Recall: a r.v. is a fn $X: S \to \mathbb{R}$

Defn: CDF For a r.v. X the CDF is

$$F(x) = \mathbb{P}(X \le x) \quad \forall x \in \mathbb{R}.$$

Defn: Identically Distributed RVs
(equal in distribution)

We say two RVs X and Y are equal in distribution if $\forall A \subset \mathbb{R}$

$$\mathbb{P}(X \in A) = \mathbb{P}(Y \in A).$$

We denote this as $X \overset{d}{=} Y$.

This doesn't mean $X = Y$ as functions.

Ex. 3 coin flips.

$X = \#$ heads and $Y = \#$ tails.

Notice: $X(HTT) = 1$ but $Y(HTT) = 2$
So X and Y are diffent RVs.

but they are equal in dist.

Ex. $\mathbb{P}(X = 1) = 3/8 = \mathbb{P}(Y = 1)$

$\mathbb{P}(X = 0) = 1/8 = \mathbb{P}(Y = 0)$

$\vdots$

Theorem: $X \overset{d}{=} Y$ iff $F_X = F_Y$.

Theorem: $X \overset{d}{=} Y$ iff $F_X = F_Y$.

CDF of X CDF of Y.

Ex. Toss coins (independently) until a H appears.

$$S = \{H, TH, TTH, TTTH, \ldots\}$$

note $|S| = \infty$

let p be the prob. of getting a H on any flip.

Define X = # flips to get a H

$s \in S$	$X(s)$
H	1
TH	2
TTH	3
TTTH	4
⋮	⋮

Q: What is the CDF of X?

$$F(x) = \mathbb{P}(X \leq x)$$

To determine F lets consider

$$\boxed{\mathbb{P}(X = x)} \text{ for } x \in \mathbb{R}$$

prob. it takes x flips to get a

{ let T_i = getting a T on i^{th}

takes x ... to get a H.

$\left\{ \begin{array}{l} \text{let } T_i = \text{getting a T on } i^{\text{th}} \\ H_i = T_i^c \end{array} \right.$

then $\quad$ "$X = x$" $= T_1 T_2 T_3 \cdots T_{x-1} H_x$

So
$$\mathbb{P}(\underbrace{X = x}_{w_x}) = \mathbb{P}(T_1 T_2 T_3 \cdots T_{x-1} H_x)$$

independence

$$= \mathbb{P}(T_1)\mathbb{P}(T_2)\cdots\mathbb{P}(T_{x-1})\mathbb{P}(H_x)$$
$$= (1-p)(1-p)\cdots(1-p)p$$
$$\boxed{= (1-p)^{x-1} p}$$

Notice that if $\omega_i =$ "$X = i$"

then "$X \le x$" $= \omega_1 \cup \omega_2 \cup \omega_3 \cup \cdots \cup \omega_x$

disjoint union

hence
$$F(x) = \mathbb{P}(X \le x) = \mathbb{P}(W_1 \cup W_2 \cup \cdots \cup W_x)$$
$$= \sum_{i=1}^{\lfloor x \rfloor} \mathbb{P}(W_i)$$

recall:
$$\sum_{i=0}^{n-1} r^i = \frac{1-r^n}{1-r}$$

$$\rightsquigarrow \; = \sum_{i=1}^{\lfloor x \rfloor} (1-p)^{i-1} p$$

$\lfloor x \rfloor$ = round x down

$$= p \sum_{i=0}^{\lfloor x \rfloor - 1} (1-p)^i$$

$r = 1-p$

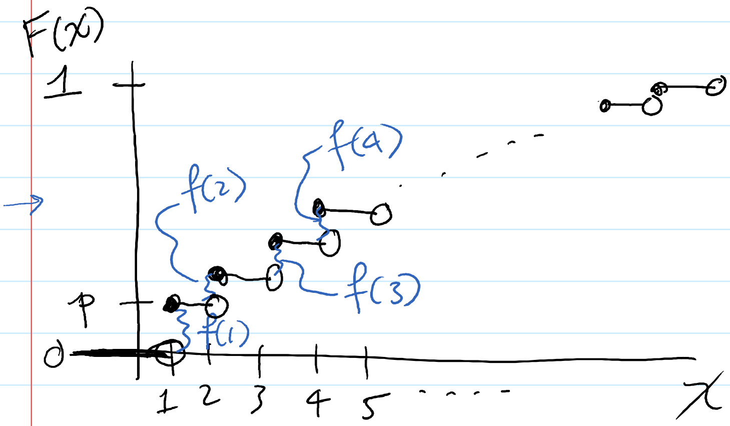
$$= \cancel{p}\,\frac{1-(1-p)^{\lfloor x \rfloor}}{\cancel{1-(1-p)}}$$

$$F(x) = 1-(1-p)^{\lfloor x \rfloor}.$$

$F(x)$

$$F(x) = 1 - (1-p)^{x}$$

for $x > 0$



We call this a <u>geometric</u> RV.

<u>Notice</u>: We derived the CDF using $\boxed{\mathbb{P}(X = x)}$

<u>Defn: Discrete/Continuous</u>

A <u>discrete</u> RVs is one whose CDF is a step function.

A <u>continuous</u> RV is one whose CDF is continuous.

<u>Defn: Probability Mass Function (PMF)</u>

For a discrete RV X, the pmf is a function $f : \mathbb{R} \to \mathbb{R}$, defined as

$$f(x) = \mathbb{P}(X = x) \quad \forall x \in \mathbb{R}.$$

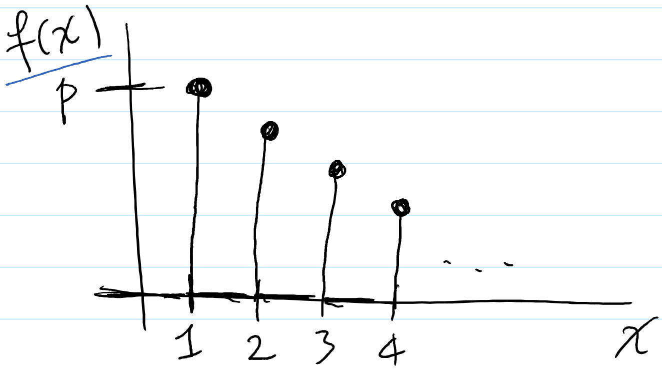
This is also called the <u>distribution</u> of X.

<u>Ex.</u> For the geometric RV,

$$f(x) = \mathbb{P}(X = x) = \begin{cases} (1-p)^{x-1} p, & x \in \mathbb{N} \end{cases}$$

$$f(x) = \mathbb{P}(X = x) = \begin{cases} (1-p)\;\; p, & x \in \mathbb{N} \\ 0, & \text{else} \end{cases}$$

"Stick plot"
"distribution of X"



<u>Theorem:</u> For discrete RVs,

$$F(x) = \sum_{i \le x} f(i)$$

$P(X \le x)$ $\qquad$ $P(X = i)$

<u>pf:</u> $"X \le x" = \bigcup_{i \le x} "X = i"$

disjoint union

hence

$$F(x) = \mathbb{P}(X \le x) = \mathbb{P}\Big(\bigcup_{i \le x} "X = i"\Big)$$
$$= \sum_{i \le x} P(X = i) = \sum_{i \le x} f(i).$$

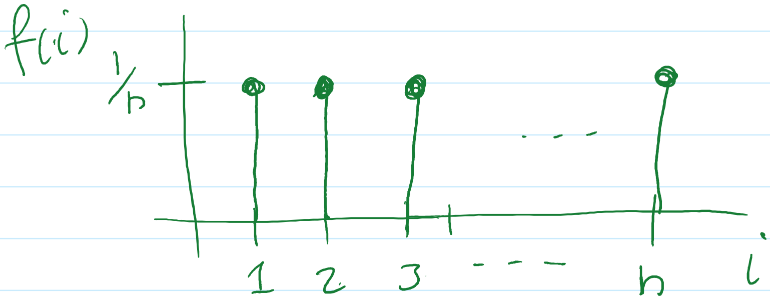
<u>Ex.</u> Say X has a discrete uniform dist. over the values $1, \dots, n$

denote

$$X \sim U(\{1, \dots, n\})$$

"dist. as"

means that $f(i) = \begin{cases} 1/n & \text{for } i = 1, \dots, n \\ 0 & \text{else} \end{cases}$



mass uniformly dist. over $1, \dots, n$

Q: what is the CDF?

$$F(x) = \sum_{i \le x} f(i) = \sum_{i=1}^{\lfloor x \rfloor} f(i) = \sum_{i=1}^{\lfloor x \rfloor} 1/n = \frac{\lfloor x \rfloor}{n}.$$

Said: $F(x) = \sum_{i \le x} f(i)$

$\mathbb{P}(X \le x)$

More generally: If $A \subset \mathbb{R}$,

over values in A

for discrete r.v.s.

Ex. $X \sim U(\{1, \dots, n\})$.

$$\mathbb{P}(2 \le X < 5)$$

$$= \mathbb{P}(X \in \{2, 3, 4\}) = \sum f(x) = 1/n + \frac{1}{n} + \frac{1}{n}$$

$x=2,3,4$

$= 3/n$.

Ex. $\mathbb{P}(X \in \{1,7,3\})$

$$= \sum_{x=1,7,3} f(x) = 3/n.$$

Ex. Roll a die 60 times (independently)

$X = \#$ of 6s I roll.

Let's derive the pmf.

$f(x) = \mathbb{P}(X=x) =$ prob. I roll x 6s in my 60 rolls.

$$f(0) = \mathbb{P}(X=0) = \underbrace{(5/6)(5/6)(5/6)\cdots(5/6)}_{60 \text{ times}} = \left(\frac{5}{6}\right)^{60}$$

choose where 6 goes

$$f(1) = \mathbb{P}(X=1) = 60\underbrace{(1/6)}_{\text{prob of } 6}\underbrace{(5/6)(5/6)\cdots(5/6)}_{59 \text{ other places}}$$

$$= 60\,(1/6)(5/6)^{59}$$

$$f(2) = \mathbb{P}(X=2) = \binom{60}{2}(1/6)(1/6)\underbrace{(5/6)\cdots(5/6)}_{58 \text{ times}}$$

$$= \binom{60}{2}(1/6)^2(5/6)^{58}$$

$= (2)(1/6)(5/6)$

General pattern:

$$\boxed{f(x) = \mathbb{P}(X = x) = \binom{60}{x}\left(\tfrac{1}{6}\right)^{x}\left(\tfrac{5}{6}\right)^{60-x}}$$

binomial coef.

We call this a Binomial RV.

If I do n experiments each w/ a yes/no answer (independently) each has a prob. of getting a "yes" of $p \in [0,1]$

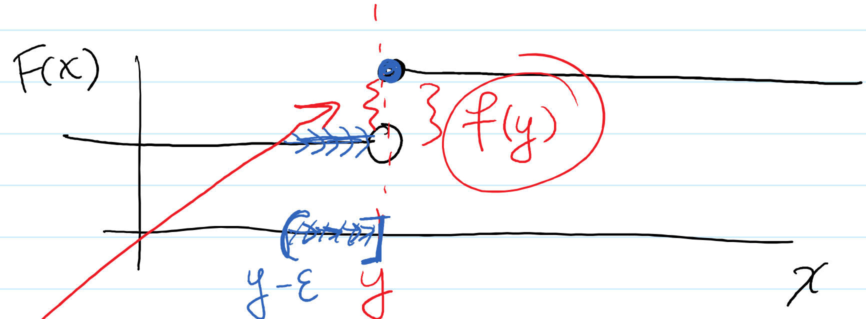
$X = \#$ of "yes" experiments.

Above:
$\to n = 60$
$\to p = 1/6$

We call X a binomial RV, denote $X \sim \text{Bin}(n, p)$.

Recall: $\boxed{\mathbb{P}(a < X \le b) = F(b) - F(a)}$

discrete RV



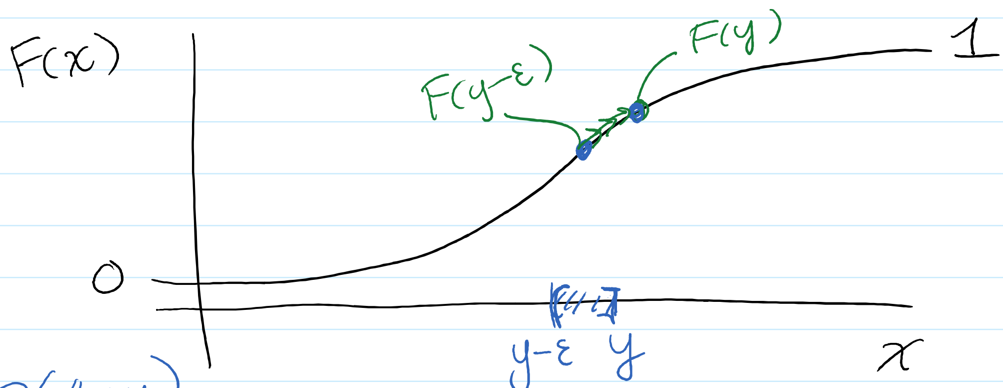
$$\mathbb{P}(X = y) = \lim_{\varepsilon \downarrow 0} \mathbb{P}(y - \varepsilon < X \le y)$$

$\varepsilon \downarrow 0$

$$= \lim_{\varepsilon \downarrow 0} \left[F(y) - F(y-\varepsilon) \right]$$

$$= \text{gap} = f(y)$$

Consider the same limit for continuous RVs



$P(X = y)$??

$$f(y) = \lim_{\varepsilon \downarrow 0} P(y - \varepsilon < X \leq y)$$

continuous

$$= \lim_{\varepsilon \downarrow 0} F(y) - F(y-\varepsilon)$$

$$= F(y) - F(y) = 0$$

Punchline! can't define a pmf like this for continuous RVs.

Can we do something similar?

Want: $\boxed{F(x) = \sum_{i \leq x} f(i)}$

Want: $\boxed{F(x) = \sum_{i \le x} f(i)}$ ⟵

Defn: Probability Density Function (PDF)

Analog of PMF for cts RVs.

The pmf is a fn $f: \mathbb{R} \to \mathbb{R}$ defined so that $\forall x \in \mathbb{R}$,

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

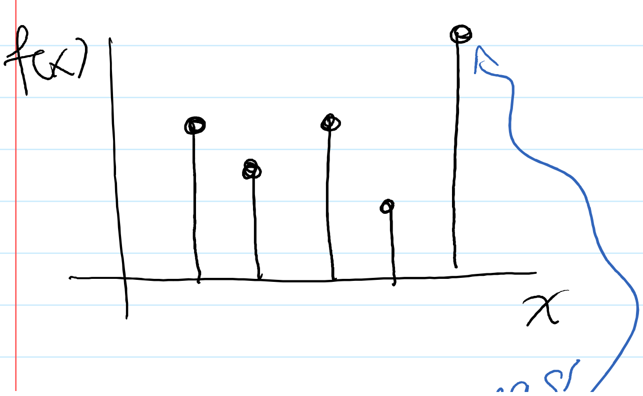
note by Fund. Theorem of Calc:

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

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

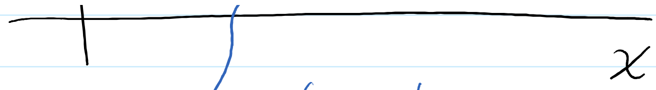
pdf = derivative of CDF.

discrete (PMF)



Continuous Case (PDF)

prob. mass at a pt. $=$ $\mathbb{P}(X$ at that pt$)$



prob. density at a pt $\neq$ a prob. of X being at that pt.