Recall from Last Class

The cumulative distribution function for any random variable X, denoted by $F(x)$, is the probability that X assumes a value less than or equal to x:

$$F(x) = Pr(X \le x)$$

The cumulative distribution function has the following properties:

- $0 \le F(x) \le 1$ for all values of x
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$
- $F(x)$ is a nondecreasing function of x

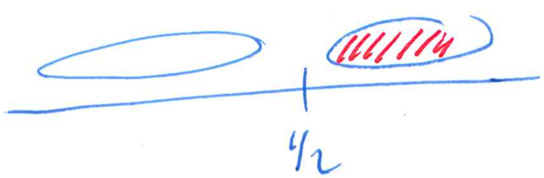
If $F(x)$ is continuous, we say X is a continuous random variable. In this case, $\boxed{Pr(X = x) = 0}$ for any single value x. (This is not necessarily the case for discrete random variables.)

ex Suppose $F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \sin x & \text{if } 0 \le x \le \pi/2 \\ 1 & \text{if } x > \pi/2 \end{cases}$ X continuous

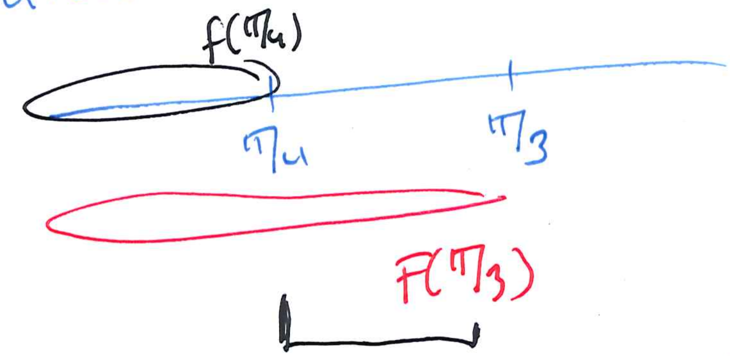
Calculate $Pr(X < 1) = Pr(X \le 1) = F(1) = \sin(1)$

$Pr(X=1) = 0$ def of CDF

Calculate $Pr(X > 1/2) = 1 - Pr(X \le 1/2) = 1 - F(1/2) = 1 - \sin(1/2)$



Calculate $Pr(\pi/4 < X < \pi/3) = Pr(X < \pi/3 \text{ AND } x > \pi/4)$



$= \underbrace{F(\pi/3)}_{\le \pi/3} - F(\pi/4)$ get rid of values $\le \pi/4$

$= \sin(\pi/3) - \sin(\pi/4) = \frac{\sqrt{3}}{2} - \frac{1}{\sqrt{2}}$

$$\Pr(X = x) = 1 - \Pr(X \neq x)$$

$$\Pr(X < x) = 1 - \Pr(X \geq x)$$

---

useful Fact: $F(x)$ is a CDF: X: continuous

$$F(b) - F(a) = \Pr(a \leq X \leq b)$$

(*) Understanding leads to memorizing

ex Suppose $F(x) = k\arctan x + c$ for some constants k, c

If F is a CDF, what are k, c?

Properties of CDF:

$\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$

$\lim_{x\to\infty} (k\arctan x + c)$

$k(\pi/2) + C = 1$

$\lim_{x\to-\infty} (k\arctan x + c)$

$= k(-\pi/2) + C = 0$

Add Eqns:

$k(\pi/2) + C = 1$

$k(-\pi/2) + C = 0$

$0k + 2c = 1$

$c = 1/2$

Subtract:

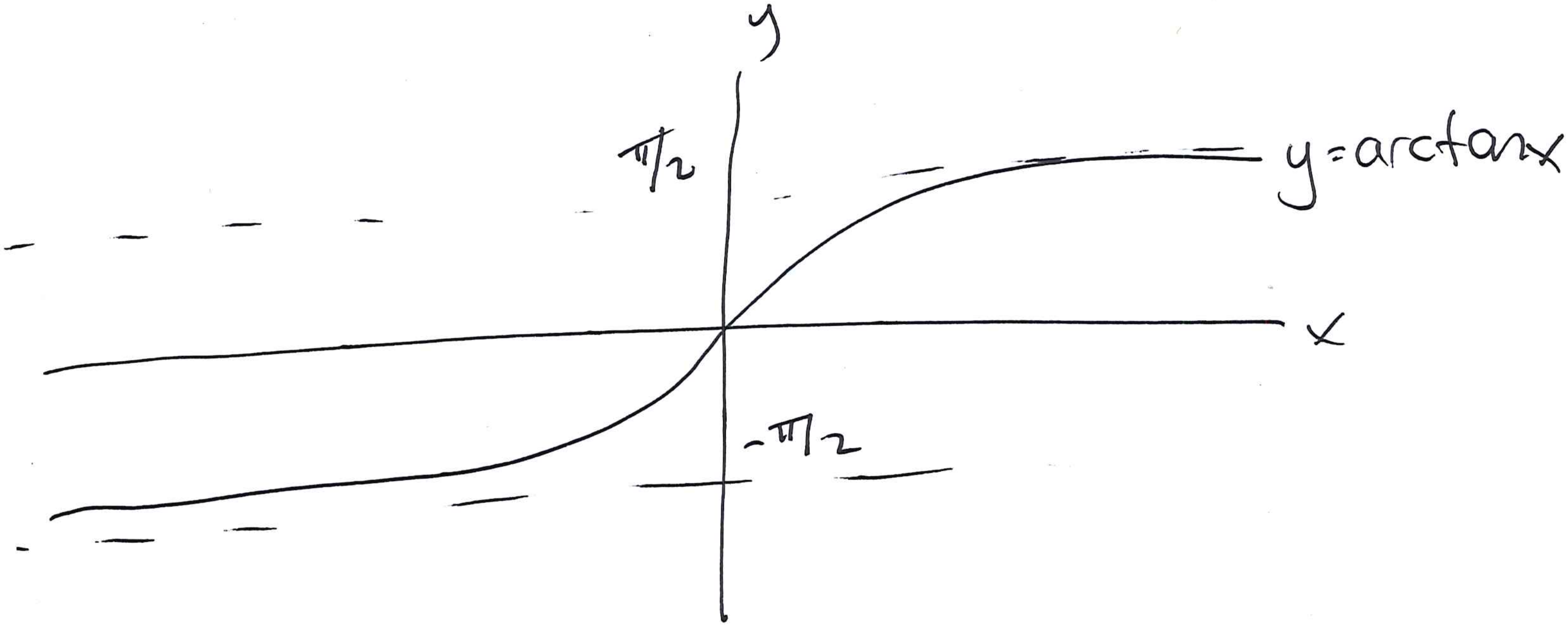
$k(\pi/2) + C = 1$

$k(-\pi/2) + c = 0$

$k(\pi/2 + \pi/2) = 1$

$\pi k = 1$

$k = 1/\pi$

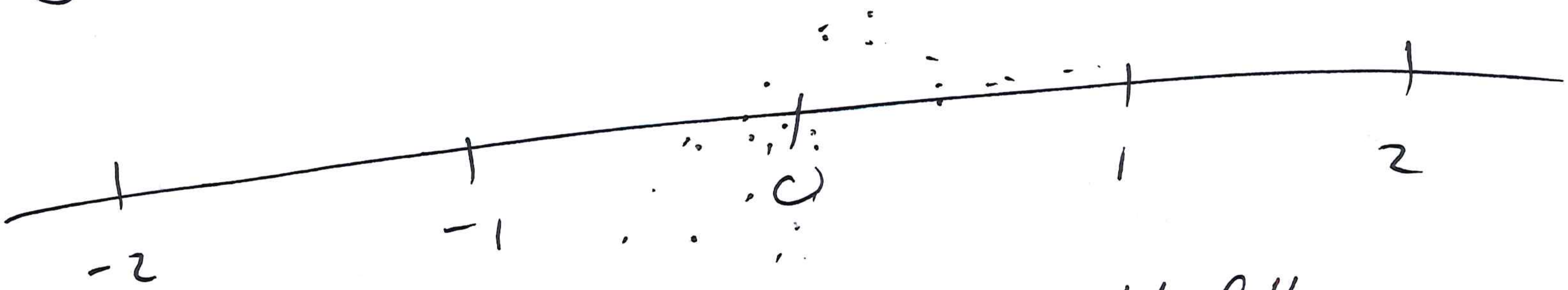


Probability Density Function

Motivation: If X continuous random variable, then $\Pr(X=x)=0$ for any value x

How can we express relative likelihoods?

eg drop pen on number line



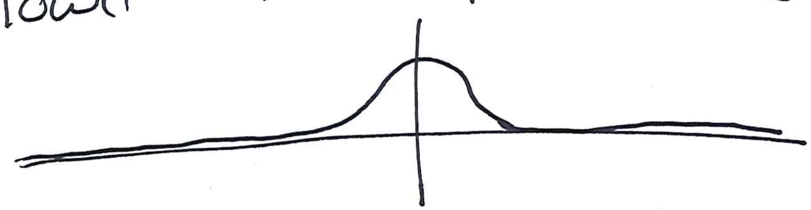
so many numbers where pen could fall

$Pr(X = 1/2) = 0$ etc

<u>But</u> they cluster around 0

So just saying $Pr(X=0) = Pr(X=10) = 0$
doesn't entirely capture behaviour.

Probability Density Function:
want to be higher in region of a cluster
lower in sparser regions

ex 

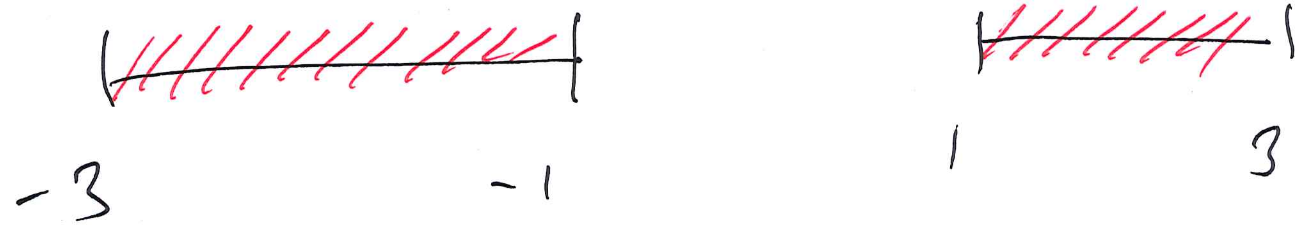
We can ask $\underbrace{Pr(a \leq X \leq a+h)}_{\text{can be} > 0} = \underbrace{F(a+h) - F(a)}_{\text{look like derivative}}$

Definition
Let $F(x)$ be the cumulative distribution function for a continuous random variable X.
The probability density function of x (PDF) is:

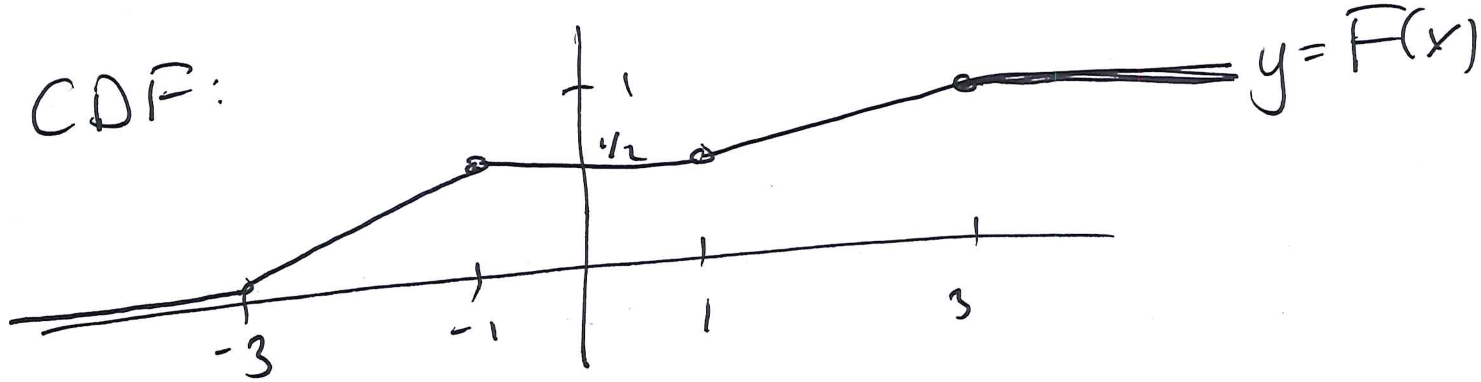
$$f(x) = \frac{d}{dx} F(x)$$

whenever the derivative exists

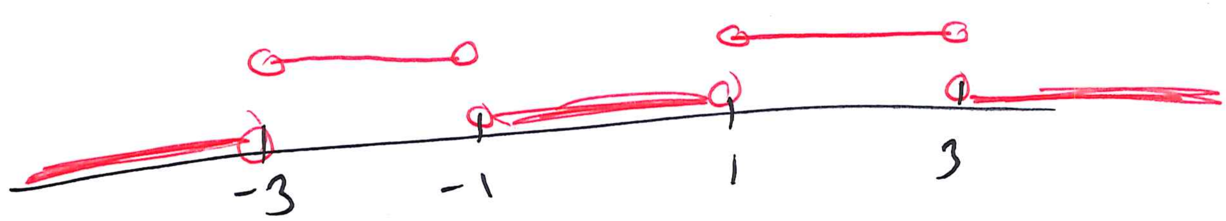
ex from before:



CDF:



PDF: $f(x) = F'(x)$



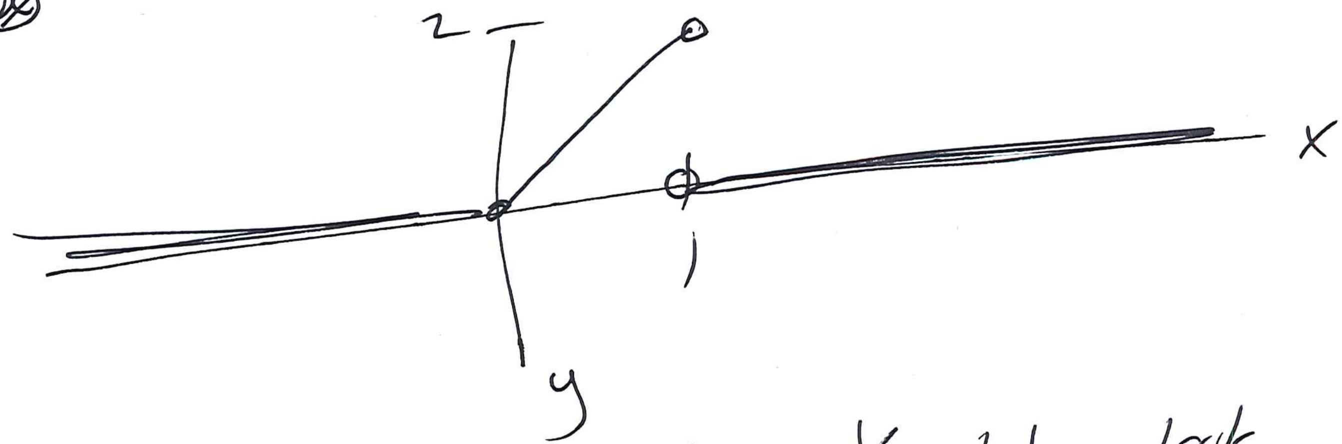
$$f(x) = \frac{d}{dx} F(x) \implies$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \qquad \left(\text{remember } \lim_{x \to -\infty} F(x) = 0\right)$$

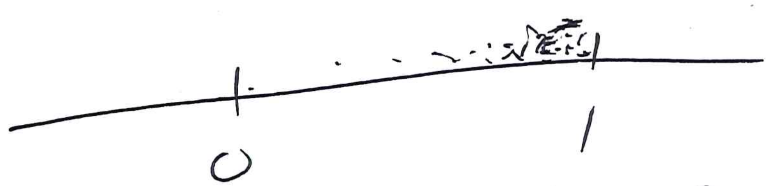
$$\Pr(a \le X \le b) = \underbrace{F(b) - F(a)}_{\text{definite } \int} = \int_a^b f(x)\,dx$$

Area under PDF on $[a, b]$ gives $\Pr(a \le X \le b)$

ex PDF $f(x)$:



What will values X takes look like?

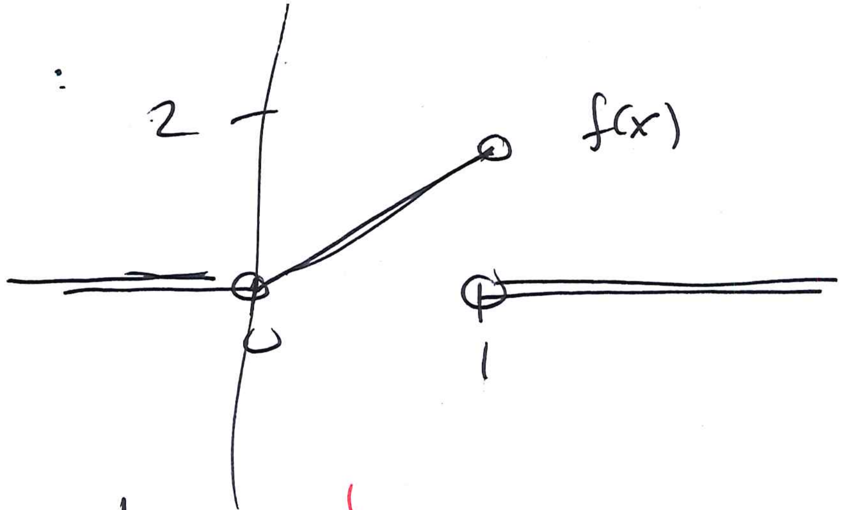


more frequent hits near 1

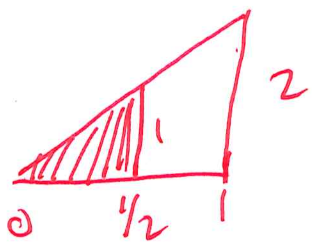
no values outside $[0,1]$

$f(x)$ is PDF: density of values X assumes

Same $f(x)$:



$$\underbrace{\Pr\left(0 \le X \le \tfrac{1}{2}\right)}_{\text{area under } f(x) \text{ on } [0,\, 1/2]} = \frac{1}{4}$$



A : $\frac{1}{2}$(base)(height)

$= \frac{1}{2}\left(\frac{1}{2}\right)(1) = \frac{1}{4}$