Monotonicity of the Sample Range of 3-D Data: Moments of Volumes of Random Tetrahedra

Stefan Kunis, Benjamin Reichenwallner and Matthias Reitzner
University of Osnabrück and University of Salzburg

Abstract

The sample range of uniform random points $X_1, \ldots, X_n$ chosen in a given convex set is the convex hull $\text{conv}[X_1, \ldots, X_n]$. It is shown that in dimension three the expected volume of the sample range is not monotone with respect to set inclusion. This answers a question by Meckes in the negative.

The given counterexample is the three-dimensional tetrahedron together with an infinitesimal variation of it. As side result we obtain an explicit formula for all even moments of the volume of a random simplex which is the convex hull of three uniform random points in the tetrahedron and the center of one facet.

2000 AMS subject classification: Primary 62H11; Secondary 52A22, 60D05.

Keywords: sample range; random convex hull; extreme points; random simplex.

1 INTRODUCTION

Choose random points $X_1, \ldots, X_n$ independently according to the uniform distribution in an interval $I \subset \mathbb{R}$. The convex hull $\text{conv}[X_1, \ldots, X_n]$ of these random points is the well-known sample range which can also be defined as the interval $[X_{[1]}, X_{[n]}]$, where $X_{[1]} \leq \cdots \leq X_{[n]}$ is the order statistic of the random points, and the endpoints $X_{[1]}, X_{[n]}$ are the extreme points of the random sample. It is trivial and immediate that the expected length of the sample range is a monotone function in $I$: Choosing uniform random points $Y_1, \ldots, Y_n$ in an interval $J \subset I$, one has

$$\mathbb{E}|\text{conv}[Y_1, \ldots, Y_n]| \leq \mathbb{E}|\text{conv}[X_1, \ldots, X_n]|,$$

where $|A|$ is the Lebesgue measure of the set $A$.

A generalization of this question to higher dimensions leads to nontrivial problems: First the definition of sample range, order statistic and extreme points is not obvious. Maybe the most natural extension of extreme points and sample range for higher dimensions is the following:

For random points $X_1, \ldots, X_n \in \mathbb{R}^d$, we define the sample range to be the convex hull $\text{conv}[X_1, \ldots, X_n]$, and the extreme points of the sample are those on the boundary of the sample range, i.e., the vertices of $\text{conv}[X_1, \ldots, X_n]$. 

The question we want to address in this paper is the following: *Is the expected volume of the sample range a monotone function in the underlying distribution?* To make this question more precise, we assume (as the most simple example) that the points are chosen according to the uniform measure in a convex set $K$. Then the monotonicity question reads as follows:

Assume that $L$, $K$ are two $d$-dimensional convex sets. Choose independent uniform random points $Y_1, \ldots, Y_n$ in $L$ and $X_1, \ldots, X_n$ in $K$. Is it true that $L \subset K$ implies

$$
E|\text{conv}\{Y_1, \ldots, Y_n\}| \leq E|\text{conv}\{X_1, \ldots, X_n\}|?
$$

Here, $|A|$ denotes the $d$-dimensional Lebesgue measure of the $d$-dimensional set $A$. The starting point for these investigations should be a check for the first nontrivial case $n = d + 1$ where the sample range is the random simplex spanned by the sample points. In this form, the question was first raised by Meckes [4] in the context of high-dimensional convex geometry.

As already mentioned, in dimension one this is immediate. It was proved by Rademacher [6] in 2012 that this is also true in dimension two. Our main result solves the three-dimensional case.

**Theorem 1.** In $\mathbb{R}^3$ the expected volume of the sample range is in general not monotone in the underlying distribution. There are three-dimensional convex sets $L \subset K$ such that

$$
E|\text{conv}\{Y_1, \ldots, Y_4\}| > E|\text{conv}\{X_1, \ldots, X_4\}|,
$$

if $Y_1, \ldots, Y_4$ are chosen uniformly in $L$ and $X_1, \ldots, X_4$ in $K$.

That the general question cannot be answered in the positive was already shown by Rademacher who, in a groundbreaking paper, gave counterexamples for dimensions $d \geq 4$ and $n = d + 1$. It remains an open problem whether there is a number $N$, maybe depending on $K$ or only on the dimension of the underlying space, such that monotonicity holds for $n \geq N$. And a suitable precise formulation of the question for non-uniform measures would also be highly interesting.

For our proof, we need to construct a pair of convex sets leading to a counterexample. A serious drawback of this approach is that one is forced to compute the expected volume of a random simplex which is known to be a notorious hard problem. In dimension two, tedious but explicit computations from the nineteenth century yielded several explicit results, but starting with dimension three, the problem turns out to be out of reach in general. The only three-dimensional convex sets where the expected volume of a random simplex is known are the ball [5], the cube [10] and the tetrahedron [1]. And in higher dimensions only the ball allows for explicit results. Since numerical computations in dimension three suggest that in the neighbourhood of the cube and the ball the expected volume of a random simplex is monotone, the only potentially tractable counterexample could be the tetrahedron and a set close to it, which also is in accordance with numerical computations by Rademacher [6], and Reichenwallner and Reitzner [7].

Already the determination of the expected volume of a random simplex in a tetrahedron $\mathcal{T} \subset \mathbb{R}^3$ was extremely hard. This question is known as Klee’s problem, and after many
attempts, erroneous conjectures and numerical estimates, Reitzner and Buchta [1] proved in a long paper that for uniform random points $X_1, \ldots, X_4$ in a tetrahedron of volume one, we have

$$
E|\text{conv}[X_1, \ldots, X_4]| = \frac{13}{720} - \frac{\pi^2}{15015} = 0.01739\ldots
$$

(4)

It seems to be out of reach to compute this expectation for any other three-dimensional convex set close to $T$. Luckily there is a wonderful alternative approach due to Rademacher, using an infinitesimal variation of convex sets, which is stated in the following Lemma.

**Lemma 1** (Rademacher [5]). For $d \in \mathbb{N}$, monotonicity under inclusion of the map

$$
K \mapsto E|\text{conv}[X_1, \ldots, X_{d+1}]|
$$

where $K$ ranges over all $d$-dimensional convex bodies and $X_i$ are iid uniform points in $K$, holds if and only if we have for each convex body $K \subseteq \mathbb{R}^d$ and for each $z \in \text{bd} K$ that

$$
E|\text{conv}[X_1, \ldots, X_{d+1}]| \leq E|\text{conv}[X_1, \ldots, X_d, z]|.
$$

Hence we get the counterexample for Theorem 1 if we succeed in computing the expectation $E|\text{conv}[X_1, \ldots, X_3, z]|$ for some $z \in \text{bd} T$. Because of symmetry, a suitable choice for $z$ should be the center $c$ of one of the facets. Yet after several attempts, we observed that computing $E|\text{conv}[X_1, \ldots, X_3, c]|$ is even more difficult than (4) and hence impossible. Nevertheless we will prove the following proposition.

**Proposition 1.** For a tetrahedron $T$ of volume one, $c$ the centroid of a facet of $T$ and $X_1, \ldots, X_3$ uniform random points in $T$, we have that

$$
E|\text{conv}[X_1, \ldots, X_3, c]| < \frac{13}{720} - \frac{\pi^2}{15015} = E|\text{conv}[X_1, \ldots, X_4]|.
$$

A combination of this result with Rademacher’s Lemma 1 yields Theorem 1. The rigorous bound in Proposition 1 is obtained by combining methods from stochastic geometry with results from approximation theory. In the background, first there is a result about the precise approximation of the absolute value function on $[-\frac{1}{3}, \frac{1}{3}]$ by suitable even polynomials, Lemma 2. To apply this in our context, we use an explicit result for all even moments of $|\text{conv}[X_1, \ldots, X_3, c]|$ which — at a first glance maybe surprisingly — is much easier to obtain than just the single first moment.

**Theorem 2.** Let $k \in \mathbb{N}$ and choose three uniform random points $X_1, \ldots, X_3$ in a tetrahedron of volume one. Then it holds:

$$
E|\text{conv}[X_1, \ldots, X_3, c]|^{2k} = \frac{8}{3^{2k-3}} \sum_{\sum k_i = 2k} (-1)^{k''} 3^{k''} \binom{2k}{k_1, \ldots, k_{18}} \prod_{i=1}^3 \frac{l_i! m_i! n_i!}{(l_i + m_i + n_i + 3)!},
$$

where the range of summation and abbreviations are given in (7). The first five even moments are given at the end of Section 2.
This paper is organized in the following way. In Section 2, we give a series representation for even moments of the volume of a random tetrahedron inside a tetrahedron where one point is fixed to be the centroid of a facet, and we use that to find an exact value for the first thirteen even moments. In Section 3, we compute an upper bound for the expected volume of our random tetrahedron, which is a rational affine combination of those even moments. This upper bound suffices to show that the tetrahedron is a counterexample.

As a general reference for results on random polytopes, we refer to the book on Stochastic and Integral Geometry by Schneider and Weil [9]. More recent surveys are due to Hug [3] and Reitzner [8].

2 EVEN MOMENTS OF THE VOLUME OF RANDOM SIMPLICES

Let $T$ be a tetrahedron of volume one and $c = (x_c, y_c, z_c)$ the centroid of one of its facets. For random points $X_1, X_2, X_3 \in T$, we write $X_i = (x_i, y_i, z_i), i = 1, 2, 3$. The volume of the simplex with vertices $X_1, X_2, X_3$ and $c$ is given by

$$|\text{conv}[X_1, \ldots, X_3, c]| = \left| \frac{1}{6} \det \begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_c & y_c & z_c & 1 \end{pmatrix} \right| = 6^{-1}|D(x_1, \ldots, z_c)|,$$

and hence by the absolute value of a polynomial $D$ of degree precisely three in the coordinates of $X_1, X_2, X_3$ and $c$. We are interested in the even moments of $|\text{conv}[X_1, \ldots, X_3, c]|$, where we get rid of the absolute value.

$$E|\text{conv}[X_1, \ldots, X_3, c]|^{2k} = 6^{2k} \int_T \int_T \int_T D(x_1, \ldots, z_c)^{2k} \quad d(x_1, y_1, z_1) \quad d(x_2, y_2, z_2) \quad d(x_3, y_3, z_3).$$

Let $T_o$ be the specific tetrahedron

$$T_o = \{(x, y, z) \in \mathbb{R}^3 : x, y, z \geq 0, x + y + z \leq 1\},$$

i.e., that with vertices $(0, 0, 0), (1, 0, 0), (0, 1, 0)$ and $(0, 0, 1)$. Note that the volume of $T_o$ is $1/6$. We choose $c_o = (1/3, 1/3, 0)$, the centroid of the facet $\{(x, y, 0) \in \mathbb{R}^3 : x, y \geq 0, x + y \leq 1\}$.

Since the expectation $E|\text{conv}[X_1, \ldots, X_3, c]|$ is invariant under volume-preserving affine transformations, we can use as a representative of a tetrahedron of volume one the tetra-
monotonicity of the sample range

Expanding the determinant, the polynomial $D$ by the Multinomial Theorem, and using the multinomial coefficient

\[ \frac{(2k)!}{k_1! \cdots k_{18}!}, \]

the (2k)-th power of it can be rewritten as

\[
D(x_1, \ldots, z_{c_0}) = 3^{-2k} \sum_{k_i=2k} (-1)^{k_i} 3^{k''} \binom{2k}{k_1, \ldots, k_{18}} (x_1 z_2)^{k_1} (x_1 z_3)^{k_2} (x_2 z_1)^{k_3} \\
\quad \times (x_2 z_3)^{k_4} (x_3 z_2)^{k_5} (x_3 z_3)^{k_6} (y_1 z_2)^{k_7} (y_1 z_3)^{k_8} (y_2 z_1)^{k_9} \\
\quad \times (y_2 z_3)^{k_{10}} (y_3 z_2)^{k_{11}} (y_3 z_3)^{k_{12}} (x_1 y_2 z_3)^{k_{13}} (x_1 y_3 z_2)^{k_{14}} \\
\quad \times (x_2 y_3 z_1)^{k_{15}} (x_2 y_3 z_2)^{k_{16}} (x_1 y_2 z_1)^{k_{17}} (x_3 y_2 z_1)^{k_{18}}. \]

\[ (6) \]
Here for abbreviation we use the following notation:

\[ k' = k_2 + k_3 + k_6 + k_7 + k_{10} + k_{11} + k_{14} + k_{15} + k_{18}, \]
\[ k'' = k_{13} + k_{14} + k_{15} + k_{16} + k_{17} + k_{18}, \]
\[ l_1 = k_1 + k_2 + k_{13} + k_{14}, \]
\[ m_1 = k_7 + k_8 + k_{15} + k_{17}, \]
\[ n_1 = k_3 + k_5 + k_9 + k_{11} + k_{16} + k_{18}, \]
\[ l_2 = k_3 + k_4 + k_{15} + k_{16}, \]
\[ m_2 = k_9 + k_{10} + k_{13} + k_{18}, \]
\[ n_2 = k_1 + k_6 + k_7 + k_{12} + k_{14} + k_{17}, \]
\[ l_3 = k_5 + k_6 + k_{17} + k_{18}, \]
\[ m_3 = k_{11} + k_{12} + k_{14} + k_{16}, \]
\[ n_3 = k_2 + k_4 + k_8 + k_{10} + k_{13} + k_{15}. \] (7)

Integration of the monomials over the tetrahedron \( T_o \) gives

\[
\int_{T_o} x^l y^m z^n \, dx \, dy \, dz = \frac{1}{l_1 + 1} B(m_i + 1, l_i + 2) B(n_i + 1, l_i + m_i + 3) \]
\[
= \frac{l_i! m_i! n_i!}{(l_i + m_i + n_i + 3)!}.
\]

Combining this with equations (5) and (6) yields

\[
\mathbb{E}[\text{conv}[X_1, \ldots, X_3, c)]^{2k} =
\]
\[
= \frac{8}{3^{2k-3}} \sum_{k_i = 2k} (-1)^{k'} 3^{k''} \binom{2k}{k_1, \ldots, k_{18}} \prod_{i=1}^{3} \frac{l_i! m_i! n_i!}{(l_i + m_i + n_i + 3)!},
\]

which is Theorem 2. We list the first five even moments of the volume of a random simplex.
Monotonicity of the Sample Range

in a tetrahedron $T$ of volume one:

$$E|\text{conv}[X_1, \ldots, X_3, c]|^2 = \frac{1}{2000} \approx 0.0005,$$

$$E|\text{conv}[X_1, \ldots, X_3, c]|^4 = \frac{43}{27 783 000} \approx 1.54771 \cdot 10^{-6},$$

$$E|\text{conv}[X_1, \ldots, X_3, c]|^6 = \frac{347}{28 805 414 400} \approx 1.20463 \cdot 10^{-8},$$

$$E|\text{conv}[X_1, \ldots, X_3, c]|^8 = \frac{2389}{14 263 395 300 000} \approx 1.67492 \cdot 10^{-10},$$

$$E|\text{conv}[X_1, \ldots, X_3, c]|^{10} = \frac{310483}{310483} \approx 3.44027 \cdot 10^{-12}.$$

3 PROOF OF THEOREM 1

As described in Section 2, the $(2k)$-th moment of $|\text{conv}[X_1, \ldots, X_3, c]|$ can be computed, with fast increasing complexity in $k$. Also note that the volume of a tetrahedron in $T$, where one vertex is fixed to be the centroid of a facet of $T$, is not larger than $1/3$. Hence, we want to approximate the absolute value function in the interval $[-1/3, 1/3]$ by a polynomial

$$P(x) = \sum_{i=0}^{n} a_i x^{2i}$$

for some $n \in \mathbb{N}$ such that $P(x) > |x|$ for all $x \in [-1/3, 1/3]$ or, equivalently, $P(x) > x$ for all $x \in [0, 1/3]$. In contrast to the classical problem of best approximation of $|x|$ by polynomials, we are interested in one-sided approximation and a certain expected value of the polynomial as objective. We use the following standard result for polynomial interpolation.

**Lemma 2.** Let $m \in \mathbb{N}$, $n = 2m + 1$, and $0 < x_0 < \ldots < x_m$ be given. Then the system of equations

$$P(x_j) = x_j \text{ and } P'(x_j) = 1 \quad \text{for } j = 0, \ldots, m$$

determines uniquely a polynomial $P(x) = \sum_{i=0}^{n} a_i x^{2i}$ with the property $P(x) \geq |x|$ for all $x \in \mathbb{R}$.

**Proof.** Let $t_j = x_j^2$ and consider the standard Hermite interpolation problem

$$Q(t_j) = f(t_j) \text{ and } Q'(t_j) = f'(t_j) \quad \text{for } j = 0, \ldots, m$$

for the functions $f(t) = \sqrt{t}$ and $Q(t) = \sum_{i=0}^{n} a_i t^i$. The condition $P'(x_j) = 1$ is equivalent to $Q'(t_j) = 1/(2\sqrt{t_j}) = f'(t_j)$. Then the interpolation error fulfills, for some $\xi \in [t_0, t_m]$, the estimate

$$f(t) - Q(t) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{j=0}^{m} (t - t_j)^2 < 0,$$

where the last inequality follows from $f^{(n+1)}(t) < 0$ for all $t$. \qed
We note in passing that for even \( n = 2m \) and \( 0 < x_0 < \ldots < x_m = 1/3 \), we only require the simple interpolation condition \( P(x_m) = x_m \) in the last point and get \( P(x) \geq |x| \) for all \( x \in [-x_m,x_m] \).

Our aim is to approximate \( E|\text{conv}[X_1,\ldots,X_3,c]| \) from above by an even polynomial \( P \) of degree \( 2n \),

\[
E|\text{conv}[X_1,\ldots,X_3,c]| \leq E P(|\text{conv}[X_1,\ldots,X_3,c]|),
\]

which holds if \( |x| \leq P(x) \) on \([-\frac{1}{3}, \frac{1}{3}] \). Moreover, the best polynomial for fixed \( n \in \mathbb{N} \) can be found via the linear optimization problem

\[
\min_P E P(|\text{conv}[X_1,\ldots,X_3,c]|) = \min_{a_i} \sum_{i=0}^{n} a_i E P(|\text{conv}[X_1,\ldots,X_3,c]|)^{2i}
\]

\[\text{s.t. } P(x) \geq x, \; x \in \left[0, \frac{1}{3}\right].\]

Please note that the constraint is infinite dimensional. Relaxing the constraint, we get a lower bound on \( E P(|\text{conv}[X_1,\ldots,X_3,c]|) \) via the finite dimensional linear program

\[
\min_P E P(|\text{conv}[X_1,\ldots,X_3,c]|) \quad \text{s.t. } P(x_\ell) \geq x_\ell, \; x_\ell \in \left[0, \frac{1}{3}\right], \; \ell = 0, \ldots, L.
\]

For \( n = 12 \) and \( L = 100 \) equidistant points \( x_\ell \in [0,1/3] \), we numerically compute via Matlab and the optimization toolbox CVX \([2]\)

\[
E P(|\text{conv}[X_1,\ldots,X_3,c]|) > 0.01746,
\]

yielding that we do not get a sufficiently precise estimate using only \( n = 12 \) even moments.

For \( n = 13 \) and \( L = 1000 \), we solved the above linear program, computed the interpolation nodes with the absolute value function numerically, and rationalized these points to

\[
\{ x_j : j = 0, \ldots, 6 = m \} = \left\{ \frac{1}{83}, \frac{1}{22}, \frac{1}{11}, \frac{2}{15}, \frac{2}{11}, \frac{5}{22}, \frac{4}{15} \right\}.
\]

Using these points for the interpolation problem in Lemma \([2]\) gives an even polynomial \( P_{\text{cert}}(x) = \sum_{i=0}^{13} a_i x^{2i} \) of degree 26 with explicitly given rational coefficients \( a_0, \ldots, a_{13} \) and the property |\( x | \leq P_{\text{cert}}(x) \). Finally, we use the even moments computed in Section \([2]\) to complete the proof of Theorem \([1]\)

\[
E|\text{conv}[X_1,\ldots,X_3,c]| \leq E P_{\text{cert}}(|\text{conv}[X_1,\ldots,X_3,c]|) = \sum_{i=0}^{13} a_i E|\text{conv}[X_1,\ldots,X_3,c]|^{2i}
\]

\[
< \frac{13}{720} \cdot \left( \frac{\pi}{15.015} \right)^2 = E|\text{conv}[X_1,\ldots,X_4]| = 0.0174600901532997.
\]
References

[1] C. Buchta and M. Reitzner, The convex hull of random points in a tetrahedron: Solution of Blaschke’s problem and more general results. J. reine angew. Math. 536 (2001), 1–29.

[2] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, version 2.1., http://cvxr.com/cvx 2014.

[3] D. Hug, Random polytopes. In: Spodarev, E. (ed.): Stochastic Geometry, Spatial Statistics and Random Fields. Lecture Notes in Mathematics 2068, pp. 205–238, Springer, Heidelberg 2013.

[4] M. Meckes, Monotonicity of volumes of random simplices. In: Recent Trends in Convex and Discrete Geometry, 2006.

[5] R. E. Miles, Isotropic random simplices. Adv. in Appl. Probab. 3 (1971), 353–382.

[6] L. Rademacher, On the monotonicity of the expected volume of a random simplex. Mathematika 58 (2012), 77–91.

[7] B. Reichenwallner and M. Reitzner, On the monotonicity of the moments of volumes of random simplices. Mathematika 62 (2016), 949–958.

[8] M. Reitzner, Random polytopes. In: Kendall W.S. and Molchanov I. (eds.): New perspectives in stochastic geometry. pp. 45–76, Oxford Univ. Press, Oxford, 2010.

[9] R. Schneider and W. Weil, Stochastic and integral geometry. Probability and its Applications (New York), Springer-Verlag, Berlin, 2008.

[10] A. Zinani, The Expected Volume of a Tetrahedron whose Vertices are Chosen at Random in the Interior of a Cube. Monatsh. Math. 139 (2003), 341–348.

Stefan Kunis, Matthias Reitzner
Institut für Mathematik
Universität Osnabrück
Albrechtstr. 28a
49076 Osnabrück, Germany
e-mail: stefan.kunis, matthias.reitzner@uni-osnabrueck.de

Benjamin Reichenwallner
Fachbereich Mathematik
Universität Salzburg
Hellbrunnerstraße 34
5020 Salzburg, Austria
e-mail: benjamin.reichenwallner@sbg.ac.at