DCP 1206: Probability

Lecture 6 — Discrete Random Variables

Ping-Chun Hsieh

October 2, 2019

Sum of 3 Cubes

- Question: Can we find integers x, y, z such that

$$x^3 + y^3 + z^3 = N, \quad 0 \leq N \leq 100$$

- $\underline{N = 0:}$ $\quad (a)^3 + (-a)^3 + 0^3 = 0$

- $\underline{N = 29:}$ $(3)^3 + (1)^3 + (1)^3 = 29$

- $\underline{N = 9m + 4 \text{ or } 9m + 5:}$ Not possible

- Unsolved case (since 1954): $N = 42$

Sum of 3 Cubes for $N = 42$

Sum of three cubes for 42 finally solved – using real life planetary computer

Press release issued: 6 September 2019

Hot on the heels of the ground-breaking 'Sum-Of-Three-Cubes' solution for the number 33, a team led by the University of Bristol and Massachusetts Institute of Technology (MIT) has solved the final piece of the famous 65-year-old maths puzzle with an answer for the most elusive number of all - 42.

The original problem, set in 1954 at the University of Cambridge, looked for Solutions of the Diophantine Equation x^3+y^3+z^3=k, with k being all the numbers from one to 100.

Beyond the easily found small solutions, the problem soon became intractable as the more interesting answers – if indeed they existed – could not possibly be calculated, so vast were the numbers required.

But slowly, over many years, each value of k was eventually solved for (or proved unsolvable), thanks to sophisticated techniques and modern computers - except the last two, the most difficult of all; 33 and 42.

Professor Andrew Booker
Image credit: University of Bristol

Share this article

Andrew Booker
(University of Bristol)

- Answer:
  - $x = -80538738812075974$
  - $y = 80435758145817515$
  - $z = 12602123297335631$

charityengine

1.3 million hours of computation

Quick Review

- What is a random variable?

- What is CDF?

This Lecture

1. Probability Mass Function (PMF)

2. Special Discrete Random Variables

- Reading material: Chapter 5.1-5.3

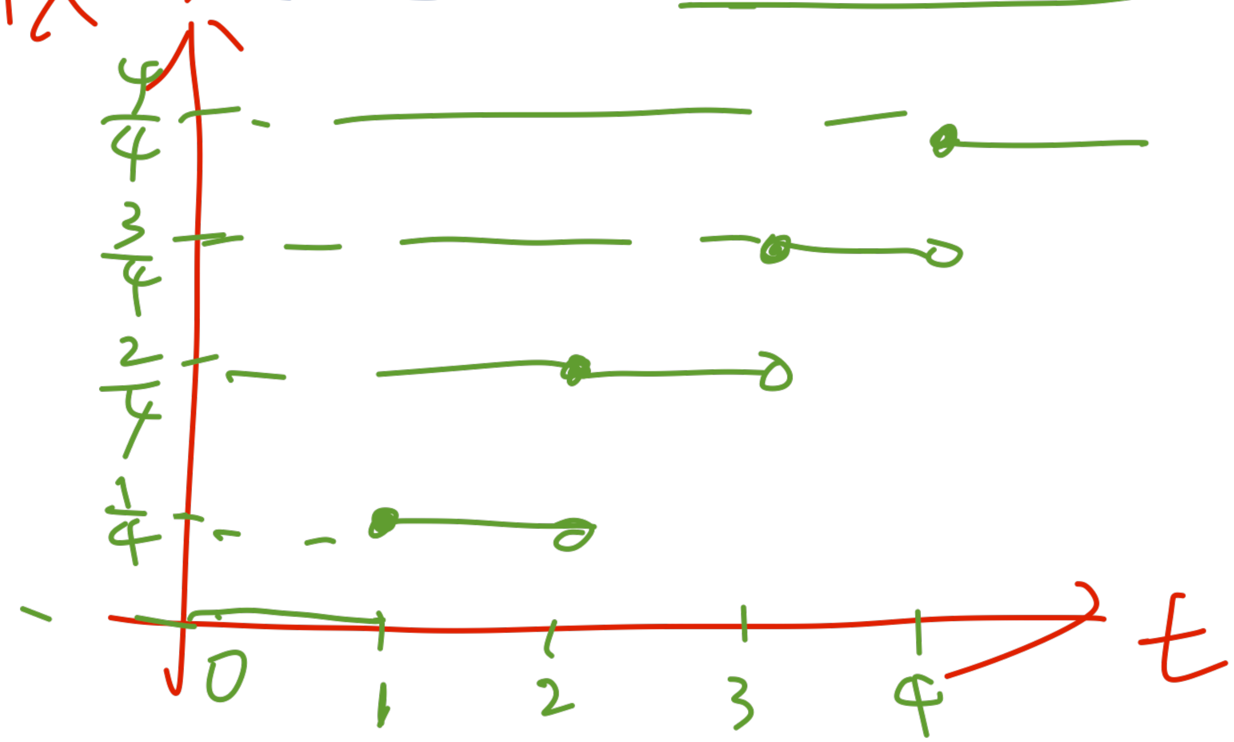
1. Probability Mass Function (PMF)

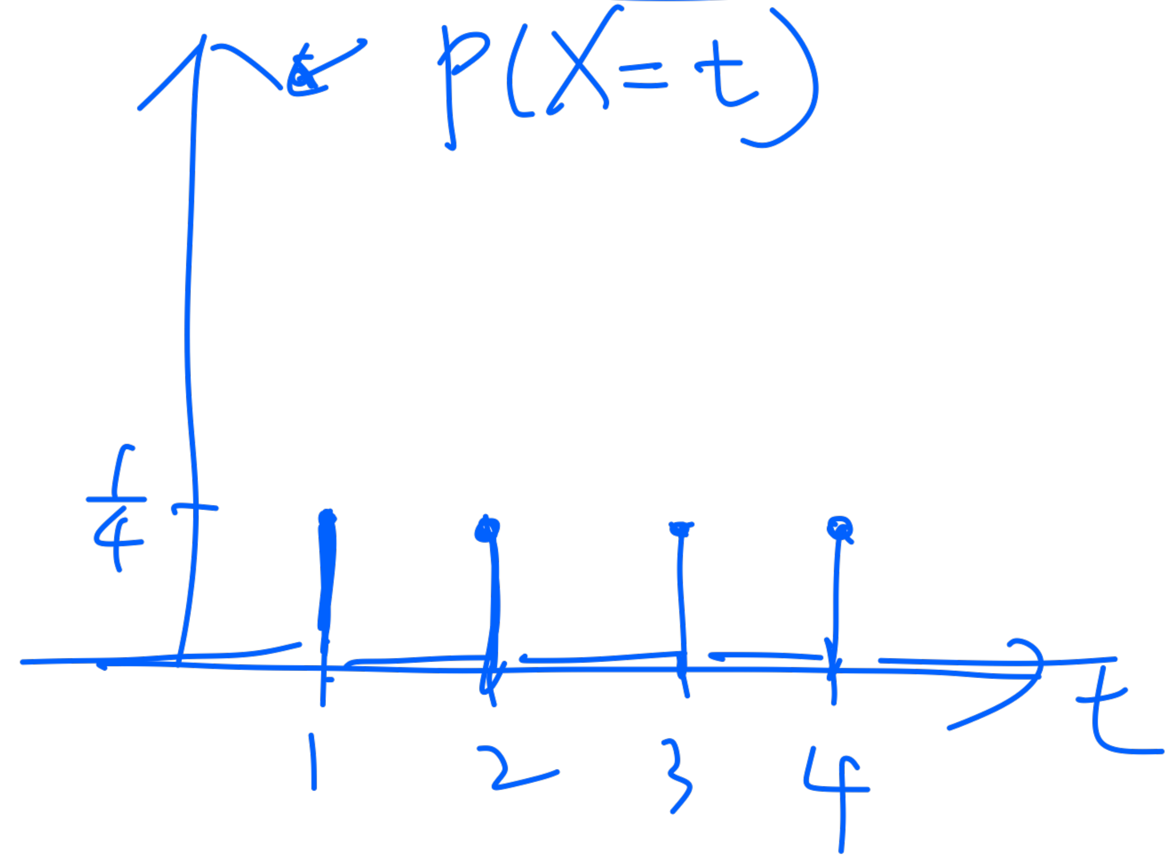
PMF: Another Way to Specify CDF of a Discrete Random Variable

- Example: Roll a fair 4-sided die once $\in \{1,2,3,4\}$
  - Define a random variable X = the number that we observe

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

$$\text{CDF} \doteq P(X\le t) \triangleq F_X(t) \qquad \text{PMF}$$





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: countable

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

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

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

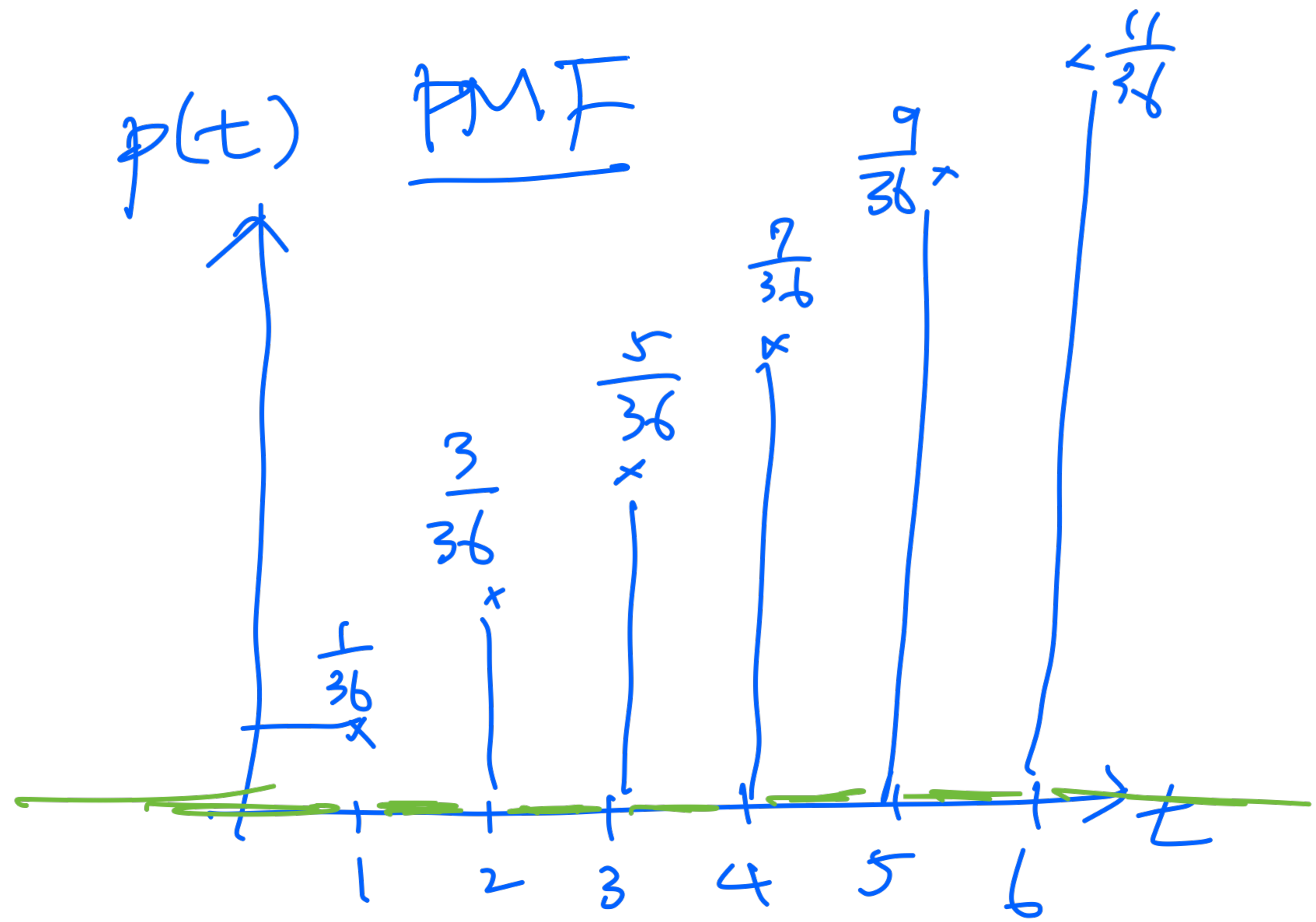
- For discrete random variables: CDF $\Leftrightarrow$ PMF

Example: From CDF to PMF

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

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

$\frac{3}{36}$ $\frac{5}{36}$



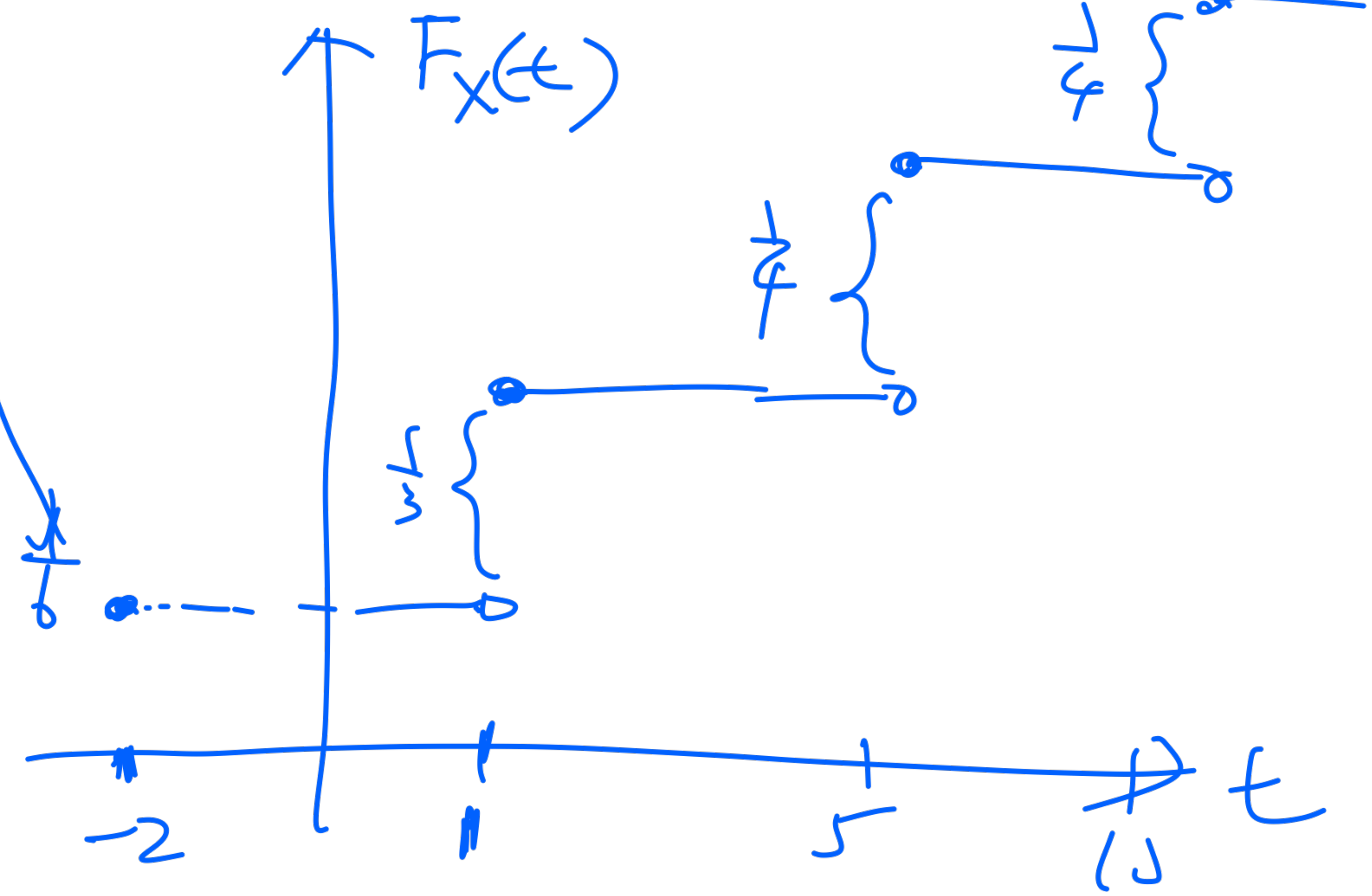
- What is the PMF of X?

Example: From PMF to CDF

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

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

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



- What is the CDF of X?

Probability Distribution

- Discrete random variables: CDF or PMF

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

2. Special Discrete Random Variables

Experiments With 2 Possible Outcomes

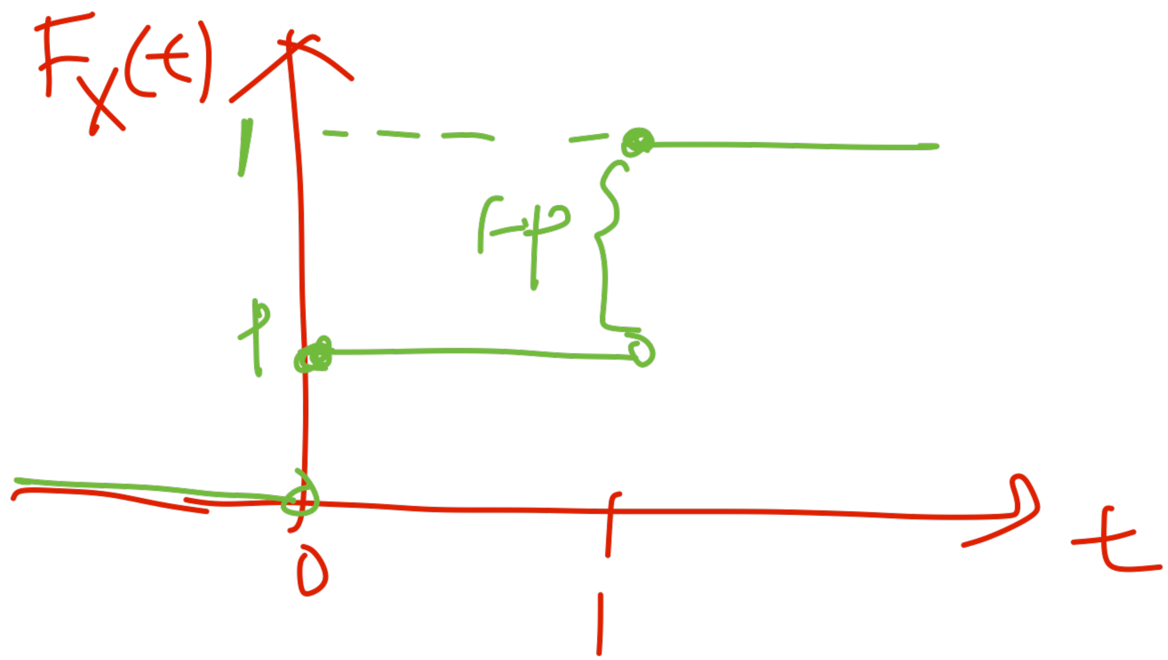
- Example: Whether an image of dog is classified correctly
- Example: Whether NCTU will merge with NYMU
- Example: Whether the 3rd student leaving the classroom wears glasses
- Example: Toss a coin once. Head or tail?
- What are the common features?
  - 1 experiment trial (no repetition) with 2 possible 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

- Example: Play with 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?

  (independent)
  - n repetitions of the same Bernoulli experiment
  - Want: how many successes in n repetitions

2. PMF of Binomial Random Variables

S / F

- 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?

$X \in \{0, 1, 2, 3, 4, 5\}$

$P(X=0) \Rightarrow \underbrace{FFFFF}$ $\Rightarrow (0.3)^5$ independence

$P(X=1) \Rightarrow$ SFFFF $\Rightarrow 5 \times 0.7 \times (0.3)^4$

FSFFF

FFSFF

FFFSF

FFFFS

$= C^5_1$

SSFFF

$P(X=2) = C^5_2 (0.7)^2 \times (0.3)^3$

$P(X=n) = C^5_n (0.7)^n \cdot (0.3)^{5-n}$

2. Binomial Random Variables (Formally)

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

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

$X \sim \text{Binomial}(n,p)$

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

$(x+y)^n = \sum_{k=0}^{n} C_k^n x^k y^{n-k}$

$x \leftarrow p$, $y \leftarrow 1-p$

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

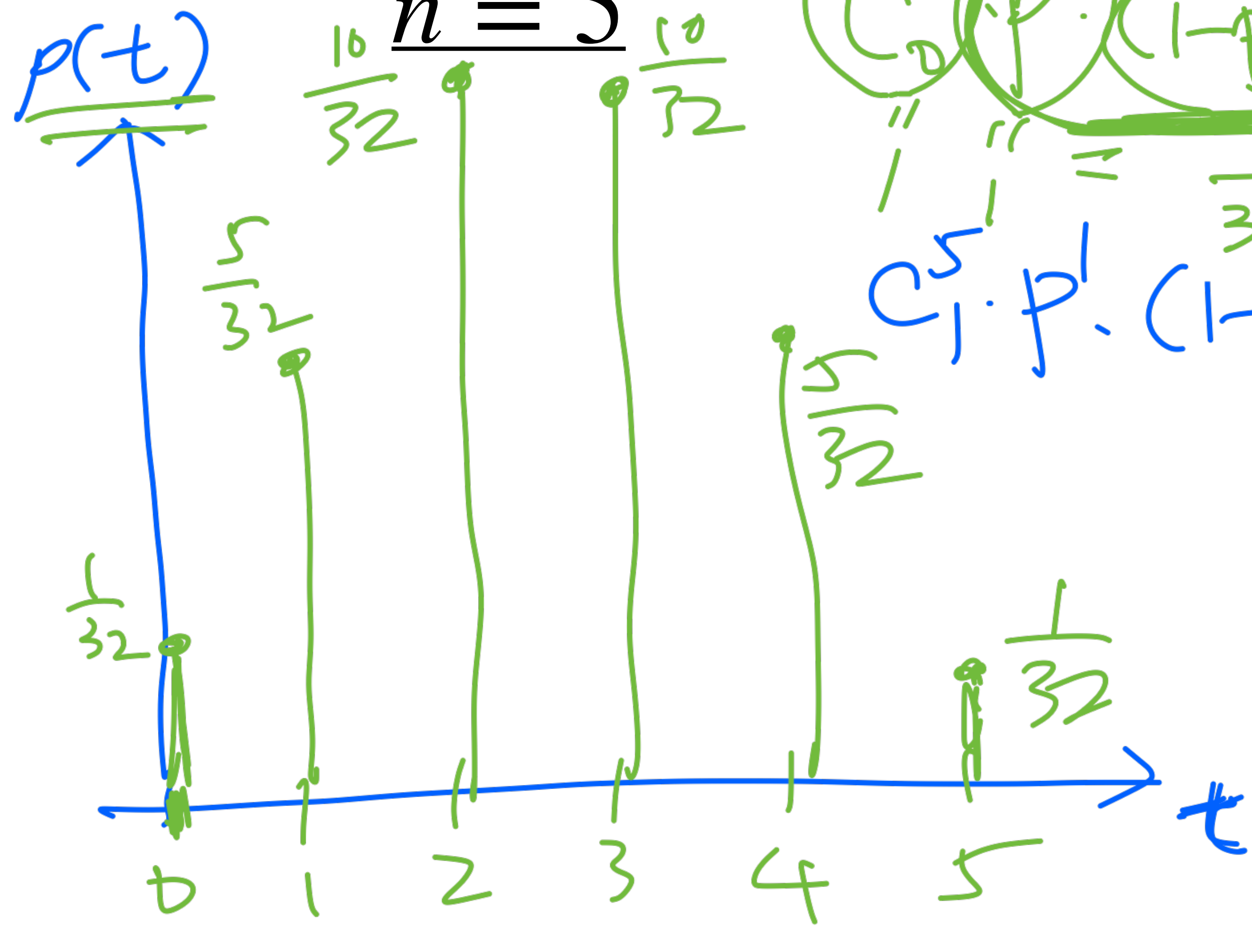
PMFs of Binomial Random Variables

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

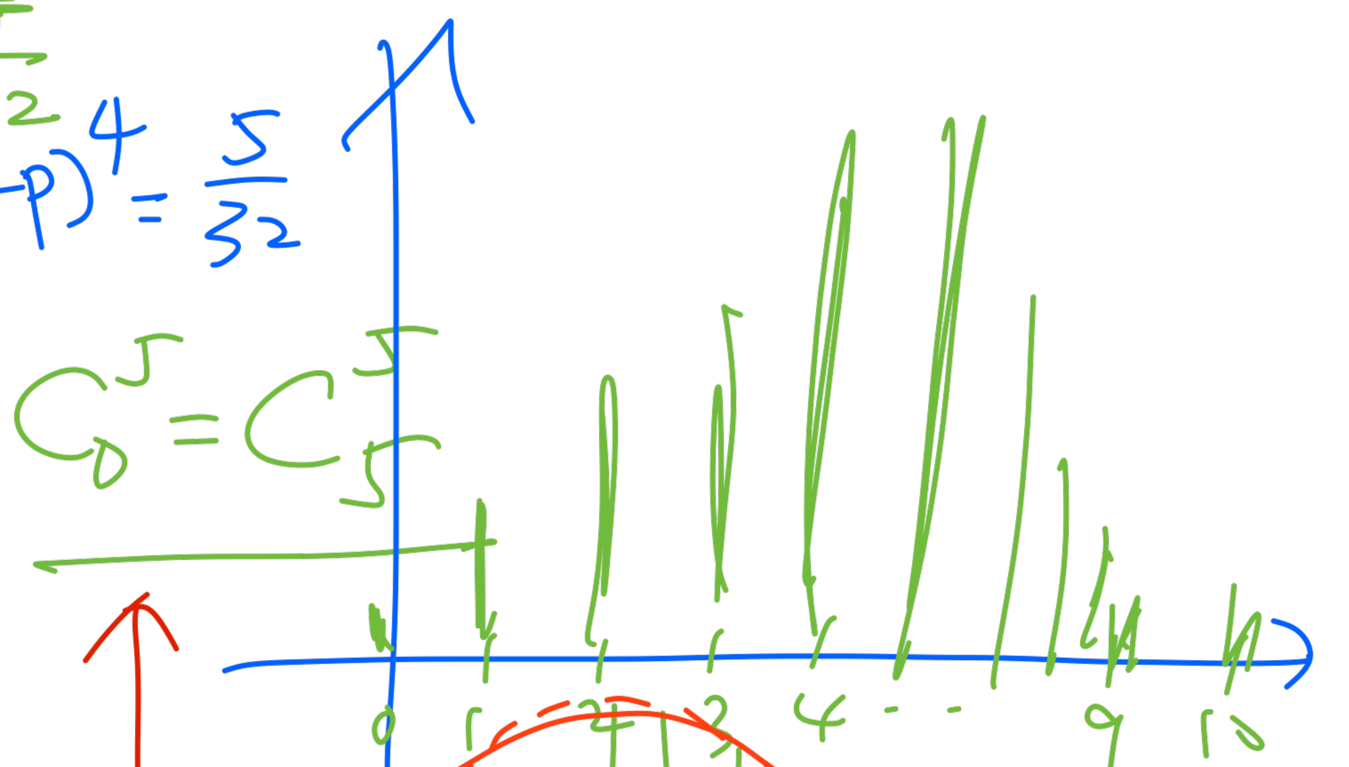
$n = 5$ $\qquad\qquad$ $n = 10$

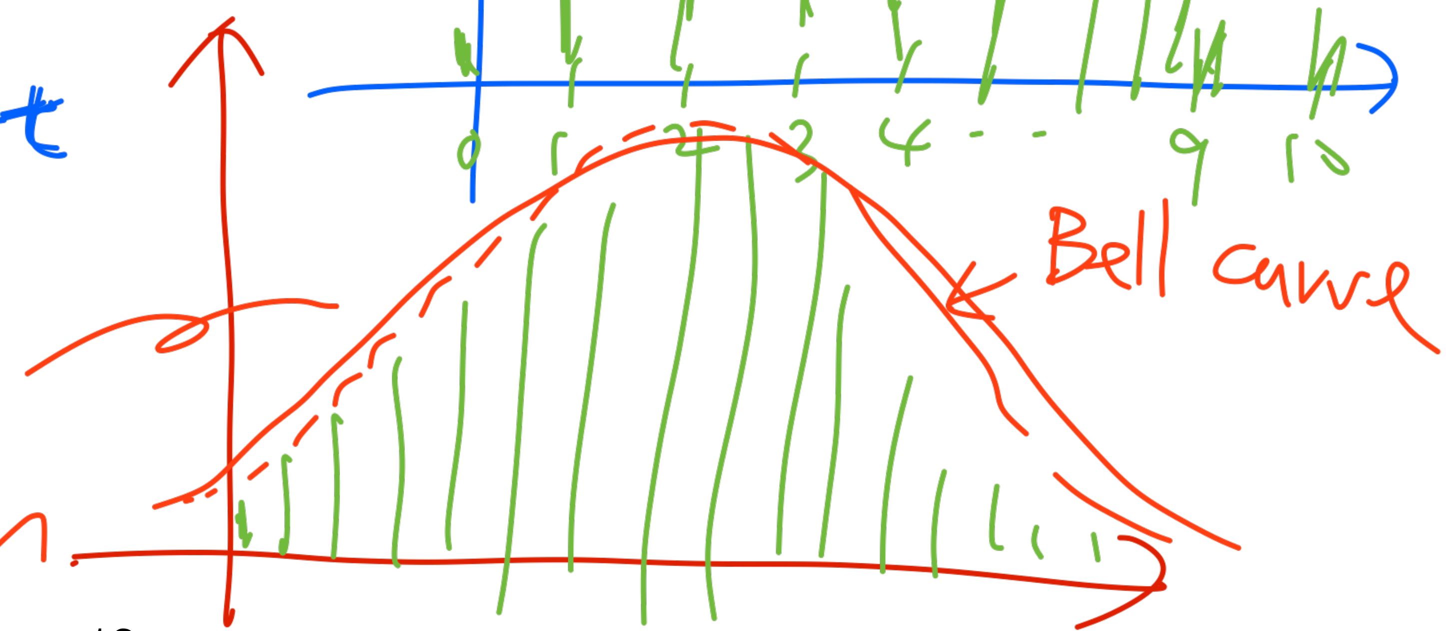
$C_0^5 \cdot p^0 \cdot (1-p)^5 = \frac{1}{32}$

$C_1^5 \cdot p^1 \cdot (1-p)^4 = \frac{5}{32}$



$C_0^5 = C_5^5$





- What if $n \to \infty$?

Central Limit theorem

Example: Ensemble Learning With Voting

r.v. $\Rightarrow X$ = # models that vote for a good decision

- 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 $p \in (0,1)$
  - Shall we choose 1 or 3 ensembles?

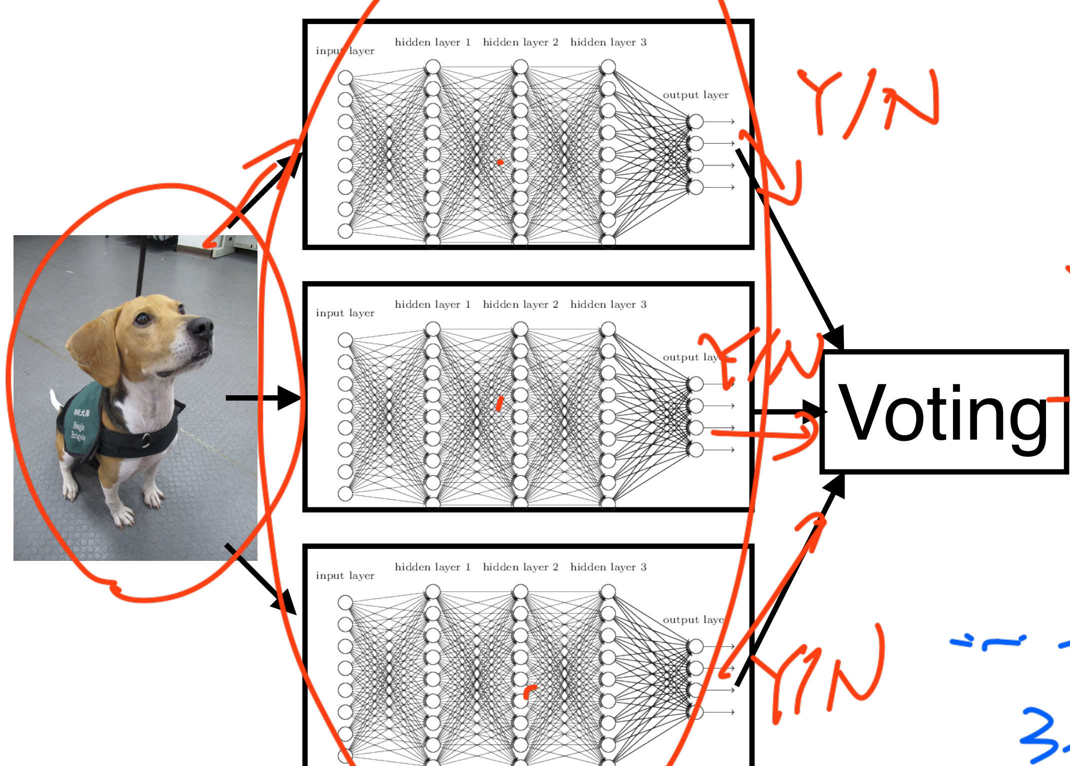
1 model: accuracy $= p$.

3 models:

$$P(\text{good decision}) = \underbrace{P(X=2) + P(X=3)}_{C^3_2 \cdot p^2 \cdot (1-p)^1 \;+\; p^3}$$

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

$$3p^2 - 2p^3 > p \iff 3p - 2p^2 > 1 \iff p > \frac{1}{2}$$



Example: A Poll of Coriander Lovers

assume $p \in (0,1)$ $\quad X \sim \text{Binomial}(N, p)$

- Example: Let p = probability that a random person likes coriander.
  - Suppose we randomly sample N people and define a random variable $X = \{$number of coriander lovers in N people$\}$ $\quad 0, 1, 2, 3, \ldots, N$
  - For a fixed integer k, under what value of p is $P(X = k)$ maximized?

$$P(X=k) = C_K^N p^K \cdot (1-p)^{N-K} = f(p)$$

$$f'(p) = C_K^N \left[K \cdot p^{K-1} \cdot (1-p)^{N-K} + p^K \cdot (N-K)(1-p)^{N-K-1} \cdot (-1) \right]$$

$$= C_K^N \cdot p^{K-1} \cdot (1-p)^{N-K-1} \cdot \left[K \cdot (1-p) + p \cdot (N-K) \cdot (-1) \right]$$

$$f'(p) = 0 \Rightarrow K - Np = 0 \Rightarrow p = \frac{K}{N}$$

$K - Np$

$f''(p) < 0$

泊松

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

$np = \lambda$

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

$P(X=k) \to \dfrac{e^{-\lambda} \cdot \lambda^k}{k!}$

Recall: $e = \lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n$

$e = 2.7\ldots$

$$P(X=k) = C^n_k \, p^k \, (1-p)^{n-k}$$

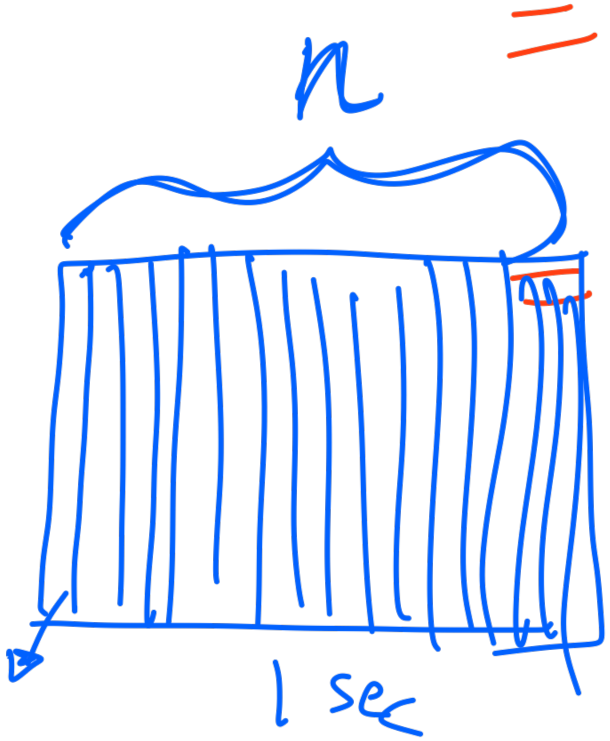
$$= \frac{n\cdot(n-1)(n-2)\cdots(n-k+1)}{k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1-\frac{\lambda}{n}\right)^{n-k}$$

k terms

$$\underbrace{\frac{n\cdot(n-1)(n-2)\cdots(n-k+1)}{n\cdot n\cdot n\cdots n}}_{\to 1} \cdot \frac{\lambda^k}{k!} \cdot \left(1-\frac{\lambda}{n}\right)^{n-k}$$

k terms

$$\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\left(1-\frac{\lambda}{n}\right)^{\frac{-n}{\lambda}\cdot(-\lambda)}}{\left(1-\frac{\lambda}{n}\right)^{k}} \to \frac{e^{-\lambda}}{1}$$



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

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

$e^{\lambda T}$

$$e^{-\lambda T} \sum_{n=0}^{\infty} \frac{(\lambda T)^n}{n!}$$

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

$$\sum_{n=0}^{\infty} \frac{e^{-\lambda T}(\lambda T)^n}{n!} = 1$$

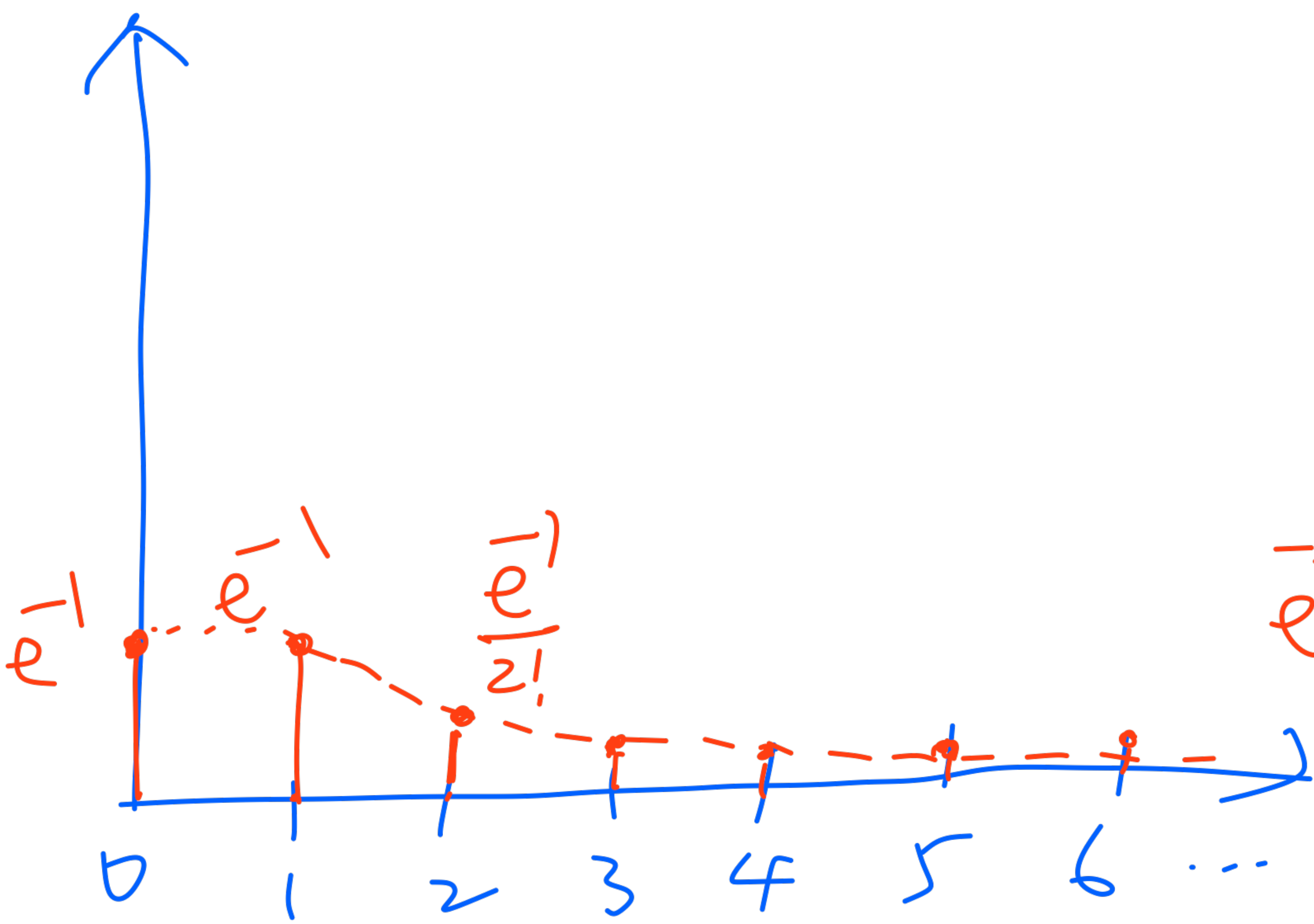
Taylor expansion: $e^x = 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$

PMFs of Poisson Random Variables

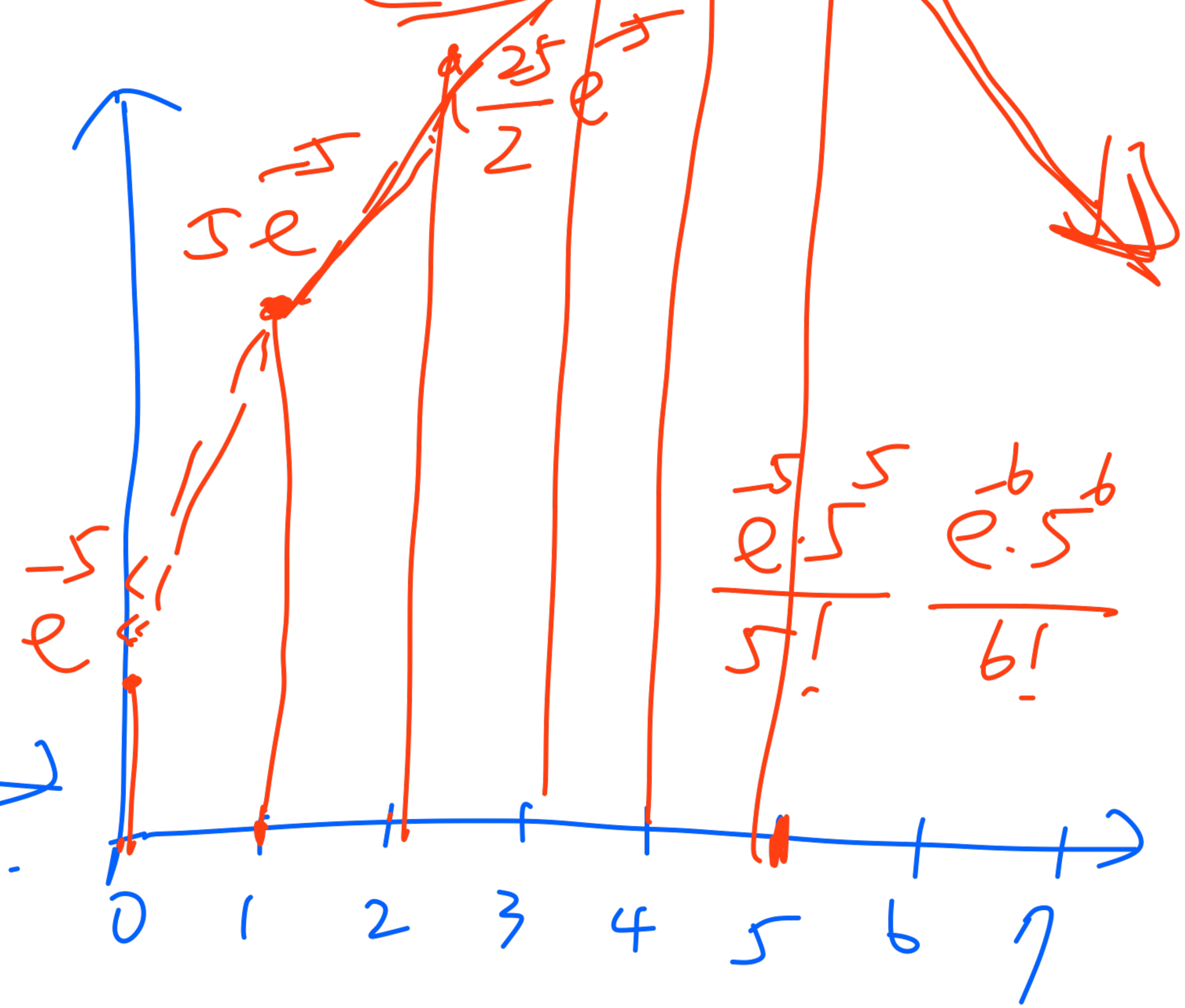
$$P(X=n) = \frac{e^{-\lambda T}\cdot(\lambda T)^n}{n!}$$

- Example: Let's plot the PMF of $X \sim$ Poisson(λT)

$\underline{\lambda T = 1}$



$\underline{\lambda T = 5}$



Example: An Interview Question by Google

- Example: Suppose we stand at the Fude temple.
    - The probability that we see at least 1 car passing through the temple in 30 minutes is 0.95.
    - What is P(we see at least 1 car in 10 mins)?

≈ 63%   rate is static

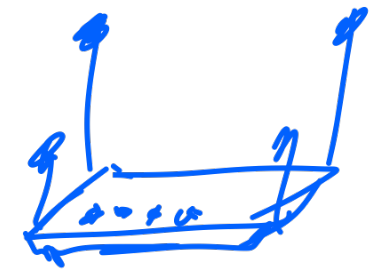
r.v. X = number of cars in 30 min, $X \sim \text{Poisson}(\lambda, T = 30\text{min})$

$$\underbrace{P(X=0)}_{= \frac{e^{-\lambda T} (\lambda T)^n}{n!},\ n \leftarrow 0} = 0.05 = e^{-\lambda T}, \quad T = 30\text{ min}$$

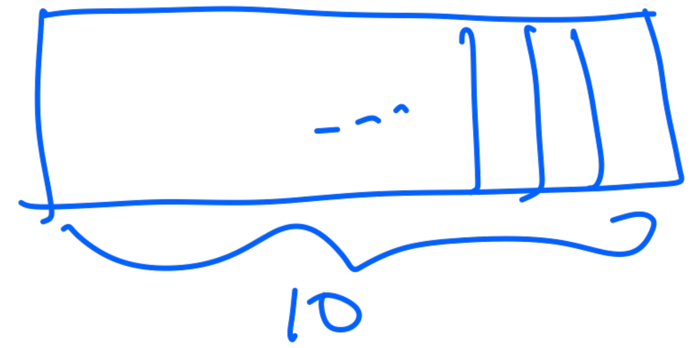
$$\boxed{P(Y=0) = e^{-\lambda T^*} = \sqrt[3]{e^{-\lambda T}} = \sqrt[3]{0.05} \approx 0.368}$$

Y = number of cars in 10 min, $Y \sim \text{Poisson}(\lambda, T^* = 10\text{ min})$

Example: Packets and A Finite Buffer



- Example: Consider a Wi-Fi access point (AP)
  - On average, 3.5 packets arrive at the AP every 1ms
  - The AP has a finite buffer of size = 10 packets
  - The AP can consume 1 packet every 1ms
  - Suppose the buffer has 1 packet at time 0
  - What is P(buffer is full at time 1ms)?



X = number packet that arrive in the 1st 1ms

$t=0$: 1 packet

$t=1$:

$X \sim \text{Poisson}(\lambda = 3.5, T = 1\text{ms})$

$\lambda T = 3.5$

$$P(X \geq 10) = \sum_{n=10}^{\infty} P(X=n)$$

$$= \sum_{n=10}^{\infty} \frac{e^{-3.5} \cdot (3.5)^n}{n!}$$

Quick Summary

1. Bernoulli random variables

1 exp, 2 outcomes, p

2. Binomial random variables → repeat Bernoulli.

limit

n trials, 2 outcomes, n, p

3. Poisson random variables

avg. fixed rate λ, observation window T

λT

4. Geometric Random Variables

- Example: Play with a claw machine, and each trial is successful with probability 0.7. What is $P(\text{get 1st toy at 10-th trial})$?

- Example: Po-Jung Wang makes a hit with probability 0.28 at each at-bat. What is $P(\text{he makes his 1st hit at 5-th at-bat})$?

- What are the common features?
  - Repetitions of the same Bernoulli experiment
  - Want: how many trials needed until the 1st success?

4. PMF of Geometric Random Variables

- Example: Play with a claw machine, and each trial is successful with probability 0.7. All trials are independent.
  - X = the number of trials until we get the first toy
  - What is the PMF of X?

4. Geometric Random Variables (Formally)

Geometric Random Variables: A random variable X is <u>Geometric with parameters p</u> if its PMF is given by

$$P(X=k)=(1-p)^{k-1}p,\ k=1,2,3,\cdots$$

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

CDF of Geometric Random Variables

$$P(X = k) = (1 - p)^{k-1}p, \ k = 1,2,3,\cdots$$

- CDF: $F_X(t) = P(X \le t)$

Geometric r.v.: Memoryless Property

- Example: Suppose $X \sim$ Geometric(p), $p \in (0,1)$
  - What is $P(X = n + m \,|\, X > m)$? $(n, m \in \mathbb{N})$
  - What is $P(X > n + m \,|\, X > m)$? $(n, m \in \mathbb{N})$

5. Negative Binomial Random Variables

- Example: Play with a claw machine, and each trial is successful with probability 0.7. What is $\underline{P(\text{get 3rd toy at 10-th trial})}$?

- Example: Po-Jung Wang makes a hit with probability 0.28 at each at-bat. What is $\underline{P(\text{he makes his 3rd hit at 5-th at-bat})}$?

- What are the common features?
  - Repetitions of the same Bernoulli experiment
  - Want: how many trials needed until the r-th success?

5. PMF of Negative Binomial Random Variables

- Example: Play with a claw machine, and each trial is successful with probability 0.7. All trials are independent.
  - X = the number of trials until we get the 3rd toy
  - What is the PMF of X?

5. Negative Binomial Random Variables (Formally)

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

$$P(X = n) = C_{r-1}^{n-1} p^r (1-p)^{n-r}, \ n = r, r+1, \cdots$$

- What if $r = 1$?
- Why is it called 'Negative Binomial'?

Why Is It Called Negative Binomial?

$$P(X=n)=C_{r-1}^{n-1}p^r(1-p)^{n-r},\ n=r,r+1,\cdots$$

- Suppose: let $k=n-r$

$$\begin{aligned}C_{r-1}^{n-1}p^r(1-p)^{n-r} &= C_k^{k+(r-1)}p^r(1-p)^k\\ &= (-1)^kC_k^{-r}p^r(1-p)^k\end{aligned}$$

Example: Modified Gambler's Ruin

- Example: Suppose two gamblers A and B play a game.
  - In each play, A wins with probability $p \in (0,1)$
  - In each play, the loser loses \$1, but the winner wins \$0
  - Initially, A has a dollars and B has b dollars
  - What is probability that B loses all the money first?

6. Discrete Uniform Random Variables

- Example: Roll a 4-sided die, and the numbers 1, 2, 3, 4 are equally likely to occur

- Example: The correct answer to an exam question: A, B, C, D are equally likely

- What are the common features?
  - 1 experiment trial (no repetition) with n equally-likely outcomes
  - Want: Whether a specific outcome occurs

6. Discrete Uniform Random Variables (Formally)

Discrete Uniform Random Variables: A random variable X is discrete uniform with parameters (a, b) ($a, b \in \mathbb{Z}$ with $a \leq b$), if its PMF is given by

$$P(X = k) = \frac{1}{b - a + 1}, \ k = a, a + 1, \cdots, b$$

Next Lecture

1. Hypergeometric random variables

2. Expectation and variance

1-Minute Summary

1. Probability Mass Function (PMF)

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

2. Special Discrete Random Variables

- Bernoulli / Binomial / Poisson
- Geometric and Negative Binomial