Last lecture

Law of Total Probability (Ch 2.10)

- Bayes formula $P(B|A) = \dfrac{P(A|B)P(B)}{P(A)}$

Independent Events/ RVs (Ch 2.4)

- Definition
- Motivation
- Examples and Facts

Agenda

Distributions

- Bernoulli (Ch 2.4.3)
- Binomial (Ch 2.4.4)

Geometric distribution (Ch 2.5)

Distributions

Common discrete RVs

Bernoulli Distribution

X is Bernoulli distribution with $\boxed{\text{parameter } p}$ if

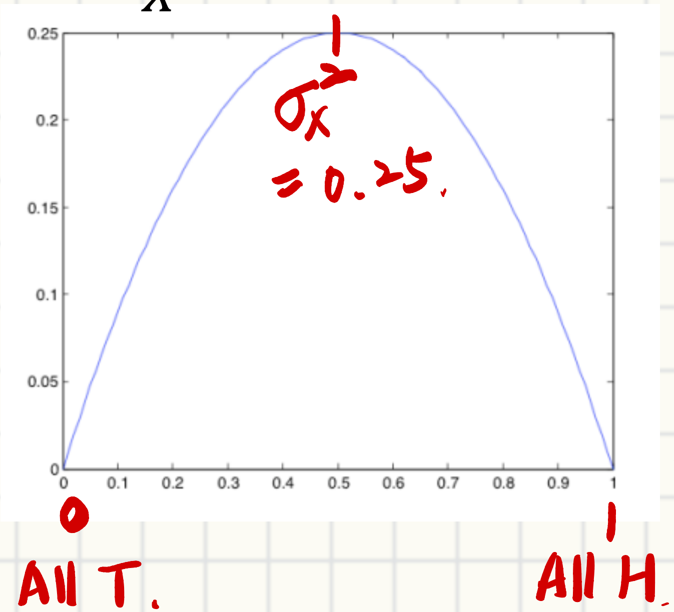
first momentum

- $P\{X=1\}=p$ and $P\{X=0\}=1-p$ $\quad p=0.5 \rightarrow$ fair coin
- "Toss a (unfair) coin with p probability Head"
- Only 2 possible outcomes, pmf contains 2 bins $\quad P_X(1)=p$, $\quad P_X(0)=1-p$
- $E[X] = 1\times P_X(1) + 0\times P_X(0) = p$
- $E[X^2] = 1^2\times P_X(1) + 0^2\times P_X(0) = p$
- $\sigma_X^2 =$

second momentum

$Var(X)$
$= E[X^2] - (E[X])^2$
$= p - p^2 = p(1-p)$

σ_X^2



p

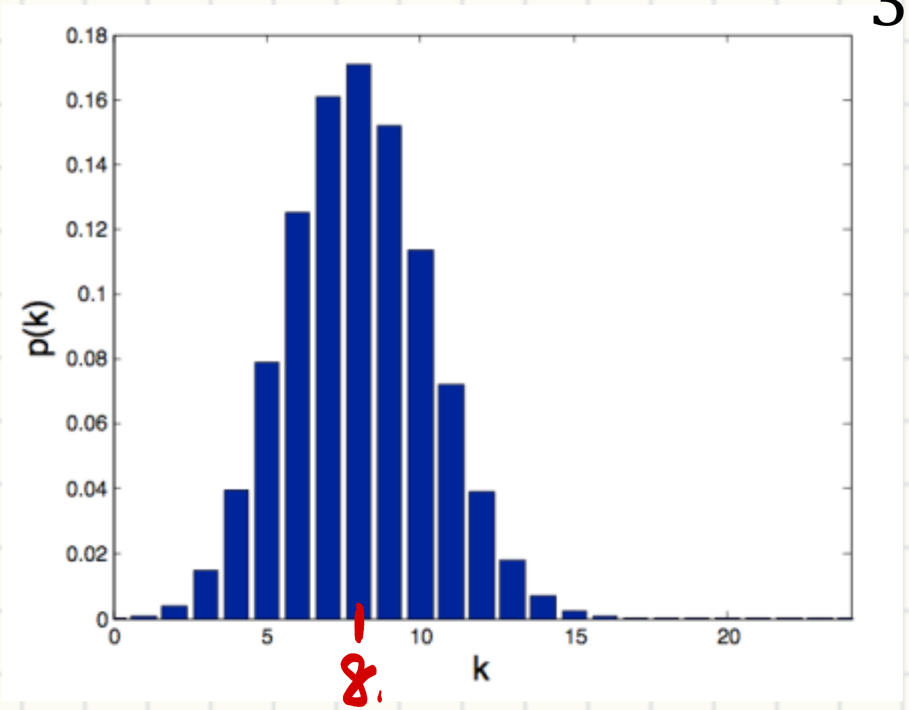
Binomial Distribution

X is binomial distribution with $\boxed{\text{parameter } (n,p)}$ if

- X is sum of n Bernoulli trials with parameter p
- ~~Draw~~ the unfair coin n times and count the Head

Toss.

$(n,p) = (24, \frac{1}{3})$



- $p_X(k) =$

$$\binom{n}{k} p^k (1-p)^{n-k}$$

(k) H. (n-k) T.

E.g. $X \sim Bi(n,p) = (4, \frac{1}{3})$

$$P_X(1) = P(\boxed{\{HTTT\}}, \{THTT\}, \{TTHT\}, \{TTTH\})$$

$$\frac{1}{3} \times \left(1-\frac{1}{3}\right)^3$$

$\times$

$$\binom{4}{1}$$

$\Downarrow$

$$\binom{4}{1}\left(p=\frac{1}{3}\right)^{k=1}(1-p)^{n-k}$$

Binomial Distribution

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof - $\sum_{k=0}^{n} p_X(k) = 1$

- $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$
- $x = \frac{p}{1-p}$

$$\left(1+\frac{p}{1-p}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{p}{1-p}\right)^k$$

$$\left(\frac{1}{1-p}\right)^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{-k}$$

$\times (1-p)^n$ on both sides.

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

Binomial mean

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Not on exam

$$E[X] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{m=0}^{n-1} \frac{n-1!}{(n-1-m)!m!} p^m (1-p)^{n-1-m}$$

$$= np \underbrace{\sum_{m=0}^{j} \binom{j}{m} p^m (1-p)^{j-m}}_{1}$$

Binomial Properties

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- Mean $E[X] =$ $np = n \times E[\text{Bern}(p)]$
- Variance $Var(X) =$ $n \times Var(\text{Bern}(p)) = np(1-p)$
  - why? n tosses are independent.

- Shape of the pmf- what is the most likely k?
  - $k^* = \lfloor (n+1)p \rfloor$

$k^* = \underset{k}{\operatorname{argmax}}\, P_X(k)$

Slido

Back to early 1900's, there's a mail fraud

- I mail N people that I can predict a series of 50-50 games
- I predict A wins in $\frac{N}{2}$ mail, B wins for the other
- Stop mailing a person after 2 wrong guess
- Say people will subscribe to my prediction if I only make at most 1 mistake in 5 guesses
- What should be N if I want to get 300 subs?

#9578467

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

3200 ? x. $\quad X \sim Bi(n=5,\ p=\frac{1}{2})$ $\quad H =$ success.

$$P_X(5) + P_X(4)$$

$$= \binom{5}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 + \binom{5}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1$$

$$= \frac{1}{32} + \frac{5}{32} = \frac{3}{16}$$

$$N \times \frac{3}{16} = 300.$$

$$\boxed{N = 1600}$$

Binomial Example – Best of K

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Team A and B play "Best of 7" games

- No tie, whoever wins 4 games out of 7 is the match winner
- E.g. $w_i = \{A, A, A, B, A\}$: the winner is A
- Let p denotes A's win rate per game
- Y denotes the number of games played, $p_Y(k) = ?$

$$P_Y(k) = P_Y(k, A \text{ wins}) + P_Y(k, B)$$

$$P_Y(k, A) = \binom{k-1}{4-1} p^4 (1-p)^{k-4}$$

$$P_Y(k, B) = \binom{k-1}{4-1} p^{k-4} (1-p)^4$$

Geometric Distributions

Geometric Distribution

of Toss on a (unfair) coin until the first Head is shown

Conduct independent Bernoulli trials of parameter p

- $\boxed{L} \triangleq$ # of trials until we get the first 1
- $p_L(1) = p$.
- $p_L(2) = P(\{T, H\}) = p(1-p)$
- $p_L(k) = p(1-p)^{k-1}$
- $P\{L > k\} = (1-p)^k$

Mean

failed first toss

$$E[L] = P\{H\} \times 1 + P\{T\} \times (E[L] + 1)$$

L

$$\mu_L = p + (1-p)[\mu_L + 1]$$

H

$$p\mu_L = 1 \qquad \mu_L = \frac{1}{p}$$

T

$\#$ 1 $+$ L

Variance

$$Var(L) = E[L^2] - \mu_L^2$$

$$E[L^2] = 1^2 \times p + (1-p)\, E[(L+1)^2]$$

$$E[L^2] = p + (1-p)\, E[L^2 + 2L + 1]$$

$$p\, E[L^2] = p + \frac{2(1-p)}{p} + 1 - p$$

$$E[L^2] = \frac{2-p}{p^2}$$

$$Var(L) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Memory trick JUST FOR Geometric dist.

$$Var(L) = \mu_L^2 - \mu_L = \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2}$$

Example

What's the expected number of rolls to get 1 to 6 at least once?

- Example series :
  $\{2, 4, 2, 3, 4, 4, 3, 5, 3, 5, 4, 4, 6, 2, 3, 3, 4, 1\}$

- $R_k \triangleq$ # roll to get the k-th unseen number

$R_1 =$ Accept any #s.

$\sim \text{Geo}(p = 1)$

$R_2 \sim \text{Geo}(p = \frac{5}{6})$ Accept $\boxed{\text{all but 1 \#}}$

$R_3 \sim \text{Geo}(p = \frac{4}{6})$

Ans.: $R_1 + R_2 + R_3 \dots + R_6$

Think of roll

$$\{\underbrace{1}_{R_1}, \underbrace{4}_{R_2}, \underbrace{1, 3}_{R_3}, \underbrace{2}_{R_4}, \underbrace{1, 3, 5}_{R_5}, \underbrace{4, 2, 6}_{R_6}\}$$

$$E[Y] = E\left[Geo(1) + Geo\left(\tfrac{5}{6}\right) + Geo\left(\tfrac{4}{6}\right) \dots Geo\left(\tfrac{1}{6}\right)\right]$$

$$= \frac{1}{1} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} \dots + 6$$

Game – Push the luck (simplified Incan-Gold)

Start a game with infinite rounds and 0 points

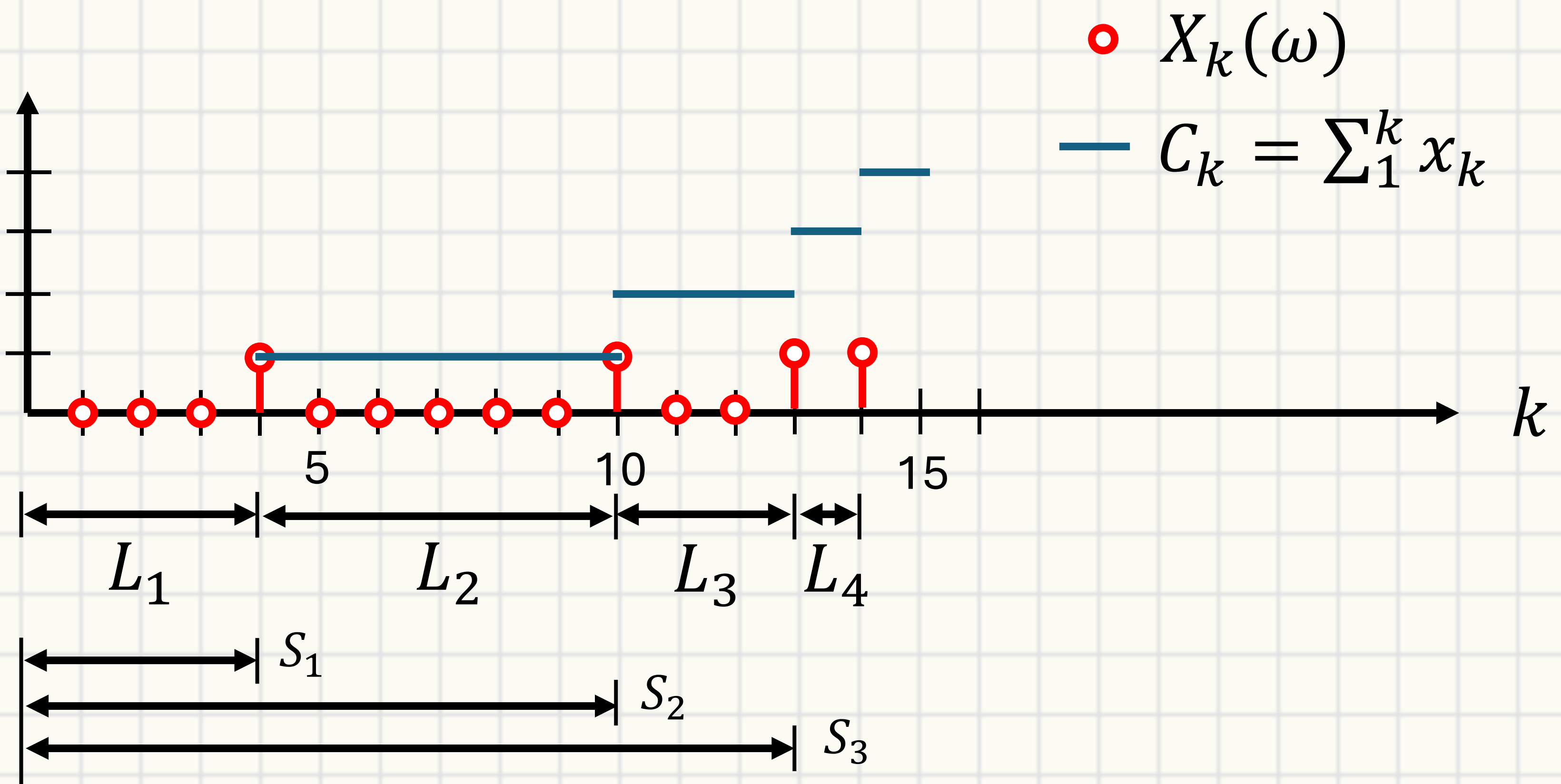
- 2 Actions per round – Go or Keep
- Go
  - $p = \frac{2}{3}$ win 1 point
  - Otherwise, lose all points
- Keep
  - Deposit the current point and end the game
- What's the best strategy?

Bernoulli Process

Bernoulli Process Definition

An infinite sequence X_1, X_2 … s.t. $X_k \sim Bern(p)$

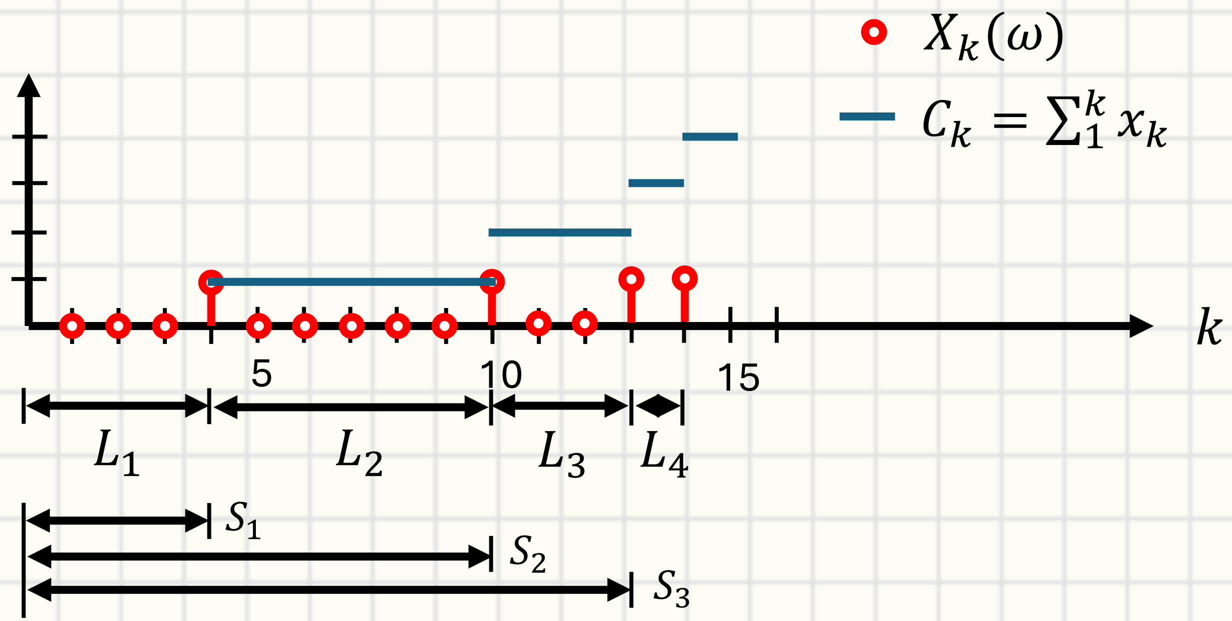
- ω is a possible outcome of the sequence
- $X_k(\omega)$ is called a "   " of outcome ω



Bernoulli Process Definition

Observe that a Bernoulli process can be defined by

1. $X_k \sim Bern(p)$
2. $C_k \sim B(k, p)$
3. $L_k \sim L(p)$
4. $S_r = \sum_1^r L_r$: # of trials required to get r ones



S_r - Negative Binomial Distribution

What is the pmf of S_r with parameter (r, p)?

- # of trials required to get r ones
- $p_S(n) =$