1179: Probability
Lecture 7 — Random Variables, CDFs & PMFs, and Discrete Random Variables

Ping-Chun Hsieh (謝秉均)

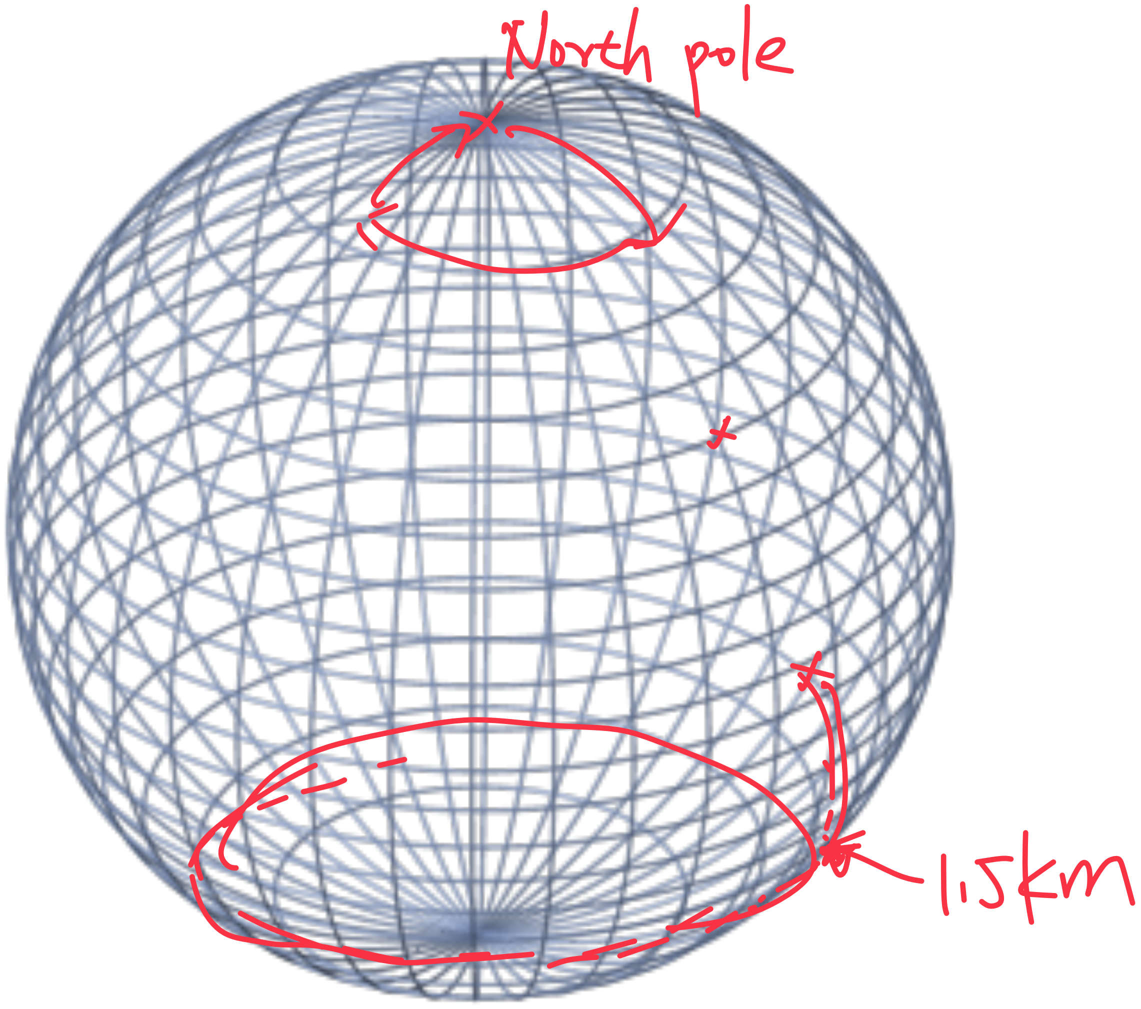
October 6, 2021

Announcement

- Weekly online office hours: 1pm-1:30pm on Wednesdays
  - https://nycu.webex.com/nycu/j.php?MTID=ma2106f2503f60807a6dedb2d5d777756 (same as the Webex link for the lectures)
- Just send me an email if you would like to meet in some other time slots or in person

An Interview Question

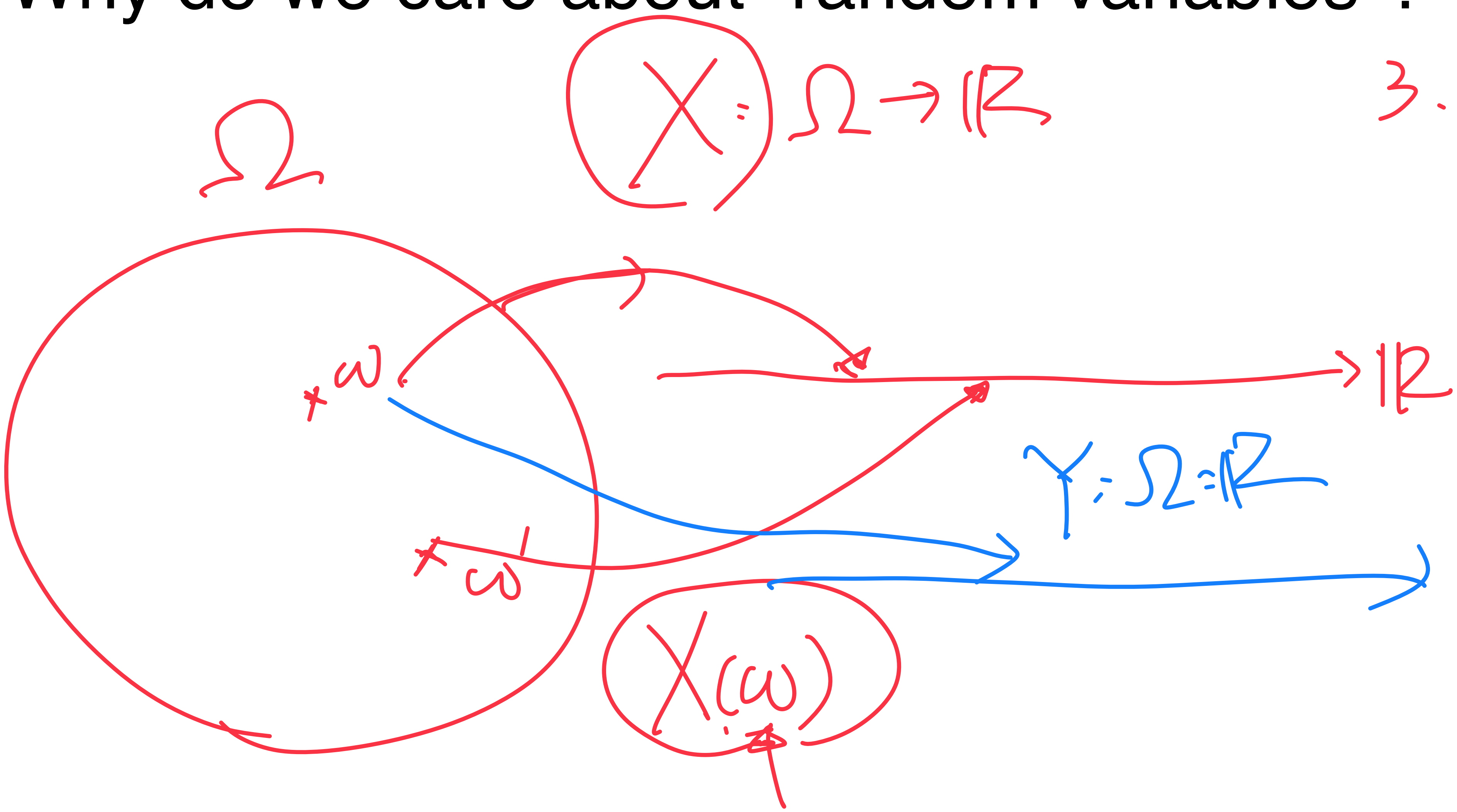
- Suppose Bill stands somewhere on the surface of the globe.
- Next, Bill goes 1.5km south, 1.5km west and then 1.5km north. It turns out that Bill returns to the same location.
- Question: Where is Bill?



Quick Review

- What is a "random variable"?
- Why do we care about "random variables"?

1. Simpler notations
2. Multiple random variables
3. R.V. capture common features of experiments



This Lecture

1. Random Variables and Cumulative Distribution Function (CDF)

2. Probability Mass Function (PMF)

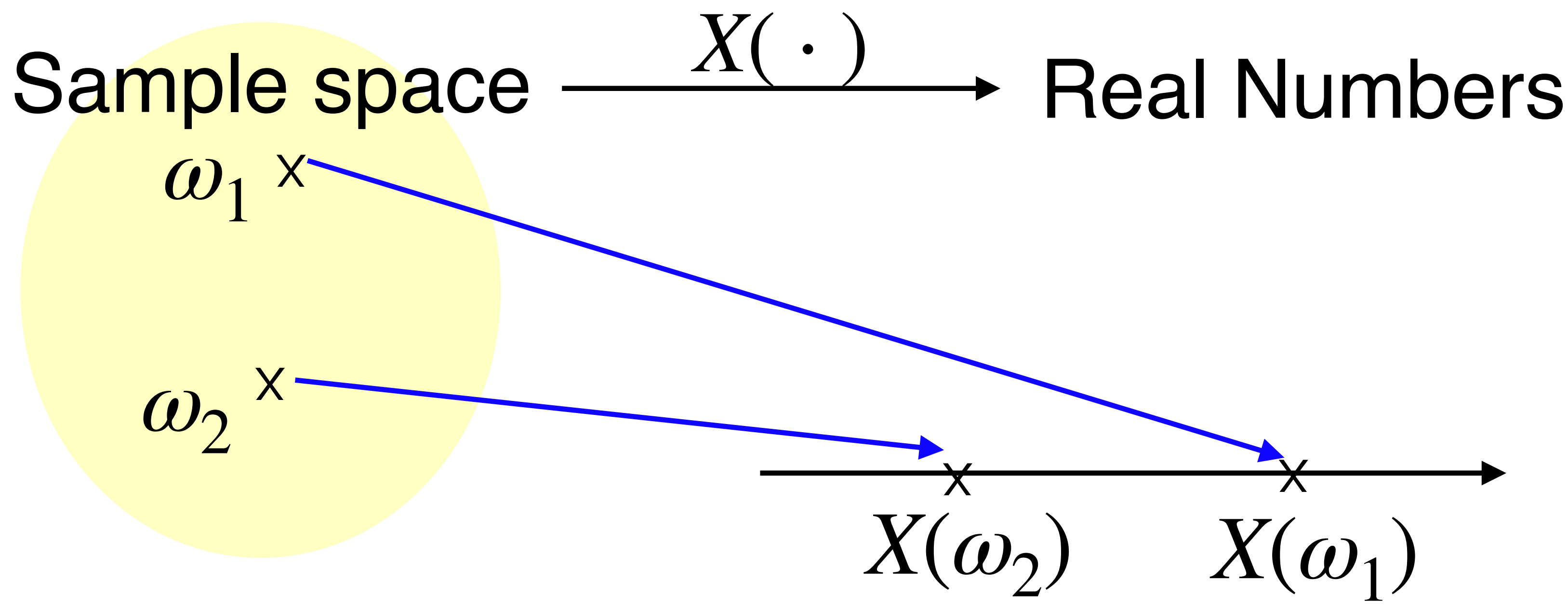
3. Special Discrete Random Variables

- Reading material: Chapter 4.1~4.3 and 5.1~5.2

1. Random Variables

What is a Random Variable (Formally)?

- Random variable: a function that maps each outcome to a real number



- Example: Whether NCTU will merge with NYMU

Ω

$X(\omega = \text{"Yes"}) = 1, \quad X(\text{"No"}) = 0$

- Example: # of people waiting in line at Shinemood

$Y(n \text{ people}) = n$

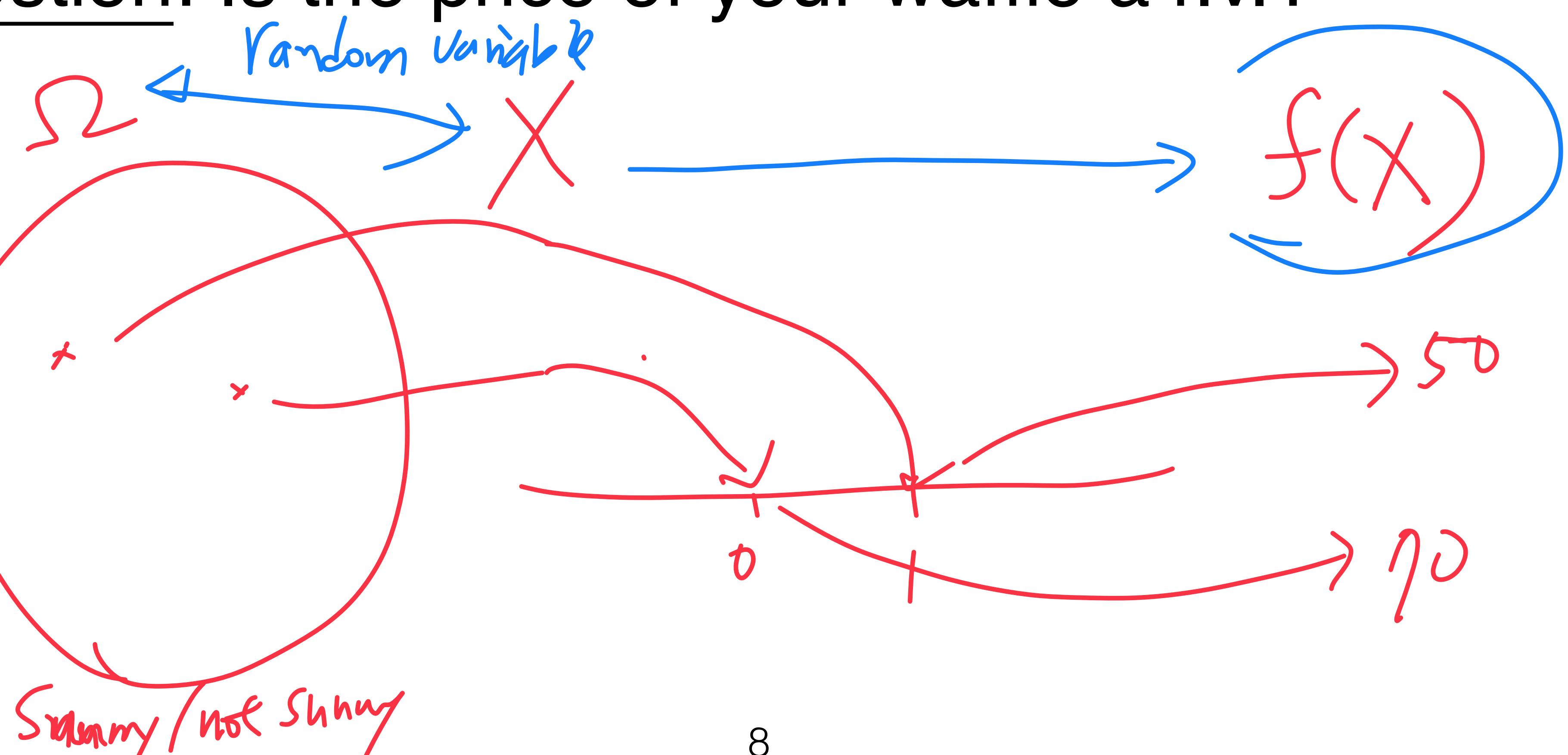
Function of a Random Variable

$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{f} \mathbb{R}$

X is a random variable
f(X) " also " " .

- Example: Buy a waffle at Shinemood
  - If it is sunny today, then you spend $50 to order a Matcha-red-bean waffle
  - Otherwise, you spend $70 to order a Fried-chicken waffle
  - Question: Is the price of your waffle a r.v.?

Yes!

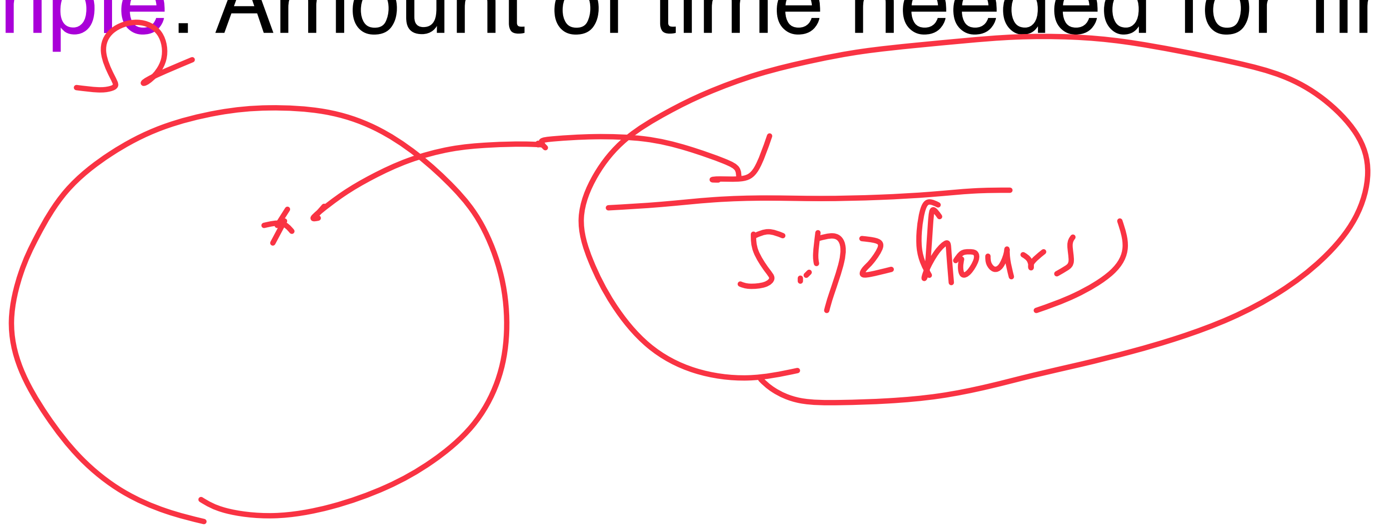


Discrete and Continuous Random Variables

- Example: # of people waiting in line at Shinemood

Discrete

- Example: Amount of time needed for finishing HW1



Continuous

Cumulative Distribution Function (CDF)

- Random variables are used to calculate the probabilities of events.

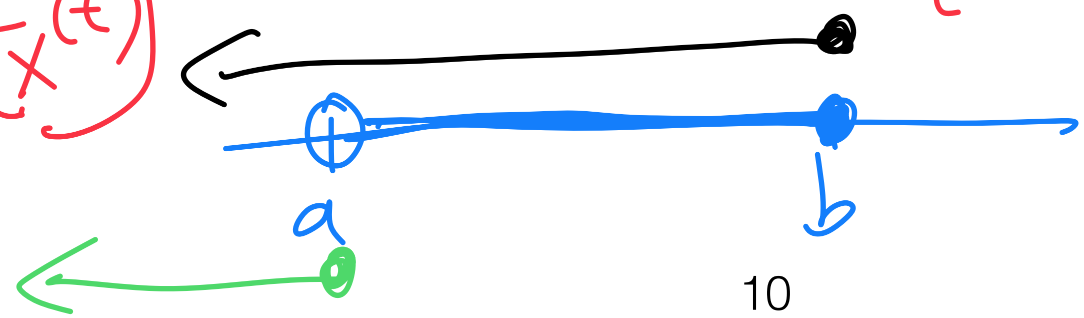
Cumulative Distribution Function (CDF): For any random variable X, the CDF of X is defined as:

$$F_X(t) = P(X \leq t), \text{ for all } t \in \mathbb{R}$$

$$= P(\underbrace{\{\omega : X(\omega) \leq t\}}_{\text{event}})$$

- What's the range of $F_X(t)$? $\in [0, 1]$
- How to use the CDF?
- Example: $P(a < X \leq b) = ?$ $P(\{\omega : a < X(\omega) \leq b\})$

CDF of X: $F_X(t)$



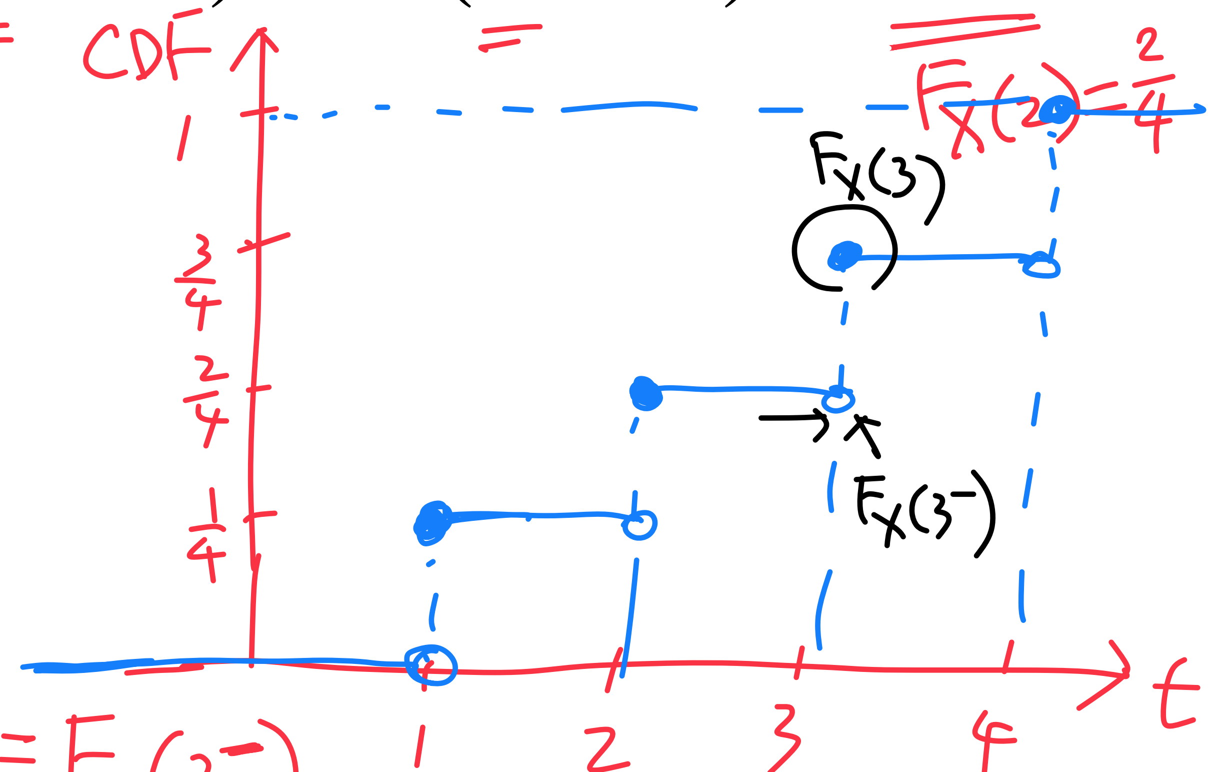
$= F_X(b) - F_X(a)$

CDF of a Discrete Random Variable

$F_X(0)=0$
$F_X(1)=\frac{1}{4}$
$F_X(1.5)=\frac{1}{4}$
$F_X(1.9999)=\frac{1}{4}$

$F_X(t) = P(X \leq t)$

- Example: Roll a fair 4-sided die
  - $P(X=1) = P(X=2) = P(X=3) = P(X=4) = 1/4$
  - What is the CDF of X?

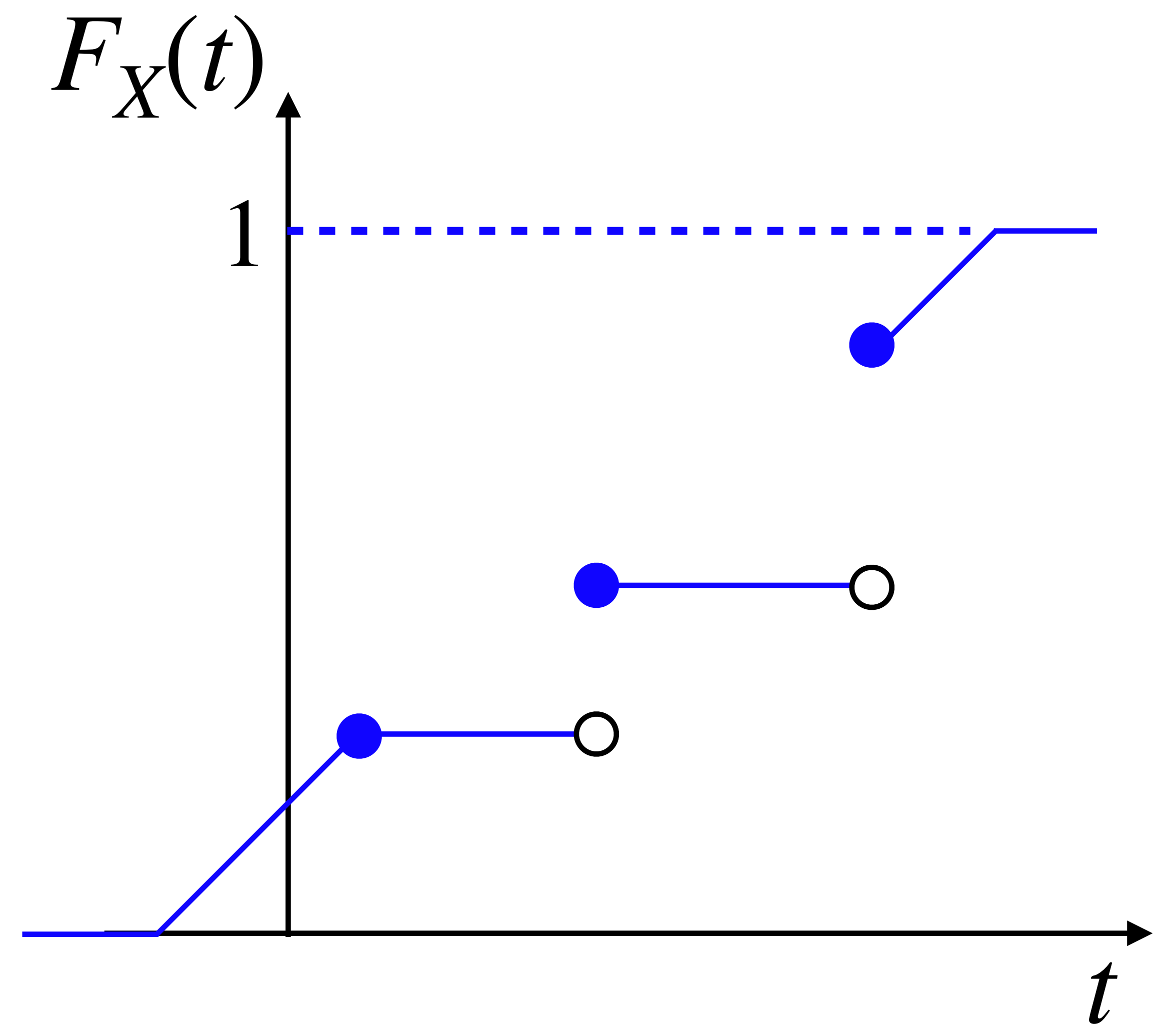


1. $P(X \leq 3) = \frac{3}{4} = F_X(3)$
2. $P(X < 3) = \frac{2}{4} = F_X(3) - P(X=3) = F_X(3^-)$
3. $P(1 < X \leq 3) = \frac{2}{4} = F_X(3) - F_X(1)$
4. $P(1 < X < 3) = \frac{1}{4} = F_X(3^-) - F_X(1)$
5. $P(X = 3) = \frac{1}{4} = F_X(3) - F_X(3^-)$

Use CDF to Find Probability of an Event (I)

$$F_X(t) = P(X \leq t), \text{ for all } t \in \mathbb{R}$$

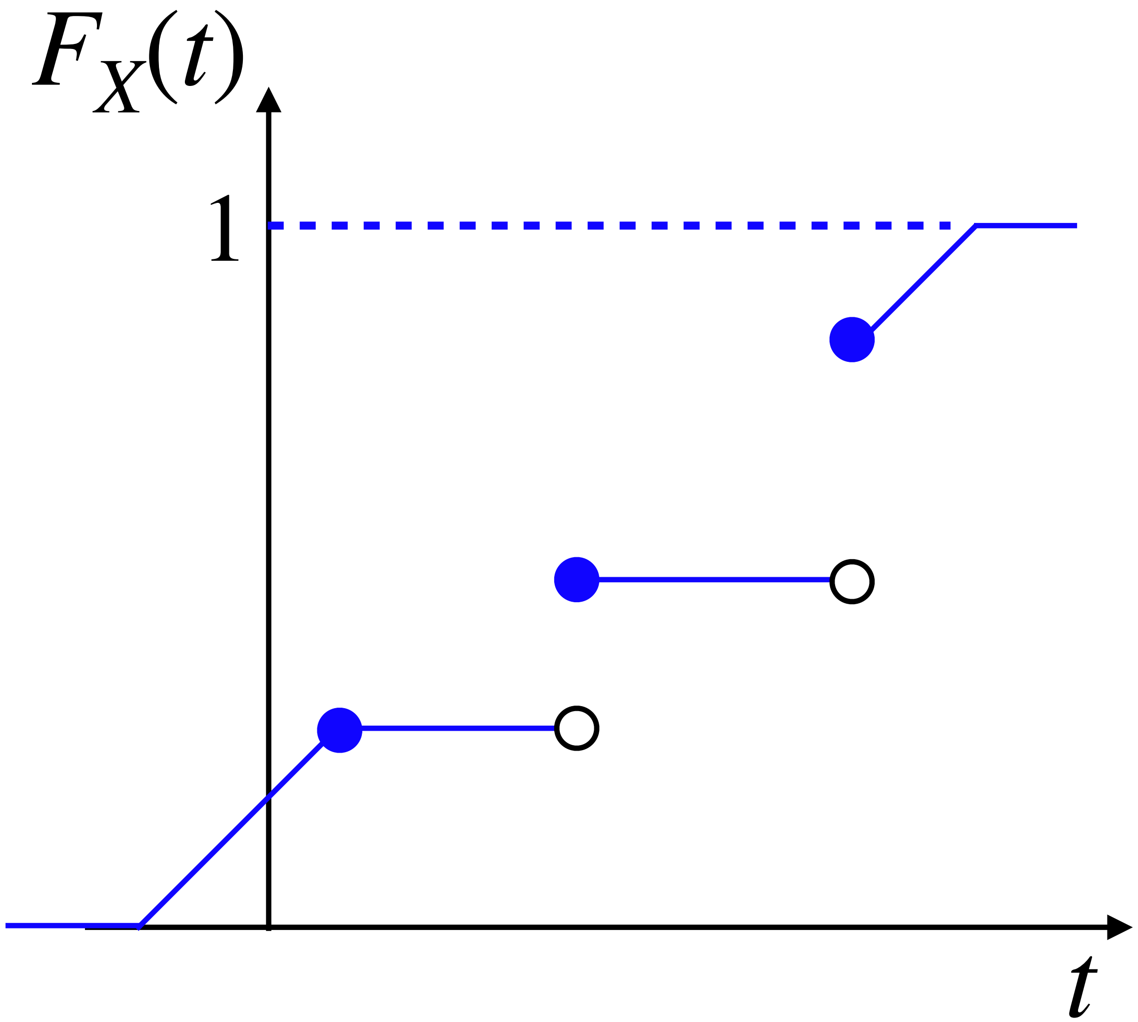
Event	Probability of the event
$X \leq a$	$F_X(a)$
$X > a$	$1 - F_X(a)$
$X < a$	$F_X(a^-) = F_X(a) - P(X=a)$
$X \geq a$	$1 - F_X(a^-)$
$X = a$	$F_X(a) - F_X(a^-)$



Use CDF to Find Probability of an Event (II)

$$F_X(t) = P(X \le t), \text{ for all } t \in \mathbb{R}$$

Event	Probability of the event
$a < X \le b$	$F_X(b) - F_X(a)$
$a < X < b$	$F_X(b^-) - F_X(a)$
$a \le X \le b$	$F_X(b) - F_X(a^-)$
$a \le X < b$	$F_X(b^-) - F_X(a^-)$



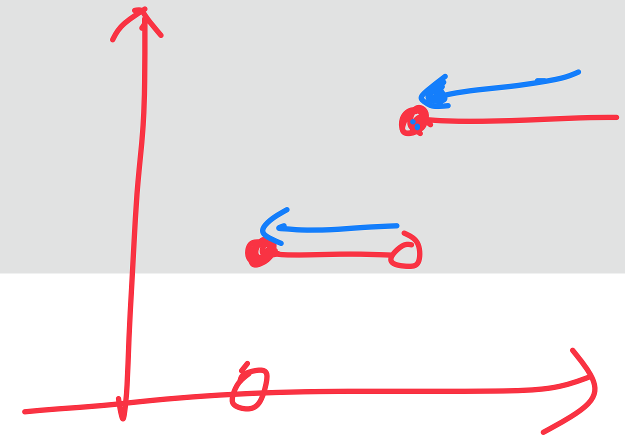
Example:Use CDF to Find Probability

- Example: The CDF of a random variable X is given by

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x/4, & 0 \leq x < 1 \\ 1/2, & 1 \leq x < 2 \\ \frac{1}{8}x + \frac{1}{2}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

1. $P(X < 2)$ =
2. $P(1 \leq X < 3)$ =
3. $P(X > 3/2)$ =
4. $P(X = 2)$ =

Properties of a Valid CDF



$$F_X(t) = P(X \le t), \text{ for all } t \in \mathbb{R}$$

Properties of a CDF:

1. $F_X(t)$ is non-decreasing

   $t' \geq t \Rightarrow F_X(t') \geq F_X(t)$

2. $\lim_{t \to \infty} F_X(t) = 1$

   $F_X(\infty) = P(X \leq \infty) = P(\{\omega : X(\omega) \leq \infty\}) = P(\Omega) = 1$

3. $\lim_{t \to -\infty} F_X(t) = 0$

   $F_X(-\infty) = P(\{\omega : X(\omega) \leq -\infty\}) = P(\emptyset) = 0$

4. $F_X(t)$ is **right-continuous** $(F_X(t^+) = F_X(t))$

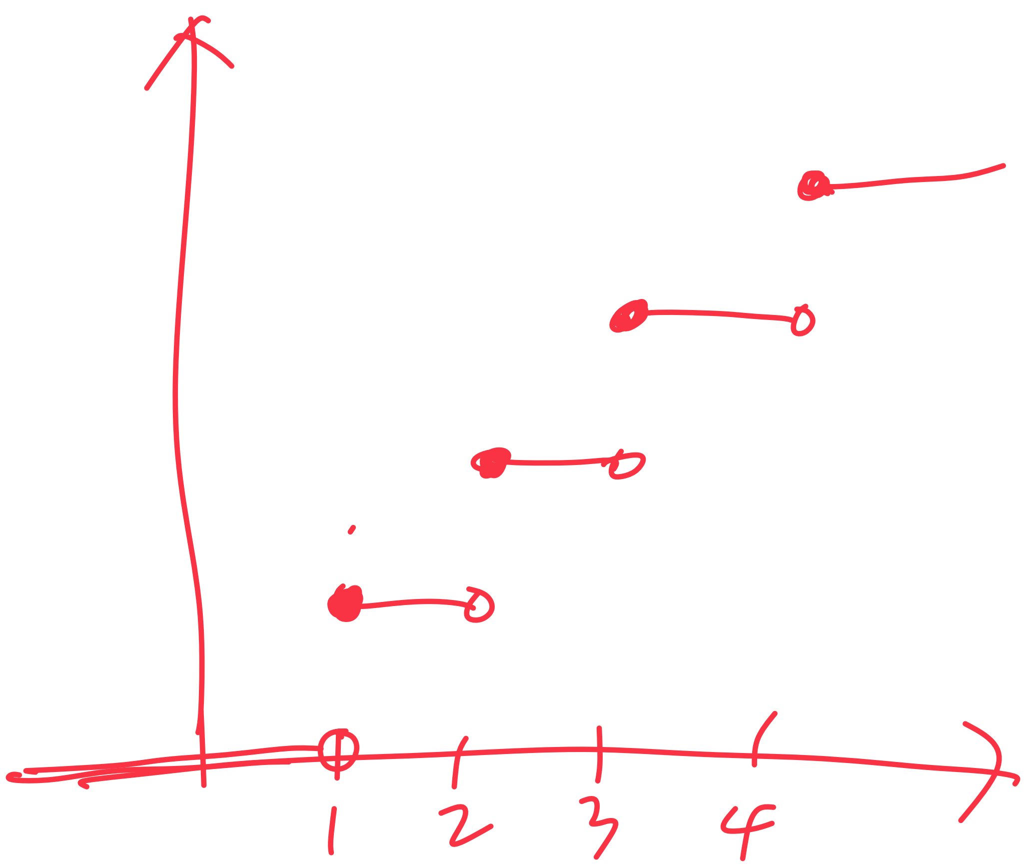
2. Probability Mass Function (PMF)

PMF: Another Way to Specify CDF of a Discrete Random Variable

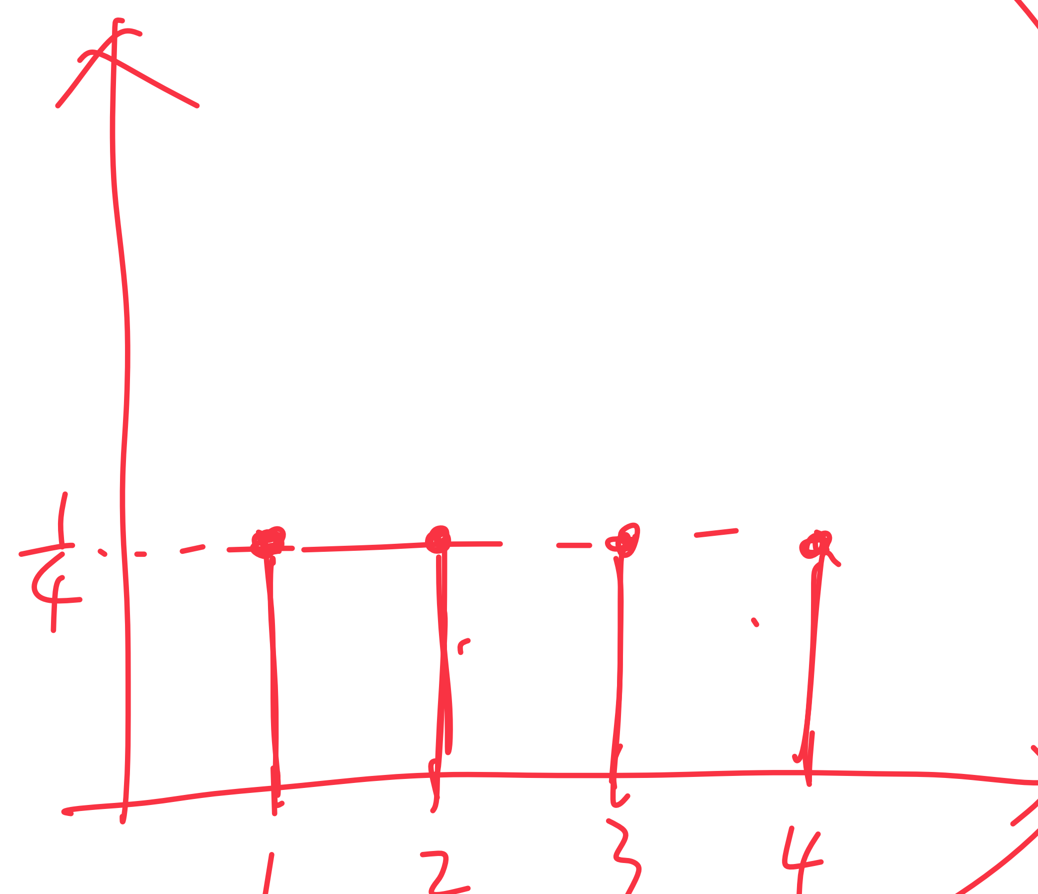
$P(X=1)=P(X=2)=P(X=3)=P(X=4)$
$= \frac{1}{4}$

- Example: Roll a fair 4-sided die once
  - Define a random variable X = the number that we observe

CDF



PMF



Probability Mass Function (PMF)

Probability Mass Function (PMF): For any discrete random variable X with possible values $\{x_1, x_2, x_3, \cdots\}$, the PMF $p(\cdot)$ of X is a function that satisfies:

finite or countably infinite

(1) $p(x_i) = P(X = x_i)$

(2) $p(x) = 0$, if $x \notin \{x_1, x_2, x_3, \cdots\}$

(3) $\displaystyle\sum_{i=1}^{\infty} p(x_i) = 1$

- For discrete random variables: CDF ⇔ PMF

Example: From CDF to PMF

- Example: Given the CDF of a discrete random variable X as:

$$F_X(t) = \begin{cases} 0, & x < 1 \\ 1/36, & 1 \le x < 2 \\ 4/36, & 2 \le x < 3 \\ 9/36, & 3 \le x < 4 \\ 16/36, & 4 \le x < 5 \\ 25/36, & 5 \le x < 6 \\ 1, & x \ge 6 \end{cases}$$

- What is the PMF of X?

PMF:

$p(1) = \frac{1}{36}$

$p(2) = \frac{3}{36}$

$p(3) = \frac{5}{36}$

$p(4) = \frac{7}{36}$

$p(5) = \frac{9}{36}$

$p(6) = \frac{11}{36}$

$p(x) = 0$ for all $x \notin \{1,2,3,4,5,6\}$

$$\Rightarrow \quad p(x) = \begin{cases} \frac{1}{36}, & \text{if } x=1 \\ \frac{3}{36}, & \text{if } x=2 \\ \frac{5}{36}, & \text{if } x=3 \\ \frac{7}{36}, & \text{if } x=4 \\ \frac{9}{36}, & \text{if } x=5 \\ \frac{11}{36}, & \text{if } x=6 \\ 0, & \text{else} \end{cases}$$

Example: From PMF to CDF

- Example: Given the PMF of a discrete random variable X as:

$$p(x) = \begin{cases} 1/6, & x = -2 \\ 1/3, & x = 1 \\ 1/4, & x = 5 \\ 1/4, & x = 10 \end{cases}$$

$$F_X(t) = \begin{cases} 0, & \text{if } t < -2 \\ \frac{1}{6}, & \text{if } -2 \le t < 1 \\ \frac{1}{6}+\frac{1}{3}, & \text{if } 1 \le t < 5 \\ \frac{1}{6}+\frac{1}{3}+\frac{1}{4}, & \text{if } 5 \le t < 10 \\ \frac{1}{6}+\frac{1}{3}+\frac{1}{4}+\frac{1}{4}, & \text{if } t \ge 10 \end{cases}$$

- What is the CDF of X?

Probability Distribution

- Discrete random variables: CDF or PMF

density

- Continuous random variables: CDF or PDF (will be discussed in the next few lectures)

2. Special Discrete Random Variables

Experiments With 2 Possible Outcomes

- Example: Whether an image of dog is classified correctly
- Example: Whether NCTU will merge with NYMU
- Example: Whether the 3rd student leaving the classroom wears glasses

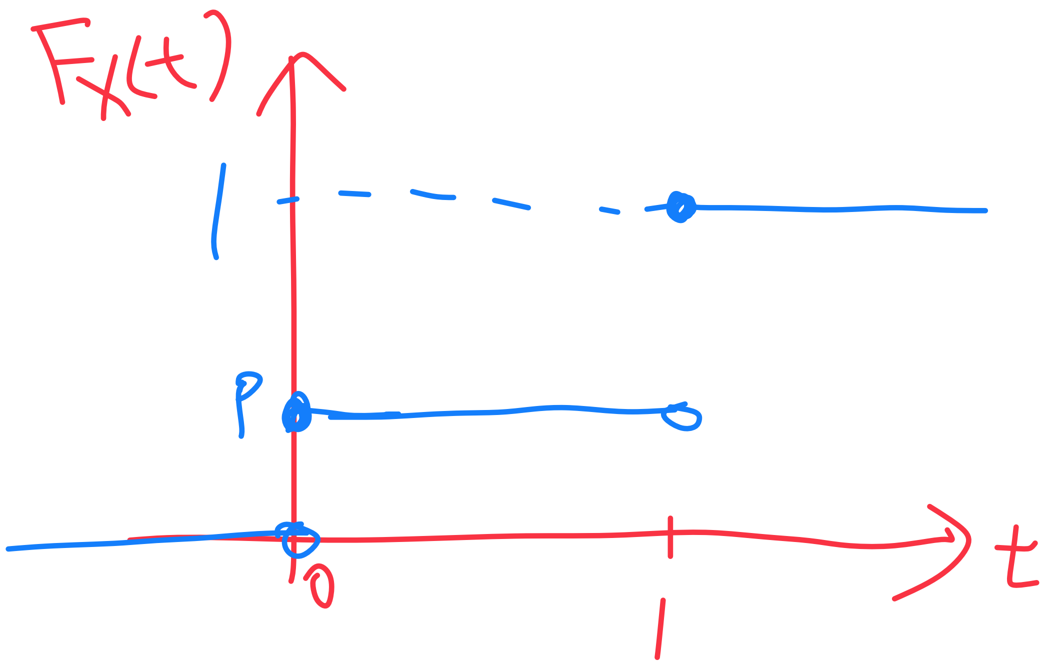
  p $\quad$ $1-p$
- Example: Toss a coin once. Head or tail? $\quad X = 0, 1$
- What are the common features?
  - 1 experiment trial (no repetition) with 2 possible outcomes
  - 1 probability parameter
  - Want: Whether a specific outcome occurs

1. Bernoulli Random Variables (Formally)

Bernoulli Random Variables: A random variable X is Bernoulli with parameter p if its PMF is given by

$$P(X = k) = \begin{cases} p, & \text{if } k = 1 \\ 1 - p, & \text{if } k = 0 \\ 0, & \text{otherwise} \end{cases}$$

- How about its CDF?



Jacob Bernoulli

2. Binomial Random Variables

(all trials are independent from each other)

- Example: Play the same claw machine for 5 times, and each trial is successful with probability 0.7. What is $P(\text{win 3 toys})$?
- Example: Stephen Curry makes 20 free throws, and each throw is good with probability 0.95. What is $P(\text{he missed 2 throws})$?
- Example: Consider n independent coin tosses with head probability p. What is $P(\text{we observe } 2 \text{ heads})$?
- What are the common features?
  - n repetitions of the same Bernoulli experiment
  - Want: how many successes in n repetitions

2. PMF of Binomial Random Variables

- Example: Play the same claw machine for 5 times, and each trial is successful with probability 0.7. All trials are independent.
  - Define a r.v. X = the number of toys we get
  - What is the PMF of X?

X can be 0,1,2,3,4,5

$$p(x) = \begin{cases} C^5_0 (0.7)^0 \cdot (0.3)^5, & \text{if } x=0 \\ C^5_1 \cdot (0.7)^1 \cdot (0.3)^4, & \text{if } x=1 \\ C^5_2 \cdot (0.7)^2 (0.3)^3, & \text{if } x=2 \\ C^5_3 \cdot (0.7)^3 \cdot (0.3)^2, & \text{if } x=3 \\ C^5_4 \cdot (0.7)^4 \cdot (0.3)^1, & \text{if } x=4 \\ C^5_5 \cdot (0.7)^5 \cdot (0.3)^0, & \text{if } x=5 \\ 0, & \text{otherwise} \end{cases}$$

2. Binomial Random Variables (Formally)

Claw machine: $X \sim \text{Binomial}(5, 0.7)$

Binomial Random Variables: A random variable X is Binomial with parameters (n, p) if its PMF is given by

$n = 5$, $p = 0.7$

$$P(X = k) = \begin{cases} C_k^n p^k (1-p)^{n-k}, & \text{if } k = 0,1,2,\cdots,n \\ 0, & \text{otherwise} \end{cases}$$

- Do we have $\sum_{k=0}^{n} P(X = k) = 1$?

$$\sum_{k=0}^{n} C_k^n p^k (1-p)^{n-k} = 1$$

Binomial expansion

$$\sum_{k=0}^{n} C_k^n p^k (1-p)^{n-k} = (p + (1-p))^n, \quad p + (1-p) = 1$$

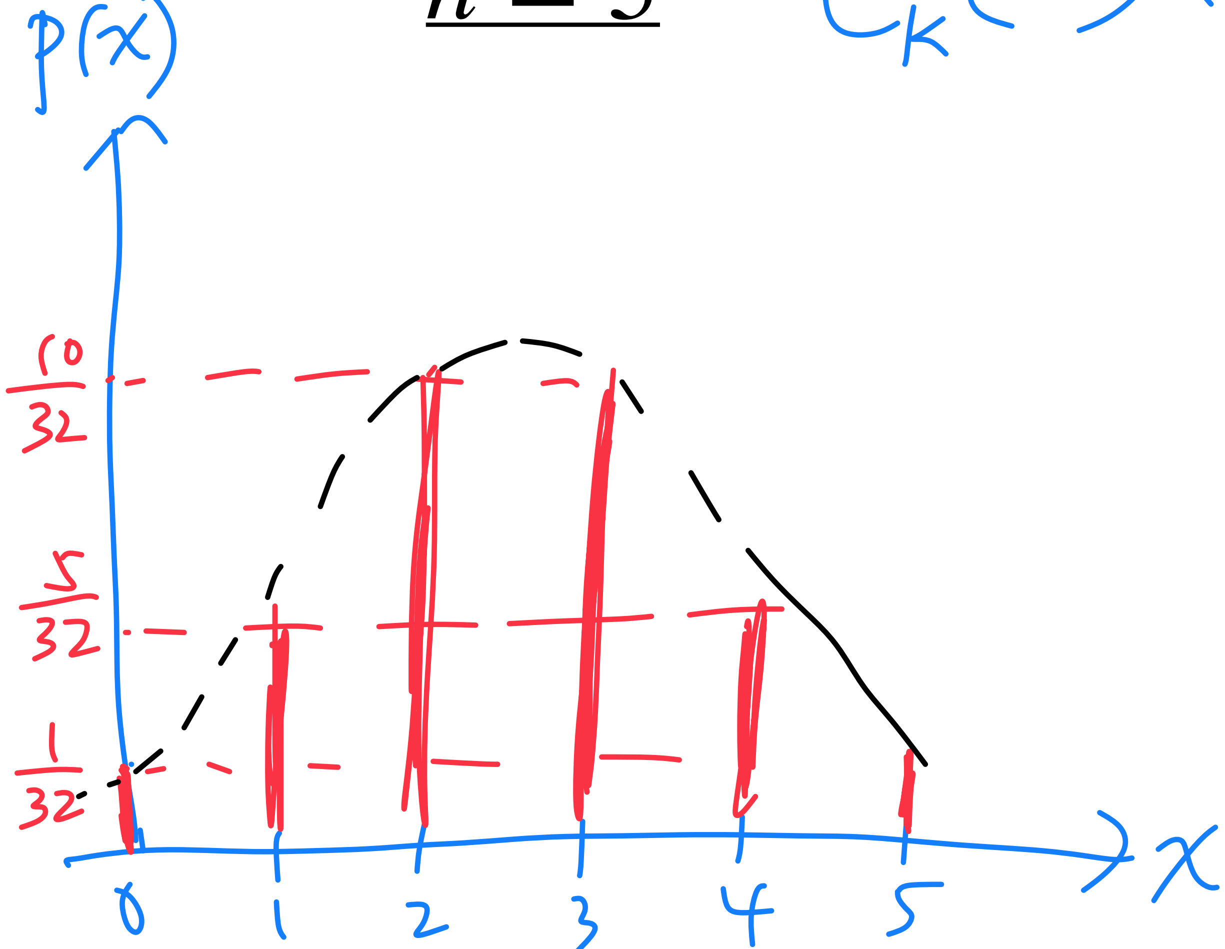
- What is a Binomial r.v. with parameter $n = 1$?

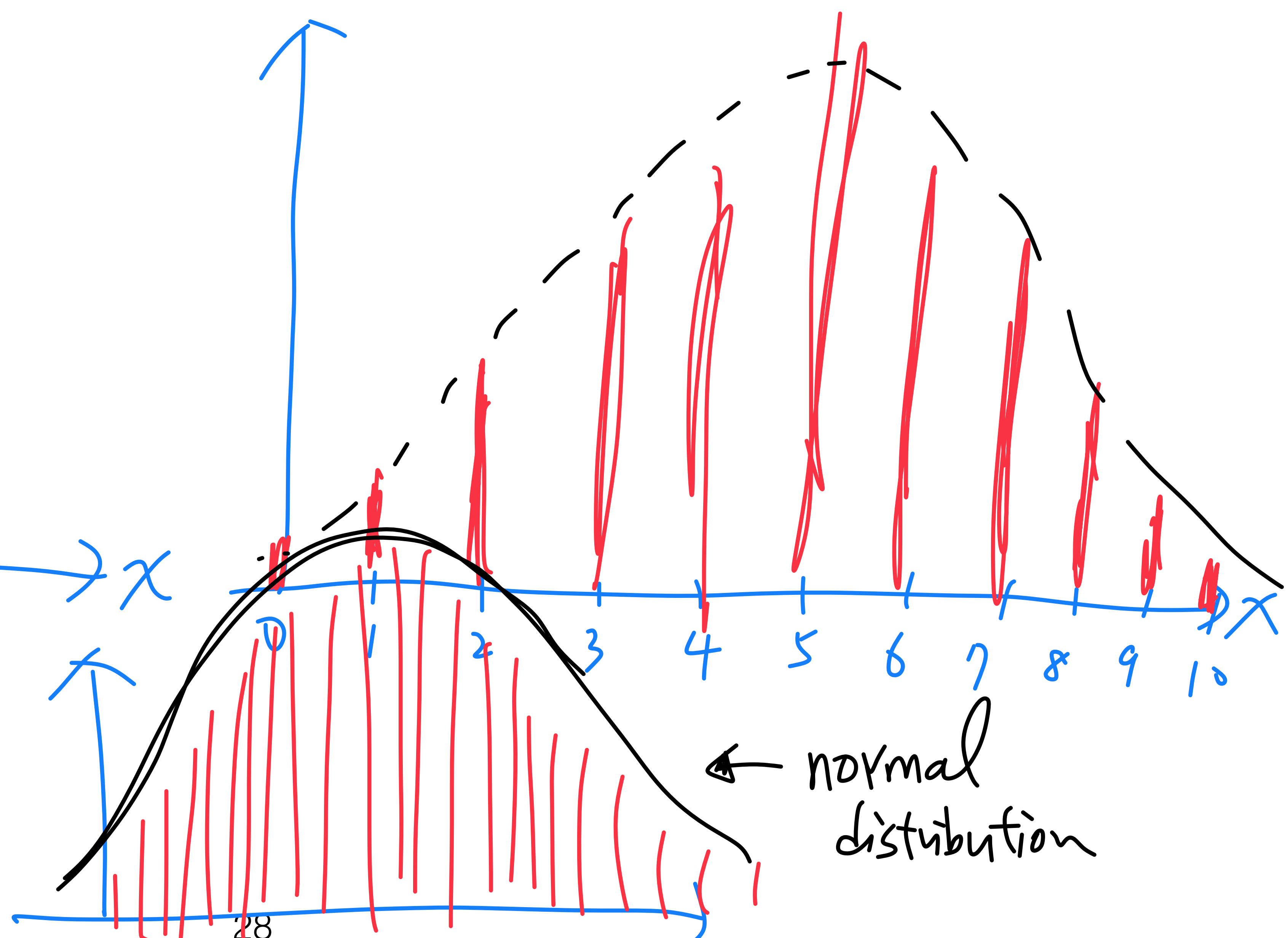
Binomial $(n=1, p) \equiv$ Bernoulli (p)

PMFs of Binomial Random Variables

- Example: Let's plot the PMF of $X \sim \text{Binomial}(n, p = 0.5)$

$n = 5$ $\quad C^5_K \cdot (0.5)^K \cdot (1-0.5)^{5-K}$ $\qquad n = 10$





- What if $n \to \infty$?

Example: Ensemble Learning With Voting

- Example: Suppose that we train an image classifier with either 1 or 3 ensembles and then apply majority voting
  - Suppose each classifier makes the correct prediction with probability p
  - Shall we choose 1 or 3 ensembles?

$(p>\frac{1}{2})$

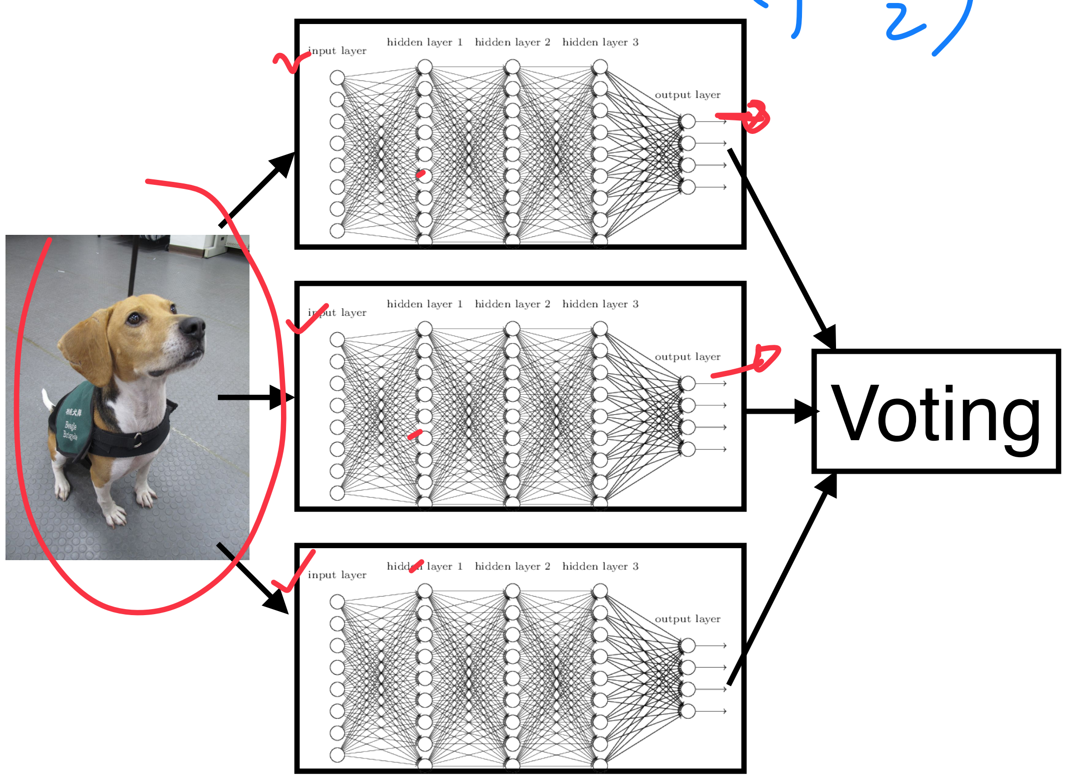
$p > p^3 + 3p^2(1-p)$

$(p<\frac{1}{2})$

① 1 ensemble: Prediction accuracy $= p$

② 3 ensembles:

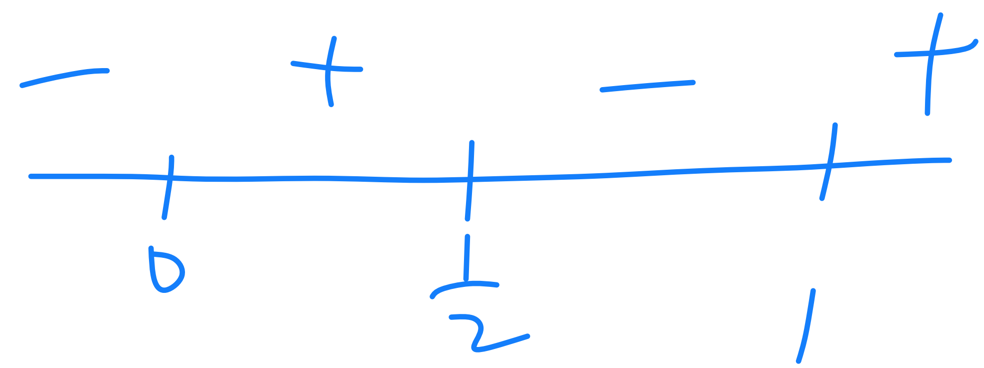
- 3 classifiers are all correct: $p \cdot p \cdot p = p^3 = C_3^3 \cdot p^3$
- 2 out of them are correct: $C_2^3 \cdot p^2 \cdot (1-p)^1 = 3p^2(1-p)$



$$P_1 > p^3 + 3p^2(1-p) = -2p^3 + 3p^2$$

$$\iff \underbrace{2p^3 - 3p^2 + p}_{} > 0$$

$$\iff p(2p^2 - 3p + 1) > 0$$



$$\iff p(2p-1)(p-1) > 0$$

$$0 \qquad \frac{1}{2} \qquad 1$$

$$\iff \cancel{p > 1} \text{ or } 0 < p < \frac{1}{2}$$

Example: A Poll of Coriander Lovers

- Example: Let p = probability that a random person likes coriander.
  - Suppose we randomly sample N people and define a random variable $X = \{\text{number of coriander lovers in } N \text{ people}\}$
  - For a fixed integer k, under what value of p is $P(X = k)$ maximized?

3. Poisson Random Variables

- Example: On average, 20 people stop by Shinemood every hour. What is $P($exactly 100 people visit Shinemood in 3 hours$)$?

- Example: On average, 1000 MayDay's concert tickets are sold every second. What is $P($all 50k tickets are sold out in 1 min$)$?

- What are the common features?
  - Average rate is known and static
  - Want: how many occurrences in an observation window?

Poisson: Limiting Case of Binomial

- Example: Consider $X \sim$ Binomial$(n,\ p = \lambda/n)$, λ is a constant
  - What is $P(X = k)$?
  - What if $n \to \infty$?

3. Poisson Random Variables (Formally)

Poisson Random Variables: Given parameters

- λ: average rate
- T: duration of the observation window

A random variable X is Poisson with parameter λT if its PMF is given by

$$P(X = n) = \frac{e^{-\lambda T}(\lambda T)^n}{n!}, \ n = 0,1,2,3,\cdots$$

- Do we have $\sum_{n=0}^{\infty} P(X = n) = 1$?

1-Minute Summary

1. Random Variables and CDF

- Function from outcomes to real numbers
- Use CDF to find the probability of an event

2. Probability Mass Function (PMF)

- An alternative way to specify the distribution of a discrete random variable

3. Special Discrete Random Variables

- Bernoulli / Binomial / Poisson