Lecture 17

Continuous Random Variables.

Let X be a random variable with uncountable range.

We say that X is <u>continuous</u> if there is a non-negative function $f(x)$ such that, for any (measurable) subset $B \subseteq \mathbb{R}$,

$$P(X \in B) = \int_B f(x)\,dx.$$

We say that $f(x)$ is the <u>probability density function</u> (pdf) of X.

Since X takes on some value in $\mathbb{R}$, it follows that

$$P(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

In particular, we have

$$P(a \le X \le b) = \int_a^b f(x)\,dx.$$

Note that, for any particular value $a \in \mathbb{R}$,

$$P(X=a) = P(a \le X \le a) = \int_a^a f(x)\,dx = 0$$

It follows that

$$P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b)$$

By the fundamental theorem of calculus

$$\frac{d}{dx}\Big(P(X \le x)\Big) = \frac{d}{dx}\Big(\underbrace{\int_{-\infty}^{x} f(t)\,dt}_{F(x).}\Big) = f(x).$$

So we have that $F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt$ is the Cumulative distribution function of X and $f(x) = F'(x)$ is the pdf.

Example: The amount of time in hours that a computer functions between breaking down is a continuous random variable with density function

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that a computer will function between 50 and 100 hours before breaking down?

First: $$1 = \int_0^{\infty} \lambda e^{-x/100}\,dx = -100\lambda e^{-x/100}\Big|_0^{\infty}$$

$$= 100\lambda - \underbrace{\lim_{t\to\infty} 100\lambda e^{-t/100}}_{=0}$$

$$= 100\lambda.$$

So $\lambda = \frac{1}{100}$.

Thus, $$P(50 \le X \le 150) = \int_{50}^{150} \frac{1}{100} e^{-x/100}\,dx$$

$$= -e^{-x/100}\Big|_{50}^{150} = e^{-1/2} - e^{-3/2}.$$

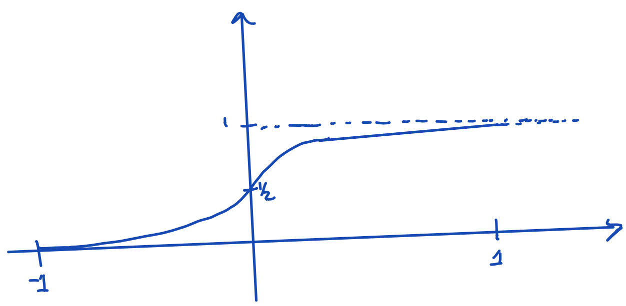
What is the probability that a computer will last greater than 100 hours?

$$P(X > 100) = \int_{100}^{\infty} \frac{1}{100} e^{-x/100}\,dx = -e^{-x/100}\Big|_{100}^{\infty}$$

$$= e^{-1}$$

<u>Ex:</u> Suppose that X is a continuous random variable with <u>cumulative distribution function</u> given by

$$F(x) = \begin{cases} 0 & \text{for } x < -1 \\ \frac{1}{2}(x+1)^2 & \text{for } -1 \le x < 0 \\ 1 - \frac{(1-x)^2}{2} & \text{for } 0 \le x < 1. \\ 1 & \text{for } x \geq 1. \end{cases}$$

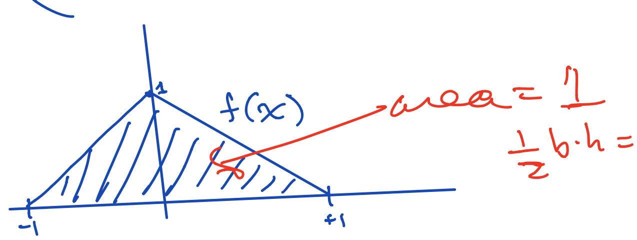


What is the pdf of X?

$$F(x) = \int_{-\infty}^{x} f(t)\,dt, \quad \text{so } \frac{d}{dx}F(x) = f(x).$$

Hence

$$f(x) = \begin{cases} 0 & \text{for } x < -1 \\ x+1 & \text{for } -1 \le x < 0 \\ 1-x & \text{for } 0 \le x < 1. \\ 0 & \text{for } x \geq 1. \end{cases}$$



Convenient property of CDF:

$$P(a \le X \le b) = \int_a^b f(x)\,dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx$$
$$= F(b) - F(a).$$

In the last example, we can find $P(-\frac{1}{2} \le X \le \frac{1}{2})$

as $\int_{-1/2}^{1/2} f(x)dx = F(1/2) - F(-1/2)$

$$= 1 - \frac{(1-1/2)^2}{2} - \frac{1}{2}(-1/2+1)^2$$

$$= 1 - \frac{1}{8} - \frac{1}{8} = 1 - \frac{2}{8} = 1 - 1/4 = 3/4.$$

Ex: Suppose that X is a random var with cdf F_X and pdf f_X. Find the pdf of $Y = e^X$.

$$F_Y(x) = P(Y \le x)$$
$$= P(e^X \le x)$$
$$= P(X \le \log(x))$$
$$= F_X(\log(x))$$

So $f_Y = \frac{d}{dx} F_X(\log(x)) = \frac{f_X(\log(x))}{x}$.

Expectation and Variance of a CRV

Recall that the expected value of a discrete random var X is $E[X] = \sum_x x\,P(X=x) = \sum_x x\,p(x)$, where $p(x) = P(X=x)$ is the pmf.

If X is a CRV, we have an analogous definition using the prob. density function:

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

Ex: For $f(x) = \begin{cases} 0 & \text{for } x \le -1 \\ x-1 & \text{for } -1 \le x \le 0 \\ 1-x & \text{for } 0 \le x \le 1 \\ 0 & \text{for } x \ge 1 \end{cases}$

What is the expected value?

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-1}^{0} x(x+1)\,dx + \int_{0}^{1} x(1-x)\,dx$$

$$= \int_{-1}^{0} x^2 + x\,dx + \int_{0}^{1} x - x^2\,dx$$

$$= \frac{x^3}{3} + \frac{x^2}{2}\Big|_{-1}^{0} + \frac{x^2}{2} - \frac{x^3}{3}\Big|_{0}^{1}$$

$$= -\left(\frac{-1}{3} + \frac{1}{2}\right) + \frac{1}{2} - \frac{1}{3}$$

$$= \frac{1}{3} - \frac{1}{2} + 1/2 - 1/3 = 0$$