Asymptotic Properties of Random Polytopes

The King Himself

July 30, 2015

Abstract

A random polytope is the convex hull of points chosen randomly in $\mathbb{R}^d$ according to some probability distribution. We examine several properties of the asymptotic behavior of random polytopes in two settings: (1) when the points are chosen uniformly and at random from a convex body K in the plane, and (2) when the points are chosen according to the standard normal distribution on the plane. In both cases, we provide insight into the asymptotics of the expected number of sides of the random polytope in the two-dimensional case. We then delve into the concept of affine perimeter and explain its relation to the question: what is the probability that n random points (chosen according to (1) or (2)) lie in convex position? We explain key parts of the computation for (1) and the barriers to extending these methods to (2), and we propose some methods which may prove fruitful in overcoming these obstacles.

1 Introduction

In [8], J. J. Sylvester discussed a class of questions, all of which pertained to "determining the chance that a system of points, each with its own specific range, shall satisfy any prescribed condition of form." Specifically, his original problem was to compute the probability that four points, chosen at random from the plane, are the vertices of a convex quadrilateral. As Sylvester himself acknowledged, this problem is not well defined, as the answer will depend on the probability distribution of the points. The question was later refined and has been the subject of a fair amount of research. Rather than focusing on exact probabilities for a specific number of points, we will instead focus on the asymptotic behavior of these probabilities and related quantities. That is, we will study what happens as the number of points becomes arbitrarily large.

Our first question concerns the expected number of sides of a random polygon. We explain the work of Rényi and Sulanke [6], who deteremined an asymptotic formula for the expected number of sides in the case that the points are chosen uniformly from a compact convex set with a non-empty interior, also known as a convex body. We then compute the same quantity for the case that the points are chosen according to the standard normal distribution on the plane.

In the final section, we turn our attention to a question more closely related to Sylvester's original problem. Following the method of Bárány in [3], we find the asymptotic behavior of the probability that n points chosen uniformly at random from a convex body are in convex position. (We refer to this as the "tail probability.") We also explore the concept of affine perimeter and its relation to the tail probability. We conclude with a discussion of the tail probability for the Gaussian case: we address the obstacles to extending Bárány's methods to the normal distribution, as well as provide suggestions for research which may help in finding a solution to the problem.

2 Computing $\mathbb{E}[V_n]$ in a convex body K

In [6], Rényi and Sulanke show how to compute the expected number of sides of a random polygon when the points are chosen uniformly at random from a convex body K. Their original paper was written in German, and to the best of the author's knowledge, it has not been translated into English. Their proof contains techniques which prove useful in other scenarios, and it is interesting in its own right, so we provide a translated version here. We address the case in which K is a convex polygon. For a translation of the case where the boundary of K is smooth and has positive curvature at all points, see [5].

K is a convex polygon with vertices $A_1, \ldots, A_r$ and $\theta_1, \ldots, \theta_r$ corresponding angles. Denote a_k to be the length of side $A_k A_{k+1}$, and F is the area of K. We choose points $P_1, \ldots, P_n$ uniformly and at random from K. Denote their convex hull by H_n, and the number of vertices of H_n by V_n.

Let $\epsilon_{ij} = 1$ when $i \neq j$ and the line segment $P_i P_j$ is one of the sides of H_n, and $\epsilon_{ij} = 0$ otherwise. Then we have

$$V_n = \sum \epsilon_{ij}.$$

Since the P_i are i.i.d, the probability

$$W_n = P(\epsilon_{ij} = 1)$$

is the same for all pairs i, j. There are $\binom{n}{2}$ such pairs, so taking the expectation of our first equation gives

$$E_n = E(V_n) = \binom{n}{2} W_n.$$

As before, for two points P_1, P_2, let F_1 be the area of the smaller part of K which is cut off by the line $P_1 P_2$. Then we get

$$W_n = \frac{1}{F^2} \int_K \int_K \left[\left(1 - \frac{F_1}{F}\right)^{n-2} + \left(\frac{F_1}{F}\right)^{n-2} \right] dP_1\, dP_2.$$

Since $F_1/F \leq 1/2$, we have

$$\frac{1}{F^2} \int_K \int_K \left(\frac{F_1}{F}\right)^{n-2} dP_1\, dP_2 \leq \frac{1}{2^{n-2}},$$

and so

$$E_n \sim \binom{n}{2} \frac{1}{F^2} \int_K \int_K \left(1 - \frac{F_1}{F}\right)^{n-2} dP_1\, dP_2.$$

Denote the area of triangle $A_{i-1}A_iA_{i+1}$ by f_i and set $f = \min\{f_i\}$. We can split our integration into three cases. We denote by C_i the set of points P_1, P_2 for which the line P_1P_2 intersects sides $A_{i_1}A_i$ and A_iA_{i+1}, and D_i for the set where P_1P_2 intersects $A_{i-1}A_i$ and $A_{i+1}A_{i+2}$. For all other choices of points, we have $1 - F_1/F \leq 1 - f/F$. Thus we have

$$E_n \sim \binom{n}{2} \frac{1}{F^2} \left(\sum_{i=1}^{r} (I_i + J_i)\right),$$

where

$$I_i = \int_{C_i} \left(1 - \frac{F_1}{F}\right)^{n-2} dP_1\, dP_2$$

and

$$J_i = \int_{D_i} \left(1 - \frac{F_1}{F}\right)^{n-2} dP_1\, dP_2.$$

Let Q_1 be the intersection of line P_1P_2 with side $A_{i-1}A_i$ and Q_2 be its intersection with side A_iA_{i+1}, and define G_{ab} to be the set of points P_1, P_2 for which $|A_iQ_1| < a$ and $|A_iQ_2| < b$. An elementary calculation yields

$$\int_{G_{ab}} dP_1\, dP_2 = \frac{a^2b^2 \sin^2\theta_i}{12}.$$

Now, the area of the triangle $Q_1A_iQ_2$ cut by line P_1P_2 with side lengths a, b (the sides which are part of the perimeter of K) is $1/2ab\sin\theta_i$, so we get

$$I_i = \int_0^{a_{i-1}} \int_0^{a_i} \left(1 - \frac{ab\sin\theta_i}{2F}\right)^{n-2} \sin^2\theta_i \frac{ab}{3}\, da\, db,$$

where a_i is the length of side A_iA_{i+1}. We now perform a change of variables. Set

$$X = a\sqrt{\frac{\sin\theta_i}{2F}}, \qquad Y = b\sqrt{\frac{\sin\theta_i}{2F}}.$$

This yields

$$I_i = \frac{4F^2}{3} \int_0^{X_i} \int_0^{Y_i} (1 - XY)^{n-2} XY\, dX\, dY,$$

where

$$X_i = a_{i-1}\sqrt{\frac{\sin\theta_i}{2F}}, \qquad Y_i = a_i\sqrt{\frac{\sin\theta_i}{2F}},$$

and also note that $\rho_i = X_i Y_i = \frac{a_{i-1} a_i \sin\theta_i}{2F} = \frac{f_i}{F} < 1$. Now we have

$$\int_0^{X_i} \int_0^{Y_i} (1-XY)^{n-2} XY \; dX \; dY = \int_0^{X_i} \int_0^{Y_i} (1-XY)^{n-2} dX dY - \int_0^{X_i} \int_0^{Y_i} (1-XY)^{n-1} \; dX \; dY.$$

We can now compute the RHS:

$$\int_0^{X_i} \int_0^{Y_i} (1-XY)^{n-1} \; dX \; dY = \int_0^{X_i} \frac{1-(1-XY_i)^n}{nX} \; dX$$

$$= \frac{1 + \frac{1}{2} + \cdots + \frac{1}{n} - \sum_{k=1}^n \frac{(1-\rho_i)^k}{k}}{n}.$$

Thus, we arrive at

$$I_i = \frac{4F^2}{3} \left[\frac{\frac{1}{2} + \cdots + \frac{1}{n}}{n(n-1)} - \frac{\sum_{k=1}^{n-1} \frac{(1-\rho_i)^k}{k}}{n(n-1)} + \frac{(1-\rho_i)^n)}{n^2} \right].$$

We now set

$$S_i = \sum_{k=1}^{\infty} \frac{(1-\rho_i)^k}{k} = \log \frac{1}{\rho_i},$$

and so we obtain

$$E_n = \frac{2}{3}(\gamma - 1 + \log n) r - \frac{2}{3} \sum_{i=1}^{r} S_i + o(1) + \frac{1}{F^2} \binom{n}{2} \sum_{i=1}^{r} J_i,$$

where γ is Euler's constant. (That is, $\gamma = \lim_{n\to\infty}(\sum_{k=1}^n 1/k - \log n)$.)

It remains to compute the integrals J_i. We extend sides $A_{i-1}A_i$ and $A_{i+1}A_{i+2}$ to their intersection at A_i^*. (The case that these lines are parallel can be treated similarly.) Define $a' = |A_i A_i^*|$, $b' = |A_{i+1} A_i^*|$, $a = |A_i Q_1|$, $b = |A_{i+1} Q_2|$, and G_{ab} as before. In this case, we again perform an elementary calculation to obtain

$$\int_{G_{ab}} dP_1 \; dP_2 = \frac{\sin^2 \gamma_i}{12} (a^2 + 2aa')(b^2 + 2bb'),$$

where γ_i is the angle at A_i^*. Thus we have

$$J_i = \int_{D_i} \left(1 - \frac{(ab + a'b + ab')\sin\gamma_i}{2F}\right)^{n-2} \frac{\sin^2\gamma_i}{3} (a+a')(b+b') \; da \; db.$$

Asymptotically, therefore, we have

$$J_i \sim \int_{D_i} \left(1 - \frac{(ab + a'b + ab')\sin\gamma_i}{2F}\right)^{n-2} \frac{\sin^2\gamma_i}{3} a'b' \; da \; db.$$

We now use the tranformation

$$a = \frac{F}{b' \sin \gamma_i} X, \qquad b = \frac{F}{a' \sin \gamma_i} Y$$

which yields

$$J_i \sim \frac{F^2}{3} \int_0^{a'_{i-1}} \int_0^{a'_{i+1}} \left(1 - \frac{X+Y}{2}\right)^{n-2} dX\, dY$$

$$= \frac{4F^2}{3n(n-1)} \left[1 - \left(1 - \frac{a'_{i-1}}{2}\right)^n - \left(1 - \frac{a'_{i+1}}{2}\right)^n + \left(1 - \frac{a'_{i-1} + a'_{i+1}}{2}\right)^n\right],$$

where

$$a'_{i-1} = \frac{a_i b' \sin \gamma_i}{F}, \qquad a'_{i+1} = \frac{a_{i+2} a' \sin \gamma_i}{F}.$$

Finally, this leads us to

$$E_n = \frac{2r}{3}(\log n + \gamma) - \frac{2}{3} \sum_{i=1}^{r} S_i + o(1).$$

3 Computing $\mathbb{E}[V_n]$ for the normal distribution

We now examine the case where the n points $P_i = (x_i, y_i)$ are chosen according to the probability density

$$f(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2},$$

i.e. the standard normal distribution on the plane. Denote their convex hull by H_n, and denote the number of vertices of H_n by V_n. Similar to the previous case, let $\epsilon_{ij} = 1$ if all of the $n-2$ points other than P_i and P_j lie on one side of the line between P_i and P_j, and 0 otherwise. Then note that $\epsilon_{ij} = 1$ precisely when the line segment between P_i and P_j is a side of H_n. Since the number of vertices of a polygon is equal to the number of sides in two dimensions, we have

$$V_n = \sum_{1 \le i < j \le n} \epsilon_{ij} \Rightarrow \mathbb{E}[V_n] = \sum_{1 \le i < j \le n} \mathbb{E}[\epsilon_{ij}].$$

Furthermore, since the points are identically distributed, we have $\epsilon_{ij} = \epsilon_{12}$ for all i, j. Also, since the ϵ_{ij} are indicator random variables, their expectation is the probability of the event they indicate. Combining these facts yields

$$\mathbb{E}[V_n] = \binom{n}{2} \mathbb{P}(P_3, \ldots, P_n \text{ lie on one side of line } \overleftrightarrow{P_1 P_2}).$$

This motivates us to find the distribution for lines generated by picking two points according to the original distribution. Let the two points which will determine the line be the random variables $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, and parameterize the resulting line by θ, p, where the line is given by

$$\overleftrightarrow{P_1P_2} = \{\mathbf{x} \in \mathbb{R}^2 \ : \ \mathbf{x} \cdot (\cos\theta, \sin\theta) = p\}.$$

Write

$$P_1 = p(\cos\theta, \sin\theta) + s(-\sin\theta, \cos\theta), \quad P_2 = p(\cos\theta, \sin\theta) + t(-\sin\theta, \cos\theta).$$

(That is, we perform a change of variables in which our new axes are in the direction of the line $\overleftrightarrow{P_1P_2}$ and the normal to this line.) In [7], it is shown that

$$dx_1\ dy_1\ dx_2\ dy_2 = |s-t| ds\ dt\ dp\ d\theta$$

and thus

$$f(P_1)f(P_2)dP_1\ dP_2 = \left(\frac{1}{2\pi}e^{-p^2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-s^2/2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\right)|s-t| ds\ dt\ dp\ d\theta.$$

As only p and θ are relevant to our final calculation, we can integrate the s and t variables out. Note that s and t are both standard normal, thus their difference is normally distributed with mean 0 and variance 2. We compute $\mathbb{E}[|Z|]$ for $Z \sim \mathcal{N}(0, 1)$:

$$\begin{aligned}
\mathbb{E}[|Z|] &= \int_{-\infty}^{\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= 2\int_0^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-u} du \\
&= \sqrt{\frac{2}{\pi}}
\end{aligned}$$

(Here we have employed the substitution $u = \frac{1}{2}z^2$.) Since $\sqrt{2}Z \sim \mathcal{N}(0, 2)$, arrive finally at

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{1}{2\pi}e^{-p^2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-s^2/2}\right)|s-t| ds\ dt = \frac{2}{\sqrt{\pi}}.$$

Thus, our distribution on lines is

$$\pi^{-3/2}e^{-p^2} dp\ d\theta.$$

Now, we need to find the normal measure of the two half-planes generated by any particular line. Due to the rotational symmetry of the normal distribution, we can assume that $\theta = 0$. Then the left half-plane L for the line $l(p, 0)$ has measure

$$\begin{aligned}\int_L f(P)dX &= \int_{-\infty}^{p} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy\right) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \int_{-\infty}^{p} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \Phi(p)\end{aligned}$$

where Φ is the cdf of the standard normal distribution. It follows that the measure of the right half plane is $1 - \Phi(p)$. Since the remaining $n - 2$ points are chosen independently, the probability of the event that they all lie in the left half-plane is $\Phi(p)^{n-2}$. Similarly, the probability that they all lie in the right half-plane is $(1-\Phi(p))^{n-2}$. Since these events are disjoint, the total probability is

$$\Phi(p)^{n-2} + (1 - \Phi(p))^{n-2}.$$

The last step is to integrate over all possible choices of p and θ:

$$\begin{aligned}\mathbb{E}[V_n] &= \binom{n}{2} \int_0^{\infty} \int_0^{2\pi} \pi^{-3/2} e^{-p^2} \left(\Phi(p)^{n-2} + (1 - \Phi(p))^{n-2}\right) d\theta \, dp \\ &= \frac{2}{\sqrt{\pi}} \binom{n}{2} \int_0^{\infty} \left(\Phi(p)^{n-2} + (1 - \Phi(p))^{n-2}\right) e^{-p^2} dp.\end{aligned}$$

Now, note that since $(1-\Phi(p)) \leq 1/2$ for $p \geq 0$, the $(1-\Phi(p))^{n-2}$ term will vanish asymptotically. Furthermore, since

$$\int_0^1 \Phi(p)^{n-2} e^{-p^2} dp \leq \Phi(1)^{n-2} \to 0$$

as $n \to \infty$, we have

$$\mathbb{E}[V_n] \sim \frac{2}{\sqrt{\pi}} \binom{n}{2} \int_1^{\infty} \Phi(p)^{n-2} e^{-p^2} dp.$$

Setting $p = \sqrt{2}z$, we have

$$\int_1^{\infty} \Phi(p)^{n-2} e^{-p^2} dp = \sqrt{2} \int_1^{\infty} \phi(z)^{n-2} e^{-2z^2} dz,$$

where ϕ is the cdf for the normal distribution $\mathcal{N}(0, 1)$ (and therefore $\phi(p) = \Phi(\sqrt{2}p)$). We can now compute the asymptotic growth rate of this final integral using lemma 2.1 from [4], which states that

$$\int_1^{\infty} \phi(z)^{\beta-\alpha} z^s e^{-\alpha z^2} dz \sim \Gamma(\alpha) 2^{\alpha-1} \pi^{\alpha/2} \beta^{-\alpha} (\log \beta)^{(\alpha+s-1)/2}.$$

In this case, $\alpha = 2$, $\beta = n$, and $s = 0$, which yields the final result

$$\mathbb{E}[V_n] \sim 2\sqrt{2\pi \log n}.$$

Note: Though we created this proof independently, the author recently discovered that this method is very similar to the one used by Rényi and Sulanke in [6]. However, there are slight differences in the two proofs, and Rényi and Sulanke's paper is in German, so we have chosen to still include this section.

4 Tail probabilties

In [3], Imre Bárány proved a connection between the probability

$p(n, K) := \mathbb{P}(n$ points chosen uniformly at random in K are in convex position)

and the affine perimeter of K. Specifically, he showed that

$$\lim_{n\to\infty} n^2(p(n, K))^{1/n} = \frac{1}{4}e^2 A^3(K),$$

where $A(K)$ is the supremum of the affine perimeters of all convex subsets $S \subseteq K$. Rather than repeating all of the technical details, we present an outline of his proof to help the reader understand the main ideas of the proof.

The following definition is due to Bárány [1]:

Definition 4.1. Let S be a convex body in $\mathbb{R}^2$ and choose a subdivision $x_1, \ldots, x_m, x_{m+1} = x_1$ of the boundary ∂S and lines l_i supporting S at x_i for all $i = 1, \ldots, m$. Denote the intersection of l_i and l_{i+1} by y_i. (If $l_i = l_{i+1}$, then set y_i to be any point between x_i and x_{i+1}. Then if T_i is the area of the triangle with vertices x_i, y_i, x_{i+1}, the **affine perimeter** of S is defined as

$$\text{AP}(S) = 2 \lim \sum_{i=1}^{m} (T_i)^{1/3},$$

where the limit is taken over a sequence of subdivisions with $\max_i |x_i - x_{i+1}| \to 0$. The limit exists and is unique since the sum is clearly bounded below by 0, and each refinement of the subdivision of the boundary decreases the sum.

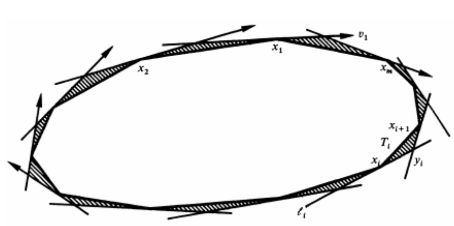

Figure 1: Figure from [1]

Refer to Figure 1. The shaded regions are the triangles mentioned in the definition. To see the connection to the tail probability, suppose that the first $m < n$ points that we have chosen are the points $x_1, \ldots, x_m$. (Clearly, if all n points are to be in convex position, these m points must be in convex position as well, so the only scenario is one of the sort depicted in the diagram.) Then when the $m + 1$'th point is chosen, it must fall in one of the shaded regions in order for $m + 1$ points to remain in convex position. As Bárány shows in [1], given that the n points are in convex position, they will tend towards a limit shape $K_0 \subset K$. (In fact, Bárány shows that K_0 is the unique subset of K which has maximal affine perimiter.) Because of this, the shaded regions in which the new points can be chosen will approximate those which arise in the definition of affine perimiter for the region K_0. The next point chosen approximates a refinement of the boundary. This is the intuitive reason that affine perimeter and tail probabilities are linked.

There is a particular portion of Bárány's argument which deserves a detailed explanation, simply due to its combinatorial beauty. Rather than "refreshing" the triangles in which we can choose a point after each one is added, we fix a set of triangles with our first choice of points, and then find the probability that the remaining points lie in convex position within these triangles. We have the following definition:

Definition 4.2. Let T be a triangle, and let P_0 and P_{k+1} be two vertices of T. We say that the points $P_1, \ldots, P_k$ form a **convex chain** in T if $P_i \in T$ and the points are in convex position.

Using this terminology, we wish to find the probability that k points chosen uniformly in a triangle are in convex position. We follow the proof of theorem 1 from [2]. First, we assume that the $P_i = (x_i, y_i)$ are placed uniformly in the unit square and find the probability that they form a convex chain in the triangle with vertices $(0, 0)$, $(1, 0)$, and $(1, 1)$. Relable the points so that the x_i are increasing, and note that, in order for the points to form a convex chain, the y_i must be nondecreasing. For any relative ordering of the y_i, with probability 1, exactly one of these is nondecreasing. Since all relative orderings are equally likely due to the uniformity of the distribution, the y_i are nondecreasing with probability $1/k!$. In this case, we say that the chain is monotone.

Now, conditioned on the fact that we have a monotone chain, we claim that all permutations of the slopes of the segments Δ_i are equally likely. Let M be the set of all monotone chains. Then there is a bijection between the events in M and the set

$$D = \{(\Delta_1, \ldots, \Delta_k) \, : \, \Delta_i = (u_i, v_i), u_i, v_i \geq 0, \sum \Delta_i = (1, 1)\},$$

i.e. the set of all sequences of slopes for a monotone chain. For a monotone chain with slopes $\{\Delta_i\}$, permute the slopes Δ_i and Δ_{i+1}.

Referring to Figure 2, we see that the vertices $P_0 \ldots, P_{i-1}, P_{i+1}, \ldots, P_{k+1}$ remain fixed, while vertex P_i is reflected through the midpoint of line segment $P_{i-1}P_{i+1}$. Since P_i is uniform in the rectangle indicated in the figure (with

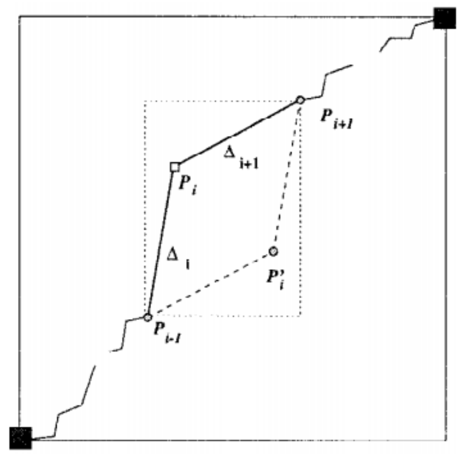

Figure 2: Figure from [2]

corners at P_{i-1} and P_{i+1}), these two choices of points are equally likely. Thus, the permuted slopes are equally likely. In order for the P_i to be in convex position, the slopes Δ_i must be nonincreasing. So, with probability 1, exactly one of the relative slope orderings gives us a convex chain. Thus, conditioned on having a monotone chain, the probability that we have a convex chain is $1/(k+1)!$. This gives a total probability of

$$\frac{1}{k!(k+1)!}$$

for the square. To find the probability for a triangle, we multiply by a factor of 2^k. This accounts for the fact that the points cannot lie in the "wrong half" of the square; all other factors remain the same. Since this probability is invariant under nondegenerate affine transformations, the probability that k points form a convex chain in any nondegenerate triangle (subject to the uniform measure) is

$$\frac{2^k}{k!(k+1)!}.$$

After some more lengthy computation, Bárány is able to show that

$$\lim_{n\to\infty} n^2 p(n,k)^{1/n} = \frac{1}{4}e^2 A^3(K),$$

where $A(K)$ is the supremum of affine perimeters taken over all convex sets $S \subset K$.

Motivated by this result, we wish to determine whether or not these techniques can yield similar formulas in the case that the points are chosen according to different log-concave probability distributions on the plane. The normal distribution is of particular interest. This may allow us to define a notion analogous to affine perimeter for a probability distribution.

There are several problems which need to be addressed at this juncture. First, in order for any of the previously mentioned techniques to apply, the points must approach some limit shape. Intuitively, if such a limit shape exists, by the symmetry of the Gaussian distribution, it seems that the limit shape must be a circle whose radius grows as a function of n. We may try rescaling the normal distribution by the inverse of this radius, but we again run into problems since Bárány's convex chain argument requires that the distribution is uniform. A feasible work-around is as follows. Suppose that the points tend to be spaced around the hypothetical limit circle in such a way that the resulting triangles in which we may have convex chains tend to be relatively flat–that is, all of the points in a given triangle lie within a small range of radii from the origin, say $[r-\epsilon, r+\epsilon]$ for ϵ suitably small. Then the probability density in these triangles will be approximately uniform, and the convex chain argument may still apply. The rest of the techniques used in the proof should be extendable without too much trouble.

Of course, should our guess about the points approaching some limit shape prove incorrect, then this line of reasoning will not work. For this reason, and because it is interesting in and of itself, we ask the question: what is the distribution of n points, chosen according to the standard normal distribution, conditioned on the fact that they lie in convex position? A naive numerical approach–rejection sampling–will not work here, as the desired event occurs with exponentially small probability in n. Thus, it quickly becomes infeasible to simulate the distribution as n grows. Some other method will be needed, and this seems a fruitful direction for further research.

References

[1] Imre Bárány. "Affine perimeter and limit shape." *Journal für die reine und angewandte Mathematik* (1997), Issue 484. 71-84.

[2] I. Bárány, G. Rote, W. Steiger, and C.-H. Zhang "A Central Limit Theorem for Convex Chains in the Square." *Discrete Comput. Geom.* (2000), Vol. 23. 3550.

[3] Imre Bárány. "Sylvester's question: the probability that n points are in convex position." *The Annals of Probability* (1999), Vol. 27, No. 4. 2020-2034.

[4] Daniel Hug, Götz Olaf Munsonius and Matthias Reitzner. "Asymptotic mean values of Gaussian polytopes." *Adv. in Appl. Probab.* (2005), Vol. 37, No. 2. 297-320.

[5] "Random Convex Polygons I." *Liviu's Math Blog* (2014). <http://liviusmathblog.blogspot.com/2014/03/random-convex-polygons-i.html>.

[6] A. Rényi and R. Sulanke. "Über die konvexe Hülle von n zufällig gewählten Punkten." *Z. Wahrsch. Verw. Geb.* (1963), Vol. 2. 7584.

[7] Sheree T. Sharpe. "Random hyperperplanes of a convex body, Sylvester's problem and Crofton's formula." The University of Georgia (2007).

[8] J. J. Sylvester. "On a Special Class of Questions on the Theory of Probabilities." *Birmingham British Assoc. Rept.* (1865). 8-9.