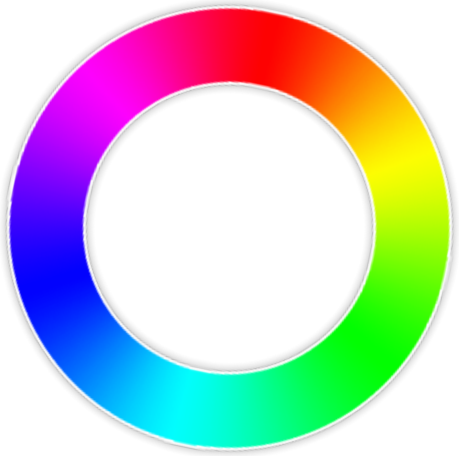
Continuous Random Variables

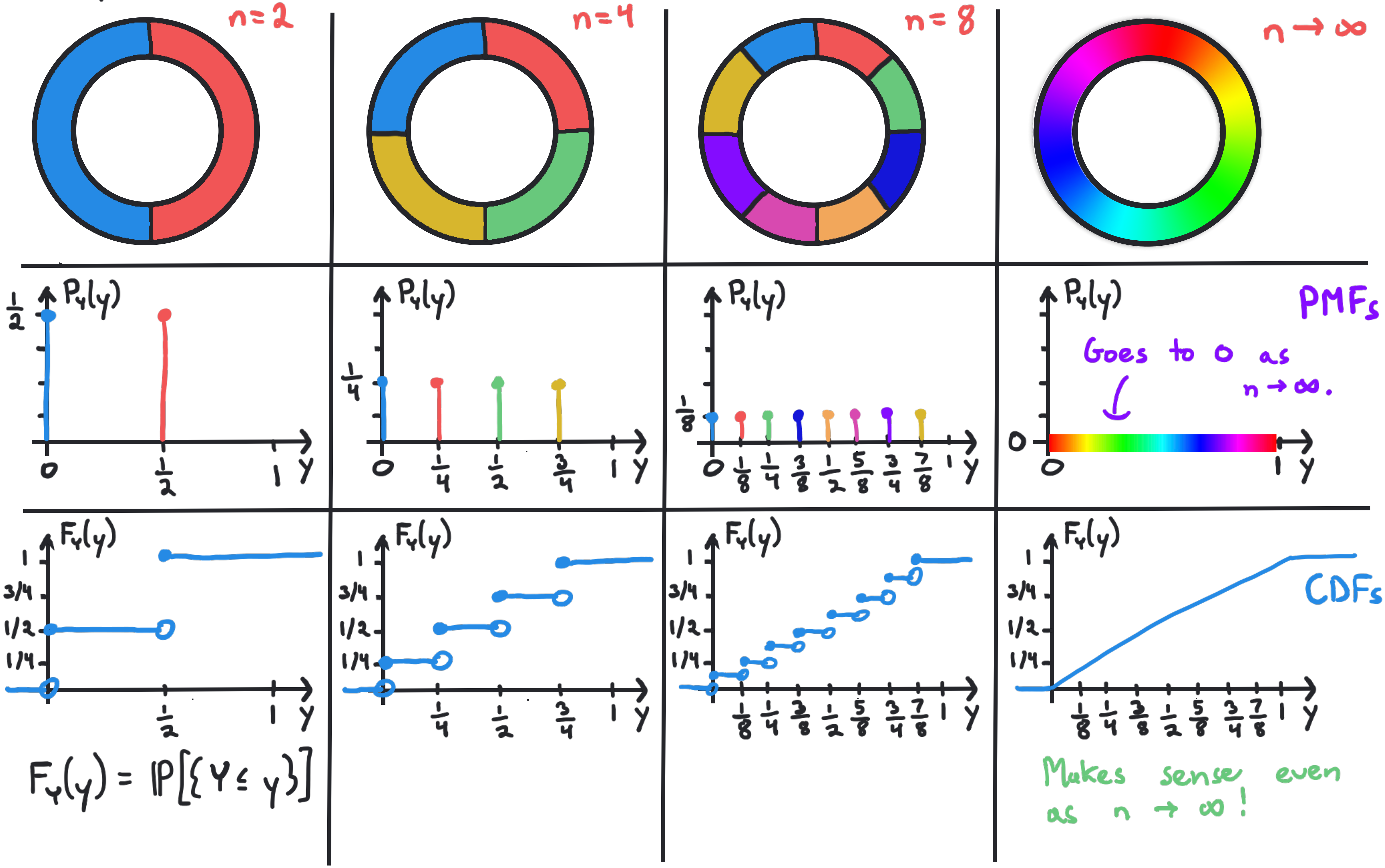
- What if the range of a random variable consists of one or more intervals?
  - → Ex: Measure a voltage between -5 to $+5$ Volts.
  - → Range is no longer discrete so we need a a new type of random variable.



- Motivating Example: → Wheel (of fortune) of diameter 1 meter.
  - → Spin the wheel very fast and record where it stops.
  - → Reasonable to model all stopping positions as equally likely.
  - → Overall, we have a random variable X that is equally likely to take any value $x \in [0,1]$.
  - → What is the probability that $X = x$ for some value $x \in [0,1]$?

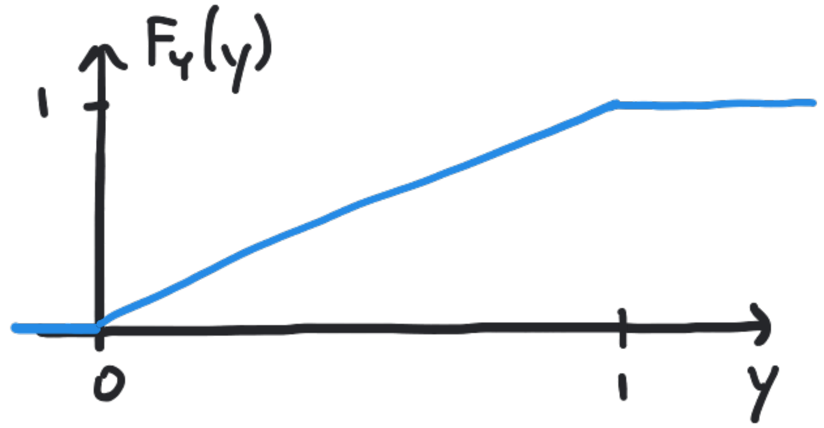
    Intuitively, this should be 0.
  - → Let's try to model this scenario using discrete random variables.

• Divide wheel into n equal-sized pieces. Let Y be the slice you land in. (Formally, let $Y = \frac{1}{n} \cdot \text{floor}(nX)$.) $R_Y = \{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\}$.



- As n increases, the PMF values tend to 0, while the CDF remains well-behaved. Specifically, as $n \to \infty$,

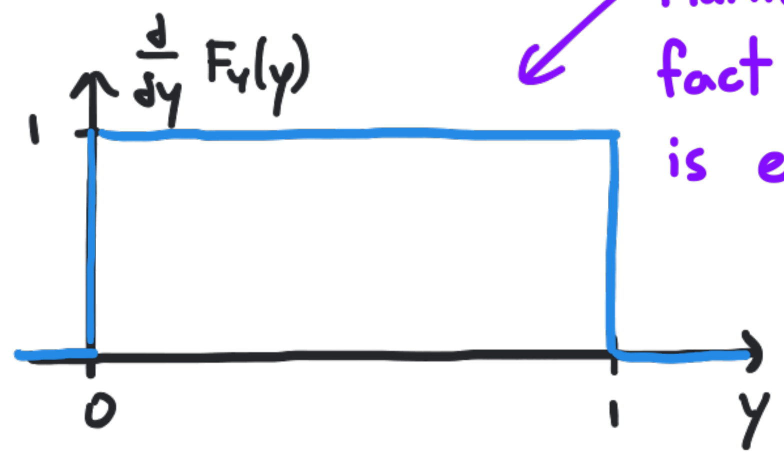
$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \le y \le 1 \\ 1 & 1 < y \end{cases}$$



→ It would be nice to have a way to represent the fact that all positions $y \in [0,1]$ are equally likely.

- <u>Idea:</u> Take the derivative of the CDF.

$$\frac{d}{dy} F_Y(y) = \begin{cases} 0 & y < 0 \\ 1 & 0 \le y < 1 \\ 0 & 1 < y \end{cases}$$



flatness captures the fact that any $y \in [0,1]$ is equally likely

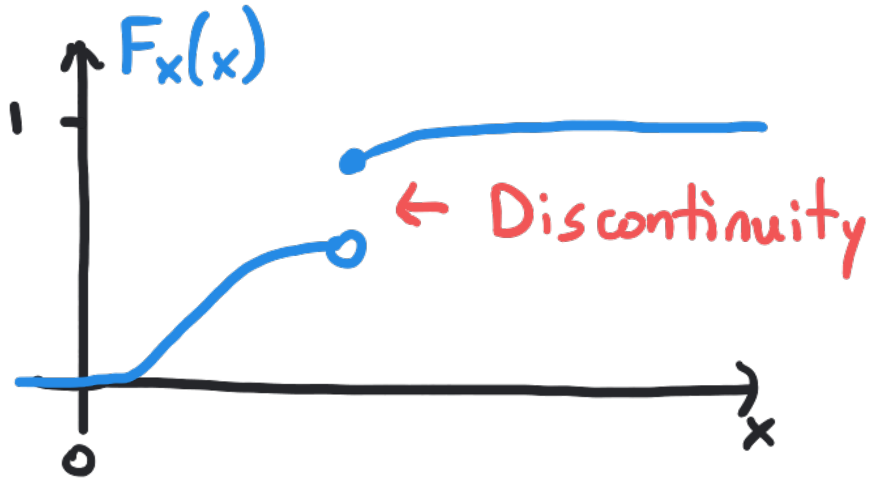
- Note that $\mathbb{P}[\{Y = y\}] = 0$ for any choice of $y \in [0,1]$.

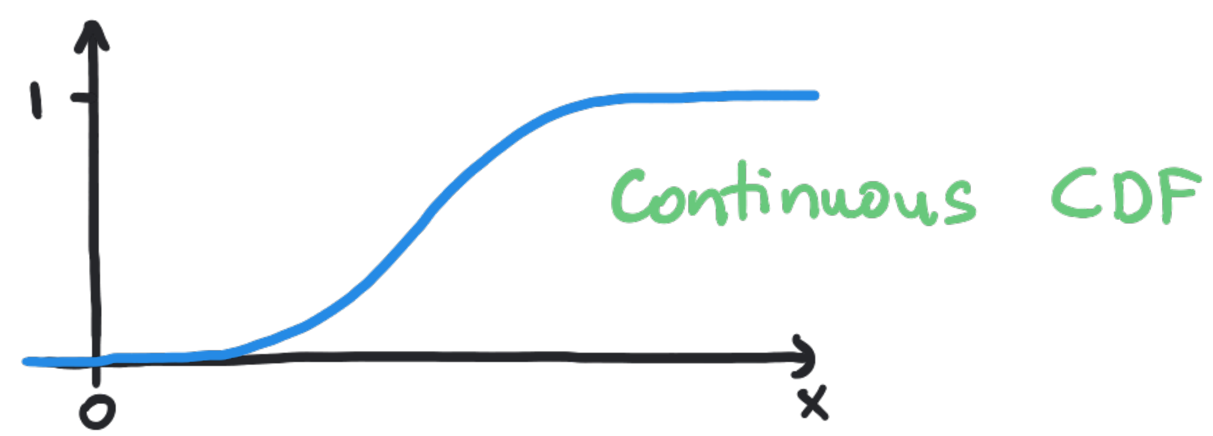
Cumulative Distribution Function

- Recall that we defined the cumulative distribution function (CDF) $F_X(x)$ as the probability that a random variable X is less than or equal to the value x,

$$\boxed{\begin{aligned} F_X(x) &= \mathbb{P}[\{\omega \in \Omega : X(\omega) \leq x\}] \\ &= \mathbb{P}[\{X \leq x\}] \quad \text{shorthand notation} \\ &= \mathbb{P}[X \leq x] \quad \text{shorthand notation} \end{aligned}}$$

- The CDF is continuous if it has no abrupt changes, $\lim_{\epsilon \to 0} F_X(x+\epsilon) = F_X(x)$ for all x.



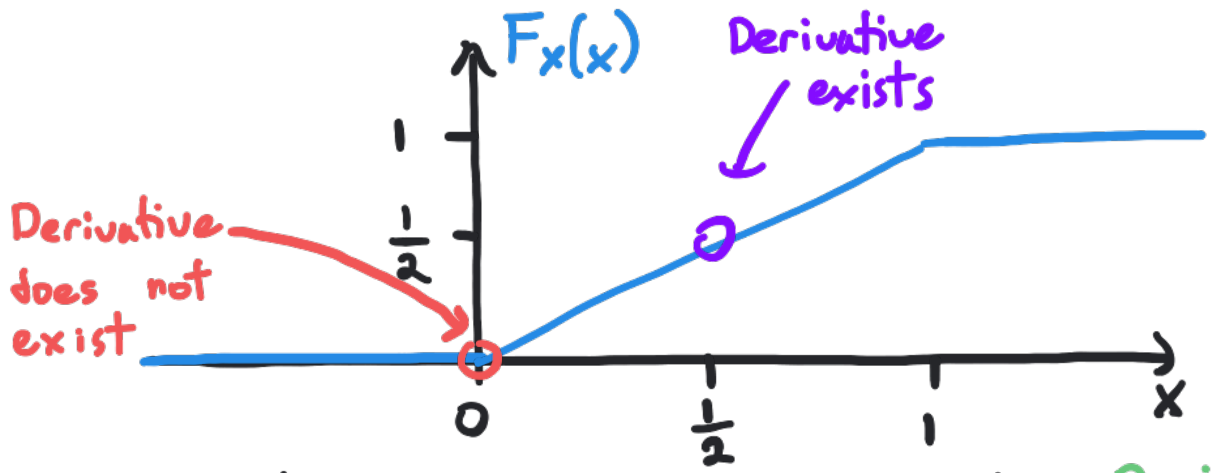


- A random variable X is continuous if it has a continuous CDF $F_X(x)$ and this CDF is differentiable almost everywhere.

  Everywhere except for a countable number of values where the derivative does not exist.

- Recall that a function $g(x)$ is differentiable at x if the derivative $g'(x) = \lim_{\epsilon \to 0} \frac{g(x+\epsilon) - g(x)}{\epsilon}$ exists.

→ Ex: $F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & 1 < x \end{cases}$



$$\lim_{\epsilon \downarrow 0} \frac{F_X(\frac{1}{2}+\epsilon) - F_X(\frac{1}{2})}{\epsilon} = \lim_{\epsilon \downarrow 0} \frac{\frac{1}{2} + \epsilon - \frac{1}{2}}{\epsilon} = 1 = \lim_{\epsilon \uparrow 0} \frac{\frac{1}{2} + \epsilon - \frac{1}{2}}{\epsilon}$$

Limit from the right — Limit from the left

Derivative exists.

$$\lim_{\epsilon \downarrow 0} \frac{F_X(0+\epsilon) - F_X(0)}{\epsilon} = \lim_{\epsilon \downarrow 0} \frac{0+\epsilon - 0}{\epsilon} = 1$$

$$\lim_{\epsilon \uparrow 0} \frac{F_X(0+\epsilon) - F_X(0)}{\epsilon} = \lim_{\epsilon \uparrow 0} \frac{0 - 0}{\epsilon} = 0$$

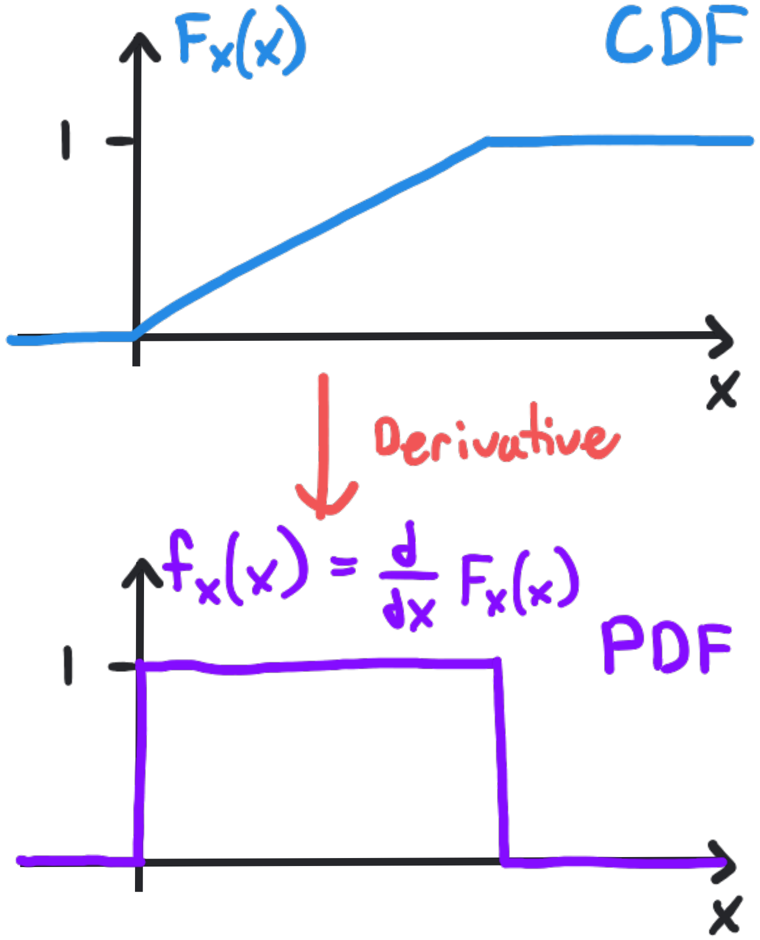
Limit from the right and from the left do not agree so the derivative does not exist at $x = 0$

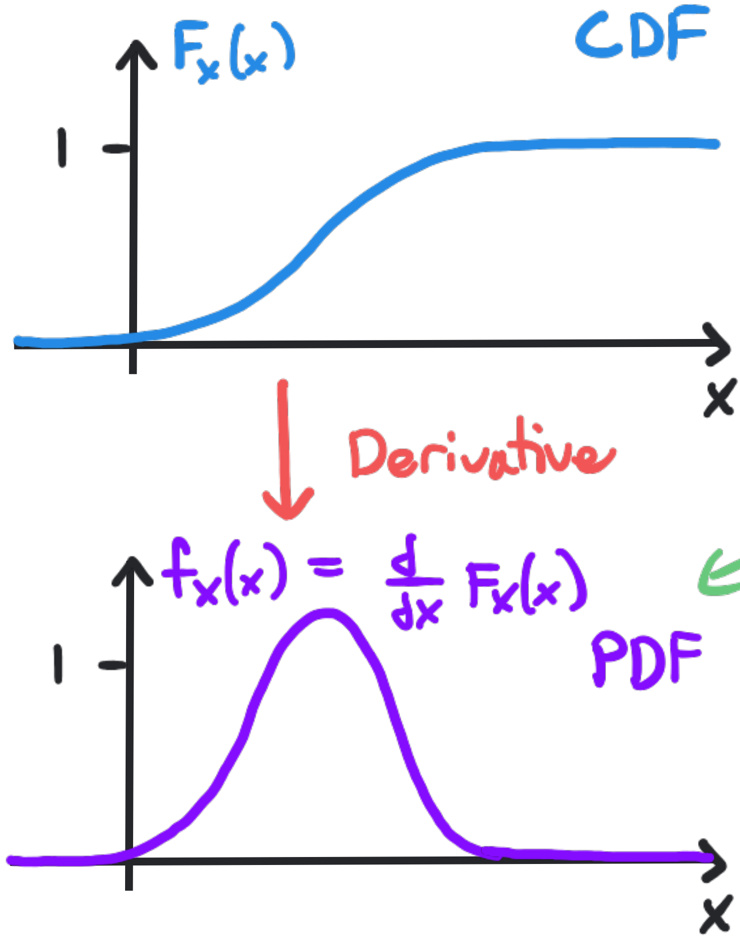
Probability Density Function

- The probability density function (PDF) $f_X(x)$ of a continuous random variable X is the derivative of its CDF $F_X(x)$.

$$\boxed{f_X(x) = \begin{cases} \dfrac{d}{dx} F_X(x) & \text{if } F_X(x) \text{ is differentiable at } x \\ \text{any non-negative number} & \text{otherwise} \end{cases}}$$

- Ex:

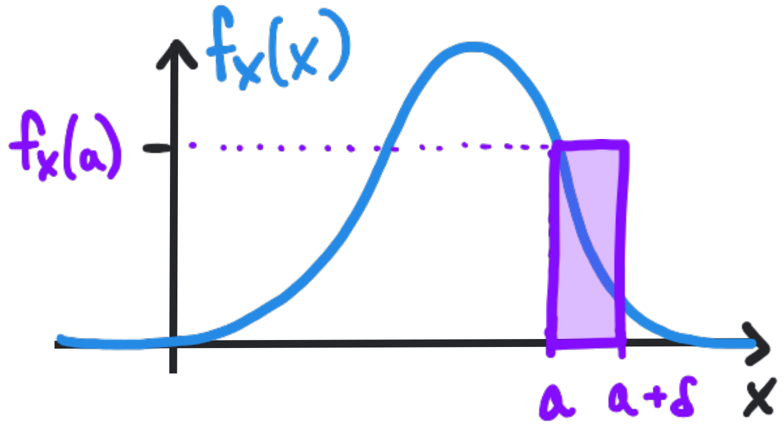




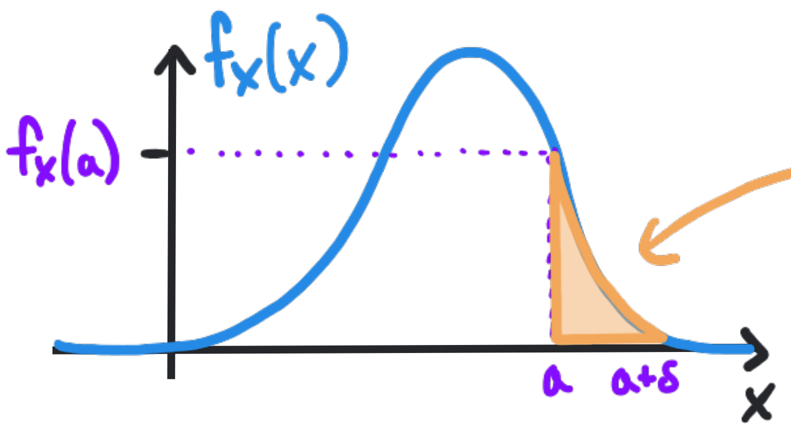
OK for PDF to exceed 1.

- The PDF $f_X(x)$ does not tell us the probability $\mathbb{P}[\{X=x\}]$, which is always equal to 0 for a continuous random variable.

- Instead, the PDF $f_X(x)$ tells us the "density" of probability around x.



Say we want to know $\mathbb{P}[\{X \in [a, a+\delta]\}]$.
We can approximate this by the area $\delta f_X(a)$.
The approximation quality improves as $\delta \to 0$.
This sounds like integration!



The exact probability is the area under the curve,

$$\mathbb{P}\left[\{X \in [a, a+\delta]\}\right] = \int_a^{a+\delta} f_X(x)\,dx .$$

- Intuition: The PDF can play the role of the PMF if we replace sums with integrals.

PDF and CDF Properties

PDF Property	Property Name	CDF Property
$f_X(x) \geq 0$	Non-Negativity	$F_X(x)$ is non-decreasing
$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$	Normalization	$F_X(+\infty) = 1,\ F_X(-\infty) = 0$
$\int_{-\infty}^{x} f_X(u)\,du = F_X(x)$ integration variable	PDF $\longleftrightarrow$ CDF	$\frac{d}{dx} F_X(x) = f_X(x)$ for x where derivative exists
$\mathbb{P}[\{a < X \leq b\}] = \int_a^b f_X(x)\,dx$	Probability of an Interval	$\mathbb{P}[\{a < X \leq b\}] = F_X(b) - F_X(a)$

- Since for a continuous random variable $\mathbb{P}[\{X = x\}] = 0$, we have

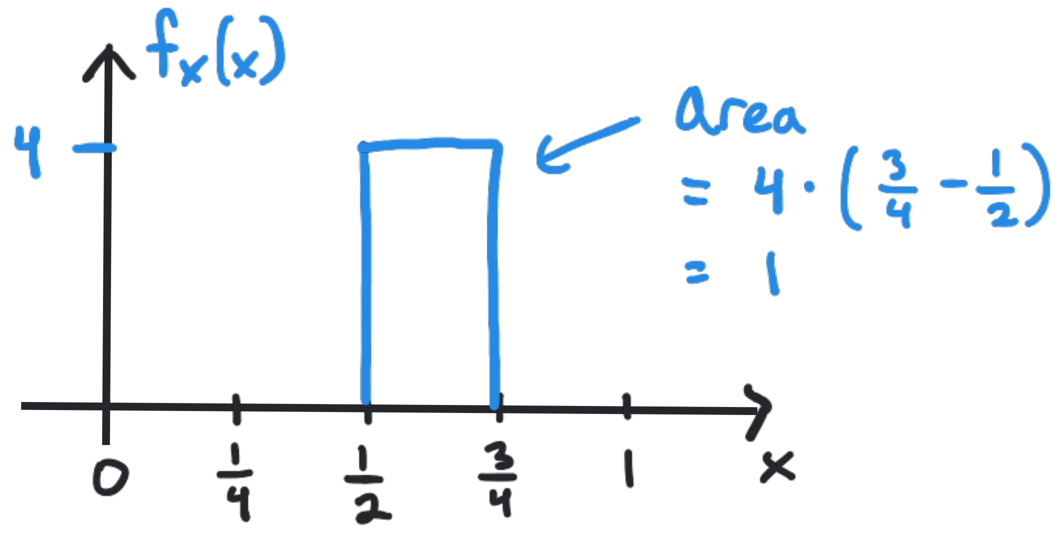
$$\mathbb{P}[\{a < X < b\}] = \mathbb{P}[\{a \leq X < b\}] = \mathbb{P}[\{a < X \leq b\}] = \mathbb{P}[\{a \leq X \leq b\}].$$

- **Example:** X is equally likely to take any value in $[\frac{1}{2}, \frac{3}{4}]$.

→ Determine the PDF.

* Easier to start with a plot.

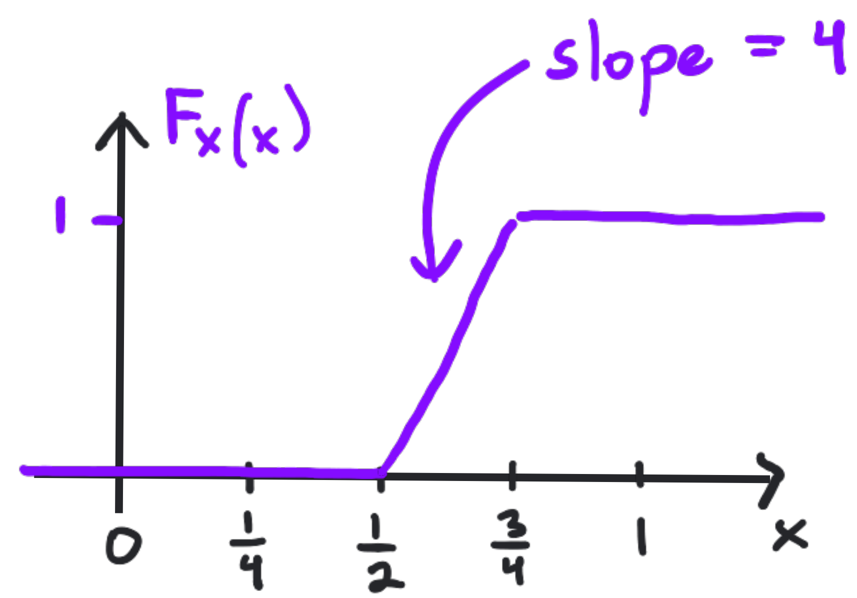
Set height so that area = 1.



$$f_X(x) = \begin{cases} 0 & x < \frac{1}{2} \\ 4 & \frac{1}{2} \le x \le \frac{3}{4} \\ 0 & \frac{3}{4} < x \end{cases}$$

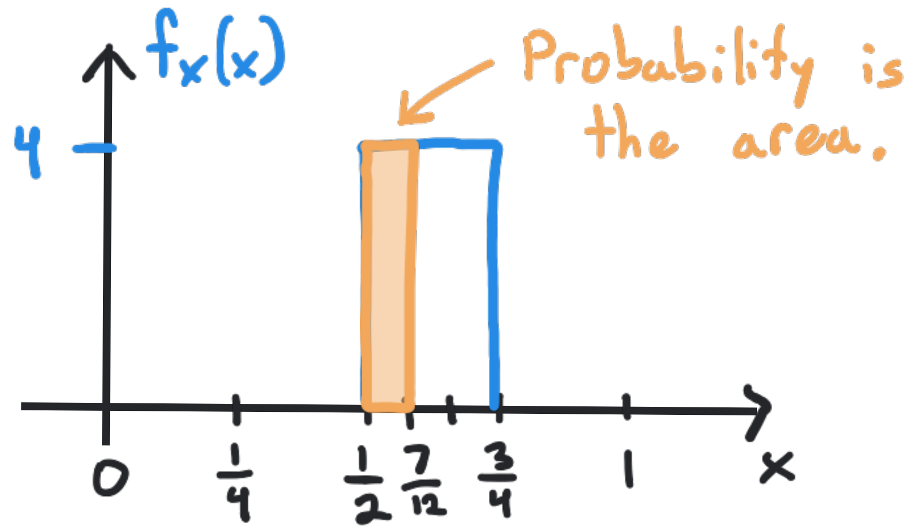
→ Determine the CDF.

* Can guess integral from plot or compute directly.



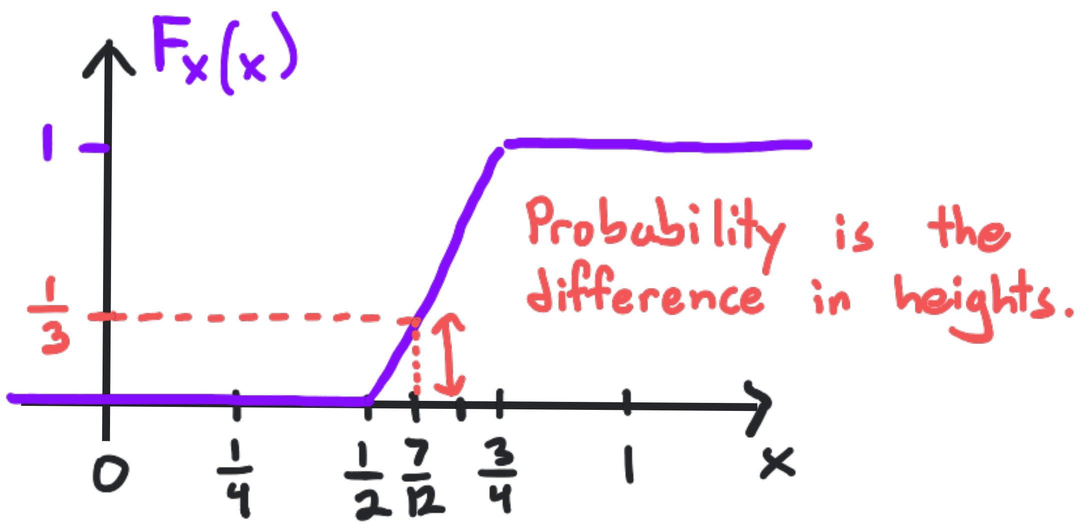
$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du = \begin{cases} \int_{-\infty}^{x} 0\,du & x < \frac{1}{2} \\ \int_{-\infty}^{1/2} 0\,du + \int_{1/2}^{x} 4\,du & \frac{1}{2} \le x \le \frac{3}{4} \\ \int_{-\infty}^{1/2} 0\,du + \int_{1/2}^{3/4} 4\,du + \int_{3/4}^{x} 0\,du & \frac{3}{4} < x \end{cases} = \begin{cases} 0 & x < \frac{1}{2} \\ 4x - 2 & \frac{1}{2} \le x \le \frac{3}{4} \\ 1 & \frac{3}{4} < x \end{cases}$$

- **Example:** X is equally likely to take any value in $[\frac{1}{2}, \frac{3}{4}]$.
  - → Determine $\mathbb{P}[\{\frac{1}{2} \le X \le \frac{7}{12}\}]$.



* PDF Calculation:

$$\int_{1/2}^{7/12} f_X(x)\,dx = \int_{1/2}^{7/12} 4\,dx = (4x)\Big|_{1/2}^{7/12}$$
$$= 4\left(\frac{7}{12} - \frac{1}{2}\right)$$
$$= 4 \cdot \frac{1}{12} = \frac{1}{3}$$



* CDF Calculation:

$$F_X\left(\frac{7}{12}\right) - F_X\left(\frac{1}{2}\right)$$
$$= \left(4 \cdot \frac{7}{12} - 2\right) - \left(4 \cdot \frac{1}{2} - 2\right)$$
$$= \frac{7}{3} - 2 = \frac{1}{3}$$

• **Example:** X is equally likely to take any value in $[\frac{1}{2}, \frac{3}{4}]$.

→ Determine $\mathbb{P}[\{\frac{5}{8} \le X \le 1\}]$.

* PDF Calculation:

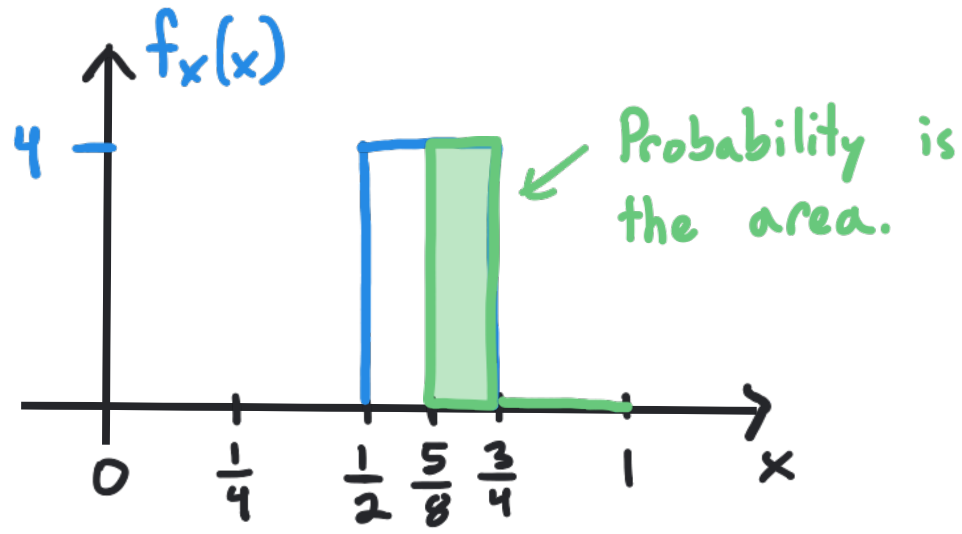
Greater than 1!

$$\int_{5/8}^{1} f_X(x)\,dx = \int_{5/8}^{1} 4\,dx = 4\left(1-\frac{5}{8}\right) = \frac{12}{8}$$

This is wrong.

$$\int_{5/8}^{3/4} 4\,dx + \int_{3/4}^{1} 0\,dx = 4\left(\frac{3}{4}-\frac{5}{8}\right) = \frac{1}{2}$$

Correctly split into two cases.

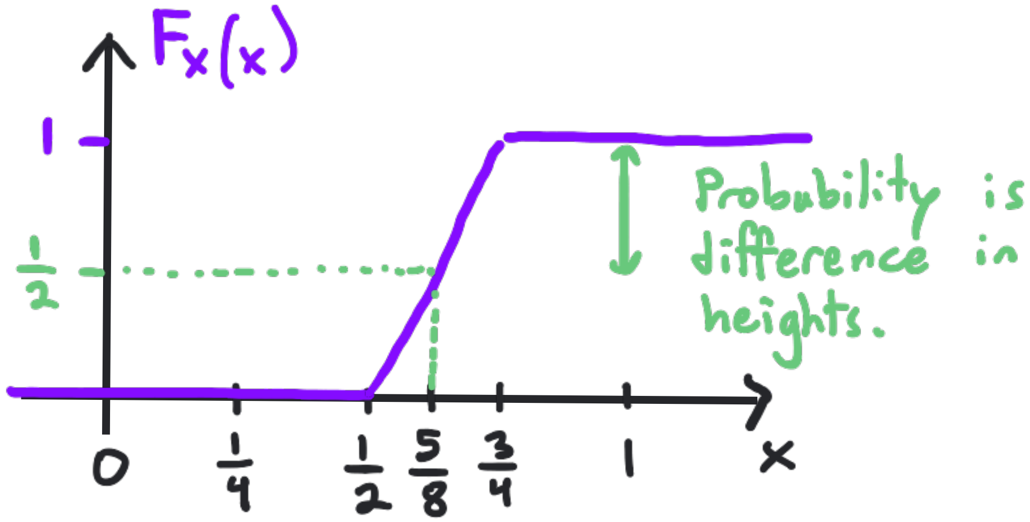


* CDF Calculation:

$$F_X(1) - F_X\left(\frac{5}{8}\right)$$

$$= 1 - \left(4\cdot\frac{5}{8} - 2\right)$$

$$= 1 - \frac{5}{2} + 2 = \frac{1}{2}$$



We don't need to be as careful in the CDF calculation since we already handled the different integration regions.