\hat{\lambda}} k'(\lambda)\,d\lambda\\ &= \int_1^{\hat{\lambda}} \frac{v'(\lambda)}{\lambda}\,d\lambda\\ &\geq \frac{1}{\hat{\lambda}}\int_1^{\hat{\lambda}} v'(\lambda)\,d\lambda\\ &= \frac{1}{\hat{\lambda}}(v(\hat{\lambda}) - v(1))\\ &= \frac{1}{\hat{\lambda}}(F(\hat{\mathbf{x}}) - F(\mathbf{x})).\end{aligned}$$

Thus $F(\hat{x}) - F(x) \le \epsilon$. Finally observe that $F(\hat{x}) \ge F(x)$, since $v'(\lambda) \ge 0$. ∎

Proof of Lemma 4.2. Let $n > n_1$, where n_1 is from Lemma 4.5. Let $\boldsymbol{x} \in [0, 1)^n$, $\boldsymbol{x} \neq \boldsymbol{0}$. By the relaxation lemma 4.6, there exists $\hat{\boldsymbol{x}} \in [0, 1)^n$, $\hat{\boldsymbol{x}} \neq \boldsymbol{0}$ such that

$$\mathbf{P}(\mathbf{Z} \in H(\boldsymbol{x})) = \mathbf{P}\left(\frac{1}{n} \sum \hat{t}_j(Z_j - \hat{x}_j) \ge -\epsilon\right)$$

and

$$F(\hat{x}) \le F(x) + \epsilon .$$

Hence by Lemma 4.5

$$\begin{aligned}\mathbf{P}(\mathbf{Z} \in H(\boldsymbol{x})) &\ge \exp\{-n(F(\hat{x}) + \epsilon)\} \\ &\ge \exp\{-n(F(\hat{x}) + 2\epsilon)\}\end{aligned}$$ ∎

D. Solid Case

The proof for the solid case is very like that for the vertex case, except that the "large" t_j no longer cause trouble.

In the exponential centering, the random variables W_j now have have the density function

$$e^w/2 \sinh t_j \qquad (-t_j \le w \le t_j) .$$

As before we find that

$$\mathbf{P}\left(\sum_{j=1}^{n} t_j(Z_j - x_j) \ge 0\right) \ge \exp\left\{-\sum_{j=1}^{n} f(x_j)\right\} e^{-\sigma} \mathbf{P}(0 \le Y \le \delta) .$$

Now $\sigma_j^2 = \mathbf{E}[Y_j^2] = 1 - t_j^2/\sinh^2 t_j$, and

$$\mathbf{E}[|Y_j|^3] \sim \left(\frac{12}{e} - 2\right) \quad \text{as } t_j \to \infty .$$

Hence the result corresponding to Lemma 4.4 holds, without the upper bound b. The result corresponding to Lemma 4.5 is now a little easier to prove—we do not introduce the set J_1. Finally we complete the proof as before. ∎

APPENDIX A: PROPERTIES OF $K(t)$

Let the random variable Z be nondegenerate and such that the moment generating function $M(t) = \mathbf{E}[e^{tZ}]$ exists for all t. Then the cumulant generating function $K(t) = \log M(t)$ is strictly for all t. For we have

$$K'(t) = \mathbf{E}[Ze^{tZ}]/M(t)\,,$$
$$K''(t) = \{\mathbf{E}[e^{tZ}]\mathbf{E}[Z^2e^{tZ}] - \mathbf{E}[Ze^{tZ}]^2\}/M(t)^2\,,$$

and it follows from the Cauchy–Schwarz inequality that $K''(t) > 0$.

Lemma A.1. *Let $a < b$, let the random variable Z take values in $[a, b]$, and suppose that for any $\delta > 0$ both $\mathbf{P}(Z > b - \delta)$ and $\mathbf{P}(Z < a + \delta)$ are positive. Then the function $t \mapsto K'(t)$ is an increasing bijection from $\boldsymbol{R}$ to (a, b).*

Proof. Since $K'(t) = \mathbf{E}[Ze^{tZ}]/\mathbf{E}[e^{tZ}]$, it follows that $K'(t) = x$ if and only if $\mathbf{E}[(Z - x)e^{tZ}] = 0$. But clearly for any t

$$a\mathbf{E}[e^{tZ}] < \mathbf{E}[Ze^{tZ}] < b\mathbf{E}[e^{tZ}]\,.$$

Thus $\mathbf{E}[(Z - a)e^{tZ}] > 0$ and so $K'(t) > a$. Similarly $K'(t) < b$.

We now know that the function $t \mapsto K'(t)$ is strictly increasing, with values contained in (a, b). Let $x \in (a, b)$. We must show that $K'(t) = x$ for some t. Without loss of generality, suppose that $x \geq \mathbf{E}[Z]$. Consider

$$g(t) = \mathbf{E}[(Z - x)e^{t(Z-x)}] = e^{-tx}\mathbf{E}[(Z - x)e^{tZ}]\,.$$

We want to show that $g(t) = 0$ for some t. Clearly $g(0) = \mathbf{E}[Z] - x \leq 0$, and so it suffices to show that $g(t) > 0$ for t sufficiently large.

Put $Y = Z - x$, so that $g(t) = \mathbf{E}[Ye^Y]$. Let $0 < \delta < b - x$. Note that $\mathbf{P}(Y > \delta) > 0$. For $t \geq 0$,

$$\begin{aligned} g(t) &= \mathbf{E}[Ye^{tY}] \\ &\geq \mathbf{E}[Ye^{tY} | Y \leq 0]\mathbf{P}(Y \geq 0) + \mathbf{E}[Ye^{tY} | Y \geq \delta]\mathbf{P}(Y \geq \delta) \\ &\geq \mathbf{E}[Y | Y \leq 0]\mathbf{P}(Y \leq 0) + \delta e^{t\delta}\mathbf{P}(Y \geq \delta)\,. \end{aligned}$$

But the former term here is a finite constant, and the latter clearly tends to ∞ as $t \to \infty$. ■

Now let us assume that Z is as in the lemma, and further that $\mathbf{E}[Z] = 0$. Thus $K'(0) = 0$.

By the lemma, there is a (unique) increasing bijection $h: (a, b) \to \boldsymbol{R}$ such that $K'(t) = x$ if and only if $h(x) = t$. Observe that $h(0) = 0$. We defined the function $f: (a, b) \to \boldsymbol{R}$ by

$$f(x) = -K(h(x)) + xh(x) = -K(t) + tK'(t)\,,$$

where $t = h(x)$. Note that

$$f'(x) = -K'(h(x))h'(x) + xh'(x) + h(x) = h(x)\,.$$

It follows that f is strictly convex, since h is strictly increasing. Note also that

$f(0) = f'(0) = 0$, and so $f(x) > 0$ for $x \neq 0$. Also, if the distribution of Z is symmetrical about 0, then f is even.

APPENDIX B: THE CONSTANTS ν

B.1. Vertex Case

Let

$$\mathbf{P}(Z=1) = \mathbf{P}(Z=-1) = \tfrac{1}{2}\,. \quad \text{Then, for all } t\,,$$

$$M(t) = \mathbf{E}[e^{tZ}] = \cosh t\,,$$

$$K(t) = \log M(t) = \log\cosh t\,,$$

$$K'(t) = \tanh t\,.$$

Hence, for $x \in (-1, 1)$,

$$h(x) = \tanh^{-1} x = \frac{1}{2}\log\left(\frac{1+x}{1-x}\right),$$

and

$$K(h(x)) = -\frac{1}{2}\log(1-x^2)\,,$$

and so

$$\begin{aligned} f(x) &= \frac{1}{2}\log(1-x^2) + \frac{1}{2}\,x\log\left(\frac{1+x}{1-x}\right) \\ &= \frac{1}{2}(1+x)\log(1+x) + \frac{1}{2}(1-x)\log(1-x)\,. \end{aligned}$$

Observe that $f(x) \to \log 2$ as $x \uparrow 1$ or as $x \downarrow -1$.

We must work out $\nu = \mathbf{E}[f(X)]$, where X is uniform on $[-1, 1]$. We have

$$\begin{aligned} \nu = \mathbf{E}[f(X)] &= \frac{1}{2}\int_0^1 \{(1+x)\log(1+x) + (1-x)\log(1-x)\}\,dx \\ &= \frac{1}{2}\int_1^2 y\log y\,dy + \frac{1}{2}\int_0^1 y\log y\,dy \\ &= \frac{1}{2}\int_0^2 y\log y\,dy \\ &= \frac{1}{2}\left\{\left[\frac{1}{2}\,y^2\log y\right]_0^2 - \int_0^2 \frac{1}{2}\,y\,dy\right\} \\ &= \log 2 - \tfrac{1}{2}\,. \end{aligned}$$

Thus $\kappa = e^{\nu} = \dfrac{2}{\sqrt{e}} \approx 1.21306\ldots$.

B.2. Solid Case

Let Z be uniform on $[-1, 1]$. Then, for all $t > 0$,

$$M(t) = \mathbf{E}[e^{tZ}] = \frac{1}{2}\int_{-1}^{1} e^{tx}\,dx = \sinh t/t$$

$$K(t) = \log(\sinh t/t)\,,$$

$$K'(t) = \coth t - 1/t\,.$$

We may consider X uniform on $[0, 1]$. Then $T = h(X)$ has density function

$$\frac{dx}{dt} = K''(t) = \frac{1}{t^2} - \frac{1}{\sinh^2 t} \quad \text{for } t \in (0, \infty)\,.$$

Hence

$$\begin{aligned}
\nu &= \mathbf{E}[f(X)] \\
&= \mathbf{E}[-K(h(X)) + Xh(X)] \\
&= \int_0^\infty (-K(t) + K'(t)t)K''(t)\,dt \\
&= \int_0^\infty (-K(t) + K'(t)t) - \frac{d}{dt}(1 - K'(t))\,dt \\
&= -[(-K(t) + K'(t)t)(1 - K'(t))]_0^\infty + \int_0^\infty tK''(t)(1 - K'(t))^2\,dt \\
&= -\frac{1}{2}\int_0^\infty t\,\frac{d}{dt}(1 - K'(t))^2\,dt \\
&= -\frac{1}{2}\left\{[t(1 - K'(t))^2]_0^\infty - \int_0^\infty (1 - K'(t))^2\,dt\right\} \\
&= \frac{1}{2}\int_0^\infty (1 - K'(t))^2\,dt \\
&= \frac{1}{2}\int_0^\infty (1 - \coth t + 1/t)^2\,dt \\
&= \frac{1}{2}\int_0^\infty \left(\frac{1}{t} - \frac{2}{e^{2t} - 1}\right)^2 dt \\
&= \int_0^\infty \left(\frac{1}{u} - \frac{1}{e^u - 1}\right)^2 du \qquad (u = 2t)
\end{aligned}$$

The integral can only be evaluated numerically, but this is not too difficult. The function $(1/u - 1/(e^u - 1))$ is numerically well-behaved provided we are careful to cancel the poles at the origin (e.g., by expanding e^u and simplifying the resulting expression). Also, for $u > 40$ (say), the integrand is very close to $1/u^2$, so $\frac{1}{40}$ is an accurate estimate of the tail of the integral from 40 onwards. Standard quadrature methods can them be employed. We obtain

$$\nu = 0.759007\ldots, \quad \text{and hence } \lambda = e^{\nu} = 2.13615\ldots$$

APPENDIX C: EXPONENTIAL CENTERING

C.1. Vertex Case

Here

$$\mathbf{P}(W_j = t_j) = \frac{e^{t_j}}{2\cosh t_j}, \qquad \mathbf{P}(W_j = -t_j) = \frac{e^{-t_j}}{2\cosh t_j}.$$

So

$$M_{W_j}(s) = \mathbf{E}[e^{sW_j}] = \cosh((1+s)t_j)/\cosh t_j .$$

Hence

$$\mathbf{E}[W_j] = t_j \tanh t_j , \qquad \mathbf{E}[W_j^2] = t_j^2 ,$$

and so

$$\sigma_j^2 = \mathrm{Var}(W_j) = t_j^2/\cosh^2 t_j .$$

Also, if $Y_j = W_j - \mathbf{E}[W_j]$, then it may be checked that

$$\mathbf{E}[|Y_j|^3] = (2\cosh t_j - \operatorname{sech} t_j)\sigma_j^3 .$$

C.2. Solid Case

Here W_j has density

$$e^{w}/2\sinh t_j \qquad \text{for } -t_j < w < t_j .$$

So, for $s > -1$,

$$K(s) = K_{W_j}(s) = \log\sinh((s+1)t_j) - \log(s+1) - \log\sinh t_j$$

$$= K'(s) = t_j \coth((s+1)t_j) - 1/(s+1) .$$

Hence

$$\mu_j = \mathbf{E}[W_j] = K'(0) = t_j \coth t_j - 1 .$$

Also

$$K''(s) = \frac{-t_j^2}{\sinh^2((s+1)t_j)} + \frac{1}{(s+1)^2},$$

and so

$$\sigma_j^2 = \mathrm{Var}(W_j) = 1 - \frac{t_j^2}{\sinh^2 t_j}.$$

Let $Y_j = W_j - \mu_j$, as before. Then

$$\begin{aligned}
2\sinh t_j \mathbf{E}[|Y_j|^3] &= \int_{-t_j}^{\mu_j} (\mu_j - w)^3 e^w \, dw + \int_{\mu_j}^{t_j} (w - \mu_j)^3 e^w \, dw \\
&= \int_0^{t_j+\mu_j} y^3 e^{\mu_j - y} \, dy + \int_0^{t_j-\mu_j} y^3 e^{\mu_j + y} \, dy \\
&= e^{\mu_j}\left\{ \int_0^{t_j+\mu_j} y^3 e^{-y} \, dy + \int_0^{t_j-\mu_j} y^3 e^{y} \, dy \right\} \\
&\sim e^{t_j - 1}\left\{ \int_0^{\infty} y^3 e^{-y} \, dy + \int_0^{1} y^3 e^{y} \, dy \right\} \qquad \text{as } t_j \to \infty, \\
&\qquad \text{since then } \mu_j = t_j - 1 + o(1) \\
&= e^{t_j}\left(\frac{12}{e} - 2\right).
\end{aligned}$$

Hence

$$\mathbf{E}[|Y_j|^3] \sim \left(\frac{12}{e} - 2\right) \qquad \text{as } t_j \to \infty.$$

REFERENCES

[1] I. Bárány and Z. Füredi, Approximation of the sphere by polytopes having few vertices, *Proc. Am. Math. Soc.*, **102**, 651–659 (1988).

[2] R. R. Bahadur, Some limit theorems in statistics, *CBMS Regional Conference Series in Applied Mathematics*, **4**, SIAM, Philadelphia, PA.

[3] M. E. Dyer, Z. Füredi, and C. McDiarmid, Random volumes in the n-cube, in Polyhedral Combinatorics, P. D. Seymour and W. Cook (Eds.), American Mathematical Society, 1990, pp. 33–38.

[4] G. Elekes, A geometric inequality and the complexity of computing the volume, *Discrete and Computational Geometry*, **1**, 289–292 (1986).

[5] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, Wiley, New York, 1971.

Received July 11, 1991