Defn: Identical 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 $X = Y$.

Ex. 3 coin flips

X = # heads

Y = # tails

these are different RVs,

$$X(HTT) = 1 \quad \text{but} \quad Y(HTT) = 2$$

However, $\boxed{X \overset{d}{=} Y.}$

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

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

Theorem!

$$X \overset{d}{=} Y \quad \text{iff} \quad F_X = F_Y .$$

F_X: CDF of X

F_Y: CDF of Y

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

$$S = \{H, TH, TTH, \dots\}$$

Let p be the prob. I get a H on any flip.

X = # flips until I get a H

$s \in S$	$X(s)$
H	1
TH	2
TTH	3
$\vdots$	$\vdots$

Q: What is the CDF of X?

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

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

We'll look at $P(X = x)$

Let H_i = i^{th} toss is a H, $T_i = H_i^c$

$$\text{“}X = i\text{”} = T_1 T_2 \cdots T_{i-1} H_i$$

independent

So

$$\underbrace{P(X = i)} = P(T_1 T_2 \cdots T_{i-1} H_i)$$
$$= P(T_1) P(T_2) \cdots P(T_{i-1}) P(H_i)$$
$$= (1-p)(1-p) \cdots (1-p) p$$
$$= \boxed{(1-p)^{i-1} p}$$

$$\text{“}X \le x\text{”} = \text{“}X=1\text{”} \cup \text{“}X=2\text{”} \cup \text{“}X=3\text{”} \cup \cdots \cup \text{“}X=x\text{”}$$

disjoint

$$F(x) = P(X \le x) = P(X=1) + P(X=2) + \cdots + P(X=x)$$
$$= \sum_{i=1}^{x} P(X=i)$$
$$= \sum_{i=1}^{x} (1-p)^{i-1} p$$
$$= p \sum_{i=1}^{x} (1-p)^{i-1}$$

Calc II:
geometric sum,

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

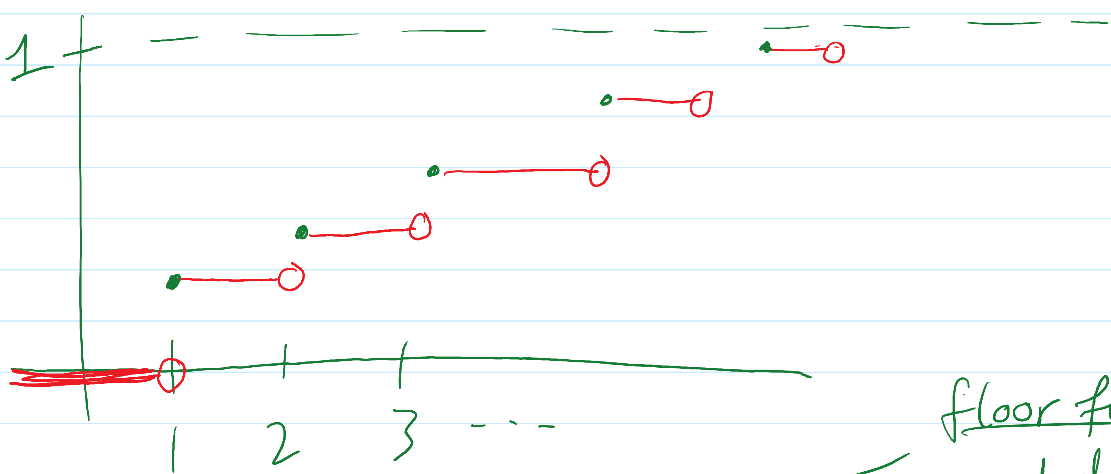
$$= p \sum_{i=1} (1-p)$$

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

$r = 1-p$
$n = x$

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

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



floor function: round down

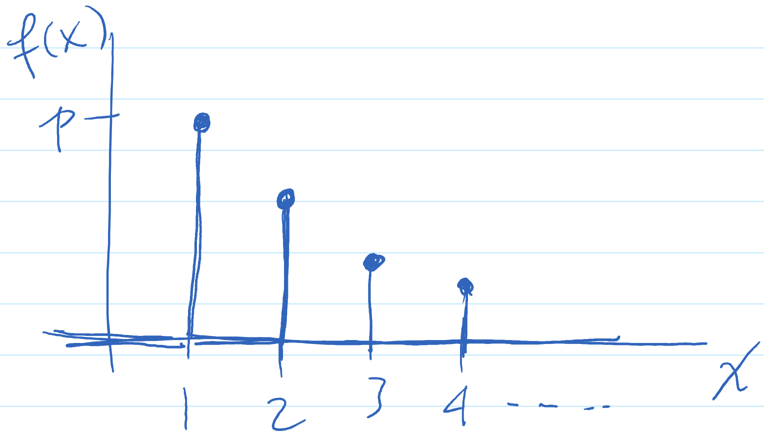
$$F(x) = \begin{cases} 0, & x < 1 \\ 1-(1-p)^{\lfloor x \rfloor}, & x \geq 1 \end{cases}$$

Defn: Discrete/Continuous RVs

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

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

Plot of $\mathbb{P}(X=x) = (1-p)^{x-1} p = f(x)$



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

For a discrete RV X, the PMF is a function $f: \mathbb{R} \to \mathbb{R}$ where

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

Also sometimes called the distribution of X.

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

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

Pf. "$\le x$" $= \bigcup_{i \le x}$ "$X = i$"

disjoint union

then

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

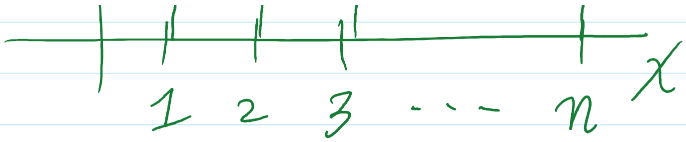
$$= \sum_{i \le x} \mathbb{P}(X = i)$$

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

<u>Ex.</u> We say X has a discrete uniform distribution over $1, \ldots, n$

<u>Notation:</u>

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



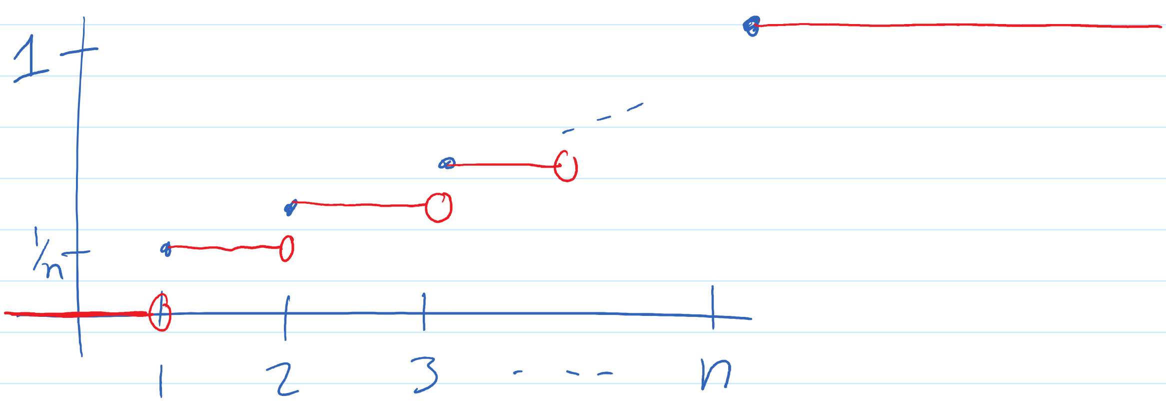
in math:

$$f(x) = \begin{cases} \frac{1}{n} & x = 1, 2, 3, \dots \\ 0 & \text{else} \end{cases}$$

Q: What's the CDF?

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

if $x = 1, 2, \dots, n$



$$F(x) = \begin{cases} 0, & x < 0 \\ \lfloor x \rfloor / n & \\ 1, & x \ge n \end{cases}$$

Saw: $P(X \le x) = F(x) = \sum_{i \le x} f(i)$

More generally:

$$P(X \in A) = \sum_{i \in A} f(i)$$

Ex. $X \sim U(\{1, \ldots, 7\})$

$$P(2 \le X \le 5) = P(X \in \{2,3,4,5\})$$

$$= \sum_{i=2}^{5} f(i)$$

$$= \sum_{i=2}^{5} 1/7 = 4/7.$$

Ex. Roll a die 60 times. (independently)

X = # of 6s I roll

What is PMF of X?

$$f(0) = P(X=0) = \underbrace{(5/6)(5/6)(5/6)\cdots(5/6)}_{60}$$

$$= (5/6)^{60}$$

$$f(1) = P(X=1) = \binom{60}{1}\left(\tfrac{1}{6}\right)\underbrace{\left(\tfrac{5}{6}\right)\left(\tfrac{5}{6}\right)\cdots\left(\tfrac{5}{6}\right)}_{59}$$

$$= \binom{60}{1}\left(\tfrac{1}{6}\right)\left(\tfrac{5}{6}\right)^{59}$$

$$f(2) = P(X=2) = \binom{60}{2}\left(\tfrac{1}{6}\right)\left(\tfrac{1}{6}\right)\underbrace{\left(\tfrac{5}{6}\right)\cdots\left(\tfrac{5}{6}\right)}_{58}$$

$$= \binom{60}{2}\left(\tfrac{1}{6}\right)^{2}\left(\tfrac{5}{6}\right)^{58}$$

$\vdots$

General pattern:

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

We call this type of RV a Binomial RV.

If I do a series of n independent experiments each with a binary outcome:

Yes/No, $\boxed{0/1}$, Success/Failure etc.

and the prob. of a 1 is p for each experiment.

and the prob. of a 1 is p for each experiment.

Let X = # of 1s

then X has a Binomial dist,

$$X \sim Bin(n, p)$$

and

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$