# Contents

# 1 Disclaimer

We have looked over this as closely as we can, but there will inevitably be errors and unclear definitions. Feel free to email us or post on Piazza so everyone can learn about corrections.

*This is meant to be a supplement to your notes, not a replacement for them.*

# 2 Proof Techniques

**Contradiction** Let's say we want to prove a statement $P$. In a proof by contradiction, we assume Not $P$ (the opposite), and show that this leads to a false statement. This shows that our assumption must have been wrong, and therefore $P$ must be true.

**Set Element Method** You can prove two sets $A, B$ are equal by showing $A \subseteq B$ and $B \subseteq A$.

**Bijection** You can prove two sets are the same size by showing that there exists a bijection between them.

**Induction** In an inductive proof, we attempt to prove an infinite number of statements indexed over the integers. We demonstrate that a *specific* case is true (called the *base case*), and prove that any true case implies another of the cases must be true. This proves that any case reachable from the base case must be true.

# 3 Proof Templates

These are meant to quickly convey what a very general case of each proof technique might look like. Do not take these as gospel or assume that, because your proof looks different, it is incorrect.

**Claim**: Proposition $p$ is true.

*Proof by contradiction.* Assume, for the sake of contradiction, that $p$ is false.

...

This assumption has lead us to prove a statement which we know to be false. Thus, our assumption cannot be true, so $p$ must be true. □

**Claim**: Sets $X$ and $Y$ are equal.

*Proof by Element Method.* We will prove that sets $X$ and $Y$ are equal by demonstrating that $X \subseteq Y$ and $X \supseteq Y$.

($\subseteq$) Consider some element $x \in X$. [Prove that $x \in Y$.] Since $x \in X \Rightarrow x \in Y$, $X \subseteq Y$.

($\supseteq$) Consider some element $y \in Y$. [Prove that $y \in X$.] Since $y \in Y \Rightarrow y \in X$, $Y \subseteq X$.

Since each set is contained within the other, they must be equal. □

**Claim**: Set $A$ is of equal cardinality to set $B$.

*Bijective Proof.* Consider the function $f : A \to B$. [Define some $f$.] We will now prove that this function is a bijection:

**Injectivity** For the sake of contradiction, assume that $f$ is not injective; that is, assume there exists some $x \neq y$ such that $f(x) = f(y)$. [Prove that evaluating $f(x)$ cannot produce the same value as evaluating $f(y)$.] Thus, $f$ must be injective.

**Surjectivity** For the sake of contradiction, assume that $f$ is not surjective; that is, assume there exists some $b$ such that, for all $a \in A$, $f(a) \neq b$. [Prove that there must be an $a$ which produces $b$.] Thus, $f$ must be surjective.

Since $f$ is both injective and surjective, it is a bijection and thus $A$ and $B$ must be of equal cardinality. □

**Claim**: The predicate $P(n)$ is true for all $n \in \mathbb{Z}$, $n \geq 1$.

*Proof by induction.* We will prove this claim by induction.

**Base Case** [Show that $P(1)$ is true.]

**Induction Hypothesis** For some fixed integer $k > 1$, assume that $P(k)$ is true.

**Induction Step** [Show that $P(k+1)$ must be true if you assume $P(k)$ is true.]

□

**Claim**: The predicate $P(n)$ is true for all $n \in \mathbb{Z}$, $n \geq 1$.

*Proof by strong induction.* We will prove this claim by induction.

**Base Case** [Show that $P(1)$ is true. You may need more base cases.]

**Induction Hypothesis** For all $1 \leq i \leq k$ for some fixed $k$, assume that $P(i)$ is true.

**Induction Step** [Show that $P(k+1)$ must be true if you assume $P(1)$ through $P(k)$ are true.]

□

# 4 Proof Writing Advice

- First and foremost, read the style guide! It contains a lot of more general advice, whereas this is just a running list of advice or ideas given in class.
- Before starting your proof, do a sanity check. Try a few small cases and maybe one larger case, to make sure you aren't trying to prove a false statement.
- Say 'for the sake of contradiction' before assuming something you know to be false. It is much clearer.
- Always consider where you are working.
- Notation is not a fixed, constant idea you must follow to the letter. Notation is intended to make your life simpler by allowing you to convey a large amount of information quickly to other people who also understand the same notation. Defining your own notation is not only allowed but encouraged – *if* it makes your proof easier to read! Defining unnecessary notation, redefining or abusing existing notation without good reason, creating confusing notation, or not defining notation you have invented is extremely bad practice.
- When doing a proof by cases, you MUST consider EVERY case. If you find yourself doing more than four cases, there is probably an easier way. Cases should be disjoint.
- The opposite of "$P$ always happens" is *not* "$P$ never happens" – it is "$P$ *can* not happen". Formally $\neg(\forall x, P) = \exists x$ such that $\neg P$.
- In (strong) induction proofs, you only *need* multiple base cases if your inductive step refers to a *constant* number of steps before the current step. Thus, for example, if you use *all* previous cases, you often do not need multiple base cases; if you use case $n-3$ to prove step $n$, you should have three base cases.
- Identify the hypothesis and conclusion. Try setting up some notation. Consider different methods of argument (direct, contrapositive, contradiction, induction, combinatoric). Outline your proof.
- For a proof by induction, try out a few base cases to see if there's a pattern.
- Temporarily avoid the hard part of a problem so that you can make progress. Solve an easier version of the problem, then gradually build up its difficulty. Break the rules, then see what works and what doesn't when you add the rules back in.
- When doing induction on graphs, consider whether you should induce on vertices or edges (or, rarely, regions). If you are inducing on vertices (edges), you should justify why your inductive step *always* reduces the number of vertices (edges) within your larger case – otherwise, your inductive hypothesis will not apply.
- When you are proving a property that applies to all graphs of a specific type, you should **always** do "build down" induction – attempting to build up almost always fails, since you would need to justify why all graphs are constructed by your induction step.

# 5 Chronological Overview

## 5.1 Set Theory

**Definition 1** (Proof)**.** A *proof* is not a complex mathematical idea or a series of alterations to an equation. A proof is just a *convincing* argument that explains why a statement (or *claim*) is true.

**Definition 2** (0-1 String)**.** A 0/1 string is a sequence of 0's and 1's written in a fixed order. The set of all 0/1 strings of length $n$ is sometimes written $\{0,1\}^n$.

**Theorem 1.** The number of 0-1 strings of length $n$ is $2^n$ for all $n \in \mathbb{Z}^+$.

**Definition 3** (Set)**.** A set is an unordered collection of distinct objects.

**Definition 4** (Subset)**.** A subset $A$ of set $B$ is a set which contains only elements that are also contained in set $B$. This is written $A \subseteq B$.[1] A set with no elements (the *empty set*) is written $\varnothing$ or $\{\}$.

**Definition 5** (Order)**.** The *order* of a finite set is the number of objects it contains.

**Theorem 2.** A set with $n$ elements has $2^n$ distinct subsets.

**Definition 6** (Proposition)**.** A *proposition* is a statement with a truth value.

**Definition 7** (Predicate)**.** A *predicate* is a proposition whose truth value is dependent on the value of a variable.

**Theorem 3.** $\sqrt{2}$ is irrational.

[1] There is a really stupid argument in formal math about whether to use $\subset$ for strict subset (i.e. $A \subset B \Rightarrow A \neq B$) or not (i.e. $A$ could equal $B$). Just be clear with your notation; this *usually* means writing $\subseteq$ when you mean that $A$ might equal $B$, but don't expect that everywhere.

**Definition 8** (Cardinality)**.** A set's *cardinality* is its 'size.' If a set $A$ is finite, its cardinality is the number of elements of $A$. If $A$ is not finite, then its cardinality is more complicated.

## Operations on Sets

**Union** $A \cup B$: all the elements that are in *either* $A$ or $B$.

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

Be careful that all elements are only included in $A \cup B$ once.

**Intersection** $A \cap B$: all the elements that are in *both* $A$ and $B$.

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

Here, we should recall the existence of the empty set ($\varnothing$), which will be produced when $A$ and $B$ share nothing in common.

**Set Difference** $A \backslash B$: all elements that are in $A$, but not in $B$.

$$A \backslash B = \{x \mid x \in A \text{ and } x \notin B\}$$

This is our first non-commutative operation: $A \backslash B$ is not necessarily equal to $B \backslash A$.

**Complement** $\bar{A} = A^C$: all things not in $A$.

$$A^C = \{x \mid x \notin A\}$$

This is ill-defined by itself! We also need to know what domain we're working in – the *universe* or *universal set*, usually denoted $U$. So, better, we should say:

$$A^C = U \backslash A$$

**Cartesian Product** $A \times B$ (read '$A$ cross $B$'). Ordered pairs containing one item in $A$ and one item in $B$.

$$A \times B = \left\{(a, b) \;\middle|\; a \in A, b \in B\right\}.$$

Keep in mind that these are ordered pairs, not sets.

**Definition 9** (Power set)**.** The *power set* of a set $X$ (written $\mathcal{P}(X)$) is the set of all subsets of $X$. It includes both the empty set and $X$ itself.

## Quantifiers

**Universal Quantifier** $\forall$: *for all*. It talks about everything in your universe.

**Existential Quantifier** $\exists$: *there exists*. There is some instance (in our universe) in which this is true.

**Definition 10** (Relation)**.** A **relation** on $A \times B$ is any subset $R \subseteq A \times B$. If you see "a relation on $A$," this implicitly means the relation is on $A \times A$.

**Definition 11** (Equivalence Relation)**.** A relation on $A$ is an *equivalence relation* if it satisfies three conditions:

**Reflexitivity** $\forall a \in A, (a, a) \in R$.

**Symmetry** $\forall a, b \in A, (a, b) \in R \Rightarrow (b, a) \in R$.

**Transitivity** $\forall a, b, c \in A, (a, b) \in Rand(b, a) \in R \Rightarrow (a, c) \in R$.

*Intuition:* Equivalence relations broaden the concept of equality. So if you have an equivalence relation, you can intuitively treat it like an equals sign.

**Definition 12** (Equivalence Class)**.** Define some set $S$ and let $R$ be an equivalence relation on $S$. For any $a \in S$, $[a]_R = \{x \in S \mid (a, x) \in R\}$. This is called $a$'s *equivalence class*, and $a$ is known as a *representative* of its equivalence class.

**Definition 13** (Partition)**.** A *partition* of a set $S$ is a collection of non-empty subsets of $S$, which we will call *blocks*, such that the union of all blocks is exactly $S$, and all blocks are pairwise disjoint. In notation:

1. $\forall x \in S, x$ is in some $B_i$.
2. $\forall i, j,\ i \neq j,\ B_i \cap B_j = \varnothing$.

**Theorem 4.** The equivalence classes of an equivalence relation $R$ on $S$ form a partition of $S$.

**Definition 14** (Function)**.** A function $f : A \to B$ is a relation on $A \times B$ such that, for all $a \in A$, there exists exactly one $b \in B$ such that $(a, b) \in f$. In the statement $f : A \to B$, we call $A$ the *domain* and $B$ the *codomain*. We also have the *image* (sometimes the *range*), which is the subset of $B$ that we use. That is, the image $I$ is $y \in B$ such that $\exists\ x \in A$ s.t. $f(x) = y$.

**Definition 15** (Injective)**.** A function $f : A \to B$ is *injective* (or *one-to-one*) if, for all $y \in B$, there exists at most one $x \in A$ such that $f(x) = y$. In other words, if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.

**Definition 16** (Surjective)**.** A function $f : A \to B$ is *surjective* (or *onto*) if, for all $y \in B$, there exists $x \in A$ such that $f(x) = y$. This also means that the image and the codomain are identical.

**Definition 17** (Bijection)**.** A function is *bijective* if it is both injective and surjective (and is called a *bijection*). This is sometimes also known as an *exact correspondence*.

**Theorem 5.** There exists some bijection $f : A \to B$ if *and only* if $|A| = |B|$.

**Definition 18** (Cardinality (continued))**.** Two sets (including infinite sets) are defined to have the same *cardinality* if there exists a bijection between them. A set which has an exact correspondence with the natural numbers is said to be *countably infinite*. Otherwise, it is *uncountably infinite*.

**Definition 19** (Countable)**.** A set $X$ is *countable* if it is either finite or there exists a bijection between $X$ and the natural numbers.

**Definition 20** (Uncountable)**.** An infinite set $X$ is uncountable if there does not exist a bijection between $X$ and the natural numbers. Note that "uncountable" is *not* a specific cardinality; it only means the set is not countable, and there are different 'sizes' of uncountable-ness (which are not in the scope of this class).

## 5.2 Number Theory

**Definition 1** (Divisibility)**.** $a|b$ ($a$ *divides* $b$) for $a, b \in \mathbb{Z}$ if $\exists\ c \in \mathbb{Z}$ s.t. $b = a \cdot c$. You may also consider that $a$ is a factor of $b$.

**Definition 2** (Prime)**.** $p \in \mathbb{Z}^+$ is prime if the only positive divisors are 1 and $p$. 1 is not prime.

**Theorem 1.** There are infinitely many prime numbers.

**Definition 3** (Division Algorithm)**.** For all $a, d \in \mathbb{Z}$, there exists a unique solution to the variables $r, q$ in the equation $a = qd + r$ under the constraint that $0 \leq r < d$.

**Definition 4** (GCD)**.** $\gcd(a, b)$ notates the *greatest common divisor* of $a$ and $b$.

## Euclidean Algorithm

1. We are trying to determine the gcd of $a$ and $d$.
2. $a = dq + r_0 \quad 0 \leq r_0 < d$
3. $d = r_0 q_1 + r_1 \quad 0 \leq r_1 < r_0$
4. $r_0 = r_1 q_1 + r_2 \quad 0 \leq r_2 < r_1$
5. Continue until $r_s = 0$. Let $s$ be the smallest integer such that $r_s = 0$.
6. If $s = 0, \gcd(a, d) = d$.
7. If $s > 0, \gcd(a, d) = r_{s-1}$.

**Theorem 2.** $\gcd(a, d) = \gcd(d, r)$ where $r$ is the unique remainder from the division algorithm.

**Definition 5** (Relatively Prime)**.** We call $a$ and $b$ *relatively prime* if $\gcd(a, b) = 1$.

**Theorem 3** (Bezout's Identity)**.** $\forall a, b \in \mathbb{Z}^+$ there exists $u, r \in \mathbb{Z}$ such that $\gcd(a, b) = au + br$. We call such a form a *linear combination.*

**Definition 6** (Mod)**.** If we have two integers $a, b$, and another positive integer $m$, then we will define a relation $R_m$ on $\mathbb{Z}$ by $aR_mb$ *if* $m \mid (a - b)$. In other words, if $\exists k$ such that $a - b = mk$. We will write $aR_mb$ as $a \equiv b \pmod{m}$.

## Properties of Mod

If $a \equiv b \pmod{m}$; $c \equiv d \pmod{m}$; and $e, n \in \mathbb{Z}$ then:

- $a + c \equiv b + d \pmod{m}$.
- $ac \equiv bd \pmod{m}$.
- $a + e \equiv b + e \pmod{m}$.
- $a^n \equiv b^n \pmod{m}$.
- $ae \equiv be \pmod{m}$.

**Theorem 4.** $ax \equiv c \pmod{m}$ has a solution iff $\gcd(a, m) \mid c$.

**Theorem 5** (Fermat's Little Theorem)**.** Let $p$ be a prime. If $\gcd(a, p) = 1$, then

$$a^{p-1} \equiv 1 \pmod{p}$$

. Equivalently, $a^p \equiv a \pmod{p}$.

**Definition 7** (Euler phi function)**.** Euler's $\phi$ (or *totient* function is defined over $n \in \mathbb{Z}^+$ such that

$$\phi(n) = \left| \left\{ k \in \mathbb{Z} \;\middle|\; 1 \leq k \leq n, \gcd(k, n) = 1 \right\} \right|.$$

**Theorem 6.** If we have two primes, $p$ and $q$, which are distinct, then $\phi(pq) = (p - 1)(q - 1)$.

**Theorem 7** (Fermat-Euler Theorem)**.** Consider $a, m \in \mathbb{Z}^+$ (not necessarily prime!). If the $\gcd(a, m) = 1$, then $a^{\phi(m)} \equiv 1 \pmod{m}$.

## RSA Encryption

RSA is a form of *public key* encryption; that is, it's one directional.

1. Choose two primes $p, q$.
2. Calculate $n = pq$.
3. Calculate $\phi(n) = (p - 1)(q - 1)$.
4. Choose some $1 < k < \phi(n)$ such that $\gcd(k, \phi(n)) = 1$.

5. Calculate $kd \equiv 1 \pmod{\phi}(n)$.
6. Release $n, k$ publicly.
7. Keep $d$ private to yourself.

For a message $m$, compute $r \equiv m^k \pmod{n}$. To decrypt, calculate $x \equiv r^d \pmod{n}$. Keep in mind that other people encrypt messages for me, and only I can decrypt; thus, in order to have two way communication, we each need to have our own separate encryption schemes.

## 5.3 Formal Logic

### Logical Operations

| $p$ | Not $p$ |
|---|---|
| T | F |
| F | T |

We also sometimes see $\neg p$ or $\overline{p}$.

| $p$ | $q$ | p And $q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | F |
| F | F | F |

Notice that And works similarly to intersection within sets. Notationally, we usually write $p \wedge q$.

| $p$ | $q$ | $p$ Or $q$ | $p$ Xor $q$ |
|---|---|---|---|
| T | T | T | F |
| T | F | T | T |
| F | T | T | T |
| F | F | F | F |

Or is also notated $p \vee q$ and Xor is $p \oplus q$.

| $p$ | $q$ | $p \Rightarrow q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | T |
| F | F | T |

$p$ Implies $q$.

| $p$ | $q$ | $p \Rightarrow q$ | $q \Rightarrow p$ | $(p \Rightarrow q) \wedge (q \Rightarrow p)$ |
|---|---|---|---|---|
| T | T | T | T | T |
| T | F | F | T | F |
| F | T | T | F | F |
| F | F | T | T | T |

We notate Iff by $p \Leftrightarrow q$.

**Definition 1** (Tautology)**.** A proposition that is always true is called a *tautology*. As an example, consider the proposition $p \vee \neg p$.

**Definition 2** (Contradiction)**.** A proposition that is never true is called a *contradiction*. As an example, consider the proposition $p \wedge \neg p$.

**Definition 3** (Boolean function)**.** A Boolean function is any function $f$ that maps some set $S$ to $\{0, 1\}$. They are the equivalent of predicates.

[Circuit notes released in full and separately.]

## 5.4 Counting

**Definition 1** (Factorial)**.** $n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1 = n!$ We also define the special case $0! = 1$.

**Definition 2** (Permutation)**.** A permutation of a set $X$ is some function $f : X \to X$ which is a bijection. It can be thought of as a *relabeling* of $X$. There are $|X|!$ permutations of $X$.

**Definition 3** (Binomial Coefficient)**.** For integers $n \geq k \geq 0$, we define the binomial coefficient

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

**Theorem 1** (Binomial Theorem)**.** $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

## Binomial Coefficient Properties

- $\binom{n}{k} = \binom{n}{n-k}$.
- $2^n = \sum_{k=0}^{n} \binom{n}{k}$.
- $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

**Theorem 2** (Inclusion-Exclusion Formula)**.** For $n$ sets $X_1 \ldots X_n$:

$$|X_1 \cup X_2 \cup \cdots \cup X_n| = \sum_i |X_i| - \sum_{i<j} |X_i \cap X_j| + \sum_{i<j<k} |X_i \cap X_j \cap X_k| - \cdots + (-1)^{n-1} |X_1 \cap X_2 \cap \cdots \cap X_n|$$

**Definition 4** (Derangement)**.** We define a permutation to be a *derangement* if it does not leave any element fixed. In other words, $f : X \to X$ such that $f(y) \neq y$ for all $y \in X$. The number of derangements $D_n$ for a set of size $n$ is:

$$D_n = n! \left(1 - \frac{1}{1} + \frac{1}{2!} - \frac{1}{3!} + \cdots - (-1)^{n-1} \frac{1}{n!}\right)$$

**Definition 5** (Pigeonhole principle)**.** If we take $k+1$ objects, and put them into $k$ boxes, some box must contain at least two objects.

**Definition 6** (Strong pigeonhole principle)**.** Consider $n$ different types of objects numbered 1 through $n$, and $n$ positive integers $m_1, m_2, \ldots, m_n$. In any collection of $m_1 + m_2 + \cdots + m_n - n + 1$ objects, there must exists some index $i$ such that the number of object $i$ within the collection is at least $m_i$.

*Intuition:* If you win a million dollars if you can flip a coin and get 100 heads or five tails, you will only need to flip the coin 104 times at most.

## 5.5 Asymptotic Counting

**Definition 1** (Big Oh Function)**.** We define $O(f(x))$ to be the set of functions that grow 'faster' than $f$. We say that $f(x) \in O(g(x))$ if $\exists\, c, k \in \mathbb{R}$ s.t. $|f(x)| \leq c|g(x)|\ \forall\, x > k$.

**Definition 2** (Big Omega)**.** Big $\Omega$ is the same as $O$ but as a lower bound instead; thus, $f(x) \in O(g(x))$ iff $g(x) \in \Omega(f(x))$.

**Definition 3** (Big Theta)**.** Big $\Theta$ is both O and $\Omega$. $f(x) \in \Theta(g(x))$, if and only if $f(x) \in O(g(x))$ and $f(x) \in \Omega(g(x))$. We say that two things which are big Theta of each other are of the *same order*.

**Definition 4** (Ceiling function)**.** We define the ceiling function $\text{ceil}(x) = \lceil x \rceil$ to be the smallest integer $y$ such that $y \geq x$.

**Definition 5** (Floor function)**.** We define the floor function $\text{floor}(x) = \lfloor x \rfloor$ to be the largest integer $y$ such that $y \leq x$.

**Definition 6** (Limit)**.** For a sequence $\{a_n\}$, we define the limit of that sequence $\lim_{n\to\infty} a_n = c$ if $\forall \epsilon > 0$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \geq n_0$, $|a_n - c| \leq \epsilon$.

**Definition 7** (Asymptotic equality)**.** We define two sequences $\{a_n\}, \{b_n\}$ if the limit $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Note that we define $\frac{0}{0} = 1$.

**Theorem 1** (Stirling's formula)**.** $n!$ is asymptotically equal to $\left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.

## 5.6 Probability

**Definition 1** (Probability space)**.** A pair $(S, p)$ where $S$ is a non-empty, finite set which we call the *sample space*, and $p$ is a *probability distribution function* $S \to \mathbb{R}$ that satisfies

1. $p(\omega) > 0 \; \forall \, \omega \in S$.
2. $\sum_{\omega \in S} p(\omega) = 1$.

**Definition 2** (Event)**.** An *event* within $S$ is any subset $E \subseteq S$. For convenience, we may write $p(E)$ for some event, which we define to be the sum of $p(\omega)$ for $\omega \in E$; additionally, $p(\varnothing) \triangleq 0$.

**Definition 3** (Uniform distribution)**.** We define the *uniform* distribution $p$ of a sample space $S$ such that $p(\omega_1) = p(\omega_2)$ for all $\omega_1, \omega_2 \in S$.

**Definition 4** (Conditional probability)**.** $p(A|B)$ or the probability of $A$ *given* $B$ is exactly $\frac{p(A \cap B)}{p(B)}$. We can think of this as 'renormalizing' our entire universe so that $B$ always happens.

**Definition 5** (Independence)**.** We call events $A, B$ *independent* if $p(A \cap B) = p(A)p(B)$. In terms of conditional probability, we get that $p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{p(A)(p(B)}{p(B)} = p(A)$.

**Definition 6** (Collective independence)**.** Let $A_1, A_2, \ldots, A_n$ be a set of events. We say that all $n$ events are independent of each other if, for all subsets $I \subseteq \{A_1, \ldots, A_n\}$, we have $p(\bigcap_{i \in I} A_i) = \prod_{i \in I} p(A_i)$. This is because pairwise independence between all pairs in a set does *not* imply independence within arbitrary subsets of that set..

**Theorem 1** (Bayes' Theorem)**.** $p(A|B) = \frac{p(B|A)p(A)}{p(B)}$.

**Theorem 2** (Law of Complete Probability)**.** For some probability space $(S, p)$, with $S = X_1 \cup \cdots \cup X_k$ such that $X_i \cap X_j = \varnothing$ for $i \neq j$. Then we have that for any event $E \subseteq S$, $p(E) = \sum_{i=1}^{k} p(E|X_i)p(X_i)$.

**Definition 7** (Extended Bayes' Law)**.**

$$p(A|B) = \frac{p(B|A)p(A)}{p(B|A)p(A) + p(B|A^C)p(A^C)}$$

**Definition 8** (Random Variable)**.** Given some probability space $(S, p)$, we define a random variable $X$ to be some function $S \to \mathbb{R}$.

**Definition 9** (Binomial Random Variable)**.** $n$ independent, repeated trials, each of which has a $p$ probability of success and a $(1-p)$ probability of failure. The probability of $k$ successes is $\binom{n}{k}p^k(1-p)^{n-k}$.

**Definition 10** (Indicator Random Variable)**.** The special case of a binomial random variable with $n = 1$. This is also known as a Bernoulli RV. It simply indicates success or failure, 0 or 1.

**Definition 11** (Expectation)**.**

$$E[X] = \sum_{\omega \in S} p(\omega)X(\omega)$$

*Intuition:* A weighted average. If you were to perform an experiment a million times, what would you see most often?

**Theorem 3** (Linearity of Expectation)**.**

$$E[X_1 + \cdots + X_n] = \sum_{i=1}^{n} E[X_i]$$

**Definition 12** (Variance)**.** The variance of a random variable $X$ is intuitively a measure of how much the random variable varies. More formally,

$$V[X] = \sum_{\omega \in S} \big(X(\omega) - E[X]\big)^2 p(\omega).$$

## Properties of Variance

- $V[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2$.
- $V[aX + b] = a^2V[X]$ for constant $a, b$.
- $V[X + Y] = V[X] + V[Y]$ if $X, Y$ are independent.

**Definition 13** (Standard Deviation)**.** The standard deviation of a random variable $X$ is $\sqrt{V[X]}$.

**Theorem 4** (Markov's Bound)**.** Consider a random variable $X$ which is non-negative.

$$\Pr[X \geq a] \leq \frac{E[X]}{a}.$$

**Theorem 5** (Chebyshev's Inequality)**.** For any random variable $X$ and any constant $a$, $\Pr\left[\left|X - E[X]\right| \geq a\right] \leq \frac{V[X]}{a^2}$.

**Theorem 6** (Weak Law of Large Numbers)**.** Consider some family of identical, independent random variables $X_1, \ldots, X_n$. We define $\mu = E[X_i]$. The Weak Law of Large Numbers states that

$$\lim_{n \to \infty} \Pr\left[\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq c\right] = 0.$$

## 5.7 Graph Theory

**Definition 1** (Graph)**.** A *graph* is a pair of sets $(V, E)$ such that $V$ is a finite set and $E$ is a set of unordered pairs of elements of $V$.

**Definition 2** (Tree)**.** A *tree* is a graph $(V, E)$ that (equivalently)...

1. is connected and has no cycles.
2. is connected and has $n - 1$ edges.
3. contains no cycles, and has $n - 1$ edges.
4. contains some unique path from $u$ to $v$ for all $u, v \in V$.
5. is connected, and deleting any $e \in E$ will disconnect it.
6. contains no cycles, and adding any $e \notin E$ forms a cycle.

**Definition 3** (Forest)**.** A *forest* is a graph containing only trees. Formally, one might say that, for any $V' \subseteq V$, the subgraph formed by $E$ on $V \backslash V'$ is a tree if it is connected.

**Definition 4** (Leaf)**.** A *leaf* is a vertex which has degree 1.

**Definition 5** (Degree)**.** The *degree* of a vertex is the number of edges that touch that vertex.

**Definition 6** (Planar graph)**.** A graph $G = (V, E)$ is *planar* if it can be drawn in the plane without crossing any edges.

**Definition 7** (Complete graph)**.** A graph with $n$ vertices is *complete* if it has all possible $(\binom{n}{2})$ edges. It is denoted by $K_n$.

**Definition 8** (Complete bipartite graph)**.** A *complete bipartite graph*, or $K_{m,l}$, can be broken into two partitions ($M$ of size $m$, and $L$ of size $l$), such that all possible edges between $M$ and $L$ are present, but no others are present.

**Theorem 1** (Euler's formula)**.** If we define $f$ to be the number of regions which a connected planar graph creates, then

$$|V| - |E| + f = 2.$$

**Theorem 2.** If $G$ is planar, and $n \geq 3$, then $e \leq 3n - 6$.

**Theorem 3.** If $G$ is planar with $n \geq 3$, and $G$ contains no triangles, then $e \leq 2n - 4$.

**Definition 9** (Subdivision)**.** A *subdivision* of a graph is produced by replacing any edge with two edges which connect in the middle of the old edge. A subdivision of a subdivision of $G$ is also a subdivision of $G$.

**Theorem 4** (Kuratowski's Theorem)**.** $G$ is planar iff it does not contain (as a subgraph) a subdivision of $K_5$ or $K_{3,3}$.

**Definition 10** (Graph Coloring)**.** A coloring of a graph is a function $\chi$ from $V \to C$, where $C$ is a (generally finite) set of 'colors'. A *proper* coloring of $G$ is a $\chi$ such that $(u, v) \in E \Rightarrow \chi(u) \neq \chi(v)$.

**Definition 11** (Chromatic Number)**.** The *chromatic number* $X(G)$ of a graph $G$ is the minimum $|C|$ such that there exists some $\chi : V \to C$ that is a proper coloring for $G$.

**Definition 12** (Hypercube)**.** An $n$-dimensional *hypercube* is a graph where every vertex is represented by a binary string of length $n$, and where two vertices are joined if their representative binary strings differ in exactly one bit.

**Definition 13** (Bipartite graph)**.** A graph is *bipartite* if it has a proper 2-coloring. In other words, a bipartite graph can be partitioned into two vertex sets such that no edge connects two vertices in the same vertex set.

**Theorem 5** (Four-Color Theorem)**.** The chromatic color of a planar graph is always at most 4.

# Index