Lecture 3

Assigning Probabilities via Size – II

In the last lecture we figured out how to assign $\mathbb{P}(A)$ via cleverly counting the number of elements in A and then dividing this number by the number of elements in S. This technique worked when the sample space was simple (i.e., S had only finitely many members and our aim was to assign each member the same chance of being the outcome when we performed the experiment once). At the very end of the last lecture we saw an example to extend the concept of "size" to deal with infinite sets, and non equally likely spaces.

Unfortunately, the counting technique breaks down when S is an infinite set. Now we present an analog of the counting method which involves infinite sample spaces and whose subregions of the same "sizes" have the same chance.

3.1 Probabilities via Lengths/Areas/Volumes

Consider the following few examples which provide the basic idea.

Example - 3.1.1 - **(Events and assigning probabilities for subsets of $\mathbb{R}$)** Let $S = [1, 5]$ be the sample space corresponding to selecting a point at random from the interval $[1, 5]$. If A is the event that the drawn point will lie between 2 and 3.5 then $A = [2, 3.5]$. These sets are shown in Figure 3.1 below. Note that the length of S is 4 (i.e., S is bounded however it has infinitely many points). The

Figure 3.1: Selecting a Point at Random from $[1, 5]$.

probability of the event A is

$$\mathbb{P}(A) = \frac{\text{size (i.e., length) of } A}{\text{size (i.e., length) of } S} = \frac{3.5 - 2}{5 - 1} = \frac{1.5}{4} = 0.375.$$

This method of assigning probabilities obeys the three axioms, namely, $\mathbb{P}(A) \geq 0$, $\mathbb{P}(S) = 1$ and

$$\mathbb{P}(A_1 \cup A_2 \cup \cdots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots,$$

for any disjoint events $A_1, A_2, \cdots$. Actually the last sentence is a bit vague.

How does one know if a subset of S is an event? Is every subset of S an event? How does one define an event when S has infinitely many points? This is the vague part in the above presentation. A fully satisfactory answer to such questions was invented around the turn of the 20-th century. Prior to that time probabilists had discovered many "paradoxes", which were in fact due to various (mis)interpretations of this issue. The tools needed to understand this feature are taught in a Real Analysis course. For now we will make a working definition of an event when S is an interval of the real line. A subset of S is an event if its length can be "*measured*".

Example - 3.1.2 - **(Events and assigning probabilities for subsets of $\mathbb{R}^2$)** Let $S = \{(x, y) : x^2 + y^2 \leq 9\}$ be the sample space corresponding to selecting a point at random from a circular dart board with radius 3 inches. Find the probability of the event, A, that the selected point lies in the square shown in Figure 3.2. The

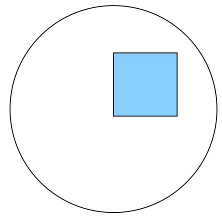


Figure 3.2: Selecting a Point at Random from a Dart Board.

length of a side of the square is 2 inches. First note that any regular geometrical shape falls in our class of events. Hence,

$$\mathbb{P}(A) = \frac{\text{size (i.e., area) of } A}{\text{size (i.e., area) of } S} = \frac{2^2}{\pi 3^2} = 0.141.$$

Again, for a planar set S, events will be those subsets whose areas can be measured. A similar definition goes for higher dimensional spaces.

Example - 3.1.3 - **(Rendezvous problem)** Two persons, Les and Leslie, agreed to meet at a specified park between noon and one O'clock tomorrow. They agreed not to wait for more than ten minutes for the other person. What is the probability that they will meet if they arrive randomly and unrelatedly during the hour.

Let U, V be the arrival times of Les and Leslie respectively. We do not know what U and V will be. All possible arrival time pairs (the sample space) can be viewed as a square. Regardless of who arrives first, if the difference of their arrival times does

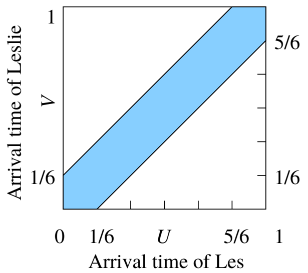


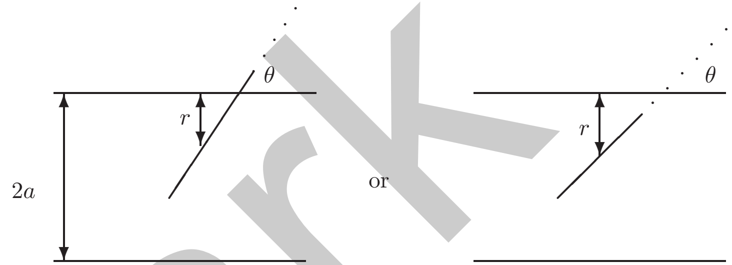
Figure 3.3: Sample Space and the Event.

not exceed $\frac{1}{6}^{\text{th}}$ of an hour (which is ten minutes) they will meet otherwise not. So, the event, that they will meet, is $|U - V| \leq \frac{1}{6}$. This event is shown in Figure 3.3. The area of the shaded region can be obtained from the areas of the two triangles, i.e., the complement of the shaded region. Hence, the probability that they will meet is

$$1 - \frac{5}{6} \times \frac{5}{6} = \frac{11}{36}.$$

Example - 3.1.4 - **(Buffon's needle problem)** George-Louis Leclerc Buffon (1707–1788) was a naturalist and he compiled 44 volumes dealing mostly with natural history. We have a grid of parallel lines, $2a$ units apart and we drop a needle of length $2l$ (where $l < a$) on the grid. What is the "probability" that the needle intersects the grid lines?

Since, $l < a$, the needle cannot intersect two lines simultaneously. Let r be the shortest distance from the center of the needle to the parallel lines. Clearly, $0 \leq r \leq a$. Let θ be the angle that the needle makes with the closest line. Thus $0 \leq \theta \leq \pi$.



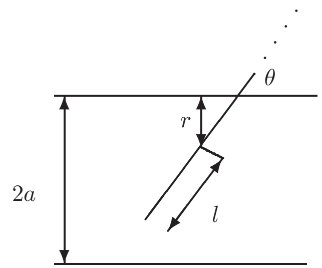
The drop of the needle is completely characterized by the pair (r, θ). That is, any drop of the needle will give a pair (r, θ) between some two lines (ignoring sliding it left or right). And conversely, any pair (r, θ) such that $0 \leq r \leq a$, $0 \leq \theta \leq \pi$ corresponds to a drop of a needle on the grid (ignoring left or right slides). The sample space could be considered as the set

$$S = \{(r, \theta) \mid 0 \leq r \leq a; \quad 0 \leq \theta \leq \pi\}.$$

The event of interest (namely, the needle crosses a line) occurs if and only if $r \leq l \sin\theta$. To see why, note that

$$\sin\theta = \frac{r}{\text{hypotenuse}} \geq \frac{r}{l}$$

if and only if the chord intersects the parallel lines.



This is why the event of interest, A, is

$$A = \{(r, \theta) \mid r \leq l \sin\theta\}.$$

Now if we consider the probability of A to be proportional to the area then the area of S is πa and the area of A is

$$\text{area of } A = \int_0^\pi l \sin\theta \, d\theta = 2l.$$

Therefore, $\mathbb{P}(A) = (2l)/(a\pi)$. Reportedly, a person named Wolf actually dropped a needle of length $2l = 36mm$ over parallel lines $2a = 45mm$ apart 5000 times. He observed that the percentage of times the needle crossed the grid was $p = 8/15.795$. From this he computed the value of π by solving

$$\frac{2l}{a\pi} = \frac{36}{22.5\pi} = \frac{8}{15.795}.$$

This gave $\pi \approx 3.159$ which is not a bad empirical estimate of π.

3.2 Non Equally Likely Spaces

Much of the contents of this section will be studied more elaborately in the later lectures under the concept of a density. We present the basic ideas with a few examples.

Example - 3.2.1 - **(From two dimensions to one dimension)** Consider $S = [0, 4]$ but we would like to assign probabilities in such a way that the regions closer to 4 are progressively more likely than regions closer to 0. For instance, consider the curve $y = x^2$ to represent this progressively increasing weights.

One way of accomplishing this is to construct a secondary sample space, say $S_2 = \{(x, y) :\ 0 \le y \le x^2,\ 0 < x < 4\}$. Take this as the sample space corresponding to selecting a point at random from the region under the parabola $y = x^2$, between $x = 0$ and $x = 4$. Finding the probability of the event in the original space, S, corresponds to selecting a point at random from S_2 and then taking its x-coordinate.

For instance, let us find the probability that a randomly chosen point lies between 3 and 4, i.e., $A = [3, 4] \subseteq S$. The corresponding event in S_2 is the shaded region shown in Figure 3.2.1. Now we may use calculus to find the required areas.

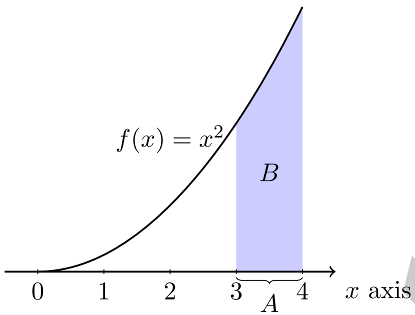


Figure 3.4: Selecting a Point at Random Under a Function.

$$\mathbb{P}(A) = \frac{\text{Area of } B}{\text{Area of } S_2} = \frac{\text{Area under } f \text{ over } A}{\text{Area under } f \text{ over } S} = \frac{\int_3^4 x^2\, dx}{\int_0^4 x^2\, dx} = \frac{37}{64} = 0.5781.$$

We have used a two dimensional concept of area to assign probabilities to one dimensional events. This allows us to introduce non-equally likely sample spaces. S_2 was equally likely but S is not. One good thing about this approach is that S need not be bounded anymore, all it needs is that S_2 has a finite area, i.e., the curve f over S has a finite area. In other words, f should be integrable of S.

Example - 3.2.2 - **(From three dimensions to two dimensions)** Let $S = \{(x, y) :\ x^2 + y^2 \le 1\}$ be the sample space corresponding to a dart board. Professional dart board throwers, unlike me, can make sure that their dart will always

land on the dart board. Furthermore, their chances of landing the dart in the bulls eye are much higher. Let us model this by using the surface of a semisphere, $f(x, y) = (1 - x^2 - y^2)^{1/2}$, describe the regions which are more likely. If bulls eye

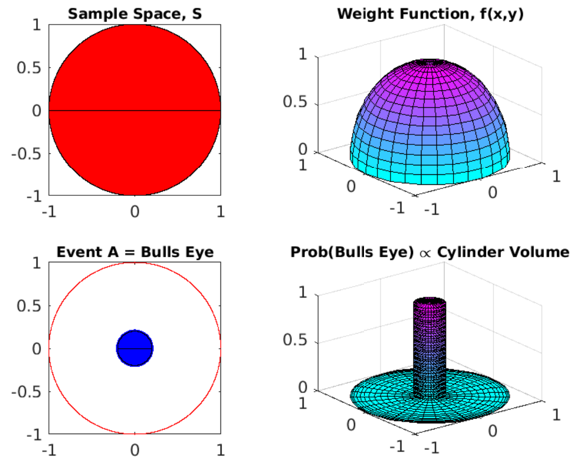


Figure 3.5: Probability of Bull's Eye.

is a circular region of radius 0.2, then the probability that the dart will land in bulls eye will be proportional to the volume of the circular cylinder under f over the bull's eye region, as shown in the lower right portion of Figure 3.5. If A is the Bull's eye region in S, then

$$\mathbb{P}(A) = \frac{\text{Volume under } f \text{ over } A}{\text{Volume under } f \text{ over } S} = \frac{\int_0^{2\pi}\int_0^{0.2} z\sqrt{1-z^2}\, dz\, d\theta}{\int_0^{2\pi}\int_0^{1} z\sqrt{1-z^2}\, dz\, d\theta} = 1 - \left(1 - (0.2)^2\right)^{3/2},$$

which is about 0.06. This example shows how we can introduce non-equally likely sample spaces in $\mathbb{R}^2$. Again, S need not be bounded anymore, all it needs is that the surface, f, over S has a finite volume.

Remark - 3.2.1 - **(Weird sets & the measurability dilemma)** The discussion of the last two examples raises some natural questions. How strange looking can the subsets of the real line be and what are their lengths?

For instance, what is the length of the set of all rational numbers in the interval $S = [0, 1]$? The answer is 0. This is because each single point has zero length and adding countably[1] many such lengths still adds up to zero. So, by complementing,

[1] A set is called "countable" if we can create a one-one correspondence of it with a subset of the positive integers. For example any set having finitely many elements is countable. The set of natural numbers is countable. The set of integers as well as the set of rational numbers are countable.

the set of all irrationals in $[0,1]$ has length 1. So, adding uncountably many zeros can add up to a number that is more than zero!

The above two sets, although hard to draw, are not so weird after all. Their lengths are defined and can be measured. There do exist sets which are so weird that their length cannot be measured. In other words, it is not our ability of measurement that is at fault but rather the concept of "length" itself cannot be defined for them. The existence of such weird sets was discovered at the turn of the 20-th century, and they are called nonmeasurable sets. Probabilists avoid dealing with these sets at all cost and so will we. The usual geometrical shapes, such as intervals, triangles, rectangles, pentagons, hexagons, cubes and spheres and other regular shapes etc. are 'safe' sets for which probabilities can be assigned.

Remark - 3.2.2 - **(Summary)** When a sample space, S, is a bounded subset, of the real line or the plane or higher dimensional space, and the experiment involves selecting a point "*at random*" from S, then we define the probability of an event as follows.

$$\boxed{\mathbb{P}(A) = \frac{\text{size of } A}{\text{size of } S}.}$$

In this idealization sliding the event A does not change is probability. In this sense this is an analog of the equilikely concept.

The nonequilikely cases are more general and will be studied systematically in the later chapters. For now, roughly speaking, we may say that for general nonequilikely sample spaces, S, whether it is bounded or not, often our model is

$$\boxed{\mathbb{P}(A) = \frac{\text{generalized size of } A}{\text{generalized size of } S} = \frac{\int_A f(x)dx}{\int_S f(x)dx},}$$

where $\int_A f$ represents the "generalized size" of A based on the weight function f, and $\int_S f$ is the corresponding "generalized size" of S. Note that the equilikely case is a special case of this, obtained by taking $f(x) \equiv 1$. The weight function, f, is related to the concept of a density to be studied later.

3.3 Exercises

Exercise - 3.3.1 - *In Example 3.1.1 find the probability that the randomly drawn point will lie between 1 and 2 or between 3.5 and 5.*

Exercise - 3.3.2 - *In Example 3.1.2 find the probability that the randomly drawn point will lie outside the disc $\{(x,y) : x^2+y^2 \le 1\}$.*

Exercise - 3.3.3 - *A point is selected at random from a rectangle $[0,2]\times[0,1]$. What is the probability that the point chosen falls below the curve $f(t) = 2t - t^2$. (Hint: draw the regions.)*

Exercise - 3.3.4 - *When a point is selected at random from a square with a side length of 8 inches, find the probability that the point falls in any of the four circular corners, as shown in Figure 3.6. The radius of each quarter circle is 2 inches.*

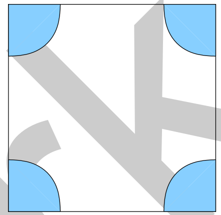


Figure 3.6: Selecting a Point at Random from a Square.

Exercise - 3.3.5 - *Let $S = \{(x,y) : x^2+y^2 \le r^2\}$ be the sample space corresponding to selecting a point at random from a circular dart board with radius r. Find the probability of the event, A, that the selected point lies in the square shown in Figure 3.7.*

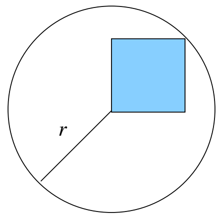


Figure 3.7: Selecting a Point at Random from a Dart Board.

Exercise - 3.3.6 - *Consider Example 3.1.3 again but now assume that the probabilistic structure of the arrival times are modeled by the weight function $f(u,v) = v^2v^2$, $0 < v < 1$ and $0 < v < 1$. What is the probability that Les and Leslie will meet.*

Exercise - 3.3.7 - *Let $S = \{(x,y) : y \le \cos x;\ -\frac{\pi}{2} < x < \frac{\pi}{2}\}$ be the sample space corresponding to selecting a point at random from the region under the function $y = \cos x$ between $x = -\frac{\pi}{2}$ and $x = \frac{\pi}{2}$. Find the probability of the events that the x-coordinate of the chosen point, when randomly chosen, will lie in*
(i) the interval $(0, \frac{\pi}{2})$, (ii) the interval $(-\frac{\pi}{4}, \frac{\pi}{4})$, (ii) the interval $(-\frac{\pi}{3}, \frac{\pi}{6})$.

Exercise - 3.3.8 - **(High dimensional loss)** *We select a point at random from $[-1,1]^n$. What is the probability that it will be in the inscribed sphere? Compute the probability for $n = 2, 3, 10, 15$. Any comments? [Hint: The volume of a unit n-sphere is $\pi^{n/2}/\Gamma(\frac{n}{2}+1)$. The symbol Γ is defined in (3.2) in Section 7.3.]*

Exercise - 3.3.9 - **(Lost hubcap)** *While driving on a road described by the function $f(x) = x^{3/2}$, I lost one of my hubcaps. The hubcap could have fallen off*