Recall

Cumulative distribution function

The *cumulative distribution function* (CDF) of a random variable X is defined by

$$F(t) \coloneqq P\{X \leq t\}, \quad \forall\, t \in \mathbb{R}$$

which has the following properties:

- Non-decreasing. • Right-continuous. • $\lim_{t\to-\infty} F(t) = 0$ and $\lim_{t\to+\infty} F(t) = 1$.

All probability questions about X can be answered in terms of CDF. In particular, for $x \in \mathbb{R}$, $P\{X < x\} = \lim_{t\to x-} F(t)$.

Continuous random variable

A random variable X is (*absolutely*) *continuous* if there exists a function, called *probability density function* (PDF), such that

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

where B is a 'measurable' set in $\mathbb{R}$. Fortunately, countable unions and intersections of intervals are 'measurable'.

Below are some facts about a **continuous** random variable X:

*Unit integral of **a** PDF.* $\int_{-\infty}^{+\infty} f(x)\, dx = 1$.

Zero probability at any point. $\forall\, x \in \mathbb{R}$, $P\{X = x\} = 0$.

Cumulative distribution function. $\forall\, t \in \mathbb{R}$, $F(t) \coloneqq \int_{-\infty}^{t} f(x)\, dx$.
For $t \in \mathbb{R}$, it follows from $F(t) = P\{X \leq t\} = P\{X < t\} = \lim_{x\to t-} F(x)$ that $F(t)$ is left-continuous, hence continuous, at t. In conclusion, the CDF of a continuous r.v. is continuous.

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

Continuous layer-cake. If X is continuous and non-negative, then $E[X] = \int_0^{+\infty} P\{X > t\}\, dt$.

LOTUS. Let $g\colon \mathbb{R} \to \mathbb{R}$. Then $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x)\, dx$.

Variance. $\mathrm{Var}(X) \coloneqq E[(X - E[X])^2] = E[X^2] - (E[X])^2$.

Affine transform. For $a, b \in \mathbb{R}$, $\begin{cases} E[aX + b] = aE[X] + b; \\ \mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}(X). \end{cases}$

Relation between PDF f and CDF F. If f is continuous at $x \in \mathbb{R}$, then $F(x)' = \frac{dF(x)}{dx} = f(x)$.

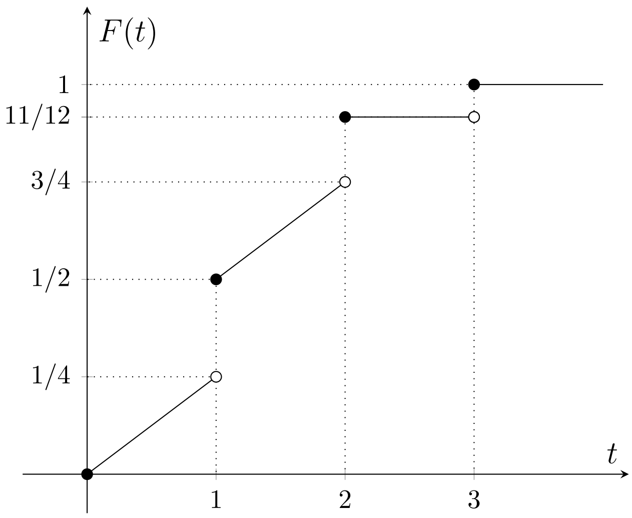
Probability computation from CDF

Example 1. Suppose a random variable X has CDF

$$F(t) = \begin{cases} 0 & t \in (-\infty, 0) \\ t/4 & t \in [0, 1) \\ 1/2 + (t-1)/4 & t \in [1, 2) \\ 11/12 & t \in [2, 3) \\ 1 & t \in [3, +\infty). \end{cases}$$

Find $P\{X = i\},\ i = 1, 2, 3$ and $P\{1 \leq X < 3\}$.

Solution. Below is the graph of $F(t)$.



Then

$$\begin{aligned} P\{X = 1\} &= P\{X \leq 1\} - P\{X < 1\} = F(1) - \lim_{t \to 1-} F(t) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}, \\ P\{X = 2\} &= P\{X \leq 2\} - P\{X < 2\} = F(2) - \lim_{t \to 2-} F(t) = \frac{11}{12} - \frac{3}{4} = \frac{1}{6}, \\ P\{X = 3\} &= P\{X \leq 3\} - P\{X < 3\} = F(3) - \lim_{t \to 3-} F(t) = 1 - \frac{11}{12} = \frac{1}{12}. \end{aligned}$$

And

$$P\{1 \leq X < 3\} = P\{X < 3\} - P\{X < 1\} = \lim_{t \to 3-} F(t) - \lim_{t \to 1-} F(t) = \frac{11}{12} - \frac{1}{4} = \frac{2}{3}.$$

□

Remark. Since the CDF of a discrete random variable should be like a step function, it follows that X in Example 1 is not discrete. On the other hand, X is not a continuous random variable either because the CDF of a continuous random variable should be continuous.

Some computations about continuous random variables

Example 2. Let X be a random variable with PDF

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

Find the value of c and the CDF of X.

Solution. Since f is a PDF, we have

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_{-1}^{1} c(1-x^2)\,dx = c(x - \frac{x^3}{3})\Big|_{-1}^{1} = \frac{4}{3}c,$$

which implies $c = \frac{3}{4}$. Recall that for $t \in \mathbb{R}$, the CDF $F(t) := \int_{-\infty}^{t} f(x)dx$.

$$\text{If } t \leq -1, \text{ then } F(t) = \int_{-\infty}^{t} f(x)dx = \int_{-\infty}^{t} 0dx = 0,$$
$$\text{If } 1 < t \leq 1, \text{ then } F(t) = \int_{-\infty}^{t} f(x)dx = \int_{-1}^{t} \frac{3}{4}(1-x^2)dx = \frac{3}{4}(t - \frac{t^3}{3} + \frac{2}{3}) = -\frac{t^3}{4} + \frac{3t}{4} + \frac{1}{2},$$
$$\text{If } t > 1, \text{ then } F(t) = P(X \leq t) = 1 - P(X > t) = 1 - \int_{t}^{\infty} 0dx = 1.$$

Thus

$$F(t) = \begin{cases} 0 & t \in (-\infty, -1] \\ -\frac{t^3}{4} + \frac{3t}{4} + \frac{1}{2} & t \in (-1, 1] \\ 1 & t \in (1, \infty). \end{cases}$$

□

Example 3. Let X be a random variable with PDF f_X. Find a PDF of random variable $Y = aX + b$ where $0 \neq a \in \mathbb{R}, b \in \mathbb{R}$.

Solution. Let F_X and F_Y denote the CDFs of X and Y respectively. For $t \in \mathbb{R}$,

$$F_Y(t) = P\{Y \leq t\} = P\{aX + b \leq t\}.$$

If $a > 0$, then $F_Y(t) = P\{X \leq \frac{t-b}{a}\} = F_X(\frac{t-b}{a})$. When F_X is differentiable at $\frac{t-b}{a}$, by chain rule

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \frac{1}{a} f_X(\frac{t-b}{a}).$$

When F_X is NOT differentiable at $\frac{t-b}{a}$, we define $f_Y(t) = \frac{1}{a} f_X(\frac{t-b}{a})$. Together, when $a > 0$, a possible PDF of Y is

$$f_Y(t) = \frac{1}{a} f_X(\frac{t-b}{a}) \quad , \forall\, t \in \mathbb{R}.$$

If $a < 0$, then $F_Y(t) = P\{X \geq \frac{t-b}{a}\} = 1 - P\{X < \frac{t-b}{a}\} = 1 - P\{X \leq \frac{t-b}{a}\} = 1 - F_X(\frac{t-b}{a})$. We omit the discussion about differentiability. By differentiation, when $a < 0$, a PDF of Y is

$$f_Y(t) = \frac{dF_Y(t)}{dt} = -\frac{1}{a} f_X(\frac{t-b}{a}) \quad , \forall\, t \in \mathbb{R}.$$

□

Remark. In Example 3, we have carefully dealt with the differentiability of a CDF in the case of $a > 0$, which is the rigorous way to think about it. However, in practice we **omit** the discussion because we know that a CDF is differentiable at **most** points. Then as in Example 3, we adjust the values on the **tiny** part of non-differentiable points to simplify the final results.

Let f be a PDF of a coninuous random variable X. After changing values of f on a **tiny** part of $\mathbb{R}$, the resulted f is still a PDF of X.

Remark. Let X be a continuous random variable and $g\colon \mathbb{R} \to \mathbb{R}$ be any function. The following example shows that we are not even sure whether $g(X)$ has a PDF. Actually, in Example 3 we have **omitted** the step to prove that $Y = aX + b$ is indeed continuous with a PDF. In practice, when the question asks for a PDF, we can take it for granted that the target PDF exists like Example 3 and Example 5.

Example 4. Let $g(x) = 0$ for all $x \in \mathbb{R}$. Then for any random variable X (including the continuous ones), $g(X)$ is the discrete random variable such that $P\{g(X) = 0\} = 1$.

Proof. Let F denote the CDF of $g(X)$. Then for $t \in \mathbb{R}$,

$$F(t) := P\{g(X) \le t\} = P\{0 \le t\} = \begin{cases} 1 & t \ge 0 \\ 0 & t < 0. \end{cases}$$

Hence $g(X)$ is the discrete random variable such that $P\{g(X) = 0\} = 1$. □

Example 5. Suppose the CDF of X is

$$F(t) = \begin{cases} 1 - e^{-t^2} & t > 0 \\ 0 & t \le 0. \end{cases}$$

Find $P\{X > 2\}$ and a PDF of X.

Solution. First

$$P\{X > 2\} = 1 - P\{X \le 2\} = 1 - F(2) = e^{-4}.$$

Then

$$\text{If } x > 0, \text{ then } \frac{dF(x)}{dx} = 2xe^{-x^2}.$$
$$\text{If } x < 0, \text{ then } \frac{dF(x)}{dx} = 0.$$

Define

$$f(x) = \begin{cases} 2xe^{-x^2} & x > 0, \\ 0 & x \le 0. \end{cases}$$

Hence $f(x)$ is a PDF of X. □