Caveats

- Problems in the slide will be independent from midterm problems
  - $P(p_1|p_1' \text{ in slide}) = P(p_1)$
- Problems are mostly from homework previous years
  - Midterm will be slightly easier than homework in general
- All numbers will be replaced by symbols in the slide
  - In midterm, you may need to compute
- We will cover top-K options from the Slido survey
  - Survey does not cover all topics
  - You still need to review all topics by yourself

Agenda & Survey result

- Bernoulli Process
  - Bernoulli/ Binomial/ Geometry/ Neg. Bi. Review
- Questions for distributions

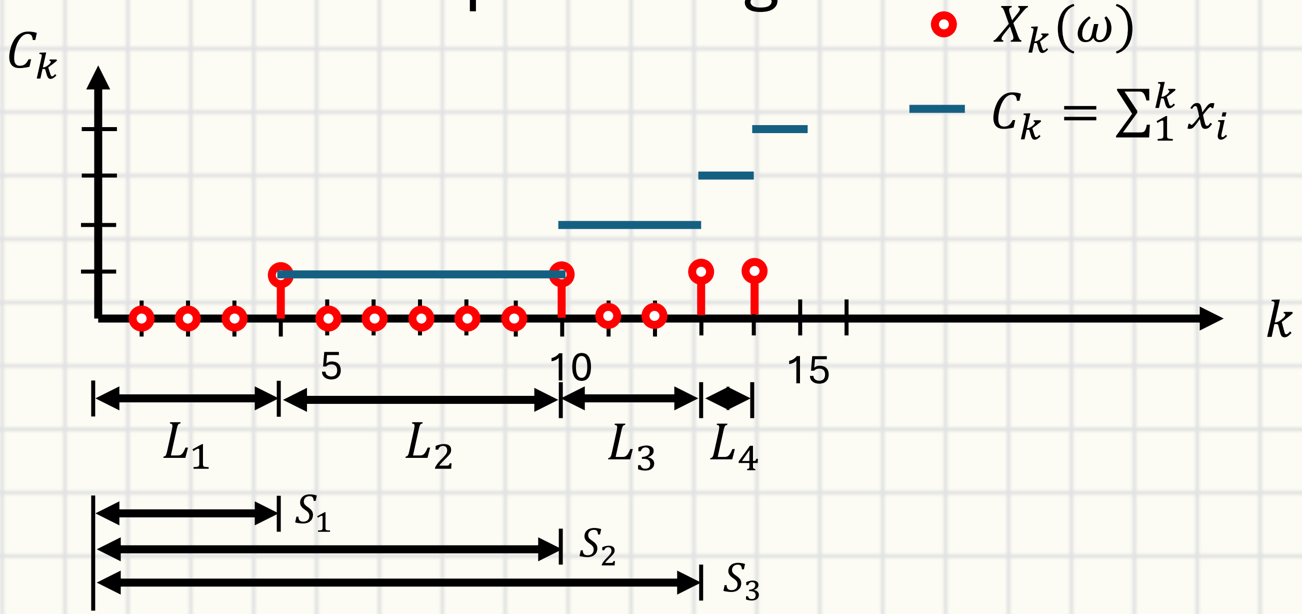
- Conditional Probability
  - Bayes and Law of Total Probability
  - Questions

Bernoulli Process

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

- Toss unfair coin many times

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



Properties

	$Bern(p)$	$Bin(n,p)$	$Poi(\lambda)$	$Geo(p)$	$NB(r,p)$
Relation	X_i	$\sum_n X_i$	$\sum_n X_i, n \to \infty$	Y_i	$\sum_n Y_i$
Mean					
Variance					
pmf					
Example	Toss a coin	Toss n times	Large n small p	Until first H	Until r^{th} H
Special		$(p+q)^n$		Memoryless	

Binomial

	$Bern(p)$	$Bin(n,p)$	$Poi(\lambda)$	$Geo(p)$	$NB(r,p)$
Mean	p	np	λ	$1/p$	r/p
Variance	$p(1-p)$	$np(1-p)$	λ	$(1-p)/p^2$	$r(1-p)/p^2$
pmf	-	$\binom{n}{k} p^k (1-p)^{n-k}$	$e^{-\lambda}\lambda^k/k!$	$(1-p)^{k-1}p$	$\binom{k-1}{r-1}(1-p)^{k-r}p^r$

An airplane has s seats but sell t tickets. Each customer has probability p to show.

- P(everyone has seats)
- Can we model the number of no-shows?

Geometry

	$Bern(p)$	$Bin(n,p)$	$Poi(\lambda)$	$Geo(p)$	$NB(r,p)$
Mean	p	np	λ	$1/p$	r/p
Variance	$p(1-p)$	$np(1-p)$	λ	$(1-p)/p^2$	$r(1-p)/p^2$
pmf	-	$\binom{n}{k} p^k(1-p)^{n-k}$	$e^{-\lambda}\lambda^k/k!$	$(1-p)^{k-1}p$	$\binom{k-1}{r-1}(1-p)^{k-r}p^r$

Play a Roulette wheel

- $\Omega = \{00, 0, 1, \dots .36\}$
- Always bet "small" $S = \{1,2, \dots 15\}$
- P(Lose first 3 bets)
- P(First win on 5^{th} bet)
- P(First win on 6^{th} bet | Lose on first bet)

Neg. Bin.

	$Bern(p)$	$Bin(n,p)$	$Poi(\lambda)$	$Geo(p)$	$NB(r,p)$
Mean	p	np	λ	$1/p$	r/p
Variance	$p(1-p)$	$np(1-p)$	λ	$(1-p)/p^2$	$r(1-p)/p^2$
pmf	-	$\binom{n}{k} p^k(1-p)^{n-k}$	$e^{-\lambda}\lambda^k/k!$	$(1-p)^{k-1}p$	$\binom{k-1}{r-1}(1-p)^{k-r}p^r$

$X \sim NB(r,p)$

- $P($r success occurs before m loss$)$?

Conditional Probability

$P(B|A) = \frac{P(A,B)}{P(A)}$ if $P(A) > 0$

If A and B are independent, $P(A|B) = P(A)$ or $P(AB) = P(A)P(B)$

Bayes - $P(B|A) = \frac{P(A,B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$

LOTP - $P(A) = \sum_i P(A|E_i)P(E_i)$

	B	B^c		
	E_1	E_2	E_3	E_4
A				
A^c				

Conditional Probability

Suppose Z_1, Z_2, Z_3 are i.i.d. Bernoulli random variables with parameter p .

- $S = Z_2 + Z_3$ if $Z_1 = 1$, and $S = Z_2 - Z_3$ if $Z_1 = 0$
- $P(S = 1)$?
- $P(Z_1 = 1 | S = 1)$?
- $P(Z_1 = 1 | S = 0)$?

	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
Z_1				
Z_1^c				

Bayes

Consider a disease test

- D denotes infected, T denotes test positive
- $P(T|D) = p$
- $P(T|D^c) = q$
- $P(D) = d$
- $P(D|T) =?$

Real-Time QA