Probability Theory

$$\mathbb{P}[A] = \frac{\#\text{ of sample points in } A}{\#\text{ of sample points in } \Omega}$$

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A]\cdot\mathbb{P}[B|A]}{\mathbb{P}[B]}$$

$\mathbb{P}[A] = \mathbb{P}[A\cap B] + \mathbb{P}[A\cap \bar{B}] = \mathbb{P}[B]\cdot\mathbb{P}[A|B] + \mathbb{P}[\bar{B}]\mathbb{P}[A|\bar{B}]$

$\mathbb{P}[A\cap B] = \mathbb{P}[A]\cdot\mathbb{P}[B]$ if A and B are independant.

Events $A_1, \dots, A_n$ are mutually independant if and only if for every subset $I \subseteq \{1, \dots, n\}$, $\mathbb{P}[\bigcap_{i\in I} A_i] = \prod_{i\in I} \mathbb{P}[A_i]$

Product Rule: $\mathbb{P}[A\cap B] = \mathbb{P}[A]\cdot\mathbb{P}[B|A]$ and more generally,

$$\mathbb{P}[\bigcap_{i=1}^{n} A_i] = \mathbb{P}[A_1]\cdot\mathbb{P}[A_2|A_1]\cdot\mathbb{P}[A_3|A_1\cap A_2]\cdots\mathbb{P}[A_n|\bigcap_{i=1}^{n-1} A_i]$$

Inclusion-Exclusion: $\mathbb{P}[\bigcup_{i=1}^{n} A_i] = \sum_{i=1}^{n}\mathbb{P}[A_i] - \sum_{i<j}\mathbb{P}[A_i\cup A_j] + \sum_{i<j<k}\mathbb{P}[A_i\cap A_j\cap A_k] - \cdots + (-1)^{n-1}\mathbb{P}[\bigcup_{i=1}^{n} A_i]$

Union-bound: $\mathbb{P}[\bigcup_{i=1}^{n} A_i] \le \sum_{i=1}^{n}\mathbb{P}[A_i]$

Bernoulli Distribution:

$$\mathbb{P}[X=i] = \begin{cases} p, & \text{if } i=1 \\ 1-p, & \text{if } i=0 \end{cases} \qquad \text{Var}(X) = p(1-p)$$

Binomial Distribution: $X \sim \text{Binom}(p)$

$$\mathbb{P}[X=i] = \binom{n}{i} p^i (1-p)^{n-i}; \quad \mathbb{E}[X] = np; \quad \text{Var}(X) = np(1-p)$$

Hypergeometric Distribution

$$\mathbb{P}[Y=k] = \binom{n}{k} \frac{\frac{B!}{(B-k)!}\frac{(N-B)!}{(N-B-(n-k))!}}{\frac{N!}{(N-n)!}} = \frac{\binom{B}{k}\binom{N-B}{n-k}}{\binom{N}{n}}$$

The joint distribution for two discrete random variables is the collection of values $\{(a,b), \mathbb{P}[X=a, Y=b]) : a \in \mathcal{A}, b \in \mathcal{B}\}$ where $\mathcal{A}$ is the set of all values taken by X and vice versa.

The marginal distribution for X is found by summing over all values of y; $\mathbb{P}[X=a] = \sum_{b\in \text{Range}(Y)} \mathbb{P}[X=a, Y=b]$

Independance for joint distributions can be seen when $\mathbb{P}[X=a, Y=b] = \mathbb{P}[X=a]\cdot\mathbb{P}[Y=b]$, $\forall a, b$

The expectation of a discrete random variable X is defined as

$$\mathbb{E}[X] = \sum_{a\in \text{Range}(X)} a\cdot\mathbb{P}[X=a]$$

Linearity of expectation: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Geometric Distribution: $X \sim \text{Geom}(p)$

$$\mathbb{P}[X=i] = (1-p)^{i-1}\cdot p$$

$$\sum_{i=1}^{\infty}\mathbb{P}[X=i] = \sum_{i=1}^{\infty}(1-p)^{i-1}p = p\sum_{i=1}^{\infty}(1-p)^{i-1} = p\cdot\frac{1}{1-(1-p)} = 1$$

$$\mathbb{E}[X] = \sum_{j=1}^{\infty}\mathbb{P}[X\ge j] = \sum_{i=1}^{\infty}(1-p)^{i-1} = \frac{1}{1-(1-p)} = \frac{1}{p}$$

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2$$
$$= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Memoryless property: $\mathbb{P}[X>n+m|X>m] = \frac{\mathbb{P}[X>n+m]}{\mathbb{P}[X>m]} = \frac{(1-p)^{n+m}}{(1-p)^m} = (1-p)^n = \mathbb{P}[X>n]$

Poisson Distribution: $X \sim \text{Poisson}(\lambda)$

$$\mathbb{P}[X=i] = \frac{\lambda^i}{i!}e^{-\lambda}$$

$$\mathbb{E}[X] = \sum_{i=0}^{\infty} i\cdot\mathbb{P}[X=i] = \sum_{i=1}^{\infty} i\frac{\lambda^i}{i!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{i=1}^{\infty}\frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda; \quad \text{Var}(X) = \lambda$$

If $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$ and they are independant, then $X+Y \sim \text{Pois}(\lambda+\mu)$

Law of the Unconscious Statistician:

$$\mathbb{E}[f(x)] = \sum_x f(x)\mathbb{P}_x[X=x]$$

$$\mathbb{E}[S_n^2] = \mathbb{E}[(X_1+\dots+X_n)^2] = \sum_{i=1}^{n}\mathbb{E}[X_i^2] + 2\sum_{i<j}\mathbb{E}[X_iX_j]$$

$$\text{Var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Uniform distribution: $X \sim \text{Uniform}\{1, \dots, n\}$

$$\mathbb{P}[X=i] = 1/n; \quad \mathbb{E}[X] = \frac{n+1}{2}; \quad \text{Var}(X) = \frac{n^2-1}{12}$$

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{cov}(X,Y)$$

Bilinearity of Covariance:

$$\text{Cov}(\sum_{i=1}^{n} a_iX_i, \sum_{j=1}^{m} b_jY_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_ib_j\,\text{cov}(X_i, Y_j)$$

$$-1 \le \text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma(X)\sigma(Y)} \le 1$$

We wish to know the value for which $\mathbb{E}[(X-x)^2]$ is minimized which happens to be $\mathbb{E}[X] = \mu$

Law of Conditional expectation:

$$\mathbb{E}[X|Y=y] = \sum_x x\cdot\mathbb{P}[X=x|Y=y]$$

Iterated Expectation: $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \sum_y \mathbb{E}[X|Y=y]\cdot\mathbb{P}[Y=y]$

Tower Rule: $\mathbb{E}[\mathbb{E}[\mathbb{E}[X|Y,Z]]]$

$\mathbb{E}[(W-w)^2|H=h]$ (MMSE estimate for weight from height)

$\mathbb{E}[W|H=h]$ can be viewed as a function of h.

We have that $\mathbb{E}_{W,H}[(W-\mathbb{E}[W|H])^2]$ is minimized w.r.t to the entire sample space.

What if we wanted a linear function that minimized this error? Introduce $\mathcal{L}(X) = mX+b$.

$$\mathcal{L}(Y|X) = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X-\mathbb{E}[X]) + \mathbb{E}[Y]$$

This is known as the linear least squares estimate of Y given X.

$$\mathbb{E}_{x,y}[(\hat{Y}, m\hat{X})^2] \Rightarrow m = \frac{\text{Cov}(\hat{X},\hat{Y})}{\text{Var}(\hat{X})}$$

Concentration Inequalities

For a non-negative r.v X, $\mathbb{P}[X\ge c] \le \frac{\mathbb{E}[X]}{c}$

For a random variable X, $\mathbb{P}[|X-\mathbb{E}[X]| \ge c] \le \frac{\text{Var}(X)}{c^2}$

$\left.\begin{aligned} &\mathbb{P}[|\hat{p}-p| \ge \varepsilon] \le \frac{\text{Var}(\hat{p})}{\varepsilon^2} \le \delta \\ &\mathbb{P}[|\hat{\mu}-\mu| \ge \varepsilon\mu] \le \delta \end{aligned}\right\}$ ε is error and δ is confidence

$\mathbb{P}[|Y| \ge c] \le \frac{\mathbb{E}[|Y|^r]}{c^r}$

Central Limit Theorem

$$\mathbb{P}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \le c\right] \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c} e^{-x^2/2}\,dx$$

Law of large numbers

As the number of trials n increase, the average probability approaches the true probability.

Continuous Probability

Properties of the PDF: A pdf for a real valued r.v X is a $f: \mathbb{R} \mapsto \mathbb{R}$
- That is non-negative: $f(x) \geq 0$ for all $x \in \mathbb{R}$
- and the total integral is equal to one.

Then the distribution of X is given by

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f(x)\,dx$$

Properties of the CDF: The CDF is given by the function $F(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f(y)\,dy$

and we also have the property that

$$f(x) = \frac{dF(x)}{dx} \text{ where } f(x) \text{ is the P.D.F}$$

Continuous analogs of discrete probability structures:

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

$$\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \left(\int_{-\infty}^{\infty} x f(x)\,dx\right)^2$$

For a joint density distribution

$$\mathbb{P}[a \leq X \leq b, c \leq Y \leq d] = \int_c^d \int_a^b f(x,y)\,dx\,dy$$

and the area of the total integral is equal to 1

Two continuous r.v are independant if for all $a \leq b$ and $c \leq d$, $\mathbb{P}[a \leq X \leq b, c \leq Y \leq d] = \mathbb{P}[a \leq X \leq b] \cdot \mathbb{P}[c \leq Y \leq d]$

additionally, the joint density of independant r.vs X and Y is $f(x,y) = f_x(x) \cdot f_y(y)$ where $f_x(x)$ are the respective marginal distributions.

Furthermore, the marginal distribution of X is given by

$$f_x(x) = \int_y f_x(x,y)\,dy \text{ and vice versa for } y.$$

Famous distributions and properties:

Exponential: $X \sim Exp(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; \text{if } x \geq 0 \\ 0 & ; \text{otherwise} \end{cases}$$

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

$$\mathbb{E}(X) = \frac{1}{\lambda}$$

$$\text{Var}(X) = 1/\lambda^2$$

Normal: $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma}$$

$$\mathbb{E}[X] = \mu$$

$$\text{Var}(X) = \sigma^2$$

If $X \sim N(0,1)$ and $Y \sim N(0,1)$ and X, Y are independant, then $Z = aX + bY \sim N(0, a^2 + b^2)$

In general, if X, Y are independant and $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$ then $Z = aX + bY \sim N(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2)$.

Uniform: $X \sim U[a,b]$

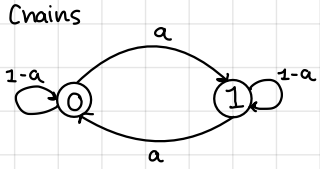
$$f(x) = \frac{1}{b-a} \text{ for } a \leq x \leq b$$

$$F(x) = \frac{x-a}{b-a} \text{ for } a \leq x \leq b$$

$$\mathbb{E}[X] = \frac{b+a}{2}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Markov Chains



X_n is the state of the markov chain at time n.

The markov chain is amnesic; at step n, it forgets what it did before getting to the current state, and its future steps only depend on its current state.

The probability transition matrix P of the above diagram is

$P(0,0) = 1-a$ $\quad P(0,1) = a$

$P(1,0) = a$ $\quad P(1,1) = 1-a$

That is, $P = \begin{bmatrix} 1-a & a \\ a & 1-a \end{bmatrix}$

and $\mathbb{P}[X_{n+1} = j \mid X_n = i, X_{n-1}, \ldots, X_0] = P(i,j)$

$\mathbb{P}[X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n] = \pi_0(i_0) P(i_0, i_1) \cdots P(i_{n-1}, i_n)$

and $\mathbb{P}[X_n = i_n] = \pi_0 P^n(i_n)$

A distribution π is invariant for the transition matrix P if

$$\pi = \pi P$$

$\pi_n = \pi_0$ if and only if π_0 is invariant

for the above markov chain, the balance equations are

$$\pi(0) = \pi(0)(1-a) + \pi(1)a$$

$$\pi(1) = \pi(0)a + \pi(1)(1-a) \rightarrow \pi(0) = 1/2, \ \pi(1) = 1/2$$

$$1 = \pi(1) + \pi(0)$$

How much time does a markov chain spend in i?

A markov chain is irreducible if it can go from every state i to every other state j, possibly in multiple steps.

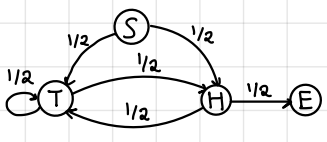
Consider a finite irreducible Markov chain with state space $\mathcal{X}$ and transition Probability matrix P. Then for any initial distribution

Fraction of time in state i

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} 1\{X_m = i\} = \pi(i), \forall i \in \mathcal{X}$$

Additionally, if the gcd of the length of all possible paths is greater than 1, then the chain is periodic with a period equal to this gcd. Else if the gcd is 1, Then the chain is said to be periodic. If the markov chain is aperiodic, then $\mathbb{P}[X_n = i] \to \pi(i), \forall i \in \mathcal{X}$, as $n \to \infty$

Hitting time:



Expected Steps until we hit E

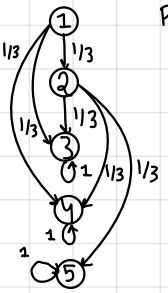
$\beta(S) = 1 + \frac{1}{2}\beta(T) + \frac{1}{2}\beta(H)$

$\beta(T) = 1 + \frac{1}{2}\beta(T) + \frac{1}{2}\beta(H)$

$\beta(H) = 1 + \frac{1}{2}\beta(T) + \frac{1}{2}\beta(E)$

$\beta(E) = 0$

Probability of A before B:



Probability of $\{3\}$ before $\{4,5\}$

$\alpha(1) = \frac{1}{3}\alpha(2) + \frac{1}{3}\alpha(3) + \frac{1}{3}\alpha(4)$

$\alpha(2) = \frac{1}{3}\alpha(3) + \frac{1}{3}\alpha(4) + \frac{1}{3}\alpha(5)$

$\alpha(3) = 1$

$\alpha(4) = 0$

$\alpha(5) = 0$

Discussion problems

X and Y are ind.

$X \sim$ Geometric(p)

$Y \sim$ Geometric(q)

$$\begin{aligned}\mathbb{P}[\min(X,Y) \geq K] &= \mathbb{P}[X \geq K \wedge Y \geq K] \\ &= \mathbb{P}[X \geq K]\,\mathbb{P}[Y \geq K] \\ &= (1-p)^{K-1}(1-q)^{K-1} = [(1-p)(1-q)]^{K-1}\end{aligned}$$

*tail probability of geometric variable.

$Z \sim$ Geometric$[1-(1-p)(1-q)]$ where $Z = \min(X,Y)$

Throw K balls into n bins. X_i = number of balls thrown into bin i.

$$\mathbb{E}[X_i] = \sum_{j=1}^{K} \mathbb{P}[\text{ball lands into } i] = \frac{K}{n}$$

$$\mathbb{E}[\#\text{ of empty bins}] = \sum_{i=1}^{n} \mathbb{E}[B_i] = \sum_{i=1}^{n} \mathbb{P}[i^{\text{th}} \text{ bins is empty}] = n\left(1-\frac{1}{n}\right)^{K}$$

(complement: probability of a ball landing in bin i; K: all K balls)

$$\begin{aligned}\mathbb{E}[\max(X,Y)] &= \mathbb{E}[X+Y-\min(X,Y)] \\ &= \mathbb{E}[X]+\mathbb{E}[Y]-\mathbb{E}[\min(X,Y)] \\ &= \frac{1}{p}+\frac{1}{p}-\mathbb{E}[\min(X,Y)] \\ &= \frac{2}{p}-\frac{1}{1-(p-1)^2}\end{aligned}$$

$$\begin{aligned}\mathbb{E}(\min(X,Y)) &= \sum_{i=1}^{\infty} \mathbb{P}[\min(X,Y) \geq i] \\ &= \sum_{i=1}^{\infty} \mathbb{P}[X \geq i]\,\mathbb{P}[Y \geq i] \\ &= \sum_{i=1}^{\infty} (1-p)^{i-1}(1-p)^{i-1} = \sum_{i=1}^{\infty} (1-p)^{2i-2} = \frac{1}{1-(1-p)^2}\end{aligned}$$

$$\mathbb{E}[X^2] = \mathbb{E}[(I_1+\ldots+I_n)^2] = \mathbb{E}\left[\sum_{i,j} I_i I_j\right] = \sum_{i}^{n} \mathbb{E}[I_i]^2 + \sum_{i \neq j} \mathbb{E}[I_i I_j]$$

if k is odd, then the k^{th} moment of the gaussian distribution is 0.

Modular Arithmetic, RSA, FLT, and Extended Euclids:

$x \equiv y \pmod m \iff m$ divides $(x-y)$

$x \equiv y \pmod m \iff x \pmod m = y \pmod m$

Theorem 6.1: If $a \equiv c \pmod m$ and $b \equiv d \pmod m$ then $a+b \equiv c+d \pmod m$

For a finite set A, $f: A \to A$ is a bijection if there is an inverse function $g: A \to A$ such that $\forall x \in A \; g(f(x)) = x$.

Bijection: every $b \in B$ has a unique pre-image $a \in A$ and $\forall a, a' \in A$, if $f(a) = f(a')$ then $a = a'$

Theorem 6.2: Let m, x be positive integers such that $\gcd(m,x) = 1$. Then x has a multiplicative inverse modulo m, and it is unique (modulo m), $ax \equiv 1$ and $bx \equiv 1$

Proof: Consider m numbers $0, x, 2x, \ldots, (m-1)x$. Assume that $ax \equiv bx \pmod m$ s.t $a \neq b$. This implies that $(a-b)x \equiv 0 \pmod m$ or $(a-b)x = km$ for $k \in \mathbb{Z}$ however, m and x are coprime (because $\gcd(m,x) = 1$) $\Rightarrow (a-b)$ is multiple of m Contradiction $(a-b) \in \{0, 1, \ldots, m-1\}$ which mean a-b cannot be multiple of 5

$\gcd(m,x) = 1$ is a necessary condition in fact Prove if $\gcd(m,x) > 1$, then x has no multiplicative inverse modulo m. **Proof:** Assume that x has an inverse a. $\Rightarrow ax \equiv 1 \pmod m$ or $ax \equiv km + 1$. $ax - km = 1 \Rightarrow d \mid ax - km \Rightarrow d \mid 1$. Contradiction. $d \neq 1$ because $d > 1$. Q.E.D.

$d = \gcd(x,y) = ax + by$ How to compute gcd? Recursion. **Theorem 6.3:** Let $x \geq y > 0$. Then $\gcd(y, x \pmod y)$. Keep recursing till gcd has a zero in it.

Extended Euclid's algorithm: Used to find multiplicative inverse. $\gcd(x,y) = ax + by$ find a and b. Recursive method is to keep track of a and b as the recursion unwinds. However, iterative version is easier to do by hand when computing larger inverses. Aggressively subtract until not possible to reduce.

Fundamental Theorem of Arithmetic: Every positive integer can be expressed uniquely in the form $p_1 p_2 \cdots p_k$ where p_i is a prime number (not necessarily unique)

Chinese Remainder Theorem: For m, n with $\gcd(m,n) = 1$, there is exactly one x mod mn that satisfies the following congruencies; $x \equiv a \pmod n$ and $x \equiv b \pmod m$.

Let $n_1, n_2, \ldots, n_k$ be positive integers that are co-prime then, $\{x \equiv a_1 \pmod{n_1}, \ldots, x \equiv a_i \pmod{n_i}, \ldots, x \equiv a_k \pmod{n_k}\}$ and $x \equiv \left(\sum_{i=1}^{k} a_i b_i\right) \bmod N$

where $b_i = \frac{N}{n_i} \left(\frac{N}{n_i}\right)^{-1}_{n_i}$ where $\left(\frac{N}{n_i}\right)^{-1}_{n_i}$ is multiplicative inverse of N/n_i w.r.t n_i.

When alice wants to send message x to bob, she computes $E(x) \equiv x^e \bmod N$

Upon recieving $y = E(x)$, Bob computes $D(y) \equiv y^d \bmod N \equiv x$

FLT: For any prime p and any $a \in \{1, 2, \ldots, p-1\}$, we have $a^{p-1} \equiv 1 \pmod p$ **Proof:** Let $S := \{1, 2, \ldots, p-1\}$. Consider set $S' := \{a, 2a, 3a, \ldots, (p-1)a \bmod p\}$. $\gcd(a,p) = 1$ therefore they are distinct from each other. None are 0 therefore there are p-1 of em. $\Rightarrow S = S'$ in different arrangement. Multiply all numbers in S to get $(p-1)! \pmod p$ and in S' to get $a^{p-1}(p-1)! \pmod p$. $\Rightarrow (p-1)! \equiv a^{p-1}(p-1)! \bmod p. \Rightarrow 1 \equiv a^{p-1} \pmod p$

RSA: Show that $(x^e)^d \equiv x \pmod N$ **Proof:** $x^{ed} - x \equiv 0 \pmod N$, following congruencies emerge $x^{ed} - x \equiv 0 \pmod p$ and $x^{ed} - x \equiv 0 \pmod q$. $x^{ed} = x^{k(p-1)(q-1)+1}$ $\Rightarrow x(x^{k(p-1)(q-1)} - 1) \equiv 0 \pmod p \to (x^{k(p-1)})^{(q-1)} \equiv 1 \pmod p$ if $p \nmid x$. In both cases where $p \nmid x$ and $p \mid x$, $x^{ed} - x \equiv 0 \pmod p$ and similar argument can be shown for q. RSA is correct Q.E.D.

Mod-pourri:

$\text{lcm}(x,y) \cdot \gcd(x,y) = x \cdot y \quad \forall x, y \in \mathbb{Z}^+$

$a^{(p-1)(q-1)} \equiv 1 \pmod{pq}$ by RSA

$1^2 \cdot 2^2 \cdot \ldots \cdot (p-1)^2 \equiv 1 \pmod p$

$x = a \pmod m$; $x = b \pmod n$ s.t $\gcd(m,n) = \gcd(a,m) = \gcd(b,n) = d$. d solutions.

$y = a/d \pmod{m/d}$ and $y = b/d \pmod{n/d}$; $x = yd + imn/d \pmod{mn}$ for $i \in \{0, \ldots, d-1\}$

How many values of $a \in \{0, \ldots, 9\}$ if $f(x) \equiv ax \pmod{10}$ a bijection? 4

If $\gcd(a,m) = x$ and $\gcd(b,m) = y$ and $\gcd(x,y) = z$, what is $\gcd(ab,m)$? xy/z

For $x, y \in \mathbb{N}$, with $x < y$, what is the smallest M, where $y \pmod x < M$? $M = y/2$

Find z s.t $xz = 5d \pmod y$ s.t $\gcd(x,y) = d$ $jd = 1$. $xz \equiv 5d \pmod y \to xz \equiv 5(ax + by) \pmod y$

Find z s.t $xz = 5d \pmod y$ s.t $\gcd(x,y) = d$ $jd \neq 1$ $xz \equiv 5d \pmod y \to xz/d = 5d/d + iy/d$

$1 \; d/d = ax/d + by/d \Rightarrow (x/d)^{-1} \equiv a \pmod{y/d} \Rightarrow \frac{x}{d} z = 5 \pmod{\frac{y}{d}} \to z \equiv 5a \pmod{y/d}$

It can be helpful to take an equation out of modspace and manipulate equations

$ax \equiv b \pmod N$ max solutions occurs when b is a factor of d.

Range of $px \pmod{pq}$ where domain is $\{0, 1, \ldots pq-1\}$. Multiples of q get mapped to 0

$xu \equiv v \pmod m$ s.t $\gcd(x,m) = y$ has a solution for u is v is a multiple of d.

find solution in terms of a, b, x, y, d. $xu = v + im \to xv/d = v/d + im/d$. $(x/d)^{-1} \equiv a \pmod{m/d}$

b/c $1 \; d/d = a^{x/d} + b^{my/d}$. $\therefore$ $a^{(x/d)v} \cdot u = a^v \pmod{m/d}$

$f(x) = ax \pmod n$ where $x \in \{0, 1, \ldots, n-1\}$ and $\gcd(a,n) = d$. Range = multiples of d (or n/d)

One useful modular arithmetic fact to know is that if $d \mid x, d \mid y$, and $d \mid n$, then $x \equiv y \pmod n \Rightarrow x/d \equiv y/d \pmod{n/d}$. since d is gcd of a, b, n, dividing through by d ensures that they are all relatively prime. $ax/d \equiv b/d \pmod{n/d} \Rightarrow x \equiv (a/d)^{-1}(b/d) \pmod{n/d}$

Polynomials and Error correcting codes:

Erasure errors: message has n packets and k are lost in transmission. Send k+n.

General Errors: message has n packets and k are corrupted while sending. Send 2k+n.

Berlekamp-Welch: We send length n message. $P(x)$: has degree $n-1$. $r_i = P(i)$ iff the recieved packet is correct. If there are $\leq k$ general errors, send $n+2k$ packets Use an Error locator polynomial to find the errors. Roots of $E(x)$ represent location of errors. Reciever does not know $E(x)$ but can find it: If there are at most k errors then $E(x)$ has degree at most k. This implies that $Q(x)$'s degree is $n+k-1$ Since $Q(x) = P(x)E(x)$ and $P(x)$'s degree is $n-1$ $Q(x)$ has deg $n+k-1 \Rightarrow n+k$ unknowns $E(x)$ has deg $k+1$ but leading coef is $1 \Rightarrow k$ unknown. $\Rightarrow n+2k$ unknowns. We have $n+2k$ points so we can solve for the unknown coeffs. After solving, $P(x) = Q(x)/E(x)$. If there are less than k general errors, there will be multiple solutions however, the resulting $P(x)$ is still the same.

Lagrange interpolation: Given $(x_1, y_1) \ldots (x_{m-1}, y_m)$, we can find a polynomial modulo p by:

$$\Delta_i = \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)} \quad \text{and} \quad p(x) = \sum_{i=1}^{d+1} y_i \Delta_i(x)$$

Notice that $\Delta_i(x)$ has deg d $\therefore$ the summation of the Δ_i's corresponds to a deg d polynomial.

Graph Theory:

An undirected graph G has Eulerian Tour iff every vertex has even deg.

A graph is a tree than it is acyclical and and connected

A graph is planar if it can be drawn without crossings. Planar graphs follow Euler's formula, $v+f = e+2$ **Proof.** Induction on edges. Trivially holds when $e=0$, $v=f=1$. Take any connected graph. If graph is a tree, then $f=1$ and $e = v-1$. If not a tree than find cycle and delete any edge which is the same as reducing e and f by one. By induction, the formula is true in the smaller graph, so it must be true in the original graph Q.E.D.

Planar graphs have faces which have atleast 3 sides. Therefore by handshake lemma $2e \geq 3f \Rightarrow e \leq 3v-6$. $K_{3,3}$ passes this test but is not planar. A graph is non-planar iff it contains K_5 and $K_{3,3}$

Dual of a graph is placing a node on each face of G and connecting nodes if its corresponding face shares an edge.

Every planar graph is five colorable. **Proof.** Trivial base case. IH: Assume that a graph with v vertices is 5-colorable. IS: prove that a graph with $v+1$ vertices is also 5 colorable. Because of euler inequality, $e \leq 3v-6$ which implies that there is a node of degree 5 or less. Consider node u with deg ≥ 5. If u deg > 5, stop here, remove u, 5 color rest of graph, add u back in and color adjacent edges. If v has deg 5 then it has 5 neighbors u_1, u_2, u_3, u_4, u_5. Change u_2's color to 4, determine the connected component with 2 and 4 colors and switch along the way. If u_4 is in component, then color swap was useless. If not, then color u_1 with missing color. Do with u_2 if doesn't work with v_2 as well, this is impossible. Path from u_2 to u_4 and u_1 to u_3 must intersect but this can't be color 1, 2, 3, or 4. By contradiction either flipping u_2 works or u_1 works. Therefore we can color u_5 with the remaing color.

Total edges in hypercube is $n2^{n-1}$ and total vertices are 2^n.

K_n has n vertices and $n(n-1)/2$ edges.

Prove that any graph with maximum degree of d is $2d-1$ colorable.

Induct on # of edges, BC: 0 edges $\Rightarrow$ 0 colorable, IH: Assume for $n = k \geq 0$ the statement holds IS: Consider a graph with $k+1$ edges. Remove an edge e. Remaining graph is degree $d' \leq d$. By IH, k edge graph is $2d'+1$ colorable. Add edge back. New edge is incident to (in worst case), 2 vertices of degree $d-1$. $\therefore$ $2(d-1)$ colors are unavailable Color edge using remaining color.

Prove a tree is d edge colorable is max degree is d. **Proof:** Induction on vertices.

Base case $(n=1)$: Only one vertex, no edges to color $\Rightarrow$ 0 colorable.

IH: Assume this holds for vertices $n = k \geq 1$.

IS: Remove a leaf from tree. Vertex neighboring leaf node has degree at most $d-1$. Color this incident vertex and entire tree by IH. One color remaining. Use this one color to color the vertex and its edge after adding back in.

Prove that all tree with atleast 2 vertices are bipartite **Proof.** Induct on v

Base case $(v=2)$: 2 vertices have 1 edge between so we can partition the 2 vertices into 2 sets of vertices (1 each)

IH: Assume if we have n vertices in a tree, then the tree is bipartite

IS: Prove that if we have $n+1$ vertices in a tree, then the tree is bipartite Remove one leaf and it corresponding edge. The new graph has n vertices $\overset{\text{by IH}}{\Rightarrow}$ it bipartite $\therefore$ it can be seperated into two sets L and R s.t. edges go only in between each set. When we add back our leaf, the edge either leads to a vertex in L or R so the vertex must be in R in L respectively.

Any graph where $|E| \leq 3|V| - 6$ is planar. False, consider $K_{3,3}$.

Any connected graph with degree strictly less than 2 is a tree. True.

Minimum amount of colors to vertex color a graph that is a simple cycle is $2 + k \bmod 2$

What are the number of edges in an n-vertex acyclical graph with k components $n-k$.

For any cut in the hypercube, the number of cut edges is atleast the size of the small side

Removing $2k$ edges from a graph with k cycles produces a graph with a minimum of $k+1$ connected components. Removing $2k$ edges from G will result in a maximum of of $2k+1$ connected components.

If we can find a vertex of degree at most k

When considering degrees and maximums and all that, make sure to think about max # of edges and min # of edges.

Stable Matching:

Most of the time, stable marriage problems involve finding counter-examples or finding contradictions.

Propose and Reject algorithm always halts.

If job j makes an offer to candidate C on the k^{th} day, then on every subsequent day C has a job offer in hand which she likes as much as J **Improvement Lemma.**

Lemma 4.3. The propose and reject algorithm terminates with a matching

Theorem 4.1. The matching produced by the algorithm is always stable.

Proof. Direct proof. Consider any couple (J,C) in the final matching and suppose that J prefers some candidate C^* to C. We will argue that C^* prefers her job J, so that (J,C^*) cannot be a rogue couple. Since C^* occurs before C in J's list, J must have made an offer to C^* before it made an offer to C. By Improvement Lemma, C^* likes her final job at least as much as J. Q.E.D.

For a given job J, the optimal candidate for J is the highest rank candidate on J's preference list that J could be paired with in any stable matching. (opposite definition for candidate). **Theorem 4.2** The matching output by the propose-and-reject algorithm is job/employer optimal.

Assume not employer optimal. Then there exists a day on which some job had its offer rejected. Suppose J was rejected by C^* in favor of J^*

Theorem 4.2 The matching output by Propose-and-Reject algorithm is job optimal. **Theorem 4.3** If a matching is employer optimal, then it is candidate pessimal.

Any stable matching is either job-optimal or candidate pessimal. False

If a stable matching is optimal for j it is pessimal for its partner C. False

There is always a job never rejected False.

There is always a candidate who never rejects a job True.

If all candidates have a single offer at the end of day k, then the algorithm stops. Assume there is a day k-i, where all candidates have one job on string. If this is the case, then the algorithm must have ended on day k-i! However this a contradiction since we assumed that the da the algorithm ended on k. This also means that in previous days there was one candidate who received no offers until day k. If he got no offers till day k, then for any day prior to day k-i, he had no offer (Lemma 2.b).

let (j,c) be a last day prop. Remaing (j,c) will still result in the same matching. False.

Prove max rejections with n job and candidates is $(n-1)^2$. Proof.

one candidate rejects no one else and the other n-1 candidates can reject n-1 jobs.

True, by definition of the algorithm, as soon as a candidate recieves a proposal J on day k+i, on beginning of day k+i+1, the candidate recieves either only J (because she has not rejected them), or a better offer J'. In both cases, on day k+i+2, the candidate will recieve J's proposal(which is atleast J). This continues until every candidate has only one in hand.

For purpose of contradiction, assume there is a day j where the candidate did recieve an offer. However, by Lemma 2(a) we know that if a candidate recieved an offer on day j then the candiate will always have a proposal for day k > j. If a candidate recieved no proposal on day j, then she recieved no offer the previous days by contradiction.

maximum number of jobs in rougue pair when one job preference is changed is 1 and n-1 for the candidates

What is number of deg 1 polys? 1

What is number of deg d polys? $(p-1)p^d$

What is number of exactly deg d poly's with d distinct roots? $(p-1)\binom{p}{d}$

What is number of exactly deg d poly's with d roots? $(p-1)\binom{d+p-1}{d}$

Remember pigeonhole pricipal for proof and well ordering pricipal.

For every natural number $n \geq 12$, it holds that $n = 4x+5y$

induction on n.

IH: Assume claim holds for all $12 \leq n \leq k$ for $k \geq 15$.

IS: Prove $n = k+1 \geq 16$. note: $(k+1)-4 \geq 12 \overset{\text{by IH}}{\Rightarrow} (k+1)-4 = 4x'+5y'$

If $x = x'+1$ and $y = y'$, we complete the proof.

Combinatoric:

How many ways ta arrange n 1's and k 0's. $\binom{n+k}{k}$

How many 19 digit tenerary bitstrings are there where no adjacent digits are similar? $3 \cdot 2^{18}$

How many 13-card bridge hand? $\binom{52}{13}$

How many 13-card bridge hands with no aces? $\binom{48}{13}$

How many 13 card bridge hands with all four aces? $\binom{48}{9}$

How many 13 card bridge hands will exactly 4 spades? $\binom{13}{4}\binom{39}{9}$

How many 49 bitstring have more 1's then 0's? 2^{48}

How many anagrams of ABCDEF if C is immediate left of E. 5!

C is on left (not nessicarily immediate? 6!/2

8 numbered balls and 25 bins. How many different ways to give? 25^8

what about 8 identical balls into 25 distinguishable? $\binom{n+k-1}{k} = \binom{32}{8}$

what about 8 identical balls into 6 distinguishable s.t none $\emptyset$? $6+\binom{8}{2}$

How many different ways to pair 20 students? $19 \times 17 \times 15 \times \ldots \times 1$.

If an object can be made by a succesion of k choices then we can count via $n_1 \times n_2 \times \ldots \times n_k$.

Pick k distinct elements but no care for order. $\frac{n!}{(n-k)!k!} = \binom{n}{k}$

Samping with replacement: $n_1 = n_2 = \ldots = n_k \Rightarrow n^k$ sequences

Sample with replacement order does not matter: $\binom{k+n-1}{n-1}$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \quad \text{: binomial theorem}$$

$$\binom{n}{k+1} = \binom{n-1}{n} + \binom{n-2}{k} + \ldots \binom{k}{k} \quad \text{: hockey stick identity}$$

$$\text{Derrangements of } n = D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

Inclusion-Exclusion Principal: $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$

Generally: $|A_1 \cup \ldots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq \{1,\ldots,n\}} |\cap_{i \in S} A_i|$

combinaterial proof $\Rightarrow$ tell a story about both sides of the equation.

How many distinct with n labeled vertices (no duplicates)?

There are $\binom{n}{2} = n(n-1)/2$ possible edges. Each edge is either present or not. So the answer is $2^{n(n-1)/2}$.

How many distinct cycles in K_n? number of vertices with atleast 3 and at most n $\Rightarrow \frac{n!}{(n-k)!}$ cycles. since every cycle is invertible/rotatable we have $\sum_{k=3}^{n} \frac{n!}{(n-k)!2k}$

Logic and propositions:

$\forall x (P(x) \wedge Q(x)) \equiv \forall x P(x) \wedge \forall x Q(x)$

$\forall x (P(x) \vee Q(x)) \not\equiv \forall x P(x) \vee \forall x Q(x)$

$\exists x (P(x) \vee Q(x)) \equiv \exists x P(x) \vee \exists x Q(x)$

$\exists x (P(x) \wedge Q(x)) \not\equiv \exists x P(x) \wedge \exists x Q(x)$

$(P(0) \wedge (\forall n \in \mathbb{N})(P(n) \Rightarrow P(n+1))) \equiv (\forall n \in \mathbb{N}, P(n))$ Induction

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

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

} make sure to check for vacuously true statements.

Computability

The problem of whether P on x uses more than M bits of memory is decidable.

The problem of whether P on x executes line n is undecidable

For every real number, there is a program k which will print out the k^{th} digit of x. False.

There is no program that takes P on x and returns whether P ever writes to memory location x.

There is a program that takes P on x and returns whether P ever writes to any memory location $i \geq n$.

There is a program that knows π and can output the k^{th} digit.

There is a program that determines if P halts on x in n steps.

Countability

The set of subsets of a countably infinite set is uncountable since there is a bijection with the powerset of $\mathbb{N}$.

The set of finite sized subsets of a countably infinite set is countable since one can assign a number to the set of subsets k for all k's (This being the maximal number.

The set of all pairs of two countably infinite sets is countable, (this is similar to the spiral trick for the rationals where we assigned the numerator a value and denominator a value a plotted it).

The rational numbers are countable.

The set of all finite graphs is countable.

There is a bijection between the powerset of the rationals and the reals.

The powerset of a countably infinite set is uncountable.

The set of ordered partitions of $\mathbb{N}$ is countable.

The set of ordered partitions of $\mathbb{Q}$ is uncountable.

The countable union of countable sets is countable.

Polynomials and More modular arithmetic (Yay !)

$P(x)$ and $Q(x)$ (degree d) have d intersections ($Q(x)-P(x)=0$)

$P(x)$ (of degree $2d$) and $Q(x)$ (degree d) have $2d$ intersections

a polynomial with roots r_1, r_2, and v at r_3 mod p is

$$v * (r_2-r_1)^{-1}(r_3-r_2)^{-1}(x-r_1)(x-r_2)$$

Every polynomial modulo a prime p is equivalent to a polynomial of degree at most $p-1$.

if $a^4 \equiv 10 \pmod p$ and $a^6 \equiv 19 \pmod p$ then $a^{10} \equiv 190 \pmod p$

$\gcd(x,y) \neq \gcd(x, x-ay)$ for $\forall x, y$

if $\gcd(a,b)=1$, $za \equiv b \pmod{ab}$ has no solution since a does not have an inverse mod ab.

There are $4 \cdot \binom{5}{2} + 20$ polynomials of degree 2 over GF(5).

A polynomial of degree d does not always have d roots mod p.

Any polynomial of degree 2 over GF(p) has 0 or 2 roots.

For distinct primes p, q, r; $\{0, 1, \ldots, N-1\}$ has $(p-1)(q-1)(r-1)$ number coprime to N

For a prime p, an x that guarentees $a^x \equiv 1 \pmod{p^2}$ for all a relatively prime to p is $p(p-1)$

$$a^{(p-1)(q-1)(r-1)} \equiv 1 \pmod{pqr}$$ by FLT

Maximum number of roots for $P(x) * Q(x)$ under GF(p) is $\max(p, r_p + r_q)$

Minimum number of roots for $P(x) * Q(x)$ is $\max(r_p, r_q)$ if $P(x)$'s roots have overlap with $Q(x)$'s roots and vice versa.

The size of $\{ay \pmod{pq} : y \in \{1, \ldots, pq-1\}\}$ when a is not a multiple of $pq-1$, is $pq-1$.

The size of $\{ay \pmod{pq} : y \in \{1, \ldots, pq-1\}\}$ when a is a multiple of p is q.

The size of $\{a \pmod n, 2a \pmod n, \ldots, (n-1)a \pmod n\}$ if $\gcd(a,n)=1$ is $n-1$.

CLT

$x \equiv 4 \bmod 7$ $\quad \gcd(7,4)=1$

$x \equiv 2 \bmod 4$ $\quad \gcd(7,5)=1$

$x \equiv 2 \bmod 5$ $\quad \gcd(4,5)=1$

mod 7	mod 4	mod 5
$x = 4 \cdot 5$	$+7 \cdot 5$	$+7 \cdot 4$

$= 20 \cdot 3 + 35 + 28 = 60 + 35 \cdot 2 + 28 = 60 + 70 + 28 \cdot 4 = 60 + 70 + 112 = 242$

$\equiv 20 \bmod 7$ $\quad 6x \equiv 4 \pmod 7$ $\quad x \equiv 24 \pmod 7$ $\quad x \equiv 3 \pmod 7$

$\equiv 6 \bmod 7$

$20x \equiv 4 \bmod 7$

$20 \bmod 7 \equiv 6 \bmod 7$

$18 \bmod 7 \equiv 4 \bmod 7$

$35 \bmod 4$

$\equiv 3 \bmod 4$

$35 \equiv 3 \bmod 4$

$6 \equiv 2 \bmod 4$

$\equiv 28 \bmod 5$

$28 \equiv 3 \bmod 5$

(2) $6 \equiv 1 \bmod 5$

(3) $9 \equiv 4 \bmod 5$

(4) $12 \equiv 2 \bmod 5$

$x \equiv 242 \pmod{140}$

$x \equiv 102 \pmod{140}$

ECC and Berlekamp Welch $\quad P(x) = E(x)Q(x)$

$P(x)$ has degree $n-1$ since n points determine a degree $n-1$ poly

We send $n+k$ packets when we have k erasures and $n+2k$ packets for k general errors.

The roots of the error polynomial $E(x)$ represent the location of corrupted packets. If there are at most k errors, then the maximum degree of $E(x)$ is k and the maximum degree of $Q(x)$ is $(n-1)+k$ $= n+k-1$ since the degree of P is $n-1$ and the degree of E is at most k. $P(x)$ can be found by $Q(x)/E(x)$

$P(0)=1 \quad P(1)=1 \quad P(2)=4$

$$\Delta_0(x) = \frac{(x-1)(x-2)}{(0-1)(0-2)} = \frac{(x-1)(x-2)}{2} \equiv 3(x-1)(x-2) \pmod 5$$

$$\Delta_1(x) = \frac{(x-0)(x-2)}{(1-0)(1-2)} = \frac{x(x-2)}{-1} \equiv 4(x^2-2x) \pmod 5$$

$$\Delta_2(x) = \frac{(x-0)(x-1)}{(2-0)(2-1)} = \frac{x(x-1)}{2} \equiv 3(x^2-x) \pmod 5$$

$$P(x) = 1\Delta_0(x) + 1\Delta_1(x) + 4\Delta_2(x)$$