What I know *after* taking CS 30

The document serves as a review of the second half of the course.

1 Probability

- **Experiments and Outcomes: Sample Space** Every time a probabilistic question is asked, figure out the *sample space*: that is, figure out what the unknown random experiment is, and what is the set of possible outcomes. Often represented by Ω.

- **Events.** Figure out the subset of outcomes you are interested in. This subset is the *event* you are interested in.

- **The Probability Distribution.** Finally, we need to figure out the function or the *probability distribution* $\mathbf{Pr} : \Omega \to [0,1]$ such that $\sum_{\omega \in \Omega} \mathbf{Pr}[\omega] = 1$. Given this distribution, we can answer what the chance/likelihood/probability of an event $\mathcal{E}$ is: it is $\sum_{\omega \in \mathcal{E}} \mathbf{Pr}[\omega]$.

  At some level, *modelling assumptions* dictate the distribution. Make as few and as natural assumptions.

- **Tree Diagrams.** For small problem, the tree diagram which starts with our state of the world and goes through all possibilities is a sure-shot way of figuring out the probabilities of all outcomes. It gets unwieldy soon, but very useful for intuition.

- **Operations on Events.**

  - Given an event $\mathcal{E}$, the negation event $\neg\mathcal{E}$ is used to denote the event that $\mathcal{E}$ doesn't take place. That is, it is simply the subset $\neg\mathcal{E} = \Omega \setminus \mathcal{E}$. Sometimes, $\neg\mathcal{E}$ is denoted as $\overline{\mathcal{E}}$.

  $$\mathbf{Pr}[\mathcal{E}] + \mathbf{Pr}[\neg\mathcal{E}] = 1$$

  - Given two events $\mathcal{E}$ and $\mathcal{F}$, the notation $\mathcal{E} \cup \mathcal{F}$ is precisely the union of the subsets in the sample space. $\mathbf{Pr}[\mathcal{E} \cup \mathcal{F}]$ captures the likelihood that at least one of the events takes place.
  - Given two events $\mathcal{E}$ and $\mathcal{F}$, the notation $\mathcal{E} \cap \mathcal{F}$ is precisely the intersection of the subsets in the sample space. $\mathbf{Pr}[\mathcal{E} \cap \mathcal{F}]$ captures the likelihood that both the events takes place.
  - Two events $\mathcal{E}$ and $\mathcal{F}$ are *disjoint* or *exclusive* if $\mathcal{E} \cap \mathcal{F} = \varnothing$. That is, they both can't occur simultaneously. A collection of events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_k$ are *mutually exclusive* if $\mathcal{E}_i \cap \mathcal{E}_j = \varnothing$ for $i \neq j$.
  - For mutually exclusive events,

  $$\mathbf{Pr}[\mathcal{E}_1 \cup \mathcal{E}_2 \cup \cdots \mathcal{E}_k] = \sum_{i=1}^{k} \mathbf{Pr}[\mathcal{E}_i]$$

– The Inclusion-Exclusion formula (for two events, aka Baby version) tells us

$$\mathbf{Pr}[\mathcal{E} \cup \mathcal{F}] = \mathbf{Pr}[\mathcal{E}] + \mathbf{Pr}[\mathcal{F}] - \mathbf{Pr}[\mathcal{E} \cap \mathcal{F}]$$

- **Conditional Probability.** For any two events $\mathcal{A}$ and $\mathcal{B}$, we have

$$\mathbf{Pr}[\mathcal{A} \mid \mathcal{B}] = \frac{\mathbf{Pr}[\mathcal{A} \cap \mathcal{B}]}{\mathbf{Pr}[\mathcal{B}]}$$

- **Chain Rule.** For any set of events $\mathcal{A}_1, \mathcal{A}_2, \cdot, \mathcal{A}_k$,

$$\mathbf{Pr}[\mathcal{A}_1 \cap \mathcal{A}_2 \cap \cdots \mathcal{A}_k] = \mathbf{Pr}[\mathcal{A}_1] \cdot \mathbf{Pr}[\mathcal{A}_2 \mid \mathcal{A}_1] \cdot \mathbf{Pr}[\mathcal{A}_3 \mid \mathcal{A}_1 \cap \mathcal{A}_2] \cdots \mathbf{Pr}[\mathcal{A}_k \mid \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_{k-1}]$$

- **The Law of Total Probability.** For any two events $\mathcal{A}$ and $\mathcal{B}$, we have

$$\mathbf{Pr}[\mathcal{A}] = \mathbf{Pr}[\mathcal{A} \mid \mathcal{B}] \cdot \mathbf{Pr}[\mathcal{B}] + \mathbf{Pr}[\mathcal{A} \mid \neg\mathcal{B}] \cdot \mathbf{Pr}[\neg\mathcal{B}] +$$

More generally, if $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ are k *mutually exclusive events* which are *exhaustive*, that is, $\sum_{i=1}^{k} \mathbf{Pr}[\mathcal{B}_i] = 1$, then

$$\mathbf{Pr}[\mathcal{A}] = \sum_{i=1}^{k} \mathbf{Pr}[\mathcal{A} \mid \mathcal{B}_i] \cdot \mathbf{Pr}[\mathcal{B}_i]$$

- **Independence.** Two events $\mathcal{A}$ and $\mathcal{B}$ are *independent* if $\mathbf{Pr}[\mathcal{A} \cap \mathcal{B}] = \mathbf{Pr}[\mathcal{A}] \cdot \mathbf{Pr}[\mathcal{B}]$.

  Be careful when figuring out when two events are independent.

- **Random Variables.** A random variable is a *function/mapping* $X : \Omega \to$ Range from the set of outcomes to a range. Usually the range is the set of natural numbers, but it could be reals, integers, etc.

- **Expectation of a Random Variable.** The expectation of a random variable is an "weighted average" defined as

$$\mathbf{Exp}[X] := \sum_{\omega \in \Omega} X(\omega) \cdot \mathbf{Pr}[\omega]$$

- **Linearity of Expectation.** For any k random variables $X_1, X_2, \ldots, X_k$, we have

$$\mathbf{Exp}[\sum_{i=1}^{k} X_i] = \sum_{i=1}^{k} \mathbf{Exp}[X_i]$$

  One cannot overstate the importance of this above fact.

- **Independent Random Variables.** Two random variables X and Y are *independent* if for any x, y in their ranges

$$\mathbf{Pr}[X = x, Y = y] = \mathbf{Pr}[X = x] \cdot \mathbf{Pr}[Y = y]$$

  k random variables $X_1, X_2, \ldots, X_k$ are *pairwise independent* if any two of them are independent. They are *mutually independent* if for any $x_1, x_2, \ldots, x_k$, we have

$$\mathbf{Pr}[X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k] = \prod_{i=1}^{k} \mathbf{Pr}[X_i = x_i]$$

- **Expectation of Product of Mutually Independent Random Variables.** If $X_1, \ldots, X_k$ are mutually independent random variables, then

$$\mathbf{Exp}[\prod_{i=1}^{k} X_i] = \prod_{i=1}^{k} \mathbf{Exp}[X_i]$$

- **Variance of a Random Variable.** Given a random variable X, the variance $\mathbf{Var}[X]$ is defined as

$$\mathbf{Var}[X] := \mathbf{Exp}[(X - \mathbf{Exp}[X])^2] = \mathbf{Exp}[X^2] - (\mathbf{Exp}[X])^2$$

  The *standard deviation* is defined as

$$\sigma(X) := \sqrt{\mathbf{Var}[X]}$$

- **Linearity of Variance for Pairwise Independent Random Variables.** Given k *pairwise* independent random variables $X_1, \ldots, X_k$, we have

$$\mathbf{Var}[\sum_{i=1}^{k} X_i] = \sum_{i=1}^{k} \mathbf{Var}[X_i]$$

- **Concentration around the mean: Chebyshev's Inequality** For any random variable X and for any $t > 0$, we have

$$\mathbf{Pr}\Big[\ |X - \mathbf{Exp}[X]| \ \geq \ t\Big] \ \leq \ \frac{\mathbf{Var}[X]}{t^2}$$

  As a corollary we get that the probability X is ***not*** in the range $\big[\mathbf{Exp}[X] - c\sigma(X), \mathbf{Exp}[X] + c\sigma(X)\big]$ is at most $\frac{1}{c^2}$.

2 Graphs

- **Notations and Definitions.**
  - *Given an edge $e = (u, v)$, the vertices u and v are the **endpoints** of e. We say e **connects** u and v. We say that u and v are **incident** to e.*
  - *Two vertices $u, v \in V$ are **adjacent** or **neighbors** if and only if (u, v) is an edge.*
  - *The **incident edges** on v is denoted using the set $\partial(v)$. So,*

    $$\partial_G(v) := \{(u, v) : (u, v) \in E\}$$

    We lose the subscript if the graph G is clear from context.
  - *Given a vertex v, the **neighborhood** of v is the set of neighbors of v. This is denoted sometimes as $N(v)$ or sometimes as $\Gamma(v)$. So,*

    $$N_G(v) := |\{(u, v) : (u, v) \in E\}|$$

    if the graph G is clear from context.
  - *The cardinality of $N_G(v)$ is called the **degree** of vertex v. We denote it using $\deg_G(v)$. This counts the number of neighbors of v. Note that,*

    $$\deg_G(v) = |N_G(v)| = |\partial_G(v)|$$

  - *A vertex v is **isolated** if its degree is 0. That is, it has no edges connected to it.*
  - *A graph $G = (V, E)$ is called **regular** if all degrees are equal, that is, $\deg_G(v) = \deg_G(u)$ for all u and v.*
  - *Given a graph $G = (V, E)$, we use $V(G)$ to denote V and $E(G)$ to denote E. This notation is useful when we are talking about multiple graphs.*

- **The Handshake Lemma.** In any graph $G = (V, E)$,

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

- **Perambulations in Graphs.** Fix $G = (V, E)$
  - *A **walk** w in G is an* alternating sequence *of vertices and edges*

    $$w = (v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$$

    such that the ith edge $e_i = (v_{i-1}, v_i)$ for $1 \le i \le k$. Both edges and vertices can repeat.
  - *A **trail** t in G is a walk with no edges repeating.*
  - *A **path** p in a graph G is a walk with no* vertices repeated.
  - *A **closed walk** is a walk whose origin and destination are the same vertex.*

- *A **circuit** is a closed trail of length at least* 1.
- *A **cycle** is a circuit with no vertex other than the source and destination repeating.*

- **Connectivity, Forests, and Trees.**
  - u is ***reachable*** from v in G if there is a walk from u to v in G. A graph G is ***connected*** if any vertex is reachable from another vertex.
  - Walk from u to v implies a path from u to v.
  - A ***forest*** is a graph with no cycles.
  - A ***tree*** is a forest which is connected.
  - **Trees have leaves.**

- **Tree Theorem.**

  Let $G = (V, E)$ be a graph. The following are equivalent statements.

  1. *G is a tree.*
  2. *G has no cycles and adding any edge to G creates a cycle.*
  3. *Between any two vertices in G there is a unique path.*
  4. *G is connected, and deleting any edge from G disconnects the graph, and the resulting graph has exactly two connected components.*
  5. *G is connected and* $|E| = |V| - 1$.
  6. *G has no cycles and* $|E| = |V| - 1$.

- **Bipartite Graphs.**

  A graph $G = (V, E)$ is ***bipartite*** if the vertex set V can be partitioned into $V = L \cup R$ and $L \cap R = \varnothing$ such that *every* edge (x, y) has exactly one endpoint in L and the other endpoint in R.

  G is bipartite $\Leftrightarrow$ G has no cycles of ***odd*** length

- **Matchings.**

  A ***matching*** $M \subseteq E$ is a subset of edges such that no two edges in M share an endpoint. In other words, a matching is a collection of *pairwise disjoint* set of edges.

- **Matchings in Bipartite Graphs: Hall's Theorem.**

  Let $G = (L \cup R, E)$ be a bipartite graph. A matching M is an L-matching if every vertex of L is an endpoint of some edge in M.

  G has an L-matching $\Leftrightarrow$ For every subset $S \subseteq L,\ |N_G(S)| \geq |S|$

3 Numbers

- **Modular Arithmetic: Definition.**
  - $a \bmod n$ is the *unique* integer $r \in \{0, 1, 2, \ldots, n-1\}$ such that a divided by n leaves remainder r.
  - The set $\{0, 1, \ldots, n-1\}$ is called the *ring of integers modulo* n, and is denoted as $\mathbb{Z}_n$ often.
  - Two integers are *equivalent modulo* n, or $a \equiv_n b$ if and only if $a \bmod n = b \bmod n$.

- **Algebra in Modular Arithmetic.** Below, a, b, c are all integers, and n is a positive integer.
  - $a \equiv_n b$ and $b \equiv_n c$ implies $a \equiv_n c$.
  - $a \equiv_n b \quad \Rightarrow \quad (a+c) \equiv_n (b+c)$.
  - $a \equiv_n b \quad \Rightarrow \quad a \cdot c \equiv_n b \cdot c$.
  - $a \equiv_n b \quad \Rightarrow \quad a^c \equiv_n b^c$ if $c > 0$.

  But **beware** that the last two implications go only in one direction.That is,

  $$a \cdot c \equiv_n b \cdot c \quad \text{doesn't necessarily imply} \quad a \equiv_n b$$

  So you can't "divide off" c from both sides. To see this, note $2 \cdot 4 \equiv_6 5 \cdot 4 \equiv_6 2$ but $2 \not\equiv_6 5$.

  Similarly,

  $$a^c \equiv_n b^c \quad \text{doesn't necessarily imply} \quad a \equiv_n b$$

  So you can't "take $1/c$th power. To see this, note $5^2 \equiv_8 3^2 \equiv_8 1$, but $5 \not\equiv_8 3$.

- **Modular Exponentiation.** A pretty fast way to compute $a^b \bmod n$.

- **Greatest Common Divisor (GCD).**
  - $\gcd(a, n)$ *is the largest number dividing both* a *and* n.
  - *Euclid's recursive algorithm to find GCD of any two numbers.*
  - ***Bezout's Theorem:*** $\gcd(a, n) = g$ *implies the existence of two integers* x, y *such that* $xa + yn = g$.
  - *The above* (x, y) *can be found by Extended GCD algorithm.*
  - *In fact,* g *is the* smallest positive *integer which can be written as* $xa + yn$.

- **Co-prime or Relatively prime numbers.**

  Two numbers a, n are co-prime or relatively prime if and only if $\gcd(a, n) = 1$. Co-prime numbers have lots of nice properties. In particular, the following facts are useful (you should be able to prove all of them using Bezout's Theorem mentioned above).

  - *If* $\gcd(a, n) = 1$, *and* $ab \equiv_n 0$, *then* $b \equiv_n 0$. *As a consequence, we get*

– *If* $\gcd(a,n) = 1$*, and* $a \cdot b \equiv_n a \cdot c$*, then* $b \equiv_n c$*.*
– *If a* prime p *divides* a *and* p *divides* b*, then* p *divides* ab*.*
– *If* $\gcd(a,n) = 1$ *and* $\gcd(b,n) = 1$*, then* $\gcd(ab,n) = 1$.

- **The Multiplicative Inverse.**

  Co-prime numbers have *inverses*; a supremely helpful fact. For any two pair of coprime numbers a and n, the *multiplicative inverse* of a in the ring $\mathbb{Z}_n$, also called the multiplicative inverse of a modulo n, is the ***unique*** element b in $\mathbb{Z}_n$ such that $ab \equiv_n 1$. We can use the Extended Euclid's GCD algorithm to compute the multiplicative inverses.

- **Fermat's Little Theorem**.

  For any prime p and number a such that $\gcd(a,p) = 1$, we have

  $$a^{p-1} = 1 \bmod p, \quad \text{or, more concisely,} \quad a^{p-1} \equiv_p 1$$

- **Public Key Cryptography.**

  A conceptual breakthrough due to Diffie and Hellman from 1976 which allowed secrets to be shared without the need for keys to be shared. Diffie-Hellman win Turing Award in 2015.

  – *Alice wants to send a message* m *to Bob.*
  – *Bob generates* **two** *keys: a* ***public*** *key* pk *which is told to all; a* ***secret*** *key* sk *which is only known to him.*
  – *Bob also publishes two algorithms* Enc *and* Dec*.*
  – *Alice uses* $\mathsf{Enc}(m,pk)$ *to get the encrypted cipher* c*.*
  – *Bob uses* $\mathsf{Dec}(c,sk,pk)$ *to decrypt the cipher.*
  – *Eve can't figure* m *out given* $\mathsf{Enc}(m,pk)$ *and* pk*.*

- **RSA protocol.**

  A fantastic algorithm implementing public key cryptography. Invented by Rivest, Shamir, Adleman in 1978. Rivest-Shamir-Adleman awarded Turing award in 2002.

  – *Bob picks two* ***large*** *primes* p,q*. Let* $N := pq$ *and* $\phi := (p-1)(q-1)$.
  – *Bob picks another number* e *such that* $\gcd(e,\phi) = 1$.
  – *Bob figures out* $d \equiv e^{-1} \bmod \phi$*, that is,* d *is the multiplicative inverse of* e *in* $\mathbb{Z}_\phi$*.*
  – *Bob's* ***public key*** *is* (e,N)*. Bob's* ***secret*** *key is* d*.*
  – *Encryption: Alice uses* (e,N) *to encrypt* $m \mapsto m^e \bmod N$*.*
  – *Decryption: Bob uses* (d,N) *to decrypt cipher* $c \mapsto c^d \bmod N$*.*