Lecture 10: Continuous Probability Spaces

CSCI2244-Randomness and Computation

March 14, 2019

1 A spinner (or, what does 'probability zero' mean?)

Figure 1 shows a game spinner. The points on the circumference of the circle are labeled by the half-open interval

$$S = \{x \in \mathbf{R} : 0 \leq x < 1\},$$

We also denote this set by $[0, 1)$.

You can think of the spinner as a continuous analogue of a die: the whole point of the spinner is that the outcomes are somehow 'equally likely'. (The spinner is the experiment simulated by a call to the random number generator `rand()`.) We have no problem saying that the event

$$E = \{x \in \mathbf{R} : 0.5 \leq x \leq 0.75\} = [0.5, 0.75].$$

depicted in Figure 2 has probability 0.25, since it occupies exactly one-quarter of the circumference of the circle.

But by exactly the same reasoning, we would say that the probability of the half-open interval $[0.5, 0.75)$, which we obtain by removing the single point 0.75, is 0.25. We thus would have

$$\begin{aligned} 0.25 &= P([0.5, 0.75]) \\ &= P([0.5, 0.75) \cup \{0.75\}) \\ &= P([0.5, 0.75)) + P(\{0.75\}) \\ &= 0.25 + P(\{0.75\}), \end{aligned}$$

because the union is disjoint. This implies

$$P(\{0.75\}) = 0,$$

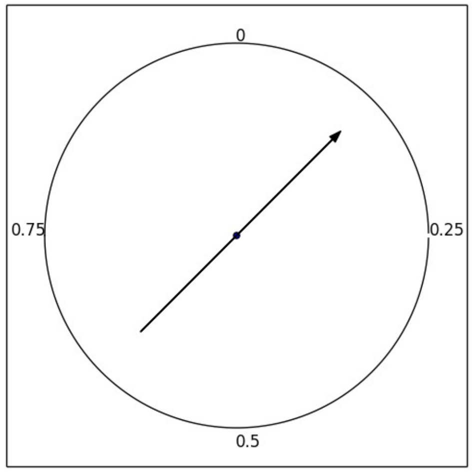

Figure 1: A spinner

and likewise the probability of *any* individual point is 0. We usually think of probability 0 as meaning that an event is impossible. But then we would be saying that it is impossible for the spinner to land on any individual point, and so, it is impossible that it lands anywhere at all!

In fact, we DO say that the probability of each individual point is 0, and we just have to be careful about how we apply the intuition that probability 0 means 'impossible'.

With this in mind, we assign to each event $E \subseteq S$ the total *length* of E. This makes sense for individual points, intervals, unions of disjoint intervals, *etc.* This is the *uniform probability distribution* on S.

A continuous sample space can be 1-dimensional like the spinner, 2-dimensional, or 3- or more-dimensional. In the uniform distribution model, the probability function on a 2-dimensional sample space S assigns to each event $E \subseteq S$ the

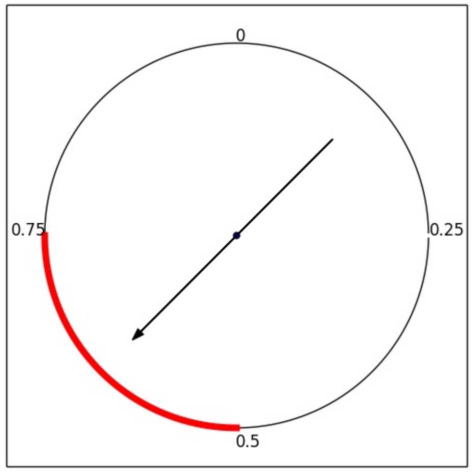

Figure 2: The spinner, showing the event $\{x : 0.5 \leq x \leq 0.75\}$

probability[1]

$$\frac{area(E)}{area(S)}.$$

We earlier gave the example of a dart thrown at a circular dartboard with radius 1 foot. The board has area π and the bulls-eye, which has radius $1/12$,, has area $\pi/144$ square feet. So in the uniform model, the bulls-eye would have probability $1/144$. There is no strong reason to believe that this the uniform distribution is a good model of reality. For instance, we might guess that a skillful player is more likely to hit the little bullseye circle than to hit any other circle that has the same area. It would *still* be the case that each individual outcome has probability 0, but that says nothing about the probability space being uniform or not.

What's different? What's the same? In a continuous probability space, the

[1] Just to be honest: One of the achievements of a branch of mathematics called *measure theory* is the discovery that it is actually *impossible* to meaningfully assign a length to every subset of $[0, 1)$, an area to every bounded subset of the plane, *etc.*, as well as to develop a theory of what constitutes a 'measurable' set. The upshot is that not every subset of a continuous sample space can be considered an event and assigned a probability. Now that you've learned this, you can forget it, because you will never encounter one of those bad sets in practice.

probability axioms are just the same as they were for discrete spaces: complementary probabilities add to 1, the probability of a pairwise disjoint union of events is the sum of the probabilities of the individual events, *etc.* What is different is that the probability function is *not* determined by the probabilities of individual outcomes, which typically are all 0.

2 Two spinners.

Now suppose our experiment consists of spinning two spinners. We can model the set of outcomes as the set of ordered pairs:

$$S = [0, 1) \times [0, 1) = \{(x, y) : 0 \le x, y < 1\}.$$

This is the unit square in $\mathbf{R}^2$. Assuming that the two spins are independent (*e.g.,* no hidden mechanism in which the position of one spinner influences that of the other) the the probability distribution is uniform. Since the square has area 1, the probability of an an event E is equal to its area.

Here is an example: What is the probability that the value on the second spinner exceeds that on the first spinner by at least $\frac{1}{2}$? We can write this event as

$$E = \{(x, y) \in S : y \ge x + 0.5\}.$$

The event is pictured in Figure 3. This right triangle with base and height both equal to $\frac{1}{2}$ has area $\frac{1}{8}$, so $P(E) = \frac{1}{8}$. Observe that if we had instead asked for the probability that the second spinner value is greater than or equal to the first spinner, the event would still be a triangle, but this time occupy half of the square, so the probability would be $\frac{1}{2}$. If we had asked for the probability that the second spinner value is *exactly* $\frac{1}{2}$ plus the first value, then the event is just a line segment, and the probability is 0.

2.1 Buffon's Needle

Here is an amusing—and at first glance, astonishing—application of these ideas. Mark off a large area on the floor with parallel lines one inch apart. Toss a one-inch-long needle in the air and let it land on the floor. Perform this experiment repeatedly, and count the proportion of trials in which the needle crosses one of the lines. This proportion will be approximately $2/\pi$. (The problem, in slightly different form, was described in the 18th century by G. L. Leclerc, Comte de Buffon, who was more famous as a naturalist than as a mathematician.)

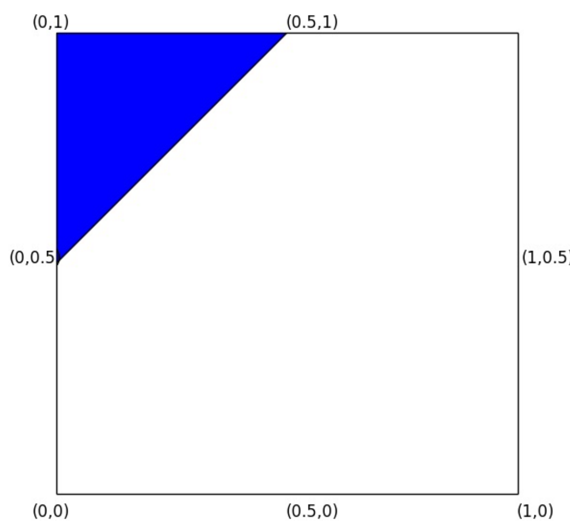

Figure 3: The unit square models the outcome of two spinners. The event pictured is ‘the value on the second spinner is at least one-half more than that on the first spinner’.

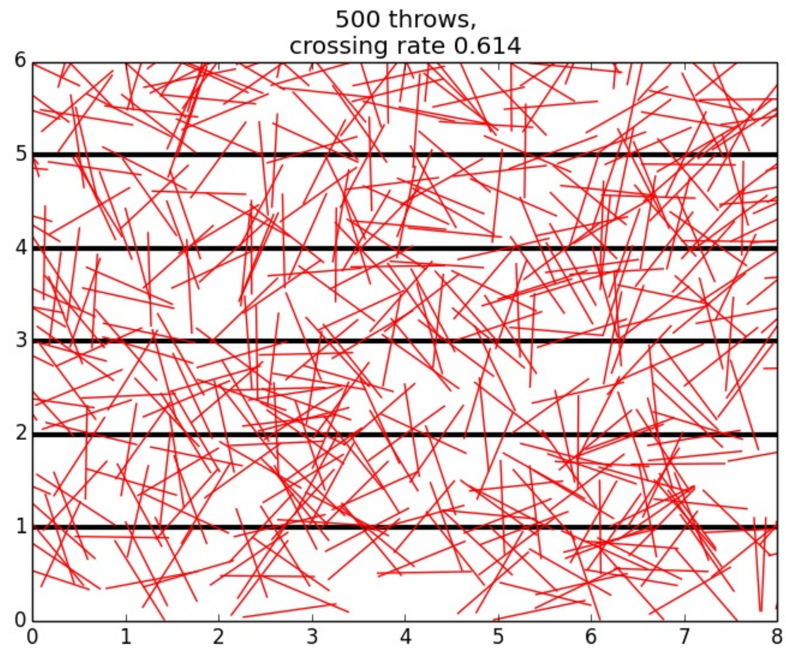

Figure 4: Buffon's Needle, 500 trials

Figure 4 illustrates a simulation of the experiment with 500 throws. In this particular run, the needles crossed the line in 61.4% of trials. The correct value of $2/\pi$ is about 0.6366.

In another run, with 60000 trials, the estimate of $2/\pi$ was 0.6389.

Let's see how this remarkable result is obtained. An outcome is where the needle lands, and this is determined by two parameters: The cooridnates of the lower extremity of the needle, and the angle $0 \leq \theta < \pi$ that the needle makes with the horizontal. Now we are only interested in whether the needle crosses one of the horizontal lines, so we don't care about the x-coordinate at all, and we only care about the integer part of the y-coordinate, that is the horizontal distance $0 \leq z < 1$ from the horizontal line immediately below.

So we can view the sample space as the rectangle

$$S = \{(\theta, z) : 0 \leq \theta < \pi, 0 \leq z < 1\}.$$

Here it is reasonable to assume a uniform distribution: we don't think that any

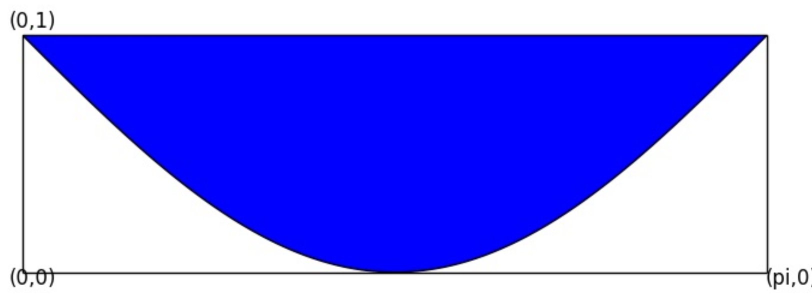

Figure 5: Sample space for Buffon Needle Problem, with the event 'the needle crosses a line' shaded

horizontal displacement z or angle θ is more likely than any other one. So this is just like the two spinners sample space, with slightly different scaling on the horizontal axis.

The upper extremity of the needle has y-coordinate $z + \sin\theta$, and thus the needle crosses the line if and only if $z > 1 - \sin\theta$. In the plot of Figure 5, this event is the shaded area.

The probability of crossing the line is thus the area of the shaded region divided by the area of the rectangle. If we turn the picture upside-down, the shaded region is just the area under the graph of $\sin\theta$ for $0 \le \theta \le \pi$, which is

$$\int_0^\pi \sin\theta d\theta = -\cos\pi - (-\cos(0)) = 1 + 1 = 2,$$

and the area of the rectangle S is π, so the probability that the needle crosses a line is $2/\pi$, as claimed.

2.2 Monte Carlo Integration

The foregoing example (minus the story about the needle) shows a method for approximating an integral: To find the area of a region X (in this case the shaded region of Figure 5), we enclose it in a region Y whose area we know (in our example, the rectangle) and select points uniformly and at random in Y. Let N be the number of points selected and N_{hits} the number of these points that are in X. Then

$$\frac{N_{hits}}{N} \approx \frac{area(X)}{area(Y)},$$

so we obtain an approximation of the area of X. Of course in the above example, we knew the integral exactly, but for peculiarly-shaped regions of integration, this method might be a practical alternative. Approximating integrals in this way is called *Monte Carlo integration*.