Probability Density Function: A random variable x has a density if we have a function $f(x)$ such that

$$P(x \leq b) = \int_{-\infty}^{b} f(x)\,dx$$

For a uniform probability measure, we find that the density equals

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

We also know that

$$P(a \leq x \leq b) = \int_{a}^{b} f(x)\,dx$$

Discrete random variables do not have density.

Ex:

$$x = \begin{cases} 0 & p = \frac{1}{3} \\ 1 & p = \frac{2}{3} \end{cases}$$

We can try to construct a density function.

$$P(x \leq 0) = \frac{1}{3}$$

We want

$$\int_{0}^{0} f(x)\,dx = \frac{1}{3}$$

This is a contradiction, because we know that the definite integral of the same point is equal to zero.

For continuous random variables, we know that

$$P(x = a) = 0$$

which allows us to have density functions.

Properties of Density Functions: Suppose x is a random variable with density $f(x)$.

1.

$$\begin{aligned} P(a \le x < b) &= P(a \le x \le b) - P(x = b) \\ &= \int_a^b f(x)\,dx - 0 \\ &= P(a \le x \le b) \end{aligned}$$

In general, excluding or including one endpoint makes no difference on the probability.

2.

$$\begin{aligned} P(-\infty < x < \infty) &= \int_{-\infty}^{\infty} f(x)\,dx \\ 1 &= \int_{-\infty}^{\infty} f(x)\,dx \end{aligned}$$

Ex:

$$f(x) = \begin{cases} \frac{1}{x^2} & x \ge 1 \\ 0 & x < 1 \end{cases}$$

We seek to prove (a). $f(x)$ is non-negative, and (b). f(x)'s integral from $-\infty$ to ∞ is one.

(a). Yes, f(x) is non-negative at all points.

(b).

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx = f(x) &= \begin{cases} \int_{-\infty}^{1} \frac{1}{x^2} & x \ge 1 \\ 0 & x < 1 \end{cases} \\ &= \frac{-1}{\infty} + 1 \\ &= 1 \end{aligned}$$

Ex:

$$f_2(x) = \begin{cases} b\sqrt{a^2 - x^2} & |x| \le a \\ 0 & |x| > a \end{cases}$$

Assuming $a > 0$.

$$y = b\sqrt{a^2 - x^2}$$
$$y^2 = b^2(a^2 - x^2)$$

Then we want

$$\begin{aligned} \int_{-\infty}^{\infty} f_2(x)\,dx &= \int_{-a}^{a} b\sqrt{a^2 - x^2}\,dx \\ &= b\int_{-a}^{a} \sqrt{a^2 - x^2}\,dx \end{aligned}$$

$$\begin{aligned} \text{Let } x &= a\sin\theta \\ dx &= a\cos\theta\,d\theta \\ \sin\theta &= -1 \end{aligned}$$

$$\begin{aligned} b\int_{-a}^{a} \sqrt{a^2 - x^2}\,dx &= b\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{a^2 - a^2\sin^2\theta}a\cos\theta\,d\theta \\ &= b\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a^2\cos^2\ d\theta \\ &= ba^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta\,d\theta \\ &= ba^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1 + \cos 2\theta}{2}\,d\theta \\ &= ba^2\left(\frac{1}{2}\theta + \frac{1}{2}\frac{\sin 2\theta}{2}\right)\Big|_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \\ &= ba^2\frac{\pi}{2} \end{aligned}$$

Thus,

$$b = \frac{2}{\pi \cdot a^2}$$

Uniform Random Variables: A random variable x is $x \sim U[a, b]$ if and only if it has density

$$f(x) = \begin{cases} c & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}(b \cdot c) - (a \cdot c) &= 1 \\ c(b-a) &= 1 \\ &= \frac{1}{b-a}\end{aligned}$$

Thus, more formally, the density is defined as

$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Ex: Let $y \sim U[-2, 5]$, what is $P(|y| \ge 1)$?

$$\begin{aligned}P(|y| \ge 1) &= P(\{y \ge 1\} \cup \{y \le -1\}) \\ &= P(y \ge 1) + P(y \le -1) \\ &= \int_1^\infty f(x)\,dx + \int_{-\infty}^{-1} f(x)\,dx \\ &= \int_1^5 f(x)\,dx + \int_{-2}^{-1} f(x)\,dx \\ &= \frac{1}{7}x\Big|_1^5 + \frac{1}{7}x\Big|_{-2}^{-1} \\ &= \frac{5}{7}\end{aligned}$$

Cumulative Distribution Functions: There are random variables that do not have density nor are fully discrete.

We may have a random variable x which may take values $\in [0, 1]$, but an additional probability to take values outside that range, such as $\frac{3}{2}$ with probability $\frac{1}{4}$.

The cumulative distribution function (cdf) of a random variable x is the function on $(-\infty, \infty)$ given by

$$F(t) = P(x \le t), \text{for } t \in (-\infty, \infty)$$

All random variables have a cdf.

Ex: Roll a die,

$$x = \begin{cases} 1 & \text{die} \ge 5 \\ 0 & \text{otherwise} \end{cases}$$

This is a Bernoulli random variable, with pmf

$$f(x) = \begin{cases} \frac{1}{3} & x = 1 \\ \frac{2}{3} & x = 0 \end{cases}$$

and cdf

$$F(t) = \begin{cases} 0 & t < 0 \\ \frac{2}{3} & t = 0 \\ \frac{2}{3} & 0 \geq t < 1 \\ 1 & t = 1 \\ 1 & t \geq 1 \end{cases}$$

Ex: Roll a die

$$x = \begin{cases} \frac{1}{6} & x = 1 \\ \frac{1}{6} & x = 1 \\ \vdots & \\ \frac{1}{6} & x = 6 \end{cases}$$

The cdf is equal to

$$F(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{6} & t = 1 \\ \frac{2}{6} & t = 2 \\ \vdots & \\ 1 & t = 6 \\ 1 & t > 6 \end{cases}$$