

Probability Distribution (Continuous random variable)

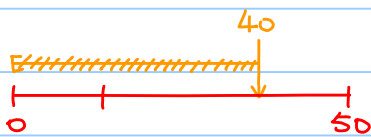
From Discrete to Continuous case :

e.g. A certain traffic light remains red for 50 seconds at a time.

Andy arrives (at random) at the light and finds it red.

Let X be the random variable denotes the waiting time of Andy.

Q: $P(\text{Andy has to wait for at least 10 sec})$?



Arrive at time x

Waiting time = $50 - x$

A: Required probability = $\frac{40}{50} = \frac{4}{5}$

$\frac{\text{length of the interval } [0, 40]}{\text{length of the interval } [0, 50]}$

Q: $P(\text{Andy has to wait for exactly 10 sec})$?

A: Required probability = $\frac{0}{50} = 0$ (strange!)

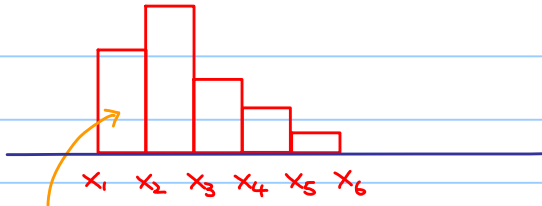
$\frac{\text{length of the interval } \{40\}}{\text{length of the interval } [0, 50]}$ ← single point.

Conclusion: We cannot use the same treatment for continuous case!

Basically, we cannot talk about the probability for a particular value.

but we can talk about the probability for an interval.

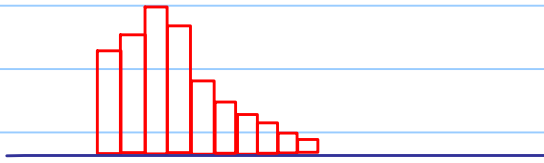
Think :



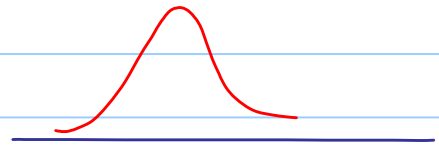
$$\text{Area} = P(x_1 \leq X \leq x_2)$$

\therefore Summing up the areas of the rectangles = 1

↓ More partitions



taking limit
 \Rightarrow



Probability Distribution (Continuous random variable)

X : Continuous random variable

A probability distribution (or probability density function pdf) is a function f such that

1) $f(x) \geq 0$

2) $\int_{-\infty}^{+\infty} f(x) dx = 1$

and under this distribution, we have

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

e.g. A certain traffic light remains red for 50 seconds at a time. Andy arrives (at random) at the light and finds it red.

Sample space = $S = [0, 50]$

Let X be the random variable that denotes the waiting time of Andy

$X: S \rightarrow \mathbb{R}$

$$X(x) = \begin{cases} 50-x & \text{if } 0 \leq x \leq 50 \\ 0 & \text{otherwise} \end{cases}$$

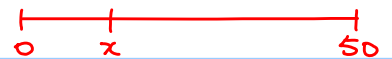


image of $X = [0, 50]$

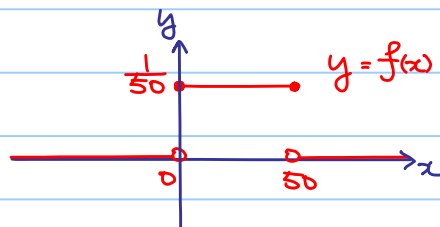
which is an uncountably infinite

$\therefore X$ is a continuous random variable.

Arrive at time x

Waiting time = $50 - x$

$$f(x) = \begin{cases} \frac{1}{50} & \text{if } 0 \leq x \leq 50 \\ 0 & \text{otherwise} \end{cases}$$



$$\begin{aligned} P(10 \leq X \leq 50) &= \int_{10}^{50} f(x) dx \\ &= \left[\frac{1}{50} x \right]_{10}^{50} = \frac{40}{50} = \frac{4}{5} \end{aligned}$$

Remark:

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

This probability density function is called a uniform distribution.

(More examples of probability density function)

Exponential Density Function

$$\text{Let } f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}, \text{ where } \lambda > 0$$

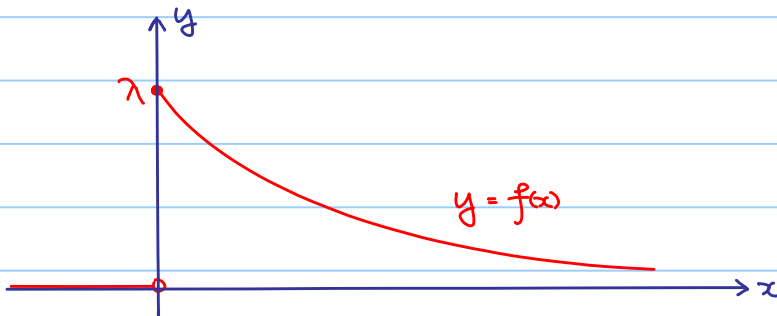
Verify f is a probability distribution!

① $f(x) \geq 0$

Easy to check!

② $\int_{-\infty}^{+\infty} f(x) dx = 1$

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= \int_0^{+\infty} \lambda e^{-\lambda x} dx \\ &= \lim_{N \rightarrow +\infty} \int_0^N \lambda e^{-\lambda x} dx \\ &= \lim_{N \rightarrow +\infty} [-e^{-\lambda x}]_0^N \\ &= \lim_{N \rightarrow +\infty} 1 - e^{-\lambda N} = 1 \end{aligned}$$



e.g. (P. 846, Example 11.2.4)

Let X be a random variable that measures the duration of cell phone calls in a certain city, and **assume** that X has an exponential distribution with density

$$f(t) = \begin{cases} 0.5e^{-0.5t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

where t denotes the duration (in minutes) of a randomly selected call.

- Find the probability that a randomly selected call lasts **no more than 1 minute**.
- Find the probability that a randomly selected call will last **at least 2 minutes**.

$$a) P(0 \leq X \leq 1) = \int_0^1 f(t) dt = \int_0^1 0.5e^{-0.5t} dt = \dots = 1 - e^{-0.5} \approx 0.3935 \quad \text{Ex.}$$

$$b) P(X \geq 2) = \int_2^{+\infty} f(t) dt = \lim_{N \rightarrow +\infty} \int_2^N 0.5e^{-0.5t} dt = \dots = e^{-1} \approx 0.3679 \quad \text{Ex.}$$

Remark:

1) $P(X \geq 2) = P(X > 2)$ **why?**

2) How to obtain the probability distribution?

Even we conduct a survey to collect all data,

we have finitely many calls (but a big number)

then the sample space = S = set of all calls

$X: S \rightarrow \mathbb{R}$ is a random variable that denotes the duration
image of S is still finite.

But it can be approximated by the X in the question.

Cumulative distribution function (cdf)

$$F(x) = P(X \leq x)$$

= Probability of getting value $\leq x$

Q: What is the relation between the pdf and cdf?

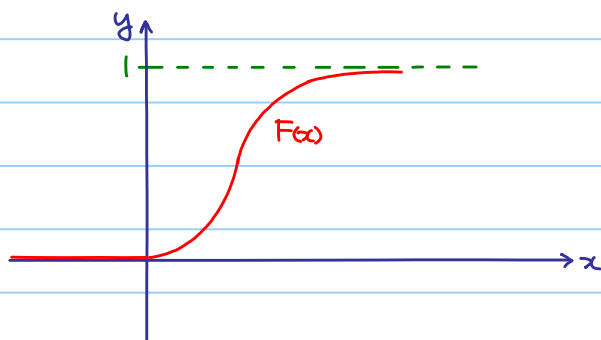
$$A: F(x) = P(X \leq x) \\ = \int_{-\infty}^x f(t) dt$$

$$\text{e.g. } f(t) = \begin{cases} 0.5e^{-0.5t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

$$F(x) = \int_{-\infty}^x f(t) dt \quad \text{if } x \geq 0 \\ = \int_0^x 0.5e^{-0.5t} dt \\ \text{Ex.} \\ = \dots = 1 - e^{-0.5x}$$

$$F(x) = \int_{-\infty}^x f(t) dt \quad \text{if } x < 0 \\ = \int_{-\infty}^x 0 dt = 0$$

$$\therefore F(x) = \begin{cases} 1 - e^{-0.5x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$



Given $f(x)$, we can find $F(x)$ by

$$F(x) = \int_{-\infty}^x f(t) dt$$

But, if $F(x)$ is given, how to find $f(x)$?

Ans: Simply by FTC,

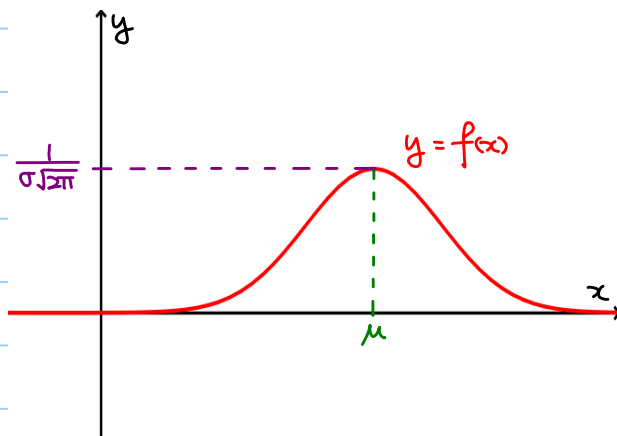
$$F'(x) = f(x) \quad (\text{if exists})$$

i.e. cdf is an antiderivative of pdf.

Normal distribution

A normal random variable with the density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



bell-shape

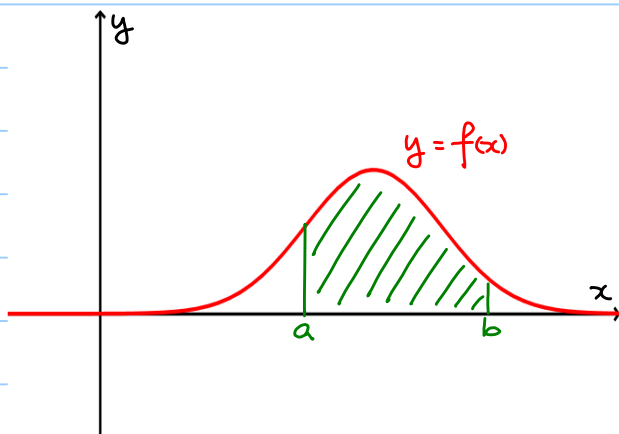
distribution of height, weight, and etc can be approximated by normal distribution.

FACT: ① $f(x) \geq 0$

② $\int_{-\infty}^{+\infty} f(x) dx = 1$ ← hard to show.

③ μ and σ are exactly the mean and standard deviation respectively (discuss later)

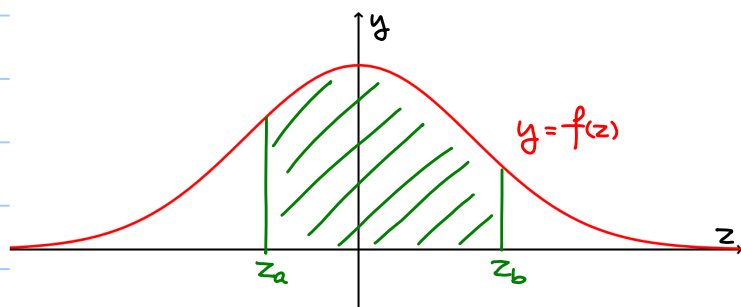
How to find $P(a \leq X \leq b) = \int_a^b f(x) dx$?



Standard score

$$\text{let } z = \frac{x - \mu}{\sigma}$$

After transformation $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$



$$z_a = \frac{a - \mu}{\sigma}$$

$$z_b = \frac{b - \mu}{\sigma}$$

$$\int_a^b f(x) dx = \int_{z_a}^{z_b} f(z) dz \quad (\text{Change of variable})$$

↑

We have a table to obtain an approximation of it.

Idea: Normal distributions with different μ, σ

↳ transformation $z = \frac{x - \mu}{\sigma}$

Make it to be the standard one ($\mu = 0, \sigma = 1$)

And then check the table

For detail, refer to section 11.4, very useful!