Lecture 10

1 Continuous random variables

Let $X : \mathbb{R} \to \mathbb{R}$ be a random variable in a probability space $(\mathbb{R}, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is a σ-algebra which includes all line segments (i.e. $[a, b], (a, b], [a, b), (a, b) \in \mathcal{F}$ for all real numbers $a \leq b$. The smallest σ-algebra satisfying this property is known as the Borel algebra). We say that X is a continuous (or absolutely continuous) random variable if there exists a function $f_X : \mathbb{R} \to \mathbb{R}$ such that

$$\mathbb{P}(X \in B) = \int_B f_X(x)dx \tag{1}$$

for every $B \in \mathcal{F}$. In particular, we must have $\mathbb{P}(\mathbb{R}) = \int_{-\infty}^{\infty} f_X(x)dx = 1$. The function f_X is called the *probability density function* of X. It is immediate from (1) that $\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx$ for all real numbers $a \leq b$. In particular, $\mathbb{P}(X = c) = \int_c^c f_X(x)dx = 0$ for every $c \in \mathbb{R}$ and so

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

for all real numbers $a \leq b$. For every $a \in \mathbb{R}$ let $F_X(a) = \mathbb{P}(X \leq a) = \int_{-\infty}^a f_X(x)dx$. The function F_X is called the *cumulative distribution function* (or CDF for brevity) of X.

Example 1: Let X be a continuous random variable whose probability density function is given by

$$f_X(x) = \begin{cases} c(1 - x^2) & \text{if } -1 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

In order to determine the value of c we use the equation $\int_{-\infty}^{\infty} f_X(x)dx = 1$. We thus have

$$1 = \int_{-\infty}^{\infty} f_X(x)dx = \int_{-1}^{1} c(1 - x^2)dx = c[x - x^3/3]|_{x=-1}^{x=1} = c[(1 - 1/3) - (-1 + 1/3)] = 4c/3$$

implying that $c = 3/4$. We can now calculate various probabilities involving X. For example

$$\mathbb{P}(X \leq 0) = \int_{-\infty}^{0} f_X(x)dx = \int_{-1}^{0} 3(1-x^2)/4dx = [3x/4-x^3/4]|_{x=-1}^{x=0} = 0-(-3/4+1/4) = 1/2.$$

Example 2: Let X be a continuous random variable with probability density function f_X and cumulative distribution function F_X. We would like to find the probability density function and the cumulative distribution function of the random variable $Y = 2X$. For every $a \in \mathbb{R}$ it holds that

$$F_Y(a) = \mathbb{P}(Y \le a) = \mathbb{P}(2X \le a) = \mathbb{P}(X \le a/2) = F_X(a/2).$$

Differentiating yields

$$f_Y(a) = \frac{1}{2} f_X(a/2).$$

1.1 Expectation and variance of continuous random variables

Recall from Probability Theory 1 that if X is a discrete random variable, then its expectation is given by $\mathbb{E}(X) = \sum_x x\mathbb{P}(X = x)$, where the sum is extended over the support of $\mathbb{P}(X = x)$. An analogous definition for the expectation of a continuous random variable would thus be $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx$.

Example 3: Let X be a continuous random variable whose probability density function is given by

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

First, note that this is indeed a probability density function as

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_0^1 2xdx = x^2|_{x=0}^{x=1} = 1.$$

Now

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 2x^2 dx = 2x^3/3|_{x=0}^{x=1} = 2/3.$$

Similarly to the case of discrete random variables, we can use the probability density function of X in order to calculate the expectation of some function of X.

Proposition 1.1. *Let X be a continuous random variable with probability density function f_X. Then, for any function $g : \mathbb{R} \to \mathbb{R}$, it holds that*

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)dx.$$

For simplicity, we will prove Proposition 1.1 only for non-negative functions g. We will first state and prove an auxiliary lemma.

Lemma 1.2. *Let Y be a non-negative continuous random variable with probability density function f_Y. Then*

$$\mathbb{E}(Y) = \int_0^{\infty} \mathbb{P}(Y > y)\, dy.$$

Proof. Since $\mathbb{P}(Y > y) = \int_y^\infty f_Y(x)dx$ holds for every $y \in \mathbb{R}$, it follows that

$$\int_0^\infty \mathbb{P}(Y > y)\, dy = \int_0^\infty \int_y^\infty f_Y(x)dxdy = \int_0^\infty \left(\int_0^x dy\right) f_Y(x)dx = \int_0^\infty x f_Y(x)dx = \mathbb{E}(Y),$$

where we obtained the second equality by interchanging the order of integration and the last equality holds since Y is non-negative. □

Proof of Proposition 1.1. Under the additional assumption that g is non-negative, it follows from Lemma 1.2 that

$$\begin{aligned}\mathbb{E}(g(X)) &= \int_0^\infty \mathbb{P}(g(X) > y)\, dy = \int_0^\infty \int_{x:g(x)>y} f_X(x)dxdy = \int_{x:g(x)>0} \left(\int_0^{g(x)} dy\right) f_X(x)dx \\ &= \int_{x:g(x)>0} g(x) f_X(x)dx = \int_{-\infty}^\infty g(x) f_X(x)dx.\end{aligned}$$

□

Proposition 1.3 (Linearity of Expectation)**.** *Let $a, b \in \mathbb{R}$ and let X be a continuous random variable. Then*

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

Proof. Let f_X denote the probability density function of X. It follows by Proposition 1.1 that

$$\mathbb{E}(aX + b) = \int_{-\infty}^\infty (ax + b) f_X(x)dx = a\int_{-\infty}^\infty x f_X(x)dx + b\int_{-\infty}^\infty f_X(x)dx = a\mathbb{E}(X) + b.$$

□

As in the case of a discrete random variable, we define the *variance* of a continuous random variable X to be $Var(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]$, and it then follows by a straightforward calculation that $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

Example 4: Let X be a continuous random variable whose probability density function is given by

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

We have seen before that $\mathbb{E}(X) = 2/3$. Using Proposition 1.1 we obtain

$$\mathbb{E}(X^2) = \int_{-\infty}^\infty x^2 f_X(x)dx = \int_0^1 2x^3 dx = x^4/2|_{x=0}^{x=1} = 1/2.$$

Therefore

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 1/2 - (2/3)^2 = 1/18.$$

Proposition 1.4. *Let $a, b \in \mathbb{R}$ and let X be a continuous random variable. Then*

$$Var(aX + b) = a^2 Var(X).$$

The proof is omitted as it is identical to the one that was given in Probability Theory 1 for discrete random variables.