

# 1 Continuous Probability Densities (Part 2)

Recall that the discrete case, a random variable's probability mass function (PMF)

$$p(x) = \mathbb{P}(X = x)$$

has the property that:

$$\mathbb{P}(X \in A) = \sum_{x \in A} p(x).$$

To find the probability density function in the continuous case, we want to find a PDF  $f$  such that

$$\mathbb{P}(X \in A) = \int_A f(x) dx.$$

We note that  $p(x)$  is a probability, namely the probability that  $X$  takes the value  $x$ . However,  $f(x)$  is **not** a probability at all; in other words,  $f(x) \neq \mathbb{P}(X = x)$ . We note that, for any  $x$ ,

$$\mathbb{P}(X = x) = \int_x^x f(u) du = 0.$$

Now, for example, if we select a random US citizen, the probability that their height is *precisely* 6.0000 feet is **0**.

Recall the spinner example that we talked about in previous class.

(Example.) Imagine we have a unit circle of circumference 1, and we mark one of the points on the circle to be the origin. Imagine that we have an arrow that points to the origin. Then, imagine we “flick” this arrow with our finger. Let  $X$  denote the position (measured counter-clockwise from 0) of the arrow when it comes to rest. One thing is for sure,  $X$  is a continuous random variable on the sample space  $\Omega = [0, 1)$ .

Assuming that we flick the arrow hard enough, it's reasonable to assume that the arrow should end up landing in a uniformly random position, i.e. that the chance of landing in some sub-interval is proportional to its length irrespective of where it is positioned along the circle. In other words, we have

$$P(a < X < b) = b - a$$

for  $a < b \in [0, 1)$ . We note that

$$\int_a^b 1 dx = b - a.$$

Therefore, in this example,  $f(x) = 1$  for  $x \in [0, 1)$  is the **probability density function** (PDF) of  $X$ . Note that the **probability density function** is the function  $f$  such that integrating it from  $a$  to  $b$  gives us the probability of the random variable being between  $a$  and  $b$ .

There's a continuum of possibilities along the circle where it can end up. The probability that it will end up in one specific place is **0**. But, there's a chance that it will be close to said specific place; in particular, the probability that it will be between  $a$  and  $b$  is the integral of the PDF between  $a$  and  $b$ .

In general, a uniform RV on an interval  $I \subset \mathbb{R}$  has PDF

$$f = \frac{1}{\text{length}(I)}.$$

Another way to find the PDF is to consider the **cumulative distribution function** (CDF) of a random variable, which is the function

$$F(x) = \mathbb{P}(X \leq x).$$

Since the spinner is uniform on  $[0, 1]$ , it is natural to expect that, in this example,

$$F(x) = \begin{cases} \mathbb{P}(X \leq x) & x \in [0, 1] \\ 0 & x < 0 \\ 1 & x > 1 \end{cases}.$$

Note that if  $x = F(x) = \int_0^x f(u)du$ , then the Fundamental Theorem of Calculus implies that

$$f(x) = 1.$$

This is called the **CDF method** of finding the PDF.

(Example.) Suppose we have a dart board with unit radius. Suppose we throw a dart at the target. The sample space is the unit disk

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}.$$

Supposing a dart lands at uniformly random position on the target, we would have the PDF

$$f(x, y) = \frac{1}{\pi}$$

for a random throw  $(X, Y)$ . Note that, because this is two-dimensional, we consider the *area* as opposed to the length of the interval.

To find the probability of landing in a certain region, we would have to integrate over that region; more specifically, we have to integrate that uniform density  $\frac{1}{\pi}$  by that region. For instance, if the “bullseye” region of the target is the center circle  $B$  of radius  $\frac{1}{5}$ , then the probability of getting a “bullseye” would be

$$\frac{\text{area}(B)}{\text{area}(D)} = \frac{\pi(1/5)^2}{\pi} = \frac{1}{25}.$$

So, we should expect approximately 1 in every 25 throws to be a “bullseye.”

Now, let  $D$  be the distance from the center to the point  $(X, Y)$  where a uniformly thrown dart lands. We note that

$$D = \sqrt{X^2 + Y^2} \in [0, 1].$$

What is the distribution of this random variable? We should not expect  $D$  to have a uniform distribution. For instance, notice that there are more points at distance  $\geq 1/2$  from the center than there are points at distance  $\leq 1/2$  from the center. So, we expect

$$\mathbb{P}(D \in [0, 1/2]) < \mathbb{P}(D \in [1/2, 1])$$

although both of these sub-intervals of  $[0, 1]$  have the same length. Simulating this on the computer, we have that the PDF is given by  $f(d) = 2d$  for  $d \in [0, 1]$ .

We can also find  $f(d)$  by the CDF method. Recall that  $f(d) = F'(d)$ . Then,

$$F(d) = \mathbb{P}(D \leq d) = \mathbb{P}(X^2 + Y^2 \leq d^2) = \frac{\pi d^2}{\pi} = d^2.$$

Therefore,  $f(d) = 2d$ .

## 1.1 Being More Formal

We now talk formally about what the above terms mean.

**Definition 1.1: Probability Density Function**

Let  $X$  be a continuous,  $\mathbb{R}$ -valued random variable. A **probability density function** (PDF) for  $X$  is an  $\mathbb{R}$ -valued, non-negative function  $f$  that satisfies

$$\mathbb{P}(a < X < b) = \int_a^b f(x)dx$$

for all  $a, b \in \mathbb{R}$ .

**Remark:** We note that, in infinitesimal notation,

$$\mathbb{P}(a < X < b) = \int_a^b f(x)dx$$

says that

$$\mathbb{P}(X \in [x, x + dx]) = f(x)dx$$

where  $dx$  is an infinitesimal. Recall that, formally, this is short-hand notation for

$$\lim_{\delta x \rightarrow 0} \frac{\mathbb{P}(X \in [x, x + \delta x])}{\delta x} = f(x).$$

From this, we can see why the CDF method works; in particular,

$$F'(x) = \frac{F(x + dx) - F(x)}{dx} = \frac{\mathbb{P}(X \in [x, x + dx])}{dx} = f(x).$$

**Theorem 1.1**

Let  $X$  have CDF  $F$  and PDF  $f$ . Then,  $F'(x) = f(x)$ .

*Proof.* Note that  $F(x) = \int_{-\infty}^x f(u)du$ . Hence, by the Fundamental Theorem of Calculus, it follows that  $F'(x) = f(x)$ .  $\square$

We note that the continuous RVs have the property that, for all  $x$ ,

$$\mathbb{P}(X = x) = \int_x^x f(u)du = 0.$$

The CDF is often useful when considering **transformations** of a RV.

(Example.) Let  $U$  be a uniform RV on  $[0, 1]$ , and consider the RV

$$X = U^2.$$

Find the PDF of  $X$ .

Let  $F_X(x)$  is the CDF of the random variable  $X$ . Then:

$$\begin{aligned}
 F_X(x) &= \mathbb{P}(X \leq x) \\
 &= \mathbb{P}(U^2 \leq x) \\
 &= \mathbb{P}(U \leq \sqrt{x}) \\
 &= \mathbb{P}_U(\sqrt{x}) \\
 &= \sqrt{x} \quad \text{Since } U \text{ is uniform.}
 \end{aligned}$$

Hence,

$$f_X(x) = \frac{1}{2\sqrt{x}}$$

for  $x \in [0, 1]$ .

(Example.) Recall that, in last class, we discussed the PDF of

$$X = M + N$$

was bound by adding together two uniformly random numbers  $M, N$  in  $[0, 1]$ . Since the random sum  $X$  can take the values between 0 and 2, but the two summands can take the values between 0 and 1, we take two cases.

1. Note that, for  $x \in [0, 1]$ , we have

$$F_X(x) = \mathbb{P}(X \leq x) = \int_0^x F_N(x-y)f_M(y)dy.$$

This gives us our CDF. Then, since  $f_M(y) = 1$  and  $F_N(x-y) = x-y$ , we find

$$F_X(x) = \int_0^x (x-y)dy = x^2 - \frac{x^2}{2} = \frac{x^2}{2}$$

and so  $f_X(x) = x$  when  $x \in [0, 1]$ .

2. If  $x \in (1, 2]$ , then

$$F_X(x) = \mathbb{P}(X \leq x) = \int_0^1 F_N(x-y)f_M(y)dy = \int_0^{x-1} dy + \int_{x-1}^1 (x-y)dy = 2x - 1 - \frac{x^2}{2}.$$

Therefore,  $f_X(x) = 2 - x$  when  $x \in (1, 2]$ .

So, we have

$$f_X(x) = \begin{cases} 0 & x < 0 \\ x & x \in [0, 1] \\ 2 - x & x \in (1, 2] \\ 0 & x > 2 \end{cases}.$$

## 1.2 Exponential Random Variable

This random variable is useful when studying events occurring at random times. For example, lifetime of a lightbulb, time until the next customer, time until the next earthquake, etc.

Recall that  $F(x) = \mathbb{P}(X \leq x)$ . Hence,

$$1 - F(x) = \mathbb{P}(X > x).$$

This is sometimes denoted by  $S(x) = 1 - F(x)$  and is referred to as the **survival function** of the random variable  $X$ . Note that if  $X$  is a random time, e.g. the lifetime of a lightbulb, then  $S(x)$  is the probability of “surviving” until time  $x$ .

A very special type of continuous random variable is the exponential random variable with rate  $\lambda > 0$ . This is the random variable with survival function

$$S(x) = e^{-\lambda x}.$$

That is, the probability of survival until time  $X$  decays exponentially with rate  $\lambda$ . Note that  $F(x) = 1 - e^{-\lambda x}$ , and so  $f(x) = \lambda e^{-\lambda x}$  for  $x \geq 0$  and  $f(x) = 0$  otherwise is its PDF.

### 1.2.1 Memoryless Property

The exponential RV is very important because it is the only continuous RV with a special property, known as the **memoryless property**. Then, given that an exponential RV has survived until time  $x$ , the probability that it survives for  $y$  amount of time longer (i.e. until time  $x + y$ ) is the same as the probability of just surviving until time  $y$ .

For example, if an exponential lightbulb has survived until time  $x$ , then given this, the probability of surviving until time  $x + y$  is the same as the probability of a brand new lightbulb.