CS 70 MT Crib Sheet DIVIT RAWAL

$A \subseteq B \Rightarrow |A| \leq |B|$ $A \subset B \Rightarrow |A| < |B|$ $\boxed{0 \in \mathbb{N}}$ Either P or $\neg P$ is T not both

always true - tautology. $P \Rightarrow Q$ only false when P is T & Q is F

equivalent to contrapos. $\neg Q \Rightarrow \neg P$. Vacuously true if P is F. $P \Rightarrow Q \equiv \neg P \vee Q$

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

$\neg(\forall x \; P(x)) \equiv \exists x \; \neg P(x)$ $\neg(\exists x \; P(x)) \equiv \forall x \; \neg P(x)$

Direct Proof: $P \Rightarrow Q$, assume P, show Q

Proof by Contraposition: Assume $\neg Q$ show $\neg P$

Proof by Cases

Pf by Contradiction: Assume $\neg P$ ... R ... $\neg R$ thus P

Induction: BC $P(0)$

IH $P(k)$ $k \geq 0$

IS $P(k+1)$

sometimes easier to prove stronger claim

Strong Induction IH $P(k)$ $\forall k \geq 0$

Well-Ordering Principle: for $S \subseteq \mathbb{N}$ & $S \neq \emptyset$, S has a smallest element

Stable Matching / Prop & Rej

- always halts
- no rogue couples / always stable match
  - J & C both prefer each other over current matching
- Improvement Lemma
  - If J makes offer to C on kth day, every day after, she has an offer at least as good as J
- always terminates w/ a matching
- Proposer Optimal
  - Candidate Pessimal

Tree - connected & acyclic; +1 edge → cycle

- $|V| = |E| + 1$; -1 edge disconnects

Bipartite Graph: $V = L \cup R$ & $E \subseteq L \times R$

degree(u) = $|\{v \in V \mid \{u, v\} \in E\}|$

Path - seq. of non-repeating ~~edges~~ vertices

Cycle - path starting & ending @ same v

Walk - path but may repeat vertices

Tour - walk starting & ending @ same v

Connected - path b/w any 2 vertices

- any graph consists of $k \leq |V|$ connected components

Eulerian uses each edge exactly once

- G has a ET iff G is even (all edges have even deg) & connected (except isolated vertices)

Euler's Formula: $v + f = e + 2$ (connected, planar) $e \leq 3v - 6$ ~~for each~~ $\forall$ planar graphs

A graph is non planar iff it contains $K_{3,3}$ or K_5. Complete graphs contain max. num. edges.

Hypercubes: each vertex is a bitstring where v_1 & v_2 have an edge iff they differ by exactly one bit pos.

- 2^n ~~edges~~ vertices $\Rightarrow n2^{n-1}$ edges

Use shrink & grow for graph proofs + ind.

mod is distributive

Euclid's Algo. For $x \geq y \geq 0$, $\gcd(x, y) = \gcd(y, x \bmod y)$

- iterate until $x \bmod y = 0$

x has an inverse mod m iff $\gcd(m, x) = 1$

Extended Euclid's Algo.

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

```
algo extended_euclid(x, y)
    if y = 0 return (x, 1, 0)
    else
        (d, a, b) := extended_euclid(y, x mod y)
        return ((d, b, a - (x div y) * b))
```

$x \text{ div } y = \lfloor x/y \rfloor$

Fundamental Thm. of Arithmetic:

$\forall n \in \mathbb{N}$, $n = p_1^{k_1} \cdots p_n^{k_n}$ for p_i prime

Chinese Remainder Thm.

For m, n coprime, $\exists! \; x$ s.t.

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

Full Version:

let $n_1, n_2, \ldots, n_k \in \mathbb{Z}_{>0}$ coprime. Then for any seq. of integers a_i $\exists! \; 0 \leq x < N = \prod_{i=1}^{k} n_i$ s.t.

$x \equiv a_1 \bmod n_1$

$\vdots$

$x \equiv a_k \bmod n_k$

$x = \left(\sum_{i=1}^{k} a_i b_i\right) \bmod N$

$b_i = \frac{N}{n_i} \left(\frac{N}{n_i}\right)^{-1}_{n_i}$ inv in mod n_i

RSA let p, q be large primes, $N = pq$

def e s.t. $\gcd(e, (p-1)(q-1)) = 1$

public key (N, e) $E(x) = x^e \bmod N$

private key $d = e^{-1} \bmod (p-1)(q-1)$ $D(y) = y^d \bmod N$

FLT: $a^{p-1} \equiv 1 \bmod p$ for p prime & $\gcd(a, p) = 1$

$\Downarrow$

$a^p \equiv p \bmod p$

Prime Num. Thm. $\pi(n) = \#$ primes $\leq n$

$\forall n \geq 17$, $\pi(n) \geq \frac{n}{\ln(n)}$

a is a root if $p(a)=0$

↳ nonzero polynomial of degree d has $\leq d$ roots

↳ $\exists!$ p of deg. d defined by $d+1$ points

Lagrange Interpolation

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

secret sharing over mod p p prime

$p(x) = c(x-a_1)\dots(x-a_d)$

$p(x)$ in $GF(m)$ has m choices for y_{d+1} given d points

ECCs

↳ $n+e$ points guards against e erasure errors → $n+2k$ for k general errors

↳ $P(i)E(i) = r_i E(i)$ $\quad E(x) = (x-e_1)\dots(x-e_k)$

$Q(x) :\equiv P(x)E(x)$ of deg $n+k-1$ & coeff of x^k in $E(x)$ is 1

↳ $Q(i) = r_i E(i) \Rightarrow P(x) = Q(x)/E(x)$

Prop Log

can move quantifiers around vars.

$a^{\mathrm{mod}\,k} \equiv a^{(k-m)}$

E_Euclid

$ax+by = \gcd(x,y)$

backtrack Euclidian Algo

Counting: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ $\quad \binom{n+k-1}{k-1}$ ways to distribute n obj among k recip.

e-euclid ex.

$17x \equiv 1 \mod 43$

Breaking RSA if know d

$de-1 = k(p-1)(q-1)$ $\quad k \leq e$

so, can find $pq=N$ by knowing d, e, k

GCD by Eu. Algo.

$43 = 2\cdot 17 + 9$

$17 = 1\cdot 9 + 8$

$9 = 1\cdot 8 + 1$

$8 = 8\cdot 1 + 0$

$\Rightarrow \gcd(43,17) = 1$

E_Euclid

$1 = 9 - 1\cdot 8$ $\quad$ w/ $8 = 17 - 1\cdot 9$

$1 = 9 - 1(17 - 1\cdot 9)$

$1 = 9 - 1\cdot 17 + 1\cdot 9$

$1 = 2\cdot 9 - 1\cdot 17$

$1 = 2\cdot(43 - 2\cdot 17)$ $\quad$ w/ $9 = 43 - 2\cdot 17$

$1 = 2\cdot 43 - 4\cdot 17 + -1\cdot 17$

$1 = 2\cdot 43 - 5\cdot 17$

s.t. $5 = 17^{-1} \mod 43$

Graph Stuff

planar:

$\sum_{i=1}^{f} s_i = 2e$ (sides of f)

$s_i \geq 3 \ \forall i$

$\Rightarrow 3f \leq 2e$

For bipartite $K_{3,3}$

$e \leq 2v-4$

odd cycles not possible in bipartite graphs

sum of deg = 2 # edges

Pf by contradiction for stable matching

Hamiltonian visits each vertex exactly once

Handshake Lemma:

$$\sum_{v\in V} \deg(v) = 2|E|$$

Eulerian tour iff $\forall v\in V$ $\deg(v)\,\%\,2 = 0$

$s \geq 3$ for planar graph

CS 70 Final Crib Sheet

DIVIT RAWAL

Countability & Computability

Injective (one-to-one): $x \neq y \Rightarrow f(x) \neq f(y)$

Surjective (onto): $\forall y\ \exists x\ \ f(x) = y$

S is countable if $\exists$ bijection b/w S & $A \subset \mathbb{N}$

$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$

$|P(S)| > |S|$ for $|S| = k$, $|P(S)| = 2^k$

Halting Problem Contradiction

Turing(P)

if TestHalt(P, P) = "yes", loop forever

else halt

Turing(Turing) is contradictory

Discrete Probability

$P[A] = \sum_{\omega \in A} P[\omega] = \frac{|A|}{|\Omega|}$ $\quad P[A|B] = \sum_{\omega \in A\cap B} P[\omega|B] = \sum_{\omega \in A \cap B} \frac{P[\omega]}{P[B]} = \frac{P[A\cap B]}{P[B]}$ $\quad P[A] = 1 - P[\bar{A}]$

$P[B|A] = \frac{P[A \cap B]}{P[A]}$ (Bayes Rule) $\quad P[A|B] = \frac{P[B|A]P[A]}{P[B|A]P[A] + P[B|\bar{A}](1-P[A])}$ (Total Prob. Rule)

For $\Omega = \bigsqcup_i A_i$ $\quad P[B] = \sum_{i=1}^n P[B\cap A_i] = \sum_{i=1}^n P[B|A_i]P[A_i]$; $\quad P[A_i|B] = \frac{P[B|A_i]P[A_i]}{P[B]} = \frac{P[B|A_i]P[A_i]}{\sum_{j=1}^n P[B|A_j]P[A_j]}$

Independent iff $P[A\cap B] = P[A] \times P[B]$ $\quad P[A \cap B] = P[A]\,P[B|A]$

Inc-Excl. $P[\bigcup_{i=1}^n A_i] = \sum_{i=1}^n P[A_i] - \sum_{i<j} P[A_i \cap A_j] + \sum_{i<j<k} P[A_i \cap A_j \cap A_k] - \cdots + (-1)^{n-1} P[A_1 \cap A_2 \cdots \cap A_n]$

Union Bound $P[\bigcup_{i=1}^n A_i] \leq \sum_{i=1}^n P[A_i]$

Random Variables

$X \sim Ber(p) \Rightarrow P[X=i] = \begin{cases} p & \text{if } i=1 \\ 1-p & \text{if } i=0 \end{cases}$ $\quad X \sim Bin(n,p) \Rightarrow P[X=i] = \binom{n}{i} p^i (1-p)^{n-i}$ for $i = 0, 1, \dots, n$

$Y \sim Hypergeometric(N, B, n) \Rightarrow P[Y=k] = \frac{\binom{B}{k}\binom{N-B}{n-k}}{\binom{N}{n}}$ $\quad P[X=a] = \sum_{b\in B} P[X=a, Y=b]$ ← marginal dist. of X

Indep. if $P[X=a, Y=b] = P[X=a]P[Y=b]$ $\forall a, b$

$E[X] = \sum_{a \in A} a \times P[X=a]$

$E[X+Y] = E[X] + E[Y]$; $E[cX] = cE[X]$

Functions of R.V.'s

$P_Y[Y=y] = \sum_{x : f(x)=y} P_X[X=x] \Rightarrow E[Y] = \sum_x f(x) P_X[X=x] \Leftrightarrow E[f(X)] = \sum_x f(x) P_X[X=x]$ (LOTUS)

$Var(X) = E[(X-\mu)^2]$ for $E[X] = \mu$ $= E[X^2] - \mu^2 = E[X^2] - E[X]^2$ & $Var(cX) = c^2 Var(X)$

For X, Y indep, $E[XY] = E[X]E[Y]$ & $Var(X+Y) = Var(X) + Var(Y)$ $\quad X, Y$ indep $\Rightarrow Cov(X,Y) = 0$

$Cov(X,Y) = E[XY] - E[X]E[Y] = E[(X-\mu_X)(Y-\mu_Y)]$ $\quad Cov(X,X) = Var(X)$

$Cov(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(X_i, Y_j)$ $\quad Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

Coupon Collecting: expected boxes for n coupons $\approx n(\ln(n) + \gamma_E)$, $\gamma_E \approx .5772$

$Corr(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \in [-1, 1]$

$X \sim Geom(p) \Rightarrow P[X=i] = (1-p)^{i-1} p$; $\quad E[X] = \sum_{i=1}^\infty P[X \geq i] = 1/p$, $Var(X) = \frac{1-p}{p^2}$

↳ Memorylessness: $P[X > n+m \mid X > m] = P[X > n]$

$X \sim Poisson(\lambda) \Rightarrow P[X=i] = \frac{\lambda^i}{i!} e^{-\lambda}$ $i \geq 0$, $E[X] = \lambda$, $Var[X] = \lambda$

$X \sim Poisson(\lambda)$, & $Y \sim Poisson(\mu) \Rightarrow X+Y \sim Poisson(\lambda + \mu)$ $\quad \lim_{n\to\infty} Binom(n, \frac{\lambda}{n}) = Poisson(\lambda)$

Markov's Ineq (only for r.v.'s ≥ 0)

$P[X \geq c] \leq \frac{E[X]}{c}$ for $c > 0$

Chebyshev's Ineq

$P[|X-\mu| \geq c] \leq \frac{Var(X)}{c^2}$

for $E[X] = \mu$ & $\sigma = \sqrt{Var(X)}$, $P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$

Law of Large Numbers: Let $X_1, X_2, \dots$ be seq of iid r.v. w/ $E[X_i] = \mu\ \forall i$. Then, $S_n = X_1 + \dots + X_n$ satisfies

$P[|\frac{1}{n}S_n - \mu| \geq \varepsilon] \to 0$ as $n \to \infty$ for $\varepsilon > 0$

Prediction

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

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

Wald's Id: $Y = X_1 + \dots + X_n$ for X_i iid

$E[Y] = E[X_1]E[N]$

$E[X]$ mimizes MSE: $E[(X-x)^2]$

LinReg: $\tilde{X} = X - E[X]$, $\tilde{Y} = Y - E[Y]$

$Var(X) = Var(\tilde{X})$, $Var(Y) = Var(\tilde{Y})$

$P[X=x|Y=y] = \dfrac{P[X=x, Y=y]}{P[Y=y]}$

ass. $b=0$ $\min E[(\tilde{Y}-m\tilde{X})^2]$ $Cov(X,Y) = Cov(\tilde{X},\tilde{Y})$

$\Rightarrow m = \dfrac{E[\tilde{X}\tilde{Y}]}{E[\tilde{X}^2]}$ $\quad L(X) = \dfrac{Cov(X,Y)}{Var(X)}(X - E[X]) + E[Y]$

$1 - (Corr(\tilde{X},\tilde{Y})^2)$ - how much the var is explained by a linear estimator given X

Continuous Probability Distributions

PDF: $f(x) \geq 0 \ \forall x \in \mathbb{R}$; $\int_{\mathbb{R}} f(x)\,dx = 1$; $P[a \leq x \leq b] = \int_a^b f(x)\,dx$ $\quad E[X] = \int_{\mathbb{R}} x f(x)\,dx$

CDF: $F(x) = P[X \leq x] = \int_{-\infty}^{x} f(z)\,dz$; $f(x) = \dfrac{dF(x)}{dx}$ $\quad Var(X) = \int_{\mathbb{R}} x^2 f(x)\,dx - \left(\int_{\mathbb{R}} x f(x)\,dx\right)^2$

Joint Dists. corr. to double integrals. Indep. if $\forall a<b, c<d$

$P[a \leq X \leq b, c \leq Y \leq d] = P[a \leq X \leq b] \cdot P[c \leq X \leq d]$

For X, Y indep. $f(x,y) = f_X(x) f_Y(y)$

$X \sim Exp(\lambda) \Rightarrow f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}$ $\quad E[X] = \dfrac{1}{\lambda}$ $\quad Var(X) = \dfrac{1}{\lambda^2}$

$P[X > t] = \int_t^\infty \lambda e^{-\lambda x}\,dx = e^{-\lambda t}$

Normal Distribution

$X \sim N(\mu, \sigma^2) \Rightarrow f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$

$E[X] = \mu$ $\quad Var(X) = \sigma^2$

$X \sim N(\mu, \sigma^2)$ then $Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$

Let $X \sim N(0,1)$, $Y \sim N(0,1)$ indep.

Then $Z = aX + bY \sim N(0, a^2+b^2)$

$\Rightarrow X \sim N(\mu_X, \sigma_X^2)$ & $Y \sim N(\mu_Y, \sigma_Y^2)$ indep.

Then, $Z = aX + bY \sim N(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$

Central Lim. Thm.:

Let $X_1, X_2, \dots$ be iid r.v. w/ $E[X_i] = \mu$

$Var(X_i) = \sigma^2 \ \forall i$ Let $S_n = \sum_{i=1}^n X_i$

Then, $\dfrac{S_n - n\mu}{\sigma\sqrt{n}} \to N(0,1)$ as $n \to \infty$

Markov Chains

- memoryless: $P[X_{n+1} = j | X_n = i] = P_{ij}$

$\pi_n = \pi_0 P^n \iff P[X_n = i_n] = [\pi_0 P^n](i_n)$ $\quad \sum_{j=1}^{K} P_{ij} = 1 \ \forall i \in K$ $\quad \pi_0 = \{\pi_0(i) \mid i \in K\}$ $\quad \sum_{i \in K} \pi_0(i) = 1$

π is invariant if $\pi = \pi P$

Markov Chain is irreducible if it can go from any i to any j in finite steps (necessary for convergence)

Period of $i = d(i) = \gcd\{n > 0 \mid P_{ii}^n > 0\}$

$\Rightarrow$ In an irreducible MC, period is same for all states

A finite irreducible MC is aperiodic if $d(i) = 1 \ \forall i$

FTMC: For any X_0 & state i $P[X_n = i] \Rightarrow \pi(i)$ as $n \to \infty$

$\exists$ unique inv. dist. & $\pi(i) > 0 \ \forall i$

Random Walks on Graphs: $\pi(v) = \dfrac{\deg(v)}{\sum_{v \in V} \deg(v)}$

First Step. Eqns. for $A \subset K$

$\beta(i) = 0$ for $i \in A$

$\beta(i) = 1 + \sum_{j \in K} P_{ij}\beta(j)$

Probability of A before B:

$\alpha(i) = \sum_j P_{ij}\alpha(j) \quad \forall i \notin A \cup B$

$\alpha(i) = 0 \quad \forall i \in B$

$\alpha(i) = 1 \quad \forall i \in A$

Sum Formulas

Arithmetic: $S_n = \dfrac{n}{2}(a + l)$ (a = first, l = last)

Geom: finite: $S_n = a\dfrac{1-r^n}{1-r}$ for $r \neq 1$

inf: $S = \dfrac{a}{1-r}$ for $|r| < 1$

What Dist to use when:

- Bernoulli - single trial w/ either success(1) or failure(0)
- Binom - fixed # of Bernoulli trials counting # successes
- Geom - # of trials to get to first success in repeated Bernoulli trials
- HyperGeom - sampling w/o replacement from finite pop. & counting # successes in fixed sample size
- Poisson - # events in a fixed interval
- Exp - time until next event in Poisson process
- Normal - symmetric bell-curve shaped data.