GEOMETRIC PROBABILITY AND INTEGRAL GEOMETRY

MEHANA ELLIS

Abstract. In this paper, I hope to cover some interesting topics, theorems, etc. in Geometric Probability/Integral Geometry. We will begin by reviewing the famous Buffon's needle problem, then learn about Poincare's Theorem/Formula, Cauchy's Formula, Crofton's Formula, kinematic density, and Santalo's theorem. I shall assume knowledge of basic geometric probability and a bit of integral geometry in this paper.

1. Buffon's Needle

We will begin by recalling the famous Buffon's Needle problem. This problem is a great first look at geometric probability, because it employs the main principals of expected value. First, let's review the main lemma regarding linearity of expectation:

Lemma 1.1. *Let X and Y be arbitrary variables, and let $c \in \mathbb{R}$. Then, we have that*

- $\mathbb{E}(cX) = c\mathbb{E}(X)$,
- $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

Buffon's needle is centred around this notion of expected value in geometric probability.

Theorem 1.2. *Suppose we have a floor made of parallel, equal-width pieces of wood. If we drop a needle onto the floor, then what is the probability that the needle will hit the line separating two different pieces?*

Proof. The needle is of the same width as the space between planks. Therefore, the value of X, where X is the number of lines the needle crosses, must either be 1 or 0. Take N needles. The number of crossings depends on N, not on the arrangement (combined, straight or bent) of the needles. Because of this, we may create a circle of circumference N and radius $R = \frac{N}{2\pi}$ out of these N needles![1] If we drop this circle, it will hit $4R$ lines ($2R$ on the left, $2R$ on the right). This means that the expected number of crossings is $4R = \frac{2N}{\pi}$, which is the same as $N \times \mathbb{E}(X)$. Therefore, we have that

$$\mathbb{E}(X) = \frac{2}{\pi}.$$

□

Remark 1.3. A trivial example of the problem is when we want to find an expected value using two needles. Let X_1 and X_2 be the number of lines they cross. We have that

$$\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = 2\mathbb{E}(X).$$

This means that the expected number of crossings is $2\mathbb{E}(X)$. Note that we get the same thing when we glue X_1 and X_2 together (in any way, including straight and bent).

Date: August 16, 2020.

[1] Okay, not quite...but very close, which works fine.

We can now state this result more formally.

Theorem 1.4. *Suppose we have a floor made of slats with width w. We drop a needle of length l on the floor. The probability P of the needle crossing a line is*

$$P = \frac{2l}{\pi w}.$$

2. Poincaré's Formula–Measuring lines which meet a curve

First, let's get a bit of background information before looking at Poincaré's Formula (note that unless otherwise stated, we will be using definitions from [Tre08]).

Definition 2.1. Loosely speaking, a function is called *smooth* if it is continuous/differentiable everywhere. If its domain has derivatives of different orders, we say the function is a $\mathcal{C}^\infty$ function. We say a function is of class $\mathcal{C}^k$ if there exist derivatives $f', f'', \ldots, f^k$ which are continuous.

The next definition won't really be used in our proof of Poincaré's Formula, but it is good to know and we shall use it later when looking at Poincaré's Formula for intersecting curves.

Definition 2.2. The *kinematic measure* is a measure of a set of lines in (p, θ) coordinates which is invariant under rigid motions. It is given by

$$dK = dp \wedge d\theta,$$

where $\wedge$ represents the wedge/exterior product.[2]

Let C be a piecewise $\mathcal{C}^1$ curve, and let L be some line in the plane. Denote the number of intersection points as $n(L \cap C)$. If L agrees with a segment that is contained in C, then we have that $n(C \cap L) = \infty$. If this is the case for C, we say that the set of lines where $n = \infty$ has a dK-measure of 0. Now we can state and prove Poincaré's Formula:

Theorem 2.3 (Poincaré's Formula)**.** *Let C be a piecewise $\mathcal{C}^1$ curve. The measure of unoriented lines which meet C is*

$$2L(C) = \int_{\{L: L \cap C \neq \varnothing\}} n(C \cap L) dK(L).$$

Proof. Let C be a $\mathcal{C}^1$ differentiable curve $Z(s) = (x(s), y(s))$. Then there exist $x(s), y(s) \in \mathcal{C}^1[0, s_0]$, where the tangent vector[3] $\dot{Z} = (\dot{x}, \dot{y})$ satisfies $\dot{Z} = 1$. We can now integrate this type of curve using the above information.

Now let's define a *flag* as the set of pairs (L, Z), where L is a line (determined by coordinate pair (p, θ)) and Z is a point on L. A subset S of the flag is the set of lines and points on these lines which touch C:

$$S = \{(L, Z); L \cap C \neq \varnothing,\ Z \in L \cap C\}.$$

The point Z is determined by an arclength coordinate q along the line from $(p \cos\theta, p \sin\theta)$:

$$\int_{\{L: L \cap C \neq \varnothing\}} n dK = \int_{\{L: L \cap C \neq \varnothing\}} \left(\sum_{Z \in L \cap C} 1 \right) dK.$$

[2]For example, the wedge product $u \wedge v$ is the square matrix $u \otimes v - v \otimes u$.

[3]The vector tangent to the curve/surface at a given point.

Note that the subset S is determined by the point $(x, y) = Z$, and L may be an unoriented line through Z with angle $0 \leq \eta < \pi$. This means we can replace the coordinates (p, θ) with (s, η), respectively for the two values. So, if $(\dot{x}, \dot{y}) = (\cos\phi(s), \sin\phi(s))$, we have

$$\tilde{p} = x(s)\cos\eta + y(s)\sin\eta.$$

Now we just need to keep replacing things, and eventually we get the following equality upon simplification:

$$d\tilde{p}d\eta = |\cos\phi(s) - \eta| ds\ d\eta.$$

Now, let's simplify, and we'll get what we want:

$$\begin{aligned}\int_{\{L:L\cap C\neq\varnothing\}} \left(\sum_{Z\in L\cap C} 1\right) dK &= \int_{\{Z:Z\in C\}} \int_{\{L:Z\in L\}} d\tilde{p}\ d\eta \\ &= \int_0^{s_0} \int_0^{\pi} |\cos(\phi(s) - \eta)| d\eta\ ds \\ &= 2\int_C ds \\ &= 2L(C).\end{aligned}$$

This completes the proof of Poincaré's Formula for lines. [Tre08] □

There is another interesting formula of Poincaré, similar to the latter but for intersecting curves. I encourage the reader to look at this in Treiburg's slides.

3. Conditional probability and Sylvester's Problem

Recall that a set $\Omega \in \mathbb{R}^2$ is *convex* if the segment $\overline{PQ}$ is contained in Ω for all $P, Q \in \Omega$. Geometric formulas for integrals hold for convex sets. We have that $n(L \cap \partial\Omega)$ is either 0 or 2 for dK, so the equation

$$L(\partial\Omega) = \int_{\{L:L\in\Omega\neq\varnothing\}} dK$$

represents the measure of unoriented lines which meet the convex set. Let's now see a probability definition before looking at the theorem.

Definition 3.1. The *conditional probability* of an event A given the event B is defined to be

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Now let's see the theorem:

Theorem 3.2 (Sylvester's Problem)**.** *Let Ω and $\omega \subset \Omega$ be a bounded convex set in the plane. Then the probability of a random line meeting ω given that it also hits Ω is*

$$P = \frac{L(\partial\omega)}{L(\partial\Omega)}.$$

Here's a corollary involving piecewise $\mathcal{C}^1$ curves and expected value, a great combination of topics we've learnt so far:

Corollary 3.3. *Let C be a piecewise $\mathcal{C}^1$ curve contained in a compact convex set Ω. Of all random lines that meet Ω, the expected number of intersections with C is*

$$\mathbb{E}(n) = \frac{2L(C)}{L(\partial\Omega)}.$$

Therefore, there exist lines that split C into at least $\frac{2L(C)}{L(\partial\Omega)}$ points. [Tre08]

Proof. We know that Ω is convex, so we have that

$$\mathbb{E}(n) = \frac{\int_{\{L:L\cap C\neq\varnothing\}} n\ dK}{\int_{\{L:L\cap\Omega\neq\varnothing\}} dK} = \frac{2L(C)}{L(\partial\Omega)}.$$

The maximum of n exceeds the average. □

4. Cauchy's Formula and the Support Function

Before we look at one of Cauchy's formulas for integrals, we must first introduce the *support function* and its properties.

Definition 4.1. The *support function*, denoted $h(\theta)$ where $\theta \in [0, 2\pi)$, is the largest p such that $L(p,\theta) \cap \Omega \neq \varnothing$. The *width* of $h(\theta)$ is

$$w(\theta) = h(\theta) + h(\theta + \pi).$$

Now let's see Cauchy's Formula:

Theorem 4.2 (Cauchy's Formula)**.** *Let Ω be a bounded convex domain. Then we have*

$$L(\partial\Omega) = \int_0^{2\pi} h(\theta)d\theta = \int_0^{\pi} w(\theta)d(\theta).$$

Proof. Simplify from $L(\partial\Omega)$ as follows:

$$\begin{aligned} L(\partial\Omega) &= \int_{\{L:L\cap\Omega\neq\varnothing\}} dK \\ &= \int_0^{2\pi}\int_0^{h(\theta)} dp\ d\theta \\ &= \int_0^{\pi} w(\theta)d(\theta). \end{aligned}$$

□

This completes the proof of Cauchy's Theorem.
Now let's look at another theorem, which gives us the formula for the area of Ω in terms of the support function:

Theorem 4.3. *For a compact, convex domain Ω with a $\mathcal{C}^2$ boundary, the area is calculated as*

$$A(\Omega) = \frac{1}{2}\int_0^{2\pi} h\ ds = \frac{1}{2}\int_0^{2\pi} h(h + \ddot{h})d\theta.$$

Proof. Let $Z(\theta)$ be the point $L(h(\theta), \theta) \cap \partial\Omega$. We have that $n(\theta) = (\cos\theta, \sin\theta)$ is the outer normal,[4] and also that $\dot{n} = (-\sin\theta, \cos\theta)$ and that $Z = hn + \dot{h}\dot{n}$. This means that

$$\dot{Z} = \dot{h}n + h\dot{n} + \ddot{h}\dot{n} - \dot{h}n = (h + \ddot{h})\dot{n}.^5$$

Therefore, $\frac{ds}{d\theta} = h + \ddot{h}$, so

$$\begin{aligned} A(\Omega) &= \int_\Omega dA \\ &= \frac{1}{2}\int_0^{2\pi} h\ ds \\ &= \frac{1}{2}\int_0^{2\pi} h(h + \ddot{h})d\theta. \end{aligned}$$

This completes the proof of this theorem. [Tre08] □

5. More advanced Buffon's Needle, finding π, Crofton's Formula and Corollary, and Bertrand's Paradox

Now that we have acquired basic knowledge of some topics in geometric probability and integral geometry, let's see a more advanced and clever method to solving the same Buffon's Needle problem as discussed earlier.

Proof of Theorem 1. Let a needle N be of length l, a line segment centred at the origin. Now, move the floor. Because $l < d$, this means the closest cracks to the origin are the only ones that the needle could touch. Let C represent the circle about the origin with radius $\frac{d}{2}$. Let's just focus on crack lines L where $\text{dist}(L, 0) \leq \frac{d}{2}$ iff $C \cap L \neq \varnothing$. This implies that $n(L \cap N) = 1$. Therefore, the probability P of needle N touching the crack is

$$P = \frac{\int_{\{L: L\cap N \neq \varnothing\}} n(L \cap N)\ dk(L)}{\int_{\{L: L \cap C \neq \varnothing\}} dK(L)} = \frac{2L(N)}{L(C)} = \frac{2l}{2\pi \cdot \frac{d}{2}} = \frac{2l}{\pi d}.$$

This completes the proof. [Tre08] □

We can use the Buffon's Needle Problem for a determination of π: if x is the number of times the needles touches a crack in n tosses, we have

$$\frac{2l}{d} \approx \frac{2l}{Pd} = \pi.$$

A fun fact about this is that mathematicians Wolf and Smith tossed 5000 and 3204 needles, respectively, and found the values of π to be about 3.1596 and 3.1553, respectively.

Let us now turn to Crofton's Formula, another theorem in integral geometry.

[4] A *normal* is a ray/vector perpendicular to some object. Furthermore, for a topological boundary (set of points in the closure of set), an *outer-pointing normal* describes the direction of the normal points.

[5] The double dot is used to denote a second derivative with respect to time. So, $\ddot{x} = \frac{d^2x}{dt^2}$, where t represents time.

Theorem 5.1 (Crofton's Formula)**.** *Let $D \in \mathbb{R}^2$ be a domain with compact closure. Let $L \in \mathbb{R}^2$ be an arbitrary line, and let $\sigma_1(L \cap D)$ be the length. Then we have*

$$\pi A(D) = \int_{\{L: L \cap D \neq \varnothing\}} \sigma_1(L \cap D) dk(L).$$

For the sake of efficiency, we won't give the proof here because it is quite long, but it is similar to those we've seen previously, involving some trigonometry and of course, many integrals.

Now let's look at Crofton's Corollary, which describes the probability of intersection of lines in Ω:

Corollary 5.2 (Crofton)**.** *Let Ω be a bounded convex domain in the plane. The probability that two arbitrary lines intersect in Ω, assuming they both meet Ω, is*

$$P = \frac{2\pi A(\Omega)}{L(\partial\Omega)^2}.$$

Note that this probability satisfies $P \leq \frac{1}{2}$, and equality holds iff Ω is a round disk.

Proof. Let $L_1(p_1, \theta_1)$ and $L_2(p_2, \theta_2)$ be arbitrary lines with invariant measure

$$dK_1 \wedge dK_2 = dp_1 \wedge d\theta_1 \wedge dp_2 \wedge d\theta_2.$$

Let $\Lambda_1 = L(p_1, \theta_1) \cap \Omega$ be a subset. Then, we can compute the average number of times a random line $L(p_2, \theta_2)$ meets Λ_1 (assuming it meets Ω) as

$$\mathbb{E}(n) = \frac{2\sigma_1(\Omega \cap L(p_1, \theta_1))}{L(\partial\Omega)}.$$

By Poincaré's Formula and Crofton's Formula, we have the probability that two lines meet:

$$\begin{aligned} P = \mathbb{E}(n) &= \frac{\int_{\{L_1: L_1 \cap \Omega \neq \varnothing\}} \int_{\{L_2: L_2 \cap \Omega \neq \varnothing\}} n(\Lambda_1 \cap L_2) dK_2 dK_1}{\int_{\{L_1: L_1 \cap \Omega \neq \varnothing\}} \int_{\{L_2: L_2 \cap \Omega \neq \varnothing\}} dK_2 dK_1} \\ &= \frac{\int_{\{L_1: L_1 \cap \Omega \neq \varnothing\}} \mathbb{E}(n) dK_1}{\int_{\{L_1: L_1 \cap \Omega \neq \varnothing\}} dK_1} \\ &= \frac{2\int_{\{L_1: L_1 \cap \Omega \neq \varnothing\}} \sigma_1(\Omega \cap L(p_1, \theta_1)) dK_1}{L(\partial\Omega) \int_{\{L_1: L_1 \cap \partial\Omega \neq \varnothing\}} dK_1} \\ &= \frac{2\pi A(\Omega)}{L(\partial\Omega)^2}. \end{aligned}$$

This completes the proof of Crofton's corollary. [Tre08] □

Next, we shall look briefly at Bertrand's Paradox, a question about the average length of a chord of Ω. The answers will vary, depending on the definition of a "random line."

Question 5.3 (Bertrand's Paradox)**.** *What is the average length of a chord of a compact convext set Ω?*

Here are some answers, where Ω is a disk of radius R:

- Uniform distance from origin and uniform angle: $\mathbb{E}(\sigma_1) = \frac{\pi R}{2}$.
- Uniform point on boundary and uniform angle: $\mathbb{E}_2(\sigma_1) = \frac{4R}{\pi}$.
- Two uniform random points on the boundary: $\mathbb{E}_3(\sigma_1) = \frac{4R}{\pi}$.

6. Kinematic density and Poincaré's Formula for intersecting curves

Let's now look briefly at kinematic density:

Definition 6.1. Let C and Γ be two piecewise $\mathcal{C}^1$ curves in the plane. We can move Γ around the plane with rigid motion:

$$\Gamma' = \mathcal{M}_{a,b,\phi}(\Gamma),$$

where $\mathcal{M}_{a,b,\phi}$ is the rotation by angle ϕ along with the translation by vector (a, b):

$$x' = x\cos\phi - b\sin\phi + a$$
$$y' = x\sin\phi + y\cos\phi + b.$$

From this, we have that *kinematic density* is the invariant measure of motions of Γ' given by

$$dK = da \wedge db \wedge d\phi.$$

Using this, we can look at Poincaré's Formula for intersecting curves:

Theorem 6.2 (Poincaré's Formula for intersecting curves)**.** *Let C and Γ be piecewise $\mathcal{C}^1$ curves. Additionally, let $n(C \cap \Gamma')$ denote the number of intersection points between C and a moving Γ'. Then, we have*

$$\int_{\{\Gamma' : C\cap\Gamma' \neq \varnothing\}} n(C \cap \Gamma')dK(\Gamma') = 4L(C)L(\Gamma).$$

The proof of this version of Poincaré's Formula is very similar to the one for lines–we introduce a flag subset, use some trigonometry, and finish using Fubini's Theorem. The reader may see Treiburgs's paper for the full proof.

7. Santaló's Formula

Finally, let's look at Santaló's formula, perhaps one of the most significant theorems employing kinematic density:

Theorem 7.1 (Santaló's Formula)**.** *Let Ω_1 and Ω_2 be convex plane domains. We assume that Ω_2' is moving in the plane with kinematic density dK_2. Then, we have*

$$\int_{\{\Omega_2' : \Omega_2' \cap \Omega_1 \neq \varnothing\}} dK_2 = 2\pi\{A(\Omega_1) + A(\Omega_2)\} + L(\partial\Omega_1)L(\partial\Omega_2).$$

[SS04]

Proof. Begin by first rotating the domain $\Omega_2' = \mathcal{M}\Omega_2$ by angle ϕ, then a translation of vector (a, b). We have that the kinematic density is

$$dK = da \wedge db \wedge d\phi.$$

Now, let's fix ϕ and let $D(\phi)$ be the *set of moving centers* (a, b) of $\Omega_2'(\phi)$ where the domains overlap, as follows:

$$\Omega_1 \cap \Omega_2'(\phi) \neq \varnothing.$$

We have that

$$f(\alpha) = h(\alpha) + g(\alpha + \pi - \phi)$$

is the support function for $D(\phi)$, $h(\alpha)$ is the support function[6] of A from the origin. for Ω_1 and $g(\alpha)$ is the support function for Ω_2.

Now we have a bunch of integrals:

$$\begin{aligned}
\mathcal{J} &= \int_{\{\Omega_2':\Omega_1\cap\Omega_2'\neq\varnothing\}} \\
&= \int_0^{2\pi}\int_{\{\Omega_2'(\phi):\Omega_1\cap\Omega_2'(\phi)\neq\varnothing\}} da\ db\ d\phi \\
&= \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} f(\alpha)[f(\alpha)+\ddot{f}(\alpha)]\ d\alpha\ d\phi \\
&= \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi}[h(\alpha)+g(\alpha+\pi-\phi)]\begin{bmatrix} h(\alpha)+g(\alpha+\pi-\phi) \\ +\ddot{h}(\alpha)+\ddot{g}(\alpha+\pi-\phi)\end{bmatrix} d\alpha\ d\phi.
\end{aligned}$$

Now we use the fact that $\int_0^{2\pi}\ddot{h}(\alpha)d\alpha = 0$, combined with Cauchy's formula, to get

$$\begin{aligned}
2\mathcal{J} = &\int_0^{2\pi}\int_0^{2\pi} h(\alpha)[h(\alpha)+\ddot{h}(\alpha)]d\alpha\ d\phi \\
&+\int_0^{2\pi}\int_0^{2\pi} g(\alpha+\pi-\phi)g(\alpha+\pi-\phi)+\ddot{g}(\alpha+\pi-\phi)d\alpha\ d\phi \\
&+\int_0^{2\pi}\int_0^{2\pi} h(\alpha)[g(\alpha+\pi-\phi)+\ddot{g}(\alpha+\pi-\phi)]d\phi\ d\alpha \\
&+\int_0^{2\pi}\int_0^{2\pi} g(\alpha+\pi-\phi)[h(\alpha)+\ddot{h}(\alpha)]d\phi\ d\alpha \\
&= 4\pi\ A(\Omega_1) + 4\pi\ A(\Omega_2) \\
&+\int_0^{2\pi} h(\alpha)[L(\partial\Omega_2)+0]d\alpha + \int_0^{2\pi} L(\partial\Omega_2)[h(\alpha)+\ddot{h}(\alpha)]\ d\alpha \\
&= 4\pi\ A(\Omega_1) + 4\pi\ A(\Omega_2) + L(\partial\Omega_1)L(\partial\Omega_2) + L(\partial\Omega_2)[L(\partial\Omega_1)+0].
\end{aligned}$$

This completes the proof of Santaló's Formula. [SS04] □

Let's finish by seeing this corollary which combines Santaló's Formula and Poincaré's Formula:

[6]Support function of a closed convex set A is the distance of supporting hyperplanes (i.e., hyperplane where the set is contained in a closed half-space bounded by the hyperplane and has at least one boundary point on the hyperplane).

Corollary 7.2. *Let Ω_1 and Ω_2 be bounded convex planar domains. The expected number of intersections of $\partial\Omega_1$ with a moving $\partial\Omega_2'$ (given that Ω_2' meets Ω_1) is*

$$\mathbb{E}(n) = \frac{4\ L(\partial\Omega_1)L(\partial\Omega_2)}{2\pi\{A(\Omega_1) + A(\Omega_2)\} + L(\partial\Omega_1)L(\partial\Omega_2)}.$$

if $\Omega_1 \cong \Omega_2$, then $\mathbb{E}(n) \geq 2$ is equal iff Ω_1 is a circle. [Tre08]

The proof of this corollary uses facts about piecewise curves and Poincaré's and Santaló's formulae.

References

[SS04] Luis Antonio Santaló Sors and Luis A Santaló. *Integral geometry and geometric probability*. Cambridge university press, 2004.

[Tre08] Andrejs Treiburgs. Integral geometry and geometric probability. page 64, 2008.