NOTE 1: PROP LOGIC

$P \to Q \equiv \neg P \vee Q$ $\quad$ $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$

$P \to Q \equiv \neg Q \to \neg P$ $\quad$ $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$

$\neg (P \wedge Q) \equiv \neg P \vee \neg Q$ $\quad$ $\neg(\forall x P(x)) \equiv \exists x \neg P(x)$

$\neg (P \vee Q) \equiv \neg P \wedge \neg Q$ $\quad$ $\neg(\exists x P(x)) \equiv \forall x \neg P(x)$

$P \vee \neg P$ always true $\quad$ False $\to$ True ✓

Note 2: Proofs / Induction

integer $\mathbb{Z} = \{..., -2, -1, 0, 1, 2\}$ $\quad$ natural $\mathbb{N}: \{0, 1, 2 ...\}$ $\quad$ $\mathbb{Q}$: rationals (can write as $\frac{a}{b}$) $\quad$ $\mathbb{R}$: reals

$a | b$ = a divides b iff int q where $b = aq$ $\quad$ * IFF $\to$ prove both directions

Direct Proof : Assume P, therefore Q $\quad$ Proof by contradiction: $\neg P \to \neg R \wedge R$

Proof by contraposition: $\neg Q \to \neg P$ $\quad$ Proof by cases: if at least one case holds

Induction
1. Base case: prove smallest case is true
2. Ind hyp: assume $n = k$ (weak) or $n \leq k$ (strong)
3. Ind step: prove $n = k+1$ true $\leftarrow$ what we want to show $\to$ cannot assume alr true & work backwards
   - ↳ try to break up into parts & use ind hyp

→ use contraposition

Pigeonhole principle: putting $n+m$ balls in n bins $\to \geq 1$ bin has ≥ 2 balls

Note 4: Stable Matching

Propose-and-reject algo always halts $\leq n^2$ days

Rogue couple: job + candidate who prefer each other

Stable matching: NO rogue couples

* possible to have matching w/ all fav jobs, matching w/ no fav candidate
* can have 1 RC, no ALL rogue couples

Improvement lemma: candidate's options improve w/ each day (use induction)

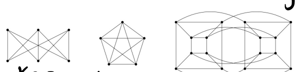
Well-ordering principle: any non-empty set has a smallest element → ("The first day that...")

→ construct a rogue couple

* proof by **induction** or **contradiction**, counting # of rejections, **2×2 case** w/ diff preferences
* algo always ends w/ a stable matching
* **job/candidate optimal:** where jobs/candidates receive highest pref of ALL stable matchings

Planar Graphs : drawn w/o crossings

nonplanar: containing $K_{3,3}$, K_5 makes a graph non-planar



faces = regions where graph subdivides plane

$\boxed{v + f = e + 2}$ Euler's formula

$3f \leq 2e \longrightarrow \boxed{e \leq 3v - 6 \text{ if planar}}$

Bipartite = edges split into 2 groups & edges

$\boxed{e \leq 2v - 4}$ if planar & bipartite

Bipartite: 2 colorable $\quad$ **Planar**: 4 colorable

d+1 colors to vertex color graph on n vertices w/ $d \geq 1$

d colors to edge color acyclic graph on n vertices w/ d

To lower deg of each vertex remove $\frac{n}{2}$ edges

Note 5: Graphs

of vertices of odd degree is even

$K_n = \frac{n(n-1)}{2}$ edges $\quad$ $\sum \deg v = 2e$

Trees: connected & no cycle $\quad$ * all trees are planar *

connected & $n-1$ edges $(n = |V|)$

connected & removing edge disconnects graph

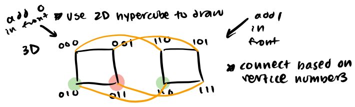
no cycles & adding edge makes a cycle

* at least 2 vertices ↓ bipartite (use induction)

Eulerian graph: every vertex = even degree

* removing an edge increases # of components by 1 *

Hypercubes: n-length bit strings that differ by exactly 1 bit



* $\boxed{n 2^{n-1} \text{ edges}}$
* bipartite $\quad$ * edge colored w/ n colors
* $\boxed{2^n \text{ vertices}}$
* $\log_2 N$ Dimensional

Eulerian walk: visits each edge ONCE

- connected & 2 vertices odd degree ← starts & ends at distinct vertices

Eulerian tour: Eulerian walk that starts & ends at same

→ all even degree and connected (* can repeat vertices but not edges)

Cycle: starting/ending at same vertex & visiting distinct vertices

connected = path between distinct vertices

complete graphs = max # of edges possible $\left(\frac{n(n-1)}{2}\right)$

* induction on edges or vertices / add or remove edge/vertex

Hamiltonian tour: visits all vertices once (* cannot repeat vertices)

# of edges in graph where v has deg d(v) is $\frac{1}{2} \underbrace{\sum_{v \in V} d(v)}_{\text{total deg}}$

# of edges w/ avg degree $d = \frac{nd}{2}$

Note 6: Modular arithmetic

$a \equiv c \bmod m$ and $b \equiv d \bmod m \to a + b \equiv c + d \bmod m$

(multiplication & addition) $\quad$ $a \cdot b \equiv c \cdot d \bmod m$

x^{-1} (modular inverse) exists mod m iff $GCD(x, m) = 1$

↳ $ax \equiv 1 \bmod m$

* reduce bases: $x^{2a} = (x^a)^2$ and $x^{2a+1} = x(x^a)^2$
* reduce exponents!

$x = 304^{2022} \equiv 0 \bmod 2$ $\quad$ $\boxed{GCD(x,y) = GCD(y, x \bmod y)}$

$x = 304^{2022} \equiv -1^{2022} \equiv 1 \bmod 5$

$x \equiv 304^{2022} \equiv 3^{2022} \equiv (3^6)^{337} \equiv 1 \bmod 7$

Fundamental thm of arithmetic: any pos int can be expressed as product of primes $\quad$ $n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_\ell$

* show $k = \ell$ and p_i are reordering of q_j

CRT: For m, n w/ $GCD(m,n) = 1$, there's exactly one x mod mn that satisfies

$x \equiv a \bmod n$ and $x \equiv b \bmod m$

* COPRIME $m_1 \cdots m_n$

→ mod M can be broken into mod m_1 and m_2 etc

$x = ua + vb \equiv a \bmod n$, $\equiv b \bmod m$ ← $u \equiv 1 \bmod n$ & $u \equiv 0 \bmod m$; $v \equiv 0 \bmod n$ & $v \equiv 1 \bmod n$

$u = m(m^{-1} \bmod n)$ & $v = n(n^{-1} \bmod m)$

Extended Euclid algorithm : $ax + by = \gcd(x, y)$

$GCD(x, m) = 1$ then x has an inverse

$GCD(x, y) = ax + by = 1 \to b$ is inverse of y mod x

* $d = ay + b(x \pmod y) = ay + b(x - \lfloor x/y \rfloor y) = bx + (a - \lfloor x/y \rfloor b) y$

$5x \equiv 3 \bmod 24 \longrightarrow GCD(5, 24)$

$5(5^{-1}) \equiv 3 \cdot 5^{-1} \equiv 3 \cdot 5 \equiv \underline{15 \bmod 24}$

$24 = 1(24) + 0(5)$
$-4 \quad 5 = 0(24) + 1(5)$
$- \quad 4 = 1(24) - 4(5)$
$1 = -1(24) + 5(5)$

$5^{-1} = 5$

* cannot assume inverse will be the inverse of other

For Fibonacci: $GCD(F_n, F_{n-1}) = 1$

Base case: $\gcd(F_1, F_0) = \gcd(1, 0) = 1$

Ind hyp: $\gcd(F_k, F_{k-1}) = 1$

Ind step: $\gcd(F_{k+1}, F_k) = \gcd(F_k + F_{k-1}, F_k) = \gcd(F_k, F_{k-1}) = 1$

$x \equiv 3 \bmod 11$
$x \equiv 7 \bmod 13$ → $x = 11a + 13b$ ✱write constraints

$3 \equiv x \equiv 13b \bmod 11$
$b \equiv 13^{-1} \times 3 \equiv 6 \times 3 \equiv 7 \bmod 11$ & $7 \equiv x \equiv 11a \bmod 13$
$a \equiv 11^{-1} \times 7 \equiv 6 \times 7 \equiv 3 \bmod 13$
$x = 11a + 13b = 11 \times 3 + 13 \times 7 = 33 + 91 \rightarrow x \equiv 124 \bmod 143$

$x \equiv x_1 \bmod m_1$
$x \equiv x_2 \bmod m_2$] $x = c_1 m_1 + c_2 m_2$

$x_1 \equiv x \equiv c_2 m_2 \bmod m_1 \rightarrow c_2 \equiv m_2^{-1} x_1 \bmod m_1$
$x_2 \equiv x \equiv c_1 m_1 \bmod m_2 \rightarrow c_1 \equiv m_1^{-1} x_2 \bmod m_2$

$x \equiv (m_1)^{-1}_{m_2} m_1 x_2 + (m_2)^{-1}_{m_1} m_2 x_1 \pmod{m_1 m_2}$

Note 7: RSA

For primes p, q → find e coprime to $(p-1)(q-1)$

- public key: $N = (pq, e)$ $N = pq$ ciphertext = encrypted msg = y
- priv key: $d = e^{-1} \bmod (p-1)(q-1)$
- encryption: $E(x) = x^e \bmod N$
- decryption: $D(y) = y^d \bmod N = x^{ed} \equiv x \bmod (p-1)(q-1)$ →

$ed \equiv 1 \bmod (p-1)(q-1) = 1 + k(p-1)(q-1)$

$x^{ed} - x = x^{1+k(p-1)(q-1)} - x = x(x^{k(p-1)(q-1)} - 1)$ → divisible by p AND q (primes) → divisible by product $pq = N$

same for q
- case 1) x not multiple of p → $x^{p-1} \equiv 1 \bmod p$ → $x^{k(q-1)(p-1)} - 1 \equiv 0 \bmod p$ ✓
- case 2) x multiple of p → x is factor so divisible by p

✱not efficient

Fermat's Little Thm: $a^{p-1} \equiv 1 \bmod p$ & $a^p \equiv a \bmod p$ — a, p must be coprime

$S = \{1, 2, \cdots, p-1\}$
$S' = \{a \bmod p, 2a \bmod p \ldots (p-1)a \bmod p\}$] Take product: $(p-1)! \bmod p$; $a^{p-1}(p-1)! \bmod p$

$(p-1)! \equiv a^{p-1}(p-1)! \equiv 1 \bmod p$ since $(p-1)!$ has inverse mod p

Prime # Thm: $\pi(n) \geq \frac{n}{\ln n}$ for $n \geq 17$

where $\pi(n)$ is # of primes $\leq n$ → # of multiplications needed is $O \log N$ bc $O \log N$ bits in N

✱find p and q → $x^e \bmod N$ and $y^d \bmod N$

✱ From d: $de - 1 = k(p-1)(q-1)$

$\frac{de-1}{k} = pq - p - q + 1$

$\phi(n) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_n}) \rightarrow a^{\phi(n)} \equiv 1 \bmod n \rightarrow a^{\phi(n)+1} \equiv a \bmod n$

Note 8: Polynomials

property 1: nonzero polynomial of degree d w/ d roots

property 2: $d+1$ pairs of points (x_i distinct) uniquely defines a polynomial of degree d

proof by contradiction
↳ assume another polynomial $q(x)$ and $r(x) = p(x) - q(x)$

Finite fields: $GF(p)$ → mod p (coefficients + var in mod p space)
↑ prime

Lagrange Interpolation: $\frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)}$ $(1,1)(2,2)(3,4)$

$\Delta_1 = \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{(x-2)(x-3)}{2}$

$\Delta_2 = \frac{(x-1)(x-3)}{(2-1)(2-3)} = \frac{(x-1)(x-3)}{-1}$

$\Delta_3 = \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{(x-1)(x-2)}{2}$

$p(x) = \sum_i y_i \Delta_i(x) = 1\Delta_1(x) + 2\Delta_2(x) + 4\Delta_3(x)$

Each pt in $GF(p)$ has p possibilities
✱use Lagrange in proofs
mod (m) → m^m diff functions

$P(x) = a_d x^d + a_{d-1} x^{d-1} + \ldots + a_1 x + a_0$

polynomial div: $p(x), q(x) \rightarrow p(x) = q(x)h(x) + r(x)$

# of polynomials of deg $\leq d$ over $GF(p)$ through $d+1-k$ points $= p^k$

Secret sharing (under GF(p) ppl) → needs to be greater than $d+1$ points

$P(0)$ = secret, $P(1) \cdots P(n)$ given to all ppl

$P(x)$ = polynomial of degree $k-1$ &

give everyone distinct pair $(i, P(i))$

✱**weighting scheme:** give diff ppl more shares

✱**diff committees:** create polynomial for each committee and dist shares
↳ need all committee secrets to reveal overall secrets

Note 9: Error Correcting Codes

- Erasure errors: k packets lost, msg length n: need to send $n+k$
  $P(x)$ of deg $n-1$ needs n points ✱LaGrange & GF mod p✱
- General errors: k packets corrupted, msg length n: need to send $n+2k$
  $\underbrace{P(i)E(i)}_{\substack{Q(i) \\ n+k \text{ unknowns}}} = \underbrace{r_i E(i)}_{k \text{ unknowns}}$ = n+2k unknowns

Berlekamp-Welch $P(x) = \frac{Q(x)}{E(x)}$

$E(x) = (x - e_1)(x - e_2) \cdots (x - e_k)$ where e_i = error location

$Q(x) = P(x)E(x)$ → degree $n+k-1$ ✱trivially true at error pts where $E(i) = 0$

$Q(i) = r_i E(i)$ for all transmitted → solve system for $E(x)$, $Q(x)$

$Q(x) = a_1 x^p + a_1 x^{p-1} + \ldots + a_p$ → solve for coeff

Note 10 Counting

Hockey stick
$\binom{n}{k+1} = \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k}{k}$

First Rule of Counting: multiply # of ways for each choice $n_1 \times n_2 \times \cdots \times n_k$

Permutations of n objects $= n!$
subsets of k from n → $\frac{n!}{k!(n-k)!} = \binom{n}{k}$

2nd Rule of counting: count ordered arrangements divide by # of ways to get unordered

	w/ rep	w/o rep
order ✓	n^k	$\frac{n!}{k!}$
order ✗	$\binom{n+k-1}{k-1}$	$\binom{n}{k}$

$\binom{n}{k} = \frac{n!}{(n-k)!k!}$ = BINOMIAL COEFFICIENT # of ways to select k from n
ORDER DOES NOT MATTER → k distinct or k out of n distinct

sampling w/ replacement: n^k (ORDER MATTERS)

Stars & Bars: n objects, k groups → n stars, $k-1$ bars → $\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$

zeroth Rule: If bijection between A & B then $|A| = |B|$

combinatorial proofs: counting same thing 2 ways

$\binom{n}{k} = \binom{n}{n-k}$: # of ways to pick k-person team from n
k ppl on team or n-k NOT on team

$\binom{n}{k}$ → ppl to form teams $\binom{k}{r}$: layers → leaders of each group

2^n → picking subsets of ppl (other side counting subsets iterating over size)

k^n → n boxes w k items

$\sum_{k=1}^{n}$ → splitting items in two w/ k items on left, n-k items on right

$(\cdot)^k$ → apply procedure k times

n·whatever
k·whatever] selecting single item

addition: OR
multiplication: AND
expand & distribute

$\binom{n}{k} k! = \frac{n!}{(n-k)!}$

Note 11:

Countable: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, $\mathbb{N} \times \mathbb{N}$, finite set
Uncountable: $\mathbb{R}$, inf length binary, $\mathcal{P}(\mathbb{N})$

$A \subseteq B$ & B is countable → A is countable
$A \supseteq B$ & A is uncountable → B is uncountable

Bijection: one to one (distinct input to output)
onto (every element in range is hit)

Set S is countable if bijection between S and $\mathbb{N}$ $\{0, 1, 2 \ldots\}$

Cantor Bernstein thm: bijection if $f: A \rightarrow B$ and $g: B \rightarrow A$ ($|A| = |B|$)
$|A| \leq |B|$ $|B| \leq |A|$

Cantor Diagonalization: prove uncountability by listing out possibilities construct new possibility diff from each listed at one place (reals, binary string)
↳ proof by contradiction

Cantor set: measure = 0 and uncountable

$A \subseteq B$ & B is countable → A is countable

$|S| = k$ finite then $P(S) = 2^k$
no bijection from S to P(S) ← uncountable
$A \supseteq B$ & A is uncountable → B is uncountable

TestHalt P(x) can NOT exist — use contradiction

def TestHalt (P, x)
  subroutine: other func] if subroutine exists: testhalt exists
  if testhalt doesn't exist: impossible for subroutine to exist

Possible to check k steps execution
impossible to check printing/running k^{th} line

✱can always get n^{th} bit of π in finite time

✱write out in words what trying to count

anagrams: $\frac{w!}{r_1! r_2! \ldots r_k!}$ ← w = # of letters; r = times letter is repeated

permutation w/ no fixed pts
Derangements: $D_n = (n-1)(D_{n-1} + D_{n-2}) = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$

Principle of Exclusion-Inclusion: $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$
↳ add/subtract all combos ↳ prove w/ induction on n

Note 13: Probability

Ω = all possible outcomes

$0 \le P(\omega) \le 1, \forall \omega \in \Omega, \sum_{\omega\in\Omega} P(\omega) = 1$

Uniform: $P(\omega) = \frac{1}{|\Omega|}, \forall \omega \in \Omega$

$P(A) = \frac{\#\text{ of points in } A}{\#\text{ of points in } \Omega} = \frac{|A|}{|\Omega|}$ $\quad P(\bar{A}) = 1 - P(A)$

Note 14: cond prob

$P(\omega|B) = \frac{P(\omega)}{P(B)}$ for $\omega \in B$

$P(A|B) = \frac{P(A\cap B)}{P(B)}$

$\boxed{\text{Bayes: } P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$

Total Prob: $P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$

$= \sum_{i=1}^{n} P(B|A_i)P(A_i)$

Independence: $P(A\cap B) = P(A)P(B)$ or $P(A|B) = P(A)$

Union bound: $P(\cup_{i=1}^n A_i) \le \sum_{i=1}^n P(A_i)$

Pairwise indep $\ne$ mutual indep

Inclusion-Exclusion: *DRAW IT OUT!!

$P[A_1 \cup A_2 \cup A_3] = P[A_1] + P[A_2] + P[A_3] - P[A_1\cap A_2] - P[A_1 \cap A_3]$

$- P[A_2 \cap A_3] + P[A_1 \cap A_2 \cap A_3]$

$P(\bigcup_{i=1}^n A_i) = \sum_{k=1}^{n} (-1)^{k-1} \left(\sum_{1\le i_1 < \dots < i_k \le n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right)$

*Symmetry: third coin in seq = first coin

Note 15: Rand Var

Random Var X: assigns sample ω to $X(\omega)$

Distribution: vals + associated probabilities

$\{(a, P(X=a)) \mid a \in A\}$

Bernoulli: used as indicator RV – coinflip/success

Binomial: $P(X=i)$ = i successes in n trials – multiple coin flips w/ success probability p (n indep Bernoullis)

$X \sim Bin(n,p), Y \sim Bin(m,p)$ indep $\rightarrow X+Y \sim Bin(n+m, p)$

Hypergeometric: $P(X=k)$ = k successes in N draws w/o rep from size N pop w/ B objs (as success)

– sampling balls w/o replacement

Joint Dist: $P(X=a, Y=b)$

↳ marginal dist: $P[A=a] = \sum_{b\in B} P[X=a, Y=b]$ (summing over y values)

↳ if indep: $P[X=a, Y=b] = P[X=a]P[Y=b] \quad \forall a,b$

$\boxed{\text{Expectation: } E[X] = \sum_{x\in X} x P[X=x]}$ weighted avg using prob

Linearity of expectation: $\boxed{E[X+Y] = E[X] + E[Y] \quad E[cX] = cE[X]}$ DOES NOT REQUIRE INDEP

*allows us to use indicators

INDICATOR

$E[I] = P(I=1)\cdot 1 + P(I=0)\cdot 0 = P(I=1)$

① Have some quantity X we want expectation

② Break X into sum of indicators

$X = X_1 + X_2 + \dots + X_n$

③ By linearity of expectation: $E[X] = E[X_1 + \dots + X_n]$

$= E[X_1] + \dots + E[X_n]$

④ Since $E[X_i] = P(X_i = 1) \rightarrow$ compute probabilities and sum

LOTUS: calculate exp value of function of RV w/o knowing dist of rand var

$\boxed{E[g(X)] = \sum_{x\in X} g(x)P(X=x)}$ $\quad E[X^2] = \sum_{x\in X^2} x^2 P(X=x)$

Note 16: var + covar

Variance: avg squared dev away from mean

$\boxed{\begin{array}{l} Var(X) = E[(X-\mu)^2] = E[X^2] - E[X]^2 \\ Var(cX) = c^2 Var(X) \\ Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y) \\ \text{↳ indep: } Var(X+Y) = Var(X) + Var(Y) \end{array}}$

standard dev: $\sigma(X) = \sqrt{Var(X)}$

Note 19: Geometric + Poisson

Geometric: $P(X=i)$ = exactly i trials until success, $X-1$ failures

↳ memoryless: $P(X > a+b \mid X > a) = P(X > b)$ waiting >b units has same probability

*probability of success doesn't depend on # of failures in prev trials

Coupon collector: n distinct items w/ equal prob, X_i = # of tries before i^{th} new item, given $i-1$ alr

$S_n = X_1 + X_2 + \dots + X_i$ before getting all items

$X_i \sim Geom\left(\frac{n-i+1}{n}\right)$ bc $i-1$ old $\rightarrow n-i+1$ new items

$E[S_n] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \frac{n}{n-i+1} = n\sum_{i=1}^n \frac{1}{i} \approx n(\ln n + 0.5772)$

Poisson: λ = avg # of successes in unit of time

↳ rare events ie # of births/hr or bus arrivals/hr

$X \sim Poisson(\lambda), Y \sim Poisson(\mu)$ INDEP $\rightarrow X+Y \sim Poisson(\lambda+\mu)$

$X \sim Bin(n, \frac{\lambda}{n})$ w/ $\lambda > 0$ constant: as $n \rightarrow \infty$, $X \rightarrow Pois(\lambda)$

Note 21: Continuous + Joint

$P[X=x \mid Y=y] = \frac{P[X=x, Y=y]}{P[Y=y]}$ $\quad E[X|Y=y] = \sum_{x\in A} x \cdot P[X=x \mid Y=y]$

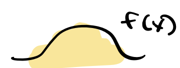
Law of Iterated Expec: $E[X] = E[E[X|Y]] = \sum_{y\in B} E[X|Y=y]P[Y=y]$

↳ $Y = X_1 + X_2 + \dots + X_n \rightarrow E[Y] = E[E[Y|N]]$

$= \sum_n E[Y|N=n]P[N=n] = \sum_n n E[X_i]P[N=n]$

$= E[X_i]\sum n P[N=n] = E[X_i]E[N]$

PDF: $f(x): \mathbb{R} \rightarrow \mathbb{R}$ $\quad \boxed{\begin{array}{l} ① f_X(x) \ge 0 \text{ for } x \in \mathbb{R} \\ ② \int_{-\infty}^{\infty} f(x)dx = 1 \end{array}}$



↳ Dist of X: $P[a \le X \le b] = \int_a^b f(x)dx$ for $a \le b$

CDF: $F_X(x) = P[X \le x] = \int_{-\infty}^{x} f_X(z)dz$ $\quad \boxed{* \frac{dF_X(x)}{dx} = f_X(x)}$

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

LOTUS: $\boxed{E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx}$

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

$= \boxed{\int_{-\infty}^{\infty} x^2 f(x)dx - \left(\int_{-\infty}^{\infty} x f(x)dx\right)^2}$

Joint Dist: $P(a \le X \le b, c \le Y \le d)$ $\quad$ ① $f_{XY}(x,y) \ge 0, \forall x,y \in \mathbb{R}$ ② $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$

$P[a \le X \le b, c \le Y \le d] = \int_c^d \int_a^b f(x,y)dxdy$ for all $a \le b, c \le d$

↳ marginal: $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$ (integrate all over y)

↳ indep: $f_{XY}(x,y) = f_X(x)f_Y(y)$

↳ cond prob: $f_{X|A}(x) = \frac{f_X(x)}{P(A)}$, $f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$

$E[X]^2 = \sum_{i=1}^n E[X_i]$

Var as sum of indicators: $E[X]^2 = E[X] + \sum_{i\ne j} E[X_i X_j]$

$Var[X] = E[X] + \sum_{i\ne j} E[X_i X_j] - E[X]^2 = \sum_{i=1}^n E[X_i]$

measure of association between X, Y

Covariance: $\boxed{\begin{array}{l} Cov(X,Y) = E[XY] - E[X]E[Y] \\ Cov(X,X) = Var(X) \end{array}}$

↳ bilinear $Cov(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(X_i, Y_j)$

↳ X, Y indep $\rightarrow Cov(X,Y) = 0$ BUT CONVERSE NOT TRUE!

$X = \begin{cases} 1 & p=0.5 \\ -1 & p=0.5 \end{cases}$ $\quad Y = \begin{cases} 1 & X=-1, p=0.5 \\ -1 & X=-1, p=0.5 \\ 0 & X=1 \end{cases}$

correlation: $corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$ $\quad -1 \le corr(X,Y) \le 1$

Note 21: continuous

Exponential distribution = continuous analog to geometric dist

- memoryless: $P(X>t+s \mid X>t) = P(X>s)$
- $P(X<Y \mid \min(X,Y)>t) = P(X<Y)$
- $X \sim Exp(\lambda_x), Y \sim Exp(\lambda_y)$ indep: $\min(X,Y) \sim Exp(\lambda_x+\lambda_y)$ & $P(X\leq Y) = \frac{\lambda_x}{\lambda_x+\lambda_y}$
- $aX \sim Exp(\frac{\lambda}{a})$

Note 17: inequalities

Markov's inequality: For nonnegative RV X

$\boxed{P[X \geq c] \leq \frac{E[X]}{c}}$ for any positive constant c

Generalized Markov: $\boxed{P(|Y| \geq c) \leq \frac{E[|Y|^r]}{c^r}}$ for $c, r > 0$

Chebyshev's inequality: $P(|X-\mu| \geq c) \leq \frac{Var(X)}{c^2}$ for $\mu = E[X]$, $c > 0$

$\hookrightarrow P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$ for $\sigma = \sqrt{Var(X)}$ $k>0$

Normal distribution/Gaussian

- if $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$: $Z = aX + bY \sim N(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2)$
- if $X \sim N(\mu, \sigma^2)$ then $Y = \frac{X-\mu}{\sigma}$
- if $Y \sim N(0,1)$ then $X = \sigma Y + \mu \sim N(\mu, \sigma^2)$
- $X \sim N(0,1)$ and $Y \sim N(0,1)$ indep standard normal random $Z = aX + bY \sim N(0, a^2 + b^2)$

Central Limit Theorem:

if $S_n = \sum_{i=1}^n X_i$, all X_i iid w/ mean μ and var σ^2

$\frac{S_n}{n} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$; $\frac{S_n - n\mu}{\sigma\sqrt{n}} \approx N(0,1)$

$$\boxed{Z_n = \frac{S_n - E(S_n)}{\sigma_{S_n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}}$$

*95% confidence: $\delta = 0.05$ confidence level

confidence intervals:

- proportions: $\boxed{P(|\hat{p} - p| \geq \varepsilon) \leq \frac{Var(\hat{p})}{\varepsilon^2} \leq \delta}$
- $\hat{p}$ = proportion of success in n trials: $Var(\hat{p}) = \frac{p(1-p)}{n}$

$\rightarrow \boxed{n \geq \frac{1}{4\varepsilon^2\delta}}$

- for means: $\boxed{P(|\frac{1}{n}S_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} = \delta}$

- Law of large numbers: $P[|\frac{1}{n}S_n - \mu| < \varepsilon] \to 1$ as $n \to \infty$ w/ all X_i's iid mean μ, variance σ^2

$S_n = X_1 + X_2 + \dots + X_n$

$\rightarrow \varepsilon = \frac{\sigma}{\sqrt{n\delta}}$, interval $= S_n \pm \frac{\sigma}{\sqrt{n\delta}}$

DISCRETE DISTRIBUTIONS

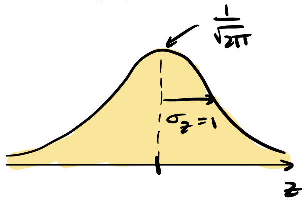
name	parameters	$P(X=k)$	$P(X\leq k)$	$E[X]$	$var(X)$	support
UNIFORM	uniform(a,b)	$\frac{1}{b-a+1}$	$\frac{k-a+1}{b-a+1}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$	$X \in [a,b]$
BERNOULLI	Bernoulli(p)	$\begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	–	p	$p(1-p)$	$X \in \{0,1\}$
BINOMIAL	Bin(n,p)	$\binom{n}{k}p^k(1-p)^{n-k}$	–	np	$np(1-p)$	$X \in \{0,1,2 \cdots\}$
GEOMETRIC	Geom(p)	$p(1-p)^{k-1}$	$1-(1-p)^k$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$X \in \{0,1,2 \cdots\}$
POISSON	Poisson(λ)	$\frac{\lambda^k e^{-\lambda}}{k!}$	–	λ	λ	$X \in \{0,1,2 \cdots\}$
HYPERGEOMETRIC	hypergeometric(N,K,n)	$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$	–	$n\frac{K}{N}$	$n\frac{K(N-K)(N-n)}{N^2(N-1)}$	$X \in \{0,1,2 \cdots\}$

CONTINUOUS DISTRIBUTIONS

name	parameters	$P(X=k)$	$P(X\leq k)$	$E[X]$	$var(X)$	support
UNIFORM	uniform(a,b)	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$X \in [a,b]$
EXPONENTIAL	Exp(λ)	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$X \in [0,\infty)$
Normal/Gaussian	$N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\Phi(x)$ use the table!!	μ	σ^2	$X \in \mathbb{R}$

Law of Total Expectation: $E(X) = E(X|A)P(A) + E(X|A^c)P(A^c)$

Gaussian RV



$E(z) = 0$

$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz = 1$

$z = \frac{x-\mu}{\sigma} = \frac{1}{\sigma}x - \frac{\mu}{\sigma}$

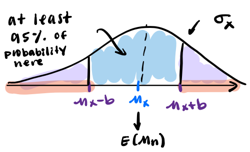
$E(z) = \frac{1}{\sigma} E(x-\mu) = 0$ ✓ constant

$Var(z) = var\left(\frac{1}{\sigma}x - \frac{\mu}{\sigma}\right) = var\left(\frac{1}{\sigma}x\right)$

$\boxed{P(|z| \le \alpha) = 2\Phi_z(\alpha) - 1}$

CLT: $\boxed{\lim_{n\to\infty} F_{z_n}(z) = \Phi_z(z) \quad \text{for every } z}$

④ Determine min # of ppl to poll



* at least $\boxed{\text{95\% confident}}$ that sample mean M_n is within $\pm\varepsilon$ of true fraction (true mean)

$1 - P(|M_n - \mu| < \varepsilon) \ge 0.95$

↓ redesign

* Probability that sample mean away from true mean within epsilon w/ at least 95% probability

$P(|M_n - \mu| \ge \varepsilon) \le 0.05$

$P(|M_n - \mu| \ge \varepsilon) \le \frac{\sigma_x^2}{n\varepsilon^2} = \frac{\mu(1-\mu)}{n\varepsilon^2} \le 0.05$

$\frac{n\varepsilon^2}{\mu(1-\mu)} \ge \frac{1}{0.05} \longrightarrow n \ge \frac{\mu(1-\mu)}{0.05\varepsilon^2}$ → since μ here, need to consider worst case

⑤ consider worst case for $\sigma_x^2 = \mu(1-\mu)$ → find μ that makes largest variance

↑ largest variance

$\frac{n\varepsilon^2}{\mu(1-\mu)} \ge \frac{1}{0.05} \longrightarrow n \ge \frac{\mu(1-\mu)}{0.05\varepsilon^2}$ (n: # of samples)

★ Partial integration

$\int_0^\infty x e^{\frac{-x^2}{2\sigma^2}} dx$

$u = -\frac{x^2}{2\sigma^2}$

$du = -\frac{x}{\sigma^2}$

$= -\sigma^2 \int_0^\infty e^u du$

$x\,dx = -\sigma^2 du$

$\boxed{\text{CRT: } x \equiv aq(q^{-1} \bmod p) + bp(p^{-1} \bmod q) \pmod{pq}}$

# of ways to have k #s that add up to n

$x_1 + x_2 + \dots + x_k = n$

$(y_1 + 1) + (y_2 + 1) + \dots + (y_k + 1) = n$

$y_1 + y_2 + \dots + y_k = n - k$

$\binom{n-1}{k-1} = \binom{n-1}{n-k}$ ways

$x_i = y_i + 1$

{ n−k stars, k−1 bars

$\boxed{\text{LOTUS continuous: } E[x^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx}$

Find density from CDF

$1 = \iint_A f(x,y)\,dx\,dy = \iint_A c\,dx\,dy = c\iint_A dx\,dy \rightarrow c = \frac{1}{\text{area}}$