Propositional Logic

Proposition: Statements that are true or false. ★

ex) $\sqrt{2}$ is irrational. → True (proposition)

$2+2=3$ → False (also a proposition!)

Johnny Depp is a good actor → not a proposition

$4+5$ → 9 (not a proposition)

$x+x$ → ?? (not a proposition)

Propositional Forms:

Conjunction (AND) $P \wedge Q$

→ P AND Q have to be true for $P \wedge Q$ to be true.

Disjunction (OR) $P \vee Q$

→ P OR Q have to be true for $P \vee Q$ to be true.

Negation (NOT) $\neg P$

→ P has to be false for $\neg P$ to be true.

Truth Tables:

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Use truth tables to prove logical equivalence!

ex) Prove $\neg(P \wedge Q)$ is equivalent to $\neg P \vee \neg Q$.

P	Q	$\neg(P \wedge Q)$	$\neg P \vee \neg Q$
T	T	F	F
T	F	T	T
F	T	T	T
F	F	T	T

$\Rightarrow$ Same Truth Table

$\rightarrow$ Equivalent!

Demorgan's Law: "Distribute and Flip" the negation ★

$$\neg(P \vee Q) \equiv \neg P \wedge \neg Q$$

$$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$$

$$(T \wedge Q) \equiv Q, \ (F \wedge Q) \equiv F$$

$$(T \vee Q) \equiv T, \ (F \vee Q) \equiv Q$$

$P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$?

Case 1: P is True.

$T \wedge (Q \vee R) \equiv (T \wedge Q) \vee (T \wedge R)$

$\downarrow \qquad\qquad\qquad \downarrow$

$(Q \vee R) \equiv (Q \vee R)$

Case 2: P is False.

$F \wedge (Q \vee R) \equiv (F \wedge Q) \vee (F \wedge R)$

$\downarrow \qquad\qquad\qquad \downarrow$

$F \equiv (F \vee F)$

$\downarrow \qquad\qquad\qquad \downarrow$

$F \equiv F$

$\Rightarrow P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$ is true!

Implication: $P \Rightarrow Q$ (If P, then Q.)

P is True, $P \Rightarrow Q$ is True $\rightarrow$ Q is True.

The statement "$P \Rightarrow Q$" is False only if P is True and Q is False!

$(P \Rightarrow Q) \wedge Q \equiv T$ does not mean $P \equiv T$! ✱

$(P \Rightarrow Q) \wedge P \equiv T$ means $Q \equiv T$! ✱

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

P	Q	$\neg P \vee Q$
T	T	T
T	F	F
F	T	T
F	F	T

$\rightarrow \neg P \vee Q \equiv P \Rightarrow Q$ ✱

Contrapositive: $\neg Q \Rightarrow \neg P$ (logically equivalent)

$P \Rightarrow Q \equiv \neg P \vee Q \equiv \neg(\neg Q) \vee \neg P \equiv \neg Q \Rightarrow \neg P$

Converse: $Q \Rightarrow P$ (not logically equivalent!)

Variables:

$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ → n is a variable, i is not.

$x > 2$ → x is a variable

n is even and the sum of two primes → n is a variable

→ These have a free variable (not propositions)

Predicate: $Q(x)$ = "boolean valued function"

Quantifiers: $\exists / \forall$

$(\exists x \in S)(P(x))$ → There exists x in S where $P(x)$ is true.

Equivalent to disjunction for all possible domains

$(\forall x \in S)(P(x))$ → For all x in S, $P(x)$ is true.

Equivalent to conjunction for all possible domains

ex) "There is a natural number that is the square of every natural number."

$(\exists y \in N)(\forall x \in N)(y = x^2)$ → False

"The square of every natural number is a natural number."

$(\forall x \in N)(\exists y \in N)(y = x^2)$ → True

→ Quantifiers are not commutative!

$\neg(\forall x \in S)(P(x)) \iff (\exists x \in S)(\neg P(x))$

LHS: Not all x in S makes $P(x)$ true.

RHS: There exists x in S such that $\neg P(x)$ is true.

→ Find a counterexample!

Proofs

Background: Integers are closed under $+$ and $\times$.

$a, b \in \mathbb{Z} \Rightarrow (a+b \in \mathbb{Z}) \wedge (a \times b \in \mathbb{Z})$.

$a \mid b$: "a divides b", $a \mid b \iff (\exists q \in \mathbb{Z})(b = aq)$

Direct Proof: Assume P. Then.... Q. ☆

ex) For any $a, b, c \in \mathbb{Z}$, if $a \mid b$ and $a \mid c$, then $a \mid (b-c)$.

Proof: Assume $a \mid b$ and $a \mid c$. → $b = aq$, $c = aq'$ $(q, q' \in \mathbb{Z})$

→ $b - c = aq - aq' = a(q - q')$ (Closedness of $\mathbb{Z}$)

$\Rightarrow a \mid (b-c)$, $(\forall a, b, c \in \mathbb{Z})$ //

ex) Let D_3 be a 3-digit natural number. For $n \in D_3$, if the alternating sum of digit of n is divisible by 11, then $11 \mid n$.

$\forall n \in D_3, (11 \mid \text{alt. digit of } n) \Rightarrow (11 \mid n)$

Proof: For $n \in D_3$, $n = 100a + 10b + c$ for some a, b, c.

alternating sum of $n := a - b + c$.

Assume $(a - b + c) = 11k$ $(k \in \mathbb{Z})$.

Add $99a + 11b$ to both sides.

$\rightarrow 100a + 10b + c = 11k + 99a + 11b = 11(k + 9a + b)$.

$\rightarrow n = 11(k + 9a + b) \Rightarrow 11 \mid n$ //

Converse: $\forall n \in D_3, (11 \mid n) \Rightarrow (11 \mid$ alt. digit of $n)$?

Proof: Assume $11 \mid n$. $\rightarrow n = 100a + 10b + c = 11k$.

Subtract $99a + 11b$ to both sides.

$\rightarrow a - b + c = 11k - 99a - 11b = 11(k - 9a - b)$.

$\rightarrow 11 \mid$ alt. sum of digits of n //

If $P \Rightarrow Q$ and $Q \Rightarrow P$, then $P \Leftrightarrow Q$. ☆

Prof by Contraposition: Assume $\neg Q$. Then... $\neg P$. ✱

ex) For $n \in \mathbb{Z}^+$ and $d|n$. If n is odd then d is odd.

→ $n = 2k+1$, $n = k'd$ ... but this is hard!

Assume $\neg Q$ that d is even. → $d = 2k$

$d|n \rightarrow n = qd = q(2k) = 2(qk) \rightarrow 2|n$

Proved $\neg P$ that n is even // $P \Rightarrow Q$

ex) For every n in $\mathbb{N}$, n^2 is even $\Rightarrow$ n is even.

$P: 2|n^2$, $Q: 2|n \longrightarrow \neg Q: n = 2k+1$, $\neg P: n^2 = 2k'+1$.

Assume $\neg Q$ that $n = 2k+1 \rightarrow n^2 = (2k+1)^2 = 4k^2+4k+1$.

$\rightarrow n^2 = 2(2k^2+2k)+1 \rightarrow k' = 2k^2+2k \rightarrow n^2 = 2k'+1$.

Proved $\neg P$ that n^2 is odd // $P \Rightarrow Q$

Proof by Contradiction: Assume $\neg P$. Then... $R \wedge \neg R$. ☆

ex) $\sqrt{2}$ is irrational. $\rightarrow \forall a,b \in \mathbb{Z}, (\frac{a}{b})^2 \neq 2$. (no common factors)

Proof: Assume $\neg P$ that $\sqrt{2}$ is rational $\rightarrow \exists a,b \in \mathbb{Z}, (\frac{a}{b})^2 = 2$.

$\rightarrow a^2 = 2b^2 \rightarrow 2|a^2 \rightarrow 2|a \rightarrow a = 2k$.

$\rightarrow (2k)^2 = 2b^2 \rightarrow 4k^2 = 2b^2 \rightarrow b^2 = 2k^2 \rightarrow 2|b^2 \rightarrow 2|b$.

$\rightarrow 2|a \wedge 2|b$. However, a and b must not have a common factor.

$\rightarrow$ Contradiction $\Rightarrow P$ is true.

ex) There are infinitely many numbers. (P)

Proof: Assume finitely many primes: $p_1, p_2, \cdots, p_k$ $(\neg P)$

consider number $q := (p_1 \times p_2 \times \cdots \times p_k) + 1$.

$\rightarrow q$ cannot be prime since $q > $ any other prime.

$\rightarrow q$ has prime divisor p $(p > 1 = R)$, $p \in \{p_i\}$.

$\rightarrow p$ divides both $x = p_1 \times p_2 \times \cdots \times p_k$ and $q \rightarrow p|(q-x)$.

$\rightarrow p \leq q - x = 1 \rightarrow p \leq 1 (\neg R) \rightarrow$ Contradiction $\rightarrow P$ is true.

We <u>did not</u> prove that q is prime because $\neg P$ is false!

Proof by Cases: P is true in either (a) or (b) or ...

ex) $x^5 - x + 1 = 0$ has no rational solution.

Lemma: if x is a solution to $x^5 - x + 1 = 0$ and $x = \frac{a}{b}$, $a, b \in \mathbb{Z}$, then both a and b are even.

Reduced Form: a and b can't both be even + Lemma

→ No rational solution

Proof of Lemma: Assume solution of the form $\frac{a}{b}$.

$\rightarrow (\frac{a}{b})^5 - \frac{a}{b} + 1 = 0 \rightarrow a^5 - ab^4 + b^5 = 0.$

Case 1) a odd, b odd: odd − odd + odd = even → not possible.

Case 2) a even, b odd: even − even + odd = even → not possible.

Case 3) a odd, b even: odd − even + even = even → not possible.

Case 4) a even, b even: even − even + even = even → possible!

However, a and b being even is not a reduced form

→ Contradiction → no rational solution.

ex) There exist irrational x and y such that x^y is rational.

Let $x = y = \sqrt{2}$.

Case 1) $x^y = \sqrt{2}^{\sqrt{2}}$ is rational. → Done!

Case 2) $x^y = \sqrt{2}^{\sqrt{2}}$ is irrational.

Then, set $x = \sqrt{2}^{\sqrt{2}}$, $y = \sqrt{2}$.

→ $x^y = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^2 = 2$ → rational!

One of the cases has to be true → Proved

(Which of the cases? No idea!)

Induction $(\forall k \in \mathbb{N})(P(k))$

Prove $P(0)$, Assume $P(k)$, Prove $P(k+1)$.

ex) For all $n \in \mathbb{N}$, $0+1+2+\cdots+n = \frac{n(n+1)}{2}$.

Base Case: $0 = \frac{0(0+1)}{2}$. ✓

I.H.: $0+1+2+\cdots+k = \frac{k(k+1)}{2}$.

Induction: $0+1+\cdots+k+(k+1) = \frac{k(k+1)}{2} + k+1$

$= \frac{k^2+k+2k+1}{2} = \frac{(k+1)(k+2)}{2} = \frac{(k+1)((k+1)+1)}{2}$ //

$P(0) \wedge (P(k) \Rightarrow P(k+1)) \Rightarrow (\forall n \in \mathbb{N})(P(n))$!

ex) $(\forall n \in \mathbb{N})(3 \mid n^3 - n)$

Base: $P(0) \to 3 \mid 0^3 - 0 \to 3 \mid 0$ ✓

I.H.: $3 \mid k^3 - k \to k^3 - k = 3q$.

Induction: $3 \mid (k+1)^3 - (k+1)$

$k^3 + 3k^2 + 3k + 1 - k - 1 = k^3 + 3k^2 + 2k$

$= (k^3 - k) + (3k^2 + 3k) = 3q + 3k^2 + 3k = 3(q + k^2 + k)$

$\Rightarrow 3 \mid (k+1)^3 - (k+1)$ //

ex) Two Color Theorem

Base: $n = 1$ (one straight line)

1) Add line 2) Get inherited colors 3) Switch color on one side

→ This algorithm works for any $P(n)$ to imply $P(n+1)$!

→ $P(1) \wedge (P(n) \Rightarrow P(n+1))$

Strengthening Induction Hypothesis:

ex) The sum of first n odd integers is a perfect square.

Stronger: The sum of first n odd integers is n^2.

Base: $1 = 1^2$. ✓

I.H.: $1+3+5+\cdots+(2k-1) = k^2$.

Induction: $1+3+5+\cdots+(2k-1)+(2k+1) = k^2+2k+1$

$= (k+1)^2 \Rightarrow P(k+1)$ //

The stronger hypothesis is true, thus the original hypothesis is also true.

ex) $2^{(2n)} \equiv 1 \pmod 3$

Base: $k=0 \rightarrow 2^0 \equiv 1 \pmod 3$ ✓

I.H: Assume $2^{2k} = 3a+1 \ (a \in \mathbb{Z})$

Induction: $2^{2(k+1)} = 2^{2k} \cdot 2^2 = 4 \cdot 2^{2k} = 4(3a+1)$

$= 12a+4 = 3(4a+1)+1 \Rightarrow 2^{2(k+1)} \equiv 1 \pmod 3$

Strong Induction:

$P(0) \wedge ((P(0) \wedge P(1) \wedge \cdots \wedge P(k)) \Rightarrow P(k+1)) \Rightarrow (\forall k \in \mathbb{N})(P(k))$

ex) Every natural number $n > 1$ can be written as a product of primes.

Base: $n = 2$. ✓

I.H.: n can be written as product of primes.

Induction: case 1) $n+1$ is prime → done!

case 2) $n+1 = a \cdot b$ $(1 < a, b < n+1)$

Strong Induction: a, b are also products of primes.

→ $n+1 =$ (factors of a)(factors of b)

Well-Ordering Theorem: Any set of natural numbers has a smallest element

→ Use to prove that if $\exists k(\neg P(k))$, then $\forall k(P(k))$ has to be false.

Sad Islanders: Use others' information to infer!

ex) For every $n \in \mathbb{N}$, $n = 4x + 5y$.

Base: $P(12), P(13), P(14), P(15)$.
$(3,0)$ $(2,1)$ $(1,2)$ $(0,3)$

Strong Induction: $P(n-4) \Rightarrow P(n)$

$(n-4 = 4x' + 5y') \Rightarrow (n = 4(x'+1) + 5y')$

Stable Matching

- n jobs and n candidates
- jobs and cands have preferences of each other

→ Produce matching such that all pairs are "stable".

Rogue Couple: $\{b, g^*\}$ for pairing S where b and g^* prefer each other to their partners in S.

ex)

Jobs

A	~~X~~	2	3
B	~~X~~	~~2~~	3
C	~~2~~	1	3

Candidates

1	C	A	B
2	A	B	C
3	A	C	B

	D1	D2	D3	D4	D5	
1	~~A~~,~~B~~	A	~~A~~,C	C	C	
2	C	~~C~~,B	B	~~B~~,A	A	
3					B	

The Propose & Reject Algorithm:

Each Day:

Each job proposes to its favorite candidate.

Each candidate rejects all but favorite proposer.

Rejected jobs cross rejecting candidate off list.

→ Then, each job finds a candidate.

Does this terminate? Yes, after max $\leq n^2$ steps.

For candidates, offers can only get better.

For jobs, proposals only get worse.

↳ Improvement Lemma

$P(k)$: "the current job is at least as good as any other offers in the future".

$P(0)$: True, the current job is the best offer.

On day $k+1$ → candidate can always choose the current job if all other offers are worse than current.

$P(k) \Rightarrow P(k+1)$!

Does every member have a pair when terminated? Yes.

Proof: Suppose a job is not matched.

Then, b has been rejected by every candidate.

That means each candidate has a job on string.

We have same number of jobs and candidates.

$\Rightarrow$ b must be on somebody's string. Contradiction. //

Is the final matching stable (no rogue couple)? Yes.

Proof: Assume there is a rogue couple (b, g^*)

b^* ——— g^* Job b proposes to g^* before g.

b - - - - g So g^* gets rejected by b (in this case)

By improvement lemma, b must prefer g^* over g!

$\Rightarrow (b, g^*)$ is not a rogue couple. Contradiction. //

Is this algorithm better for jobs or for candidates?

x's matching is optimal if x's partner is its best partner for any stable matching. Pessimal if its partner is worst out of all stable matching scenarios.

A matching is job-optimal if all jobs have optimal pairs.

... applies to candidate-optimal, job-pessimal, cand-pessimal.

(stability does not imply first on the list but rather best out of possible stable pairings)

Theorem: Job proposal - Candidate reject results job optimal.

Proof: Assume a job is not optimal. ($b \not\leftrightarrow g$)

Then, there is a stable pairing S where ($b \leftrightarrow g$).

Let t be the first day b gets rejected by g.

Then, b^* knocks off b in day t. (g prefers $b^* > b$)

By t, b^* likes g at least as much as optimal cand.s.

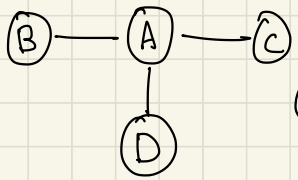
Thus, b^* prefers g to its partner g^* in S.

$\Rightarrow$ Rogue couple for S, S cannot exist. Contradiction.

Also, this algorithm is candidate-pessimal.

Graphs

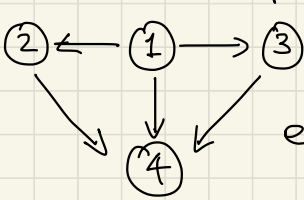
$G=(V,E)$. V := set of vertices. $E \subseteq V \times V$:= set of edges.



$V=\{A,B,C,D\}$. $E=\{\{A,B\},\{A,C\},\{A,D\}\}$

CS70 generally does not cover multigraphs.

Directed Graphs: Direction of edges exist.



$V=\{1,2,3,4\}$, $E=\{(1,2),(1,3),(1,4),(2,4),(3,4)\}$

ex) One-way street, tournament, social network

Concepts - neighbors, adjacent, degree, incident, in/out degree

u is a neighbor of v if $(u,v) \in E$.

Edge $\{u,v\}$ is incident to u and v.

Degree is the number of incident edges.

The sum of vertex degrees $= 2|E|$ (Handshake Lemma)

Path: A sequence of edges. ex) $\{(v_1,v_2),(v_2,v_3),\cdots,(v_{k-1},v_k)\}$

Cycle: Path from v_1 to v_{k-1} + edge(v_{k-1},v_1)

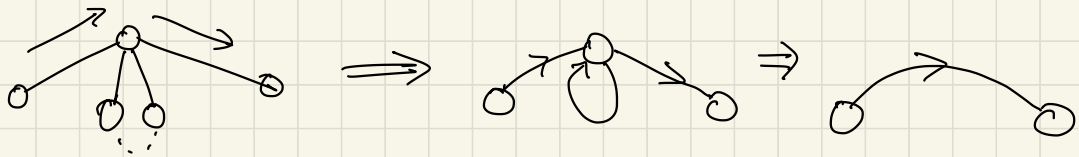
Walk: A sequence of edges with possible repeated vertex or edge.

Tour: Cycle for walks.

Connectivity: u and v are connected if there is a path between u & v.
A graph is connected if all pair of vertices are connected.

Eulerian Tour: A tour that visits each edge exactly once.

An undirected graph has an Eulerian tour iff. all vertices have even vertices and are connected.



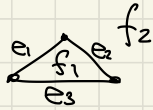
Take a walk from V to V. (this exists b/c connected, d is even)
Then, we can "smooth out" any uncovered edges.

Trees: Connected graph without a cycle.

Planar Graph: A graph that can be drawn in the plane without edge crossings.

Euler's Formula: ($v = |V|, e = |E|, f = \#$ of faces)

* $v + f = e + 2$ for any planar graph.



For a simple graph with $V \geq 3$, consider face adjacencies.

Each face is adjacent to at least 3 edges.

→ At least $\geq 3f$ face adjacencies

Also, each edge is adjacent to 2 faces.

→ $2e$ face adjacencies $\Rightarrow \underline{3f \leq 2e} \rightarrow f \leq \frac{2}{3}e$

Plug into Euler: $v + \frac{2}{3}e \geq e + 2 \rightarrow \underline{e \leq 3v - 6}$

(This is a necessary condition to be a planar graph!)

ex) K_5: $e = 10, v = 5 \rightarrow 10 \leq 15 - 6 = 9 \rightarrow$ False, cannot be planar

$K_{3,3}$: $e = 9, v = 6 \rightarrow 9 \leq 18 - 6 = 12 \rightarrow$ True, but not planar!

Actually, all cycles in $K_{3,3}$ is of even length → it is incident to ≥ 4 faces → $4f \leq 2e$ ... and able to prove nonplanar.

Proof of Euler's Formula: Induction on e.

Base: $e=0, v=f=1$. $1+1=0+2$ ✓

Induction: if it is a tree $\to e=V-1, f=1 \to (e+1)+1=e+2$ ✓

If it is not a tree $\to$ Find a cycle. Remove an edge.

This "joins" two faces. New graph has $(e-1)$ edges, $(f-1)$ faces.

I.H: $v+(f-1)=(e-1)+2 \Rightarrow v+f=e+2$ //

Graph Coloring: Each edge's endpoints have different colors.

Six Color Theorem: Recall $e \le 3v-6$ for any $v \ge 2$ planar graph.

Also, $\sum_{v \in V} \deg(v) = 2e$. $\to$ Average degree $= \frac{2e}{V} \le \frac{2(3V-6)}{V} \le 6 - \frac{12}{V}$.

$\Rightarrow \exists$ a vertex with deg < 6 (or at most 5).

Remove that vertex. Then inductively color remaing graph.

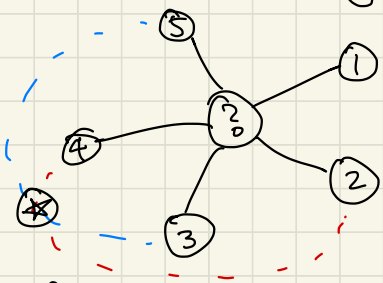
There is always a leftover sixth color to color the partition when the vertex is added back in.

Five Color Theorem:

Observation: Connected components of vertices with two colors in a legal coloring can switch colors.

(b/c other edges do not care about those two colors)

Consider a deg=5 vertex with all different colorings:



① Pick two vertices. If they are not connected by those two colors, switch only one of them. → Done!

If they are, consider another pair...?

That pair cannot ALSO be connected be cause of the assumption of planar graphs (what is the color of the vertex where the two paths are bound to cross? Contradiction.)

⇒ Thus, at least one pair is disconnected (recolored) into two same colors. ⇒ Fill in the leftover color! //

Complete Graph: K_n, All vertices are neighbors.

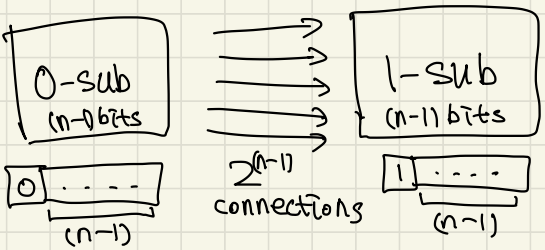
Hypercubes: $G=(V,E)$, $|V|=\{0,1\}^n$

$|E|=\{(x,y)|$ x and y differ by one bit$\}$

Then, to reach any vertex from another, max n steps.

2^n vertices, $n2^{n-1}$ edges.

Every hypercube can be divided into 2 subcubes.

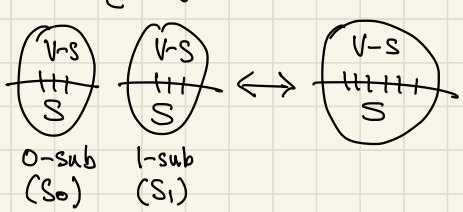


Theorem: For any subset S where $|S|\leq|V|/2$, S has more than $|S|$ edge connecting it to $V-S$.

$(S, V-S)$ is cut. $(E\cap S\times(V-S))\rightarrow$ cut edges

Base: $n=1\rightarrow V=\{0,1\}$ 0 — 1

Induction:



Case 1) $|S_0|\leq|V_0|/2, |S_1|\leq|V_1|/2$

Modular Arithmetic

Theorem: if $d|x$ and $d|y$, then $d|(y-x)$.

($x=ad, y=bd \Rightarrow d|(b-a)d \Rightarrow d|(y-x)$)

Theorem: $n \geq 2$, n can be described by product of primes.

"X is congruent to Y modulo m" or "$X \equiv Y \pmod m$"

$\iff (X-Y)$ is divisible by m ($m|(x-y)$).

$\iff x = y + km$ for some integer k.

* $\oplus$, $\ominus$, and $\otimes$ can be done with any equivalent x and y.

→ if $a \equiv c \pmod m$ and $b \equiv d \pmod m$, $a+b \equiv c+d \pmod m$,

$a-b \equiv c-d \pmod m$, and $a \cdot b \equiv c \cdot d \pmod m$.

$$\text{mod}(x, m) \longrightarrow \{0, 1, \cdots, m-1\} = x - \left\lfloor \frac{x}{m} \right\rfloor m$$

Thm: If $\gcd(x, m) = 1$, then x^{-1} s.t. $x^{-1}x \equiv 1 \pmod{m}$ exists

Proof: set $S = \{0x, 1x, 2x, \cdots, (m-1)x\}$ contains all distinct modulom. If not distinct, then $(a-b)x = km$.

However, $(a-b) < m$, so $m \nmid x$. → contradiction!

→ set S has to contain 1 as its element → inverse!

In fact, if $\gcd(x, m) = 1$, $f(a) = xa \pmod{m}$ is a bijection.

(Unique pre-image and same size)

Thm: If $\gcd(x, m) \neq 1$, then x has no multiplicative inverse (mod m).

Assume a is x^{-1}, or $ax = 1 + km \rightarrow ax \equiv 1 \pmod{m}$

$x = n \cdot d$ and $m = l \cdot d$ for $d > 1$ (b/c $\gcd(x, m) \neq 1$.)

$\rightarrow a \cdot n \cdot d = 1 + k \cdot l \cdot d \rightarrow d(an - kl) = 1$

Then $d = 1$ & $(an - kl) = 1$ → contradiction.

How to find inverses? How to chek existence of inverse?

→ find $gcd(x, m)$. $gcd(x, m) = 1 \iff$ inverse exists.

Is there a smarter way than checking $\{1, \cdots, (m-1)\}$?

Lemma: if $d|x$ and $d|y$, then $d|\, mod(x, y)$.

$mod(x, y) = x - \lfloor \frac{x}{y} \rfloor y = r = x - s \cdot y \ (s \in \mathbb{Z})$

$\rightarrow (kd) - (s)(ld) = d(k - sl) \rightarrow d|\, mod(x, y)$!

Lemma2: if $d|y$ and $d|\, mod(x, y)$, then $d|x$ and $d|y$.

$\Rightarrow gcd(x, y) = gcd(y, mod(x, y))$ ☆

Proof: x and y have the same set of divisors as x and $mod(x, y)$

Euclid's Algorithm:

euclid(x, y):= if $(y = 0)$: return x.

else: return euclid$(y, mod(x, y))$.

Base: $gcd(x, 0) = x$.

Induction: $mod(x, y) < y \leq x$ when $x \geq y$. By strong induction,

euclid$(y, mod(x, y))$ = euclid(x, y).

Theorem: euclid(x,y) uses $2n$ "divisions" where $n \simeq \log_2(x)$.

→ much faster than trying $y = \{2, \cdots, \lfloor y/2 \rfloor\}$.

Proof: First arg decreases by factor of at least 2 after two recursive calls. → Case analysis.

Case1) $y < x/2$. first argument is y. → True in one recursive call.

Case2) $y \geq x/2$. $\Rightarrow$ mod(x,y) $\leq x/2$. → euclid(y, mod(x,y))

Extended Euclid's Algorithm - Finding Inverse.

For any x,y there are integers a,b such that

$ax + by = d$ where $d = gcd(x,y)$.

What is the multiplicative inverse of $x \pmod m$?

$ax + bm \equiv 1 \pmod m \Rightarrow \underline{a \equiv x^{-1} \pmod m}$.

ext-gcd(x,y): = if (y=0): return (x, 1, 0).

else: (d,a,b) := ext-gcd(y, mod(x,y)).

return (d, b, a − floor(x/y) · b).

Base: ext-gcd(x,0) returns (d=x, 1, 0) with $x = 1(x) + 0(y)$.

IH: ext-gcd(y, mod(x,y)) returns (d, a, b) with

$d \equiv ay + b \pmod{(x,y)} = ay + b(x - \lfloor \frac{x}{y} \rfloor y)$

$= bx + (a - \lfloor \frac{x}{y} \rfloor b)y \rightarrow (d, b, (a - \lfloor \frac{x}{y} \rfloor b))$.

Chinese Remainder Theorem: Find $x = a \pmod m$, $x = b \pmod n$ where $\gcd(m,n) = 1$. There is a unique solution $x \pmod{mn}$.

Proof of existence: Consider $u = n(n^{-1} \pmod m)$.

$u = 0 \pmod n$, $u = 1 \pmod m$.

Consider $v = m(m^{-1} \pmod n)$. $v = 1 \pmod n$, $v = 0 \pmod m$.

Let $x = au + bv$. $x = a \pmod m$ (b/c $bv = 0 \pmod m$, $au = a \pmod m$)

$x = b \pmod n$ (b/c $au = 0 \pmod n$, $bv = b \pmod n$). exists!

Uniqueness: Assume two solutions x and y.

$\rightarrow (x - y) \equiv 0 \pmod m$ and $(x - y) \equiv 0 \pmod n \rightarrow mn \mid (x - y)$

$\rightarrow x - y \geq mn \rightarrow x, y \notin \{0, \dots, (mn - 1)\} \rightarrow$ contradiction!

Fermat's Little Theorem: for prime p, and $a \not\equiv 0 \pmod p$,

$a^{(p-1)} \equiv 1 \pmod p$.

Proof: Consider set $S = \{a\cdot 1, \dots, a\cdot(p-1)\}$. → all distinct.

→ $(a\cdot 1)(a\cdot 2)\cdots(a\cdot(p-1)) \equiv 1\cdot 2\cdot\cdots\cdot(p-1) \pmod p$.

→ $a(1)\cdot a\cdot(2)\cdots\cdot a(p-1) \equiv a^{p-1}(1\cdot 2\cdot\cdots\cdot(p-1))$

$\equiv 1\cdot 2\cdot\cdots\cdot(p-1) \pmod p$. Since $2\cdots(p-1)$ all have inverses, $a^{p-1} \equiv 1 \pmod p$. //

Public Key Cryptography

Bijection: one-to-one onto function

$f(x) = ax \pmod m$ if $\gcd(a, m) = 1$ → bijection

bijection between $(a \bmod m, b \bmod n)$ & $x \bmod mn$ (CRT)

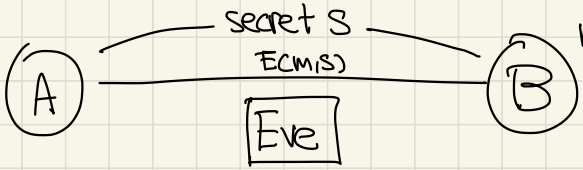
↳ in fact, closed under addition & multiplication! (Isomorphism)

XOR: Exclusive OR (1 if two numbers are different)

↳ also mod addition (mod 2)

* $A \oplus B \oplus B = A$

Cryptography



$m = D(E(m,s), S)$

ex) One-time pad: Bitwise XOR m with s with |m| bits.

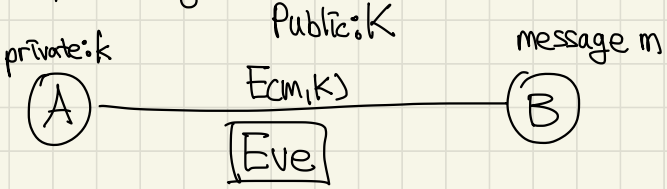
$E(m,s) = m \oplus S$, $D(x,s) = x \oplus S$ → works b/c $m \oplus S \oplus S = m$.

→ Secure b/c X can be any message for some S!

Disadvantages: secret shared, uses up one-time pad.

→ less and less secure each time used...

Public Key:



$m = D(E(m,K), k)$

Everyone knows K, everyone can encode, but only Alice can decode with k! → Ideal PKC.

Realistically, making brute-force extremely slow works.

RSA

Pick two large primes p, q. $N := p \cdot q$.

Choose e relatively prime to $(p-1)(q-1)$.

Compute $d = e^{-1} \pmod{(p-1)(q-1)}$. → Private key

Announce $N (= p \cdot q)$ and $e \rightarrow K(N, e)$ is pubic key.

Encoding: $\text{mod}(X^e, N)$. Decoding: $\text{mod}(y^d, N)$.

Does $D(E(m)) = m^{ed} = m \pmod{N}$??

ex) $p = 7, q = 11. \rightarrow N = 77. (p-1)(q-1) = 60.$

choose $e = 7.$ ($\gcd(60, 7) = 1$) $\rightarrow d = e^{-1} = -17 = 43 \pmod{60}$.

Public Key: $(77, 7)$, message choices: $\{0, \ldots, 76\}$.

message: 2. $\rightarrow E(2) = 2^e = 2^7 = 128 = \underline{51} \pmod{77}$.

$D(51) = 51^{43} \pmod{77} \rightarrow$ ... not a good idea.

Repeated Squares: $43 = 32 + 8 + 2 + 1 \rightarrow 101011$ in binary.

$51^{(32+8+2+1)} = 51^{32} \cdot 51^8 \cdot 51^2 \cdot 51^1 \pmod{77}$

$51^1 = 51 \pmod{77}$. $51^2 \equiv 60 \pmod{77}$. $(51^2)^2 \equiv 60^2 \equiv 58 \pmod{77}$

→ only $\lceil \log_2 43 \rceil$ computations, then multiply everything.

(n) (n^2) $\Rightarrow O(n^3)$.

Correctness: want $(m^e)^d = m^{ed} = m \pmod N$. always.

$d = e^{-1} \pmod{(p-1)(q-1)} \iff ed = k(p-1)(q-1) + 1$

Then, consider FLT: $a^{p-1} \equiv 1 \pmod p$ if $p \nmid a$.

$\rightarrow a^{k(p-1)} \equiv 1 \pmod p$, $a^{k(p-1)+1} \equiv a^1 \cdot (a^{(p-1)})^k \equiv a \pmod p$!

Does this imply ... $a^{k(p-1)(q-1)+1} \equiv a \pmod{pq}$?

$\rightarrow a^{k(p-1)(q-1)+1} \equiv a \pmod p \equiv a \pmod q$.

$\rightarrow (a^{k(p-1)(q-1)+1} - a)$ is a multiple of p and $q \rightarrow \equiv a \pmod{pq}$!

$\Rightarrow D(E(x)) = x^{ed} \equiv x \pmod{pq}$

$(x^{ed} = x^{k(p-1)(q-1)+1} \equiv x \pmod{pq})$

Key construction:

Prime Number Theorem: if $\pi(N)$ is the # of primes $\leq N$, for all $N \geq 17$, $\pi(N) \geq N/\ln(N)$.

$\rightarrow$ a random # has $\simeq \frac{1}{\ln(N)}$ chance to be prime.

Then, choose e, compute d. $\Rightarrow$ all logarithmic time!

Secret Sharing

Share secret among n people.

Any $(k-1)$ ppl knows nothing. Any k ppl knows secret.

Polynomials

$P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0 \Rightarrow \{a_d, a_{d-1}, \cdots, a_0\}$

$P(x)$ contains (a, b) if $P(a) = b$.

Polynomials with modulo p: $a_i \in \{0, \cdots, p-1\}$,

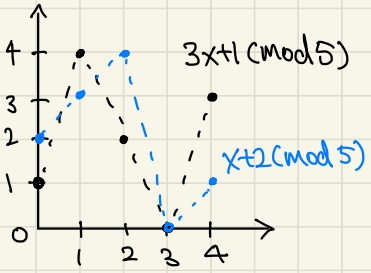
$P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0 \pmod{p}$

for $x \in \{0, \cdots, p-1\}$.

Line: $P(x) = a_1 x + a_0 = mx + b$

Parabola: $P(x) = a_2 x^2 + a_1 x + a_0 = ax^2 + bx + c$

$P(x) = 3x + 1 \pmod 5$, $P(x) = x + 2 \pmod 5$



$\rightarrow 3x + 1 \equiv x + 2 \pmod 5$

$\rightarrow 2x \equiv 1 \pmod 5 \rightarrow x \equiv 3 \pmod 5$ ($2^{-1} \pmod 5$)

↳ Found an intersection!

Fact: Exactly 1 (degree $\le d$) polynomial contains $d+1$ points.

→ 2 points define a line. 3 points define a parabola.

→ Also works with modulo prime p!

Secret $s \in \{0, \cdots, p-1\}$. → "k out of n" scheme:

① Choose $a_0 = s$ and random $a_1, \cdots, a_{k-1}$.

② $P(x) = a_{k-1}x^{k-1} + a_{k-2}x^{k-2} + \cdots + a_0$.

③ Share $(i, P(i) \pmod p)$ to i-th person $(0 \sim n)$

→ Robustness: any k shares gives secret.

know k points → unique $P(x)$ → find $P(0)$

→ Secrecy: $\le k-1$ points give no information.

Given points $(x_1, y_1), (x_2, y_2), \cdots, (x_k, y_k)$

→ set up system of equations with $\{a_0, \cdots, a_{k-1}\} \pmod p$

Interpolation: $a_2x^2+a_1x+a_0$ hits $(1,2),(2,4),(3,0)$.

- Find $\Delta_1(x)$ containing $(1,1),(2,0),(3,0)$.

→ Try $(x-2)(x-3) \pmod 5$. → value 0 at 2 and 3, 2 at 1.

→ "Divide" by 2 (or multiply by 3) → $\Delta_1(x)=3(x-2)(x-3) \pmod 5$

- $\Delta_2(x)=4(x-1)(x-3) \pmod 5$, contains $(1,0),(2,1),(3,0)$.

- $\Delta_3(x)=3(x-1)(x-2) \pmod 5$, contains $(1,0),(2,0),(3,1)$.

→ $P(x)=2\cdot\Delta_1(x)+4\cdot\Delta_2(x)+0\cdot\Delta_3(x)$.

→ $P(x)=2x^2+x+4 \pmod 5$.

Fields: $\oplus, \otimes$, $\oplus$ identity $\otimes$ identity. $\otimes$ inverse except identity

ex) reals, rationals, complex numbers (o) integers (x)

modulo p → addition, multiplication, inverses ✓

Delta polynomial: $\Delta_i(x)=\begin{cases}1; & x=x_i\\ 0; & x=x_j, i\neq j\\ ?; & \text{otherwise.}\end{cases}$

→ $y_1\Delta_1(x)$ contains (x_1,y_1). $y_2\Delta_2(x)$ contains (x_2,y_2).

⇒ $y_1\Delta_1(x)+y_2\Delta_2(x)$ contains both (x_1,y_1) and (x_2,y_2)!

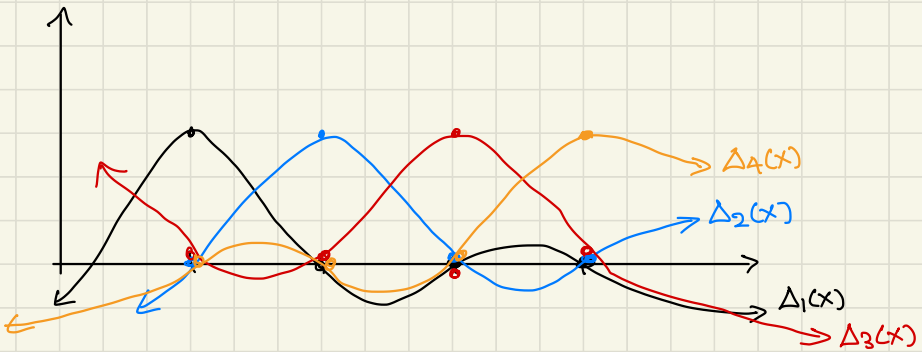
→ extend to any length of polynomials

$P(x)=y_1\Delta_1(x)+y_2\Delta_2(x)+\cdots+y_n\Delta_n(x)$.

$$\Delta_i(x) = \frac{\prod_{j\neq i}(x-x_j)}{\prod_{j\neq i}(x_i-x_j)} = \prod_{i\neq j}(x-x_j)\cdot\prod(x_i-x_j)^{-1}$$ ✶

Numerator is 0 at $x_j \neq x_i$. "Denominator" makes $\Delta_i(x)=1$ at $x_j = x_i$.

$\rightarrow P(x) = y_1\Delta_1(x) + \cdots + y_{d+1}\Delta_{d+1}(x) \Rightarrow$ degree d polynomial!



Uniqueness: degree d polynomial has at most d roots.

Assume 2 solutions $P(x), Q(x)$. $R(x)=P(x)-Q(x)$ has $d+1$ roots with d degrees $\rightarrow$ Contradiction!

$P(x)$ has root a iff $\frac{P(x)}{(x-a)}$ has remainder 0 $\rightarrow P(x)=(x-a)Q(x)$.

P has d roots, then $P(x)=C(x-r_1)(x-r_2)\cdots(x-r_d)$.

# of degree d polynomials over $GF(m)$? $\rightarrow m^{(d+1)}$

Erasure Codes: Some information gets lost in transmission.

n-packet message, channel that loses k packets

→ must send n+k packets! (any n packets constructs n messages)

① Choose prime $p \geq 2^b$ for packet size b.

② $P(x) = m_{n-1}x^{n-1} + m_{n-2}x^{n-2} + \cdots + m_0 \pmod{p}$ ($p \geq n+k$)

③ Send $P(1), P(2), \ldots, P(n+k)$.

→ any n of n+k gives polynomial $P(x)$ and the message!

ex) Make $P(x)$ s.t. $P(1)=1, P(2)=4, P(3)=4 \pmod 7$

$$P(1) = a_2 + a_1 + a_0 = 1 \pmod 7$$
$$P(2) = 4a_2 + 2a_1 + a_0 = 4 \pmod 7$$
$$P(3) = 2a_2 + 3a_1 + a_0 = 4 \pmod 7$$

$$\Rightarrow P(x) = 2x^2 + 4x + 2$$

→ Send $(i, P(i))$: $(1,1), (2,4), (3,4), (4,0), (5,2), (6,0)$

→ Recieved: $(1,1), (2,4), (6,0)$ → Reconstruct $P(x)$ with 3 points.

Corrupt Codes: k packets are corrupted (altered) in transmission.

→ Reed-Solomon Code:

① Make $P(x)$ of degree $n-1$, $P(1)=m_1, \ldots, P(n)=m_n$.

② Send $P(1), \ldots, P(n+2k)$.

→ Recieve values $R(1), \ldots, R(n+2k)$.

$P(i)=R(i)$ for at least $n+k$ points i, and $P(x)$ is a unique degree $n-1$ polynomial that contains $\geq n+k$ points.

ex) $3,0,6 \rightarrow P(x)=x^2+x+1 \pmod 7$, $P(1)=3$, $P(2)=0$, $P(3)=6$.

$P(4)=0$, $P(5)=3$. → Recieved $R(2)=1$, other points same.

$P(i)=R(i)$ for $n+k=3+1=4$ points. Brute Force?

If $Q(x)$ constructed by n points satisfy k more points, $P(x)=Q(x)$!

... But too expensive (exponential runtime)

Berlekamp-Welsh: Multiply both sides by 0 if $P(i) \neq R(i)$.

$E(x)$:= "error polynomial", $(x-e_1)(x-e_2)\cdots(x-e_k) = x^k + \cdots + b_0$

$Q(x) = E(x)P(x) = E(x)R(x)$ for every point!

Q: $d:(n+k-1)$, E: $d:k$, P: $d:(n-1)$, E: $d:k$, R: $d:0$

The first coefficient of E is always 1 → k unknowns

Q has degree of $(n+k-1)$ → $(n+k)$ unknowns

→ knowing $(n+2k)$ points reveals coefficients of $Q(x)$ and $E(x)$!

⇒ $P(x) = \frac{Q(x)}{E(x)}$ (long division) → $\{P(0), P(1), \cdots, P(n)\}$

Uniqueness: If $\exists\, Q'(x), E'(x)$, then $\frac{Q'(x)}{E'(x)} = \frac{Q(x)}{E(x)} = P(x)$.

→ $Q'(x)E(x) = Q(x)E'(x)$ → degree $(n+2k-1)$

They also agree on $n+2k$ points, while $E(x)$ & $E'(x)$ have at most k roots. → $\frac{Q'(x)}{E'(x)}$ and $\frac{Q(x)}{E(x)}$ agrees on n points → equal!

Counting

First Rule of Counting (Product Rule): Objects made by choosing from n_1, then n_2, then $\dots$ n_k, the number of objects is $n_1 \times n_2 \times \dots \times n_k$.

Function Mapping: $f: S \to T$ has $|T|^{|S|}$ mappings (in general).

Polynomials of degree d modulo p? $\to p^{d+1}$

Permutations: sampling size k from n numbers $\to \frac{n!}{(n-k)!}$ choices

How many one-to-one functions $f: S \to S$? $\to |S|!$ mappings

Combinations: When order doesn't matter $\to \frac{n!}{(n-k)!\,k!}$ a.k.a. $\binom{n}{k}$

(Find permutations, then divide by any duplicate instances of k)

Second Rule of Counting: When order doesn't matter, count the number of ordered objects then divide by number of orderings.

ex) Anagram of ANAGRAM? $\to$ Ordered set: $7!$ but...

A's are all identical $\to$ 3 As $\to$ $3!$ duplicate countings $\to$ $\frac{7!}{3!}$

	With replacement	W/o replacement
Order matters	n^k	$\frac{n!}{(n-k)!}$
Order does not matter	$\binom{n+k-1}{n}$	$\binom{n}{k}$

Sum Rule: Can sum over disjoint sets.

ex) 2 indistinguishable jokers in 54 card deck.

How many 5 card hands?

$$\binom{52}{5} + \binom{52}{4} + \binom{52}{3}$$

↑ no joker ↑ one joker ↑ two jokers

Distinguishable jokers: $\binom{52}{5} + 2\binom{52}{4} + \binom{52}{3} = \binom{54}{5}$

Combinatorial Proofs: examples.

$\binom{n}{k} = \binom{n-1}{k-1} + \cdots + \binom{k-1}{k-1}$ → consider subset where i-th element is first chosen

$[1 \cdots i \underbrace{\cdots n}_{\geq k}]$ → then choose $(k-1)$ elements from $(n-i)$

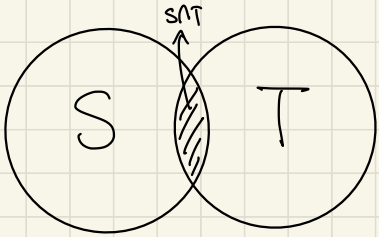
→ $\binom{n-i}{k-1}$ such subsets, $i \in [1, n-k+1]$ → $\sum_{i=1}^{(n-k+1)} \binom{n-i}{k-1} = \binom{n-1}{k-1} + \cdots + \binom{k-1}{k-1}$

$2^n = \binom{n}{n} + \binom{n}{n-1} + \cdots + \binom{n}{0}$. → consider subset of set S, $|S| = n$.

For each element, there are 2 possibilities: include/exclude.

→ 2^n subsets. Or, count subsets of size i → $\binom{n}{i}$ → $\sum_{i=0}^{n} \binom{n}{i} = 2^n$.

Inclusion/Exclusion: $|S \cup T| = |S| + |T| - |S \cap T|$ ✡



Countability

S is countable if there is a bijection between S and some subset of $\mathbb{N}$. (Enumerable $\Rightarrow$ Countable) ★

ex) $\mathbb{Z}^+$ is countable. $f: \mathbb{N} \to \mathbb{Z}^+$, $f(x) = x - 1$

ex) All binary strings $B = \{0,1\}^*$. $B = \{\phi, 0, 1, 00, 01, 10, 11, 000, \ldots\}$

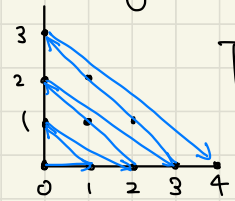
If n bits, it appears before position $2^{(n+1)}$.

ex) All rational numbers $\mathbb{Q}$? Any two fractions has another in btwn.

→ can't list in "order". Consider domain $(\mathbb{N} \times \mathbb{N})$, size $(\infty)^2$??

→ enumerate in such a way: (0,0), (1,0), (0,1), (2,0), (1,1), (0,2)...

→ This gives all points in the 1st quadrant.



The pair (a,b) is in first $\simeq \frac{(a+b+1)(a+b)}{2}$ element of the list (triangle area) $\Rightarrow \mathbb{N}$

$\frac{a}{b} \in \mathbb{Q}$, $a, b \in \mathbb{N}$, $\gcd(a,b) = 1$. → (a,b) is a subset of $\mathbb{N} \times \mathbb{N}$.

$\Rightarrow$ Rationals are countably infinite! (alternate positive and negative)

Real numbers... $|\mathbb{R}| = |\mathbb{N}|$??

Consider the reals in $[0,1]$. Each has a decimal representation.

→ Can we enumerate all reals in this range?

Diagonalization: If countable, $\exists$ listing L, $\forall x \in \mathbb{R}, x \in L$. However...

L							
0	**5**	4	7	8	4	1	...
1	2	**8**	9	6	3	5	...
2	3	1	**1**	8	5	9	...
3	1	3	7	**4**	9	0	...
4	0	2	4	7	**5**	6	...

→ $d = 0.69856...$

Construct a "diagonal number" such that i-th digit is different from the i-th digit of the i-th number → guaranteed to be different from every number in L.

→ L cannot exist due to contradiction ⇒ $\mathbb{R}$ is uncountably infinite! //

Set of all subsets of $\mathbb{N}$? Assume countable, construct listing L.

Define a diagonal set, D: if i-th set in L does not contain i, $i \in D$.

D is different from every set in L. → $D \notin L$, $D \in \mathbb{N}$.

→ L does not contain all elements in $\mathbb{N}$. → Contradiction.

⇒ Power set of $\mathbb{N}$ is uncountably infinite! ($|\mathbb{N}_0| < |2^{\mathbb{N}_0}|$)

$|[0,1]|$ vs $|\mathbb{R}|$? $f: \mathbb{R}^+ \to [0,1]$. $f(x) = \begin{cases} x + \frac{1}{2} & (0 \le x \le \frac{1}{2}) \\ \frac{1}{4x} & (\frac{1}{2} < x) \end{cases}$

one-to-one: $x \neq y \Rightarrow f(x) \neq f(y)$ → Bijection. → $|[0,1]| = |\mathbb{R}|$.

Continuum Hypothesis: no infinite set S s.t. $|\mathbb{N}_0| < |S| < |2^{\mathbb{N}_0}|$.

Halting Machine: HALT(P, I) := outputs whether P(I) halts

(P is a text string, which can be used as an input to HALT)

How do we implement HALT? wait until it stops?

→ HALT does not exist. Consider diagonalization.

Proof: Assume ∃ HALT(P, I). → ∃ a text string HALT.

define Turing(P) :=

if (HALT(P, P) == "halts"): loop infinitely.

else: halt immediately. → ∃ a text string Turing.

→ Turing(Turing)?

if Turing(Turing) halts → HALT(Turing, Turing) halts.

→ Turing(Turing) loops indefinitely.

if Turing(Turing) loops indefinitely → Turing(Turing) halts!

⇒ HALT cannot exist (if it did, Turing can exist, which is false.)

Another view: Any program if a fixed length string. → enumerable.

HALT	P_1	P_2	P_3	...
P_1	(H)	L	H	
P_2	L	(L)	H	...
P_3	H	H	(H)	
⋮				

⋮ Turing(P_1) → L

Turing(P_2) → H ...

HALT - diagonal. Turing - is not HALT.

→ Turing is not on list L. Turing is not a program. → HALT cannot exist!

Probability

Random Experiment: One fair coin $\to \Omega = \{H, T\}$, $Pr[H] = Pr[T] = \frac{1}{2}$.

An unfair coin: $Pr[H] = p \in (0,1)$, $Pr[T] = 1-p$

Two fair coins $\to \Omega = \{HH, HT, TH, TT\} = \{H, T\}^2$

↳ Likelihood: $\frac{1}{4}$ for each possibility.

Ω is the set of possible outcomes. *

Each outcome ω has a non-negative probability and add up to 1. *

$\to 0 \le Pr[\omega] \le 1$, $\sum_{\omega \in \Omega} Pr[\omega] = 1$.

Flipping a fair coin n times: $\Omega = \{H, T\}^n$, $|\Omega| = 2^n$.

↳ Likelihood: $\frac{1}{2^n}$ each

Uniform Probability Space: $Pr[\omega] = \frac{1}{|\Omega|}$ for all $\omega \in \Omega$.

Probability of exactly one head in two coin flips?

→ Sum the $Pr[\omega]$ of all ω with one head → $\{HT, TH\}$

An event E is a subset of Ω; $E \subset \Omega$

Probability of E: $Pr[E] = \sum_{\omega \in E} Pr[\omega]$ **

Stirling Formula: $n! \simeq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, $\binom{2n}{n} \simeq \frac{4^n}{\sqrt{\pi n}}$

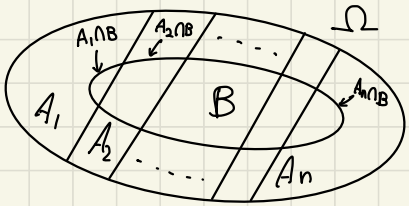
Additivity: if disjoint ($A \cap B = \emptyset$), $Pr[A \cup B] = Pr[A] + Pr[B]$

if $(A_1, \dots, A_n)$ are pairwise disjoint, $Pr[A_1 \cup \dots \cup A_n] = \sum_{i=1}^{n} Pr[A_i]$

$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$ (P.I.E.)

$Pr[A_1 \cup \dots \cup A_n] \leq \sum_{i=1}^{n} Pr[A_i]$ (Union Bound)

If $(A_1, \dots, A_n)$ is a partition of Ω, $Pr[B] = \sum_{i=1}^{n} Pr[B \cap A_i]$



(Law of Total Probability)

All of $\omega \in B$ is in exactly one of the sets.

Conditional Probability: "B given A", $Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$ ★

ex) Toss 3 balls into 3 bins → $\Omega = \{1,2,3\}^3$, $\omega = (\text{bin of ball 1}, \dots)$

A = "1st bin empty" B = "2nd bin empty" $Pr[A|B]$?

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{1/27}{8/27} = 1/8$

prior choice probability, conditional probability

Product Rule: $Pr[B|A] = \frac{Pr[B \cap A]}{Pr[A]} \Rightarrow Pr[A] \cdot Pr[B|A] = Pr[B \cap A]$ ★

$Pr[A \cap B \cap C] = Pr[(A \cap B) \cap C] = Pr[A \cap B] \cdot Pr[C | A \cap B]$

$Pr[\bigcap_{i=1}^{n} A_i] = \prod_{i=1}^{n} Pr[A_i | \bigcap_{j=1}^{i-1} A_j]$

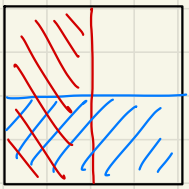
Correlation: $Pr[A \cap B] > Pr[A] \cdot Pr[B]$, NOT causality

Bayes' Rule: $Pr[A \cap B] = Pr[A|B]Pr[B] = Pr[B|A]Pr[A]$

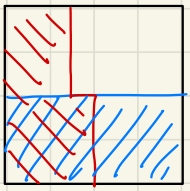
$\Rightarrow Pr[A|B] = \frac{Pr[B|A]Pr[A]}{Pr[B]}$ ★★

Independence: $Pr[A \cap B] = Pr[A] \cdot Pr[B]$, $Pr[A|B] = Pr[A]$

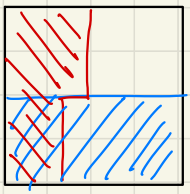
→ Knowing B doesn't reveal any information about A!



independent



positive corr.

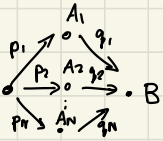


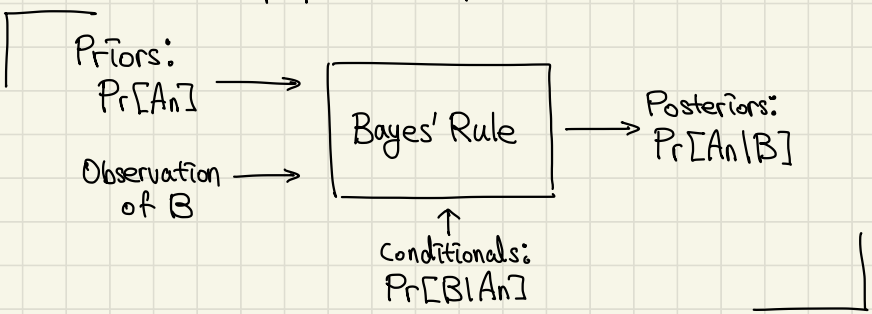
negative corr.

Bayes' General Case: $p_n = Pr[A_n]$, $q_n = [B|A_n]$, A is a partition

$Pr[A_n \cap B] = p_n q_n$, $Pr[B] = p_1 q_1 + \cdots + p_N q_N$

$$\Rightarrow Pr[A_n|B] = \frac{p_n q_n}{\sum_{m=1}^{N} p_m q_m} = \frac{Pr[B|A_n] \cdot Pr[A_n]}{Pr[B]} \quad \star$$





Pairwise Independence: $A = (H, *)$ $B = (*, H)$ $C = \{(H,T), (T,H)\}$

A, C are independent; B, C are independent; $A \cap B$, C are <u>not</u> independent

Mutual Independence: Events $A_1, \ldots, A_N$ are mutually ind. if $\star$

$$Pr[\cap_{k \in K} A_k] = \prod_{k \in K} Pr[A_k] \text{ for all } K \subseteq \{1, \ldots, N\}$$

Random Variables

Random Variable: X assigns a real number $X(\omega)$ for all $\omega \in \Omega$.

(random variable induces a partition; $A_y = \{\omega \in \Omega \mid X(\omega) = y\} = X^{-1}(y)$)

Distribution: $\{(a, Pr[X=a]) \mid a \in A\}$, A is the range of X

Expectation: $E[X] = \sum_a a \cdot Pr[X=a] = \sum_\omega X(\omega) \cdot Pr[\omega]$

Binomial Distribution: number of heads in n flips $\rightarrow Pr[X=i]$

i heads out of n flips $\rightarrow \binom{n}{i}$ outcomes $\rightarrow Pr[X=i] = \binom{n}{i} (p)^i (1-p)^{n-i} \rightarrow B(n,p)$

$\hookrightarrow E[X] = n \cdot p$

ex) corruption possibility of p, sending $n+2k$ packets.

$Pr[\text{at most } k \text{ corruptions}] = \sum_{i \le k} \binom{n+2k}{i} p^i (1-p)^{n+2k-i}$.

Geometric Distribution: $Pr[X=n] = (1-p)^{n-1} p$, $n \ge 1$ (flip until heads)

$\rightarrow \sum_{n=1}^{\infty} Pr[X=n] = \sum_{n=1}^{\infty} (1-p)^{n-1} p = p \sum_{n=1}^{\infty} (1-p)^{n-1} = p \sum_{n=0}^{\infty} (1-p)^n = p \cdot \frac{1}{1-(1-p)} = \frac{p}{p} = 1.$

$E[X] = \sum_{n=1}^{\infty} n \, Pr[X=n] = \sum_{n=1}^{\infty} n (1-p)^{n-1} p.$

$E[X] = p + 2(1-p)p + 3(1-p)^2 p \cdots$

$(1-p)E[X] = \quad (1-p)p + 2(1-p)^2 p \cdots$

$\rightarrow pE[X] = \sum_{n=1}^{\infty} (1-p)^n p = 1$

$\rightarrow \underline{E[X] = \frac{1}{p}}$

Poisson Distribution: How many arrive at McDonalds in an hour?

Average is λ. $\rightarrow$ Distribution? (Assume arrivals are independent)

Cut an hour into n intervals of length $\frac{1}{n}$.

$\Pr[\text{Two arrivals}] = (\frac{\lambda}{n})^2 \rightarrow$ small if n is large $\rightarrow$ binomial

Flip a coin n times. $\Pr[H] = \frac{\lambda}{n}$. $\rightarrow X \sim Bi(n, \frac{\lambda}{n})$.

Poisson Distribution is X as $n \rightarrow \infty$.

$$\Pr[X=m] = \binom{n}{m} p^m (1-p)^{n-m} = \frac{n!}{m!(n-m)!} \left(\frac{\lambda}{n}\right)^m \left(1-\frac{\lambda}{n}\right)^{n-m}$$

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

$$\lim_{n\to\infty} \frac{\lambda^m}{m!}\left(1-\frac{\lambda}{n}\right)^n = \frac{\lambda^m}{m!} e^{-\lambda}$$

$$E[X] = \sum_{m=1}^{\infty} m \times \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} = e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^{(m+1)}}{m!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!}$$

$$= e^{-\lambda} \cdot \lambda \cdot e^{\lambda} \text{ (Taylor expansion)} = \lambda.$$

Joint Distribution: $\{(a,b,\Pr[X=a,Y=b] \mid a\in A, b\in B\}$ where A and B are possible values of X and Y, respectively.

$$\sum_{a\in A,\, b\in B} \Pr[X=a,Y=b] = 1$$

Marginal for X: $\Pr[X=a] = \sum_{b\in B} \Pr[X=a,Y=b]$

Conditional: $\Pr[X=a \mid Y=b] = \frac{\Pr[X=a,Y=b]}{\Pr[Y=b]}$

Independent Random Variables: $\Pr[Y=b,X=a] = \Pr[Y=b]$ for all a,b.

Also, $\Pr[X=a,Y=b] = \Pr[X=a]\cdot\Pr[Y=b]$

Linearity of Expectation: $E[X]+E[Y] = E[X+Y]$, $E[cX] = cE[X]$

Proof: $E[X] = \sum_{\omega\in\Omega} X(\omega)\Pr[\omega]$.

$$\underline{E[X+Y]} = \sum_{\omega\in\Omega}(X(\omega)+Y(\omega))\Pr[\omega]$$
$$= \sum_{\omega\in\Omega} X(\omega)\Pr[\omega] + \sum_{\omega\in\Omega} Y(\omega)\Pr[\omega] = \underline{E[X]+E[Y]} \;//$$

Indicators: Event A, $X(\omega) = \begin{cases} 1 & \text{if } \omega\in A \\ 0 & \text{if } \omega\notin A \end{cases}$ is the indicator of A.

$\Pr[\omega=1] = \Pr[A]$, $\Pr[X=0] = \Pr[\bar{A}]$. $\rightarrow E[X] = \Pr[A]$.

(also denoted as $1\{\omega\in A\}$ or $1_A(\omega)$, $X = 1_A$)

ex) Roll a die n times. $X_m := $ # of pips on roll m.

$X = X_1 + \cdots + X_n = $ (total # of pips in n rolls)

$E[X] = E[X_1] + \cdots + E[X_n] = n \cdot E[X_1] = \underline{n \cdot \frac{7}{2}}$

ex) Hand out assignment to n students. $X = $ # of correct matches

$Pr[X=m]$? $X = X_1 + \cdots + X_n$ where $X_m = 1$ {student m gets their own HW back}

$Pr[X_i = 1] = \frac{1}{n}$. $E[X] = E[X_1 + \cdots + X_n] = n \cdot E[X_1] = n \cdot \frac{1}{n} = 1$.

Coupon Collector's Problem: get random coupon from n until collected all

Outcome := sequence of numbers $[1, n]$ $X :=$ length of outcome

$X = {}_0G(p) \to Pr[X=n] = (1-p)^{n-1}(p)$, $E[X] = \frac{1}{p}$.

$X_1 :=$ time to get 1 coupon $= 1$. $E[X_1] = 1$.

$X_2 :=$ time to get 2 different coupons. $Pr[\text{new coupon} \mid \text{one coupon}] = \frac{n-1}{n}$

$E[X_2] = \frac{1}{p} = \frac{n}{n-1}$. $\to E[X_i] = \frac{n}{n-i+1}$

$\sum_{i=1}^{n} E[X_i] = \frac{n}{n} + \cdots + \frac{n}{1} = n\left(\frac{1}{n} + \cdots + 1\right) \simeq \overbrace{n(\ln(n) + \gamma)}^{= H(n)}$

Let $Y = g(X)$. Assume distribution of X is known. $E[Y]$?

$\Pr[Y=y] = \Pr[X \in g^{-1}(Y)]$ ($g^{-1}(x) = \{x \in \mathbb{R} \mid g(x)=y\}$) → Hard

$E[g(X)] = \sum_x g(x)\Pr[X=x]$ ★ (Law of the Unconscious Statistician)

↳ $E[g(X)] = \sum_\omega g(X(\omega))\Pr[\omega] = \sum_x \sum_{\omega \in X^{-1}(x)} g(X(\omega))\Pr[\omega]$

$= \sum_x \sum_{\omega \in X^{-1}(x)} g(x)\Pr[\omega] = \sum_x g(x) \sum_{\omega \in X^{-1}(x)} \Pr[\omega] = \sum_x g(x)\Pr[X=x]$

Geometric Distribution's Memoryless Property: $X \sim Geo(p)$.

$\Pr[X > m+n \mid X > n] = \Pr[X > m]$ $(m, n > 0)$

→ Intuitively, n coin flips before starting m flips should not affect it.

$\Pr[X > n] = (1-p)^n$ → $\Pr[X > m+n \mid X > n] = \frac{\Pr[X > m+n \text{ and } X > n]}{\Pr[X > n]}$

→ $\Pr[X > m+n \mid X > n] = \frac{(1-p)^{m+n}}{(1-p)^n} = (1-p)^m = \Pr[X > m]$

Variance

Variance: $\sigma^2(X) = \mathrm{var}[X] = E[(X - E[X])^2]$ ★ ($\sigma(X)$: Std.Dev)

$\mathrm{Var}[X] = E[(X-E[X])^2] = E[X^2 - 2XE[X] + E[X]^2]$

$= E[X^2] - 2E[X]E[X] + E[X]^2 = E[X^2] - E[X]^2$ ★

ex) $\Pr[X=i] = \frac{1}{n}$ for $i = \{1, \dots, n\} \to E[X] = \frac{n+1}{2}$.

$E[X^2] = \sum_{i=1}^{n} i^2 \Pr[X=i] = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{1+3n+2n^2}{6}$

$\to \mathrm{Var}(X) = E[X^2] - E[X]^2 = \frac{1+3n+2n^2}{6} - \frac{n^2+2n+1}{4} = \frac{n^2-1}{12}$.

ex) $X \sim \mathrm{Geo}(p) \to \Pr[X=n] = (1-p)^{n-1}p$, $E[X] = \frac{1}{p}$

$E[X^2] = p + 4p(1-p) + 9p(1-p)^2 + \cdots$

$-(1-p)E[X^2] = -[p(1-p) + 4p(1-p)^2 + \cdots]$

$\to pE[X^2] = p + 3p(1-p) + 5p(1-p)^2 + \cdots$

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

$-(p + p(1-p) + p(1-p)^2 + \cdots)$

$= 2E[X] - 1 \to E[X^2] = \frac{2-p}{p^2}$

$\mathrm{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$.

ex) "# of students that get their homework back" from n students

$X = X_1 + X_2 + \cdots + X_n$ (X_i is indicator of i-th student)

$E[X^2] = \sum_i E[X_i^2] + \sum_{i \neq j} E[X_i X_j]$.

$E[X_i^2] = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0] = Pr[X_i = 1] = \frac{1}{n}$

$E[X_i X_j] = 1 \times Pr[X_i = 1 \wedge X_j = 1] + 0 \times Pr[\text{else}] = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$

$\rightarrow E[X^2] = n \cdot \frac{1}{n} + (n)(n-1)\frac{1}{n(n-1)} = 2$.

$\rightarrow Var(X) = E[X^2] - E[X]^2 = 2 - 1 = 1$.

Properties of variance:

① $Var(cX) = c^2 Var(X)$ ② $Var(X+c) = Var(X)$

If $Pr[X=a, Y=b] = Pr[X=a] Pr[Y=b]$ (independent), ✱

$\underline{E[X \cdot Y] = E[X] \cdot E[Y]}$, $\underline{Var(X+Y) = Var(X) + Var(Y)}$

Proof: Assume $E[X] = E[Y] = 0$. $\rightarrow E[XY] = E[X]E[Y] = 0$.

$\rightarrow Var(X+Y) = E[(X+Y)^2] = E[X^2] + \cancel{2E[XY]}^{\,0} + E[Y^2]$

$= E[X^2] + E[Y^2] = Var(X) + Var(Y)$

ex) Variance of binomial distribution $X \sim Bi(p)$

$X_i = \begin{cases} 1 & \text{if } i\text{-th flip is heads} \\ 0 & \text{otherwise} \end{cases} \rightarrow E[X^2] = p, Var(X_i) = p - E[X]^2$

X_i and X_j are independent. $= p - p^2 = p(1-p)$

$\rightarrow Var(X) = Var(X_1 + \cdots + X_n) = np(1-p)$

ex) $X \sim Poi(\lambda) \rightarrow Pr[X=m] = \frac{\lambda^m}{m!} e^{-\lambda}$

$X \sim Poi(\lambda) = \lim_{n\to\infty} Bi(\frac{\lambda}{n}) \rightarrow pn = \lambda \rightarrow Var(X) = \lambda(1-\frac{\lambda}{n})$

$\rightarrow Var(X) = \lim_{n\to\infty} \lambda - \frac{\lambda^2}{n} = \lambda.$

$E[X^2] = Var(X) + E[X]^2 = \lambda + \lambda^2.$

Covariance: $cov(X,Y) := E[(X-E[X])(Y-E[Y])]$ ✶

$Cov(X,Y) = E[XY] - E[X]E[Y]$

Correlation: $Cor(X,Y) := \frac{cov(X,Y)}{\sigma(X)\sigma(Y)}$ ✶

Theorem: $-1 \le Cor(X,Y) \le 1.$

$cov(a+bX, c+dY) = bd\, cov(X,Y)$

$Cov(A+B, C+D) = cov(A,C) + cov(A,D) + cov(B,C) + cov(B,D)$

$X \perp Y \Rightarrow cov(X,Y) = 0$, but not the converse!

Inequalities

Markov's Inequality: Assume $f: \mathbb{R} \to [0, \infty)$ is non decreasing.

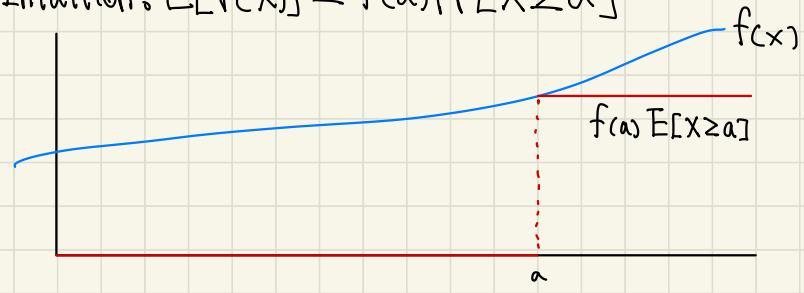
Then, $\Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)}$ for all a s.t. $f(a) > 0$. *

Proof: $1\{X \geq a\} \leq \frac{f(X)}{f(a)}$. If $X < a$, $0 \leq \frac{f(X)}{f(a)}$.

If $X \geq a$, $1 \leq \frac{f(X)}{f(a)}$ since f is non decreasing.

$\to E[1\{X \geq a\}] \leq \frac{E[f(X)]}{E[f(a)]} \to \Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)}$

Intuition: $E[f(X)] \geq f(a) \Pr[X \geq a]$



ex) $X \sim Geo(p)$. $f(x) = x \to \Pr[X \geq a] = \frac{E[X]}{a} = \frac{1}{ap}$.

$f = x^2 \to \Pr[X \geq a] = \frac{E[X^2]}{a^2} = \frac{2-p}{a^2 p^2}$, both are larger than $(1-p)^{a-1}$.

Chebyshev's Inequality: $\Pr[|X - E[X]| > a] \leq \frac{Var(X)}{a^2}$ *

$\to$ Let $Y := |X - E[X]|$, $f(y) = y^2$. $\to \Pr[Y \geq a] = \frac{E[f(Y)]}{f(a)} = \frac{Var(X)}{a^2}$.

ex) How likely is that the fraction of heads differ from 50%?

Define $Y_n = \frac{X_1 + \cdots + X_n}{n} \to Pr[|Y_n - 0.5| \geq 0.1]$

$\to Pr[|Y_n - 0.5| \geq 0.1] \leq \frac{Var(Y_n)}{0.1^2} = 100 Var(Y_n)$

$= \frac{100}{n} \cdot Var(X_1) \leq \frac{100}{4n}$ $(Var(X_1) = p(1-p) \leq \frac{1}{4}) = \frac{25}{n}$

As $n \to \infty$, $Pr[|Y_n - 0.5| \geq 0.1] \to 0$

Weak Law of Large Numbers: $Pr[|\frac{X_1 + \cdots + X_n}{n} - \mu| \geq \varepsilon] \to 0$ as $n \to \infty$

$Pr[|\frac{X_1 + \cdots + X_n}{n} - \mu| \geq \varepsilon] \leq \frac{Var(\frac{X_1 + \cdots + X_n}{n})}{\varepsilon^2} = \frac{n Var(X_1)}{n^2 \varepsilon^2} = \frac{Var(X_1)}{n \varepsilon_0} \to 0$ as $n \to \infty$

Confidence Interval: An interval $[a, b]$ is a 95% confidence interval for an unknown quantity θ if $Pr[\theta \in [a, b]] \geq 95\%$.

Interval $[a, b]$ is calculated via observations. $(a = a(X_1, \ldots, X_n))$

$|A_n - p| \leq \varepsilon \iff p \in [A_n - \varepsilon, A_n + \varepsilon]$.

ex) For coin flips, $Pr[|A_n - p| \geq \varepsilon] \leq 5\%$ $\to$ chebyshev $\to \varepsilon = \frac{2.25}{\sqrt{n}}$.

$\to [A_n - \frac{2.25}{\sqrt{n}}, A_n + \frac{2.25}{\sqrt{n}}]$ is a 95% interval for p.

(we don't know p, but at least $Var(X_n) = p(1-p) \leq \frac{1}{4}$.)

$\to Pr[|A_n - \mu| \leq \frac{4.5\sigma}{\sqrt{n}}] \geq 95\%$

Linear Regression

"Best" guess is Mean Squared Error.

i.e. the value of a that minimizes $E[(Y-a)^2]$ is $a = E[Y]$.

Proof: Let $\hat{Y} := Y - E[Y]$. $\rightarrow E[\hat{Y}] = E[Y] - E[Y] = 0$.

$\rightarrow E[c \cdot \hat{Y}] = 0$. Now, $E[(Y-a)^2] = E[(Y - E[Y] + E[Y] - a)^2] = E[(\hat{Y} + c)^2]$

$= E[\hat{Y}^2 + 2\hat{Y}c + c^2] = E[Y^2] + E[c^2] \geq E[\hat{Y}^2]$

$\rightarrow E[(Y-a)^2] \geq E[(Y-E[Y])^2] \quad \forall a. \rightarrow \min_a E[(Y-a)^2] = E[Y]$.

Now, how do we use another observation X to improve the guessing of Y?

Conditional Expectation: $E[Y|X] = g(X)$ (a function of x) where

$g(x) := E[Y|X=x] := \sum_y y \cdot Pr[Y=y, X=x] = \sum_\omega Y(\omega) Pr[\omega | X=x]$.

Properties: (a) $X \perp Y \Rightarrow E[Y|X] = E[Y]$

(b) $E[aY + bZ|X] = aE[Y|X] + bE[Z|X]$

(c) $E[Y \cdot h(X)|X] = h(X)E[Y|X]$

(d) $E[h(X)E[Y|X]] = E[h(X)Y]$

(e) $E[E[Y|X]] = E[Y]$

cd> says $E[h(X)(Y-E[Y|X])|X]=0$.

(estimation error $Y-E[Y|X]$ is orthogonal to every function $h(X)$ of X)

MMSE: function $g(X)$ s.t. it minimizes $E[(Y-g(X))^2]$. $g(X)=E[Y|X]$.

a simpler function? $g(X)=a+bX$. (linear function)

$$LLSE[Y|X]=\hat{Y}=E[Y]+\frac{Cov(X,Y)}{Var(X)}(X-E[X])$$

$$E[(Y-L[Y|X])^2]=Var(Y)-\frac{Cov(X,Y)^2}{Var(X)}$$

Hashing

Balls in bins: throwing m balls into $n > m$ bins.

Theorem: $\Pr[\text{No Collision}] \simeq e^{(-\frac{m^2}{2n})}$ for large n.

Proof: $A_i :=$ no collision for ith ball thrown.

$\Pr[A_i | A_1 \cap A_2 \cap \cdots \cap A_{i-1}] = (1 - \frac{i-1}{n})$.

No collision $= A_1 \cap \cdots \cap A_m \rightarrow \Pr[\text{No collision}] = (1-\frac{1}{n}) \cdots (1 - \frac{m-1}{n})$

$\rightarrow \underline{\ln(\Pr[\text{No Collision}])} = \sum_{k=1}^{m-1} \ln(1 - \frac{k-1}{n}) \simeq \sum_{k=1}^{m-1} (-\frac{k}{n}) = -\frac{1}{n} \cdot \frac{m(m-1)}{2} \simeq \boxed{\frac{-m^2}{2n}}$ ✱

$X :=$ # of collisions = # of pairs i and j s.t. ball i and j are in the same bin.

$X_{ij} := 1\{\text{ball } i \text{ and } j \text{ are in same bin}\} \rightarrow E[X_{ij}] = \frac{1}{n}$.

$E[X] = \frac{m(m-1)}{2} E[X_{ij}] \simeq \frac{m^2}{2n}$

ex) Checksums: m files, b bits of checksum. $\Pr[\text{collision}] \leq 10^{-3}$?

Claim: $b \geq 2.9 \ln(m) + 9$. Proof: $n = 2^b$ (# of checksums)

$\frac{m^2}{2n} = 10^{-3} \rightarrow \frac{m^2}{2^{b+1}} = 10^{-3}$

Coupon Collector's Problem. $\Pr[\text{miss one specific coupon}] \simeq e^{-\frac{m}{n}}$

$\Pr[\text{miss } \sim] = \underbrace{(1-\frac{1}{n})\cdots(1-\frac{1}{n})}_{m} = (1-\frac{1}{n})^m = ((1-\frac{1}{n})^n)^{\frac{1}{n}\cdot m} \simeq e^{-\frac{m}{n}}$

$\Pr[\text{miss any coupon}] \simeq ne^{-\frac{m}{n}}$

Load Balancing: m balls in n bins → balance "loads" of balls

For simplicity: n balls in n bins → load 1 (on average), n (worst)

max load with probability $\geq 1-\delta$? ($\delta = \frac{1}{n^c}$ for today.)

For each of n balls, choose random bin → X_i balls in bin i.

$\Pr[X_i \geq k] \leq \sum_{S\subseteq[n],|S|=k}^{\text{every subset size } k} \Pr[\text{balls in } S \text{ all choose bin } i]$

$\Pr[\text{balls in } S \text{ all choose bin } i] = (\frac{1}{n})^k$, $\binom{n}{k}$ subsets S.

$\to \Pr[X_i \geq k] \leq \binom{n}{k}(\frac{1}{n})^k \leq \frac{n^k}{k!}(\frac{1}{n^k}) = \frac{1}{k!} \leq \frac{1}{n^2}$

$\Pr[\text{any } X_i \geq k] \leq n \cdot \frac{1}{n^2}$

Sum of Poisson R.V.s: For $X \sim P(\lambda)$, $\Pr[X=i] = e^{-\lambda}(\frac{\lambda^i}{i!})$

For $X \sim P(\lambda)$, $Y \sim P(\mu)$, distribution of $X+Y$? $X+Y \sim P(\lambda+\mu)$.

$P(\lambda) = \lim_{n\to\infty} B(n, \frac{\lambda}{n}) \to P(\lambda+\mu) = \lim_{n\to\infty} B(n, \frac{\lambda+\mu}{n}) \to X+Y$!

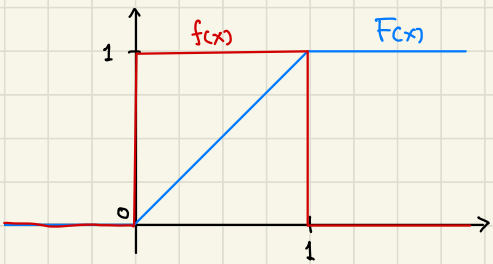
Continuous Probability

Choose a real number X, uniformly at random in $[0,1]$.

$\Pr[X \in [a,b]] = b-a$, $\forall 0 \le a \le b \le 1$. ($X \in [a,b] \subseteq \Omega$ is an event!)

In continuous probability, $\Pr[\omega]$ is usually 0. Instead, use events.

Uniform Distribution: $\Pr[X \le x] = X$ for $X \in [0,1]$. $F(x) = \Pr[X \le x]$.



$\Pr[X \in (a,b]] = F(b) - F(a)$

→ F specifies the probability of all events.

$f(x) = \frac{d}{dx} F(x) = 1\{X \in [0,1]\}$. → $F(b) - F(a) = \int_a^b f(x)\,dx$.

→ $\Pr[X \in A] = \int_A f(x)\,dx$ (Probability mass over A) → $f(x) \ge 0$, $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

Nonuniform Distribution: $f(x) = 2x \cdot 1\{0 \le x \le 1\}$ → $F(x) = \Pr[X \le x] = x^2$.

$\Pr[X \in (x, x+\varepsilon]] = \int_x^{x+\varepsilon} f(u)\,du \simeq f(x) \cdot \varepsilon$.

General Random Choice in $\mathbb{R}$: Let $F(x)$ be a nondecreasing function.

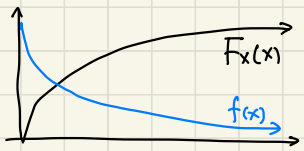
$F(-\infty) = 0$, $F(\infty) = 1$, $\Pr[X \in (a,b]] = F(b) - F(a)$.

Let $f(x) = \frac{d}{dx} F(x)$. $\Pr[X \in (x, x+\varepsilon]] = F(x+\varepsilon) - F(x) \simeq f(x) \cdot \varepsilon$.

$f(x)$ is the "local probability by unit length", i.e. "density".

Expo(λ): Limit of a geometric RV.

$S(t) = \Pr[X > t] = e^{-\lambda t}$ for $t > 0$.



$F(x) = 1 - S(x) \to F'(x) = f(x) = -S'(x) = \lambda e^{-\lambda x}$ $\{x \geq 0\}$

$\to F_X(x) = \begin{cases} 0 & (x < 0) \\ 1 - e^{-\lambda x} & (x \geq 0) \end{cases}$

Multiple Continuous RV: $f_{X,Y}(x,y)$ for $x, y \in \mathbb{R}$ where

$f_{X,Y}(x,y)\,dx\,dy = \Pr[X \in (x, x+dx), Y \in (y, y+dy)]$ → joint PDF

marginal distribution: $f_X(x) = \int f_{X,Y}(x,y)\,dy$

Independent Continuous RV: $\Pr[X \in A, Y \in B] = \Pr[X \in A] \cdot \Pr[Y \in B], \forall A, B.$

$\iff f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$

Conditional Density: $f_{X|Y}(x,y)$

Conditional Prob.: $\Pr[X \in A \mid Y \in B] = \dfrac{\Pr[X \in A, Y \in B]}{\Pr[Y \in B]}$

$\Pr[X \in [x, x+dx] \mid Y \in [y, y+dy]] = \dfrac{f_{X,Y}(x,y)\,dx\,dy}{f_Y\,dy}$

$f_{X|Y}(x,y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} = \dfrac{f_{X,Y}(x,y)}{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx}$

Expectation: $E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$

ex) $X = Uni[0,1] \to E[X] = \int_0^1 x \cdot 1\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$

$X = Expo(\lambda) \to f_x(x) = \lambda e^{-\lambda x} \to E[X] = \int_0^\infty x \lambda e^{-\lambda x} dx$

$= \cancel{\lambda e^{-\lambda x}\Big|_0^{\infty}}^{\to 0} - \int_0^\infty -e^{\lambda x} dx = \frac{1}{\lambda} e^{-\lambda x}\Big|_0^\infty = \boxed{\frac{1}{\lambda}}$

Normal Distribution: $Y = N(\mu, \sigma^2)$, $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$

Let $X = N(0,1)$. Then, $Y = \mu + \sigma X = N(\mu, \sigma^2)$.

Central Limit Theorem: For $X_i \sim N(\mu, \sigma^2)$, $S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{\sum X_i - n\mu}{\sigma\sqrt{n}}$.

Then, $S_n \to N(0,1)$ as $n \to \infty$.

Markov Chains

Finite Markov Chain: A finite set of states $X = \{0, \dots, K\}$

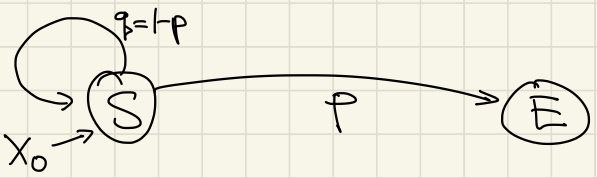
A probability distribution π_0 on X $\pi_0(i) \geq 0$, $\sum_i \pi_0(i) = 1$.

Transition probabilities $P(i,j)$ for $i, j \in X$

$Pr[X_0 = i] = \pi_0(i)$, $Pr[X_{n+1} = j \mid X_0, \dots, X_n = i] = P(i,j)$

ex) flip a coin until we get heads. average flips?

$X_0 = S$ (start), $X_n = S$ if no heads yet, $X_n = E$ if heads already.



$\beta(S)$:= average time until E. $\beta(S) = 1 + q \cdot \beta(S) + p \cdot 0$.

$\rightarrow \beta(S) = 1 + (1-p)\beta(S) \rightarrow p \cdot \beta(S) = 1 \rightarrow \beta(S) = \frac{1}{p}$.

N := number of steps until E, starting from S.

N' := " , after second visit to S.

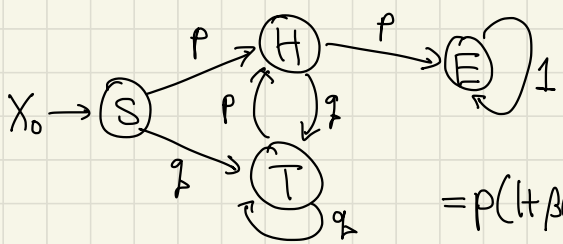
$Z := 1\{\text{first flip} = 0\} \rightarrow N = 1 + (1-Z) \cdot N' + Z \cdot 0$

$\rightarrow E[N'] = E[N] = \beta(S) \rightarrow \beta(S) = 1 + (1-p)\beta(S) + p \cdot 0$

ex) flip a coin until two consecutive heads. average flips?

$X_0 = S$, $X_n = E$ (already got two consecutive heads)

$X_n = T$ (last flip was tails) $X_n = H$ (last flip was heads)



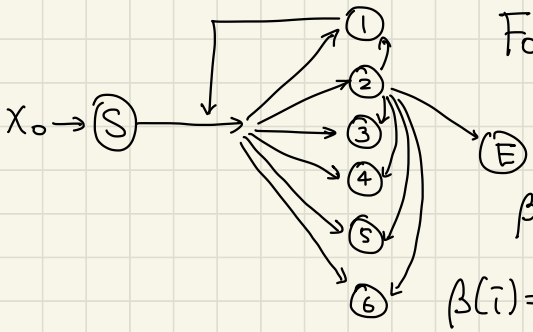
$\beta(S) = 1 + p\beta(H) + q\beta(T)$

$= Pr[H] \cdot E[\beta(S)|H] + Pr[T] \cdot E[\beta(S)|T]$

$= p(1 + \beta(H)) + q(1 + \beta(T)) = \underline{1 + p\beta(H) + q\beta(T)}$

$\underline{\beta(H) = 1 + p \cdot 0 + q \cdot \beta(T)}$, $\underline{\beta(T) = 1 + p \cdot \beta(H) + q \cdot \beta(T)}$

ex) roll die until sum of last two rolls is 8. average rolls?



For $X_0 = 1$, "start over"

For $X_0 = 2 \dots 6$, $\frac{1}{6}$ chance to end.

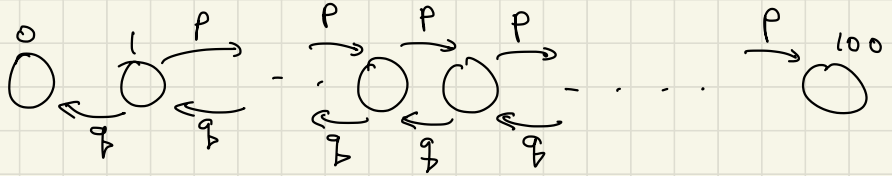
$\beta(S) = 1 + \frac{1}{6}\sum_{j=1}^{6} \beta(j) = \beta(1)$

$\beta(i) = 1 + \frac{1}{6}\sum_{(j=1\dots6,\, j \neq 8-i)} \beta(j)$ $(i = 2 \dots 6)$

$\rightarrow \beta(2) = \cdots = \beta(6) =: \gamma$, $\beta(S) = \beta(1)$.

$\rightarrow \underline{\beta(S) = 1 + \frac{5}{6}\gamma + \frac{1}{6}\beta(S),\ \gamma = 1 + \frac{1}{6}\beta(S) + \frac{4}{6}\gamma} \rightarrow \beta(S) \simeq 8.4.$

ex) $Pr[\text{win } \$1] < 0.5$. start with \$10. $Pr[\text{reach } \$100 \text{ before } \$0]$?



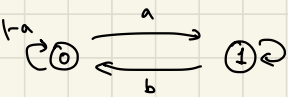
$\alpha(n) :=$ probability of reaching 100 before 0 in state $n = \{0, 1, \cdots, 100\}$.

$\alpha(0) = 0. \alpha(100) = 1. \alpha(n) = p \cdot \alpha(n+1) + q \cdot \alpha(n-1)$

Distribution of X_n: π_n is a distribution over states for X_n.

Stationary distribution – $\pi = \pi P$.

$Pr[\text{entering } i] = \sum_{i,j} P(j,i)\pi(j) = Pr[\text{leaving } i] = \pi_i$.

ex) 

$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$

$[\pi(1) \ \ \pi(2)] = [\pi(1) \ \ \pi(2)] \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = [(1-a)\pi(1) + b\pi(2) \quad a\pi(1) + (1-b)\pi(2)]$

Irreducibility: A Markov Chain is irreducible if all states are reachable from all states.

Theorem: A finite irreducible Markov Chain has a unique invariant distribution.

Then, for all i, $\frac{1}{n}\sum_{m=0}^{n-1} 1\{X_m = i\} \to \pi(i)$ as $n \to \infty$. (fraction of time)

Periodicity: gcd(lengths of all closed walks in an irreducible chain)

periodicity is 1 → aperiodic, otherwise → periodic

Theorem: $X_n :=$ irreducible, aperiodic MC with invariant distribution π.

Then, $\pi_n \to \pi(i)$ as $n \to \infty$.