Continuous Probability Continued

In Lecture 11, we introduced continuous random variables. For example, let’s consider $X \sim \text{Unif}[0,1]$ which is a random variable that takes on continuous values. Then,

1. $P(X = a) = 0 \forall a$
2. $P(a < X \leq b) = b - a$, for $0 \leq a \leq b \leq 1$

In general, the probability that the random variable belongs to an interval of length $\Delta$ can be approximated by the probability density function (pdf) $f(x)$:

$$P(x < X \leq x + \Delta) \approx f(x)\Delta.$$  

Recall that probability values don’t have units. However, as can be seen from the previous approximation, the density of the random variable does have a unit. $\Delta$ has the same unit as the random variable itself. Therefore, the unit of $f$ should be the inverse of the unit of $\Delta$ so that the probability is unitless. For example, the unit of $f$ is $m^{-1}$ if the random variable is measured in meters.

Another Example: Throwing Darts

Suppose you throw a dart and it lands uniformly at random on a target which is a disk of unit radius. What is the probability density function of the distance of the dart from the center of the disk?

Let $X$ be the distance of the dart from the center of the disk. We first calculate the probability that $X$ is between $x$ and $x + \delta$. If $x$ is negative or greater than or equal to 1, this probability is zero, so we focus on the case that $x$ is between 0 and 1. The event in question is that the dart lands in the ring (annulus) shown in Figure 1. Since the dart lands uniformly at random on the disk, the probability of the event is just the ratio of the area of the ring and the area of the disk. Hence,

$$P(x < X \leq x + \delta) = \frac{\pi((x + \delta)^2 - x^2)}{\pi(1)^2} = x^2 + 2\delta x + \delta^2 - x^2 = 2\delta x - \delta^2.$$  

Using

$$f(x) = \lim_{\delta \to 0} \frac{P(x < X \leq x + \delta)}{\delta},$$  

from the last lecture, we can now compute the probability density function of $X$:

$$f(x) = \lim_{\delta \to 0} \frac{P(x < X \leq x + \delta)}{\delta} = f(x) = \lim_{\delta \to 0} \frac{2\delta x - \delta^2}{\delta} = 2x.$$
Figure 1: The sample space is the disk of unit radius. The event of interest is the ring.

Figure 2: (a) The probability density function and (b) the cumulative distribution function of the distance $X$ from the target center.

Summarizing, we have

$$f(x) = \begin{cases} 
0 & \text{for } x \leq 0; \\
2x & \text{for } 0 \leq x < 1; \\
0 & \text{for } x \geq 1.
\end{cases}$$

It is plotted in Figure 2(a). Note that although the dart lands uniformly inside the target, the distance $X$ from the center is not uniformly distributed in the range from 0 to 1. This is because an ring farther away from the center has a larger area than an ring closer to the center with the same width $\delta$. Hence the probability of landing in the ring farther away from the center is larger.
Cumulative Distribution Function

Let us re-interpret equation (1) in the dart throwing example above. In words, we are saying:

\[
P(x < X \leq x + \delta) = \frac{\text{area of ring}}{\text{area of target}} = \frac{(\text{area of disk of radius } x + \delta) - (\text{area of disk of radius } x)}{\text{area of target}} = P(X \leq x + \delta) - P(X \leq x).
\]

This last equality can be understood directly as follows. The event \( A \) that \( X \leq x + \delta \) (dart lands inside disk of radius \( x + \delta \)) can be decomposed as a union of two events: 1) the event \( B \) that \( X \leq x \) (dart lands inside disk of radius \( x \)), and 2) the event \( C \) that \( x < X \leq x + \delta \) (dart lands inside ring). The two events are disjoint. (See Figure 1.) Hence,

\[
P(A) = P(B) + P(C)
\]
or

\[
P(x < X \leq x + \delta) = P(X \leq x + \delta) - P(X \leq x),
\]

which is exactly the same as above.

Clearly, the reasoning leading to (3) has nothing much to do the particulars of this example but in fact (3) holds true for any random variable \( X \). All we needed are the facts that \( A = B \cup C \) and \( B \) and \( C \) are disjoint events, and the facts are true in general.

Substituting (3) into (2), we obtain:

\[
f(x) = \lim_{\delta \to 0} \frac{P(X \leq x + \delta) - P(X \leq x)}{\delta}.
\]

What does this equation remind you of? To make things even more explicit, let us define the cumulative distribution function

\[
F(x) = P(X \leq x).
\]

Then we have:

\[
f(x) = \lim_{\delta \to 0} \frac{F(x + \delta) - F(x)}{\delta} = \frac{d}{dx}F(x).
\]

The function \( F \) has a name: it is called the cumulative distribution function of the random variable \( X \) (sometimes abbreviated as cdf). It is called “cumulative” because at each value \( x \), \( F(x) \) is the cumulative probability up to \( x \). Note that the cumulative distribution function and the probability density function of a random variable contains exactly the same information. Given the cumulative distribution function \( F \), one can differentiate to get the probability density function \( f \). Given the probability density function \( f \), one can integrate to get the cumulative distribution function:

\[
F(x) = \int_{-\infty}^{x} f(a) \, da.
\]
Therefore the cdf and pdf contain the same information since they can be obtained from each other. So strictly speaking, one does not need to introduce the concept of cumulative distribution function at all. However, for many problems, the cumulative distribution function is easier to compute first and from that one can then compute the probability density function. Note that the cumulative distribution applies to both continuous and discrete random variables. That is because the cumulative distribution is a well-defined probability. In contrast, the probability density function applies only to continuous random variables. For discrete random variables, we should be using the probability mass function.

Summarizing, we have:

**Definition 13.1 (Cumulative Distribution Function):** The cumulative distribution function for a random variable $X$ is a function $F : \mathbb{R} \to \mathbb{R}$ defined to be:

$$F(x) = P(X \leq x). \quad (6)$$

Its relationship with the probability density function $f$ of $X$ is given by

$$f(x) = \frac{d}{dx} F(x), \quad F(x) = \int_{-\infty}^{x} f(a)da.$$ 

The cumulative distribution function satisfies the following properties:

1. $0 \leq F(x) \leq 1$
2. $\lim_{x \to -\infty} F(x) = 0$
3. $\lim_{x \to \infty} F(x) = 1$

The cdf of the distance $X$ of the dart from the origin is plotted in Figure 2 (b).

**Expectation and variance of a continuous random variable**

By analogy with the discrete case, we define the expectation of a continuous r.v. as follows:

**Definition 13.2 (Expectation):** The expectation of a continuous random variable $X$ with probability density function $f$ is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx.$$ 

Note that the integral plays the role of the summation in the discrete formula $\mathbb{E}[X] = \sum aP(X = a)$. Since variance is really just another expectation, we can immediately port its definition to the continuous setting as well:

**Definition 13.3 (Variance):** The variance of a continuous random variable $X$ with probability density function $f$ is

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \left( \int_{-\infty}^{\infty} xf(x)dx \right)^2.$$ 