Lecture Notes on Combinatorics: Draft

Garrett J. Kepler

Updated April 6, 2026

Chapter 1

What is combinatorics?

One great property of the field of combinatorics is most people have already encountered combinatorial questions in their own life.

Considering the optimal path we should take to get groceries and visit the coffee shop then return home. Now that we have the groceries, figuring out which set of dishes we can cook for the party when we've invited a vegetarian, a non-vegetarian, and someone allergic to peanuts. Now that we know what to cook, determining the best drinks paired with which dishes.

As simple of an example party planning is, we have described three combinatorial problems already: minimal path, dish enumeration, and drink/dish combinations.
Many of the questions we ask in combinatorics focus on counting of some sort.

- Enumeration: How many ways. . . ?
- Existence: Does there exist. . . ?
- Extremal: What is the largest/smallest possible. . . ?
- Expectation: What is the expected number of. . . ?

1.0.1 Week's Warmups

Question 1.0.1.
How many ways are there to assign 3 students 3 distinct Math 325 projects? 4 students 4 distinct projects? n students n distinct projects?

Question 1.0.2.
How many people must attend the first day of lecture in Math 325 to guarantee 2 people either both have met before or both have not met before?

Question 1.0.3.
How many people must attend the first day of lecture in Math 325 to guarantee 3 people either have all met before or have all not met before?

These preliminary warmup problems are classic combinatorial questions. The first, a question of enumeration, we may have a straightforward way to answer already:

Answer 1.0.4.
Well, we could assign student A project 1. Or, we could assign them project 2. Or, we could assign them project 3. That's 3 ways. Okay, for student B, we could assign project 1. Or, we could assign them project 2. Or,. . . This is exhausting listing them out. I'm bored.

What we are going to learn in this class is methods to avoid this one by one counting. Instead, we will "count". Rather than assigning a specific student to a specific project one by one, we'll deal with the collection of students and the collection of projects in their entirety:

Answer 1.0.5.

Given a student, we assign them 1 of the 3 projects. Since we want distinct assignments, the next student must be assigned one of the other 2 projects. Likewise, the final student must be assigned the leftover project. Using what we will later refer to as the multiplication principle*, we have* $3 \cdot 2 \cdot 1 = 6$ *ways to do this.*

Questions 1.2 and 1.3 ask questions of guaranteed substructures that we will discuss later in the Ramsey Theory section.

Chapter 2

Problem Solving Basics & Set Theory

We will start with questions of *enumeration* and dedicate a fair amount of time to this concept.

Definition 2.0.1 (Enumeration)**.**
The problem of *enumeration* is to count the total number of objects with a specified property.

Example 2.0.2.
A few basic examples:

- How many ways are there to assign n students to m projects?
- How many distinct Sudoku puzzles are there?
- How many total (potentially isomorphic) graphs are there are on n nodes?

$\triangle$

This questions can often be difficult to answer given a certain object with a desired property. However, the philosophical approach we often take is: **simplify**!

Our approach to these problems can be summarized in a joke told to me by a faculty member at WSU:

A mathematician walks into a room. In one corner, he sees an empty bucket. In another corner, he sees a sink with a water faucet. In yet another corner, he sees a fire. Leaping into action, he picks up the bucket, fills it up at the faucet, and douses the fire.
The next day, the same mathematician returns to the room. Once again, he sees a fire in the same corner. But, this time the bucket rests next to the fire full of water. Once more he leaps into action: he picks up the bucket, dumps the water in the sink, places the bucket empty where it was from the day before and leaves. As he exits, he exclaims victoriously "I've reduced it to a previously solved problem!"

Our first goal will be to mimic this proverbial mathematician's mentality: *simplify*! Once we've simplified, we'll ask ourselves: "can we generalize this?"

2.0.1 Simplification

Question 2.0.3.
Consider the following problems:

1. How many whole numbers are between 1 and 325 inclusive?
2. How many whole numbers are between 28 and 352 inclusive?
3. How many even numbers are there between 28 and 352 inclusive?

4. How many numbers between 28 and 352 are divisible by 3?

On the one hand, these are four distinct problems. On the other, you may be able to see there is a process we can use to answer all of them. One main approach we often utilize in combinatorics is the following: to count what we need to count, reframe the problem as something we already know how to count. To quote a great mathematician:

"*I have reduced it to a previously solved problem!*"

Consider our first problem: "How many whole numbers are between 1 and 325 inclusive?"

- 1 is our 1^{st}, 2 is our 2^{nd}, and so on, 325 is our 325^{th}. So, there must be 325 such numbers.

This was so easy to solve! What we want to do now is use this solved problem to answer another. Consider our second problem: "How many whole numbers are between 28 and 352 inclusive?"

- We begin counting: 28 is our 1^{st} number, 29 is our 2^{nd}, and so on, so 352 must be our $?^{th}$.

What do we put here? No idea. We only know how to count numbers between 1 and 325. Let's instead transform this counting problem to the one we know how to count:

- We may notice 28 is 27 away from 1. Likewise, 29 is 27 away from 2 and so on.
- So, take each number $28, 29, 30, \ldots, 352$ and subtract 27: $1, 2, 3, \ldots, 325$.
- Amazing! We already know this answer: there are 325 such numbers.

One issue you may already have thought of: "how do I know there are the same amount of numbers between 1 and 325 as 28 and 352 by simply subtracting 27?" That is a great point. We don't yet know. However, this is the beauty of a *bijection*. Something we'll talk about: if we have a bijection between two sets, we know they are the same size. More on the formalities later!

"*I have generalized the method to solve a previously unsolved problem!*"

For now, consider the third problem: "How many even numbers are there between 28 and 352 inclusive?". Again, we don't know how to count this. But, on the second problem, we transformed to the numbers $1, 2, 3, \ldots, 325$ and made our problem easier. Let's try something similar:

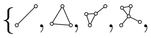
- Take each number, $28, 30, 32, \ldots, 350, 352$, and divide by 2: $14, 15, 16, \ldots, 175, 176$.
- They may look different, but we can transform similarly to problem two. Subtract 13 and we get $1, 2, 3, \ldots, 162, 163$.
- There are 163 such numbers, so there must be 163 such even numbers between 28 and 352.

Now that we have a specific method of solution, we might reformat this method to a more general problem. We may want to count the numbers between m and n ($m < n$).
In our second and third problem, we took each number and found a way to obtain $\{1, 2, 3, \ldots, n\}$ so that we could enumerate the solutions. If we want to take our 1^{st} number m and send it the first natural number 1, we might subtract $m - 1$ (since $m - (m - 1) = m - m + 1 = 1$). For each of these numbers, we get $m - (m - 1), m + 1 - (m - 1), m + 2 - (m - 1), \ldots, n - 1 - (m - 1), n - (m - 1)$ which really is $1, 2, 3, \ldots, n - m, n - m + 1$. As such, we obtain a count of $n - m + 1$ integers between m and n ($m < n$). What's left to prove is that the process of subtracting $m - 1$ here guarantees the two sets of numbers are the same size. The next section will discuss how to use a *bijection* to do so.

2.1 Set Theory Basics

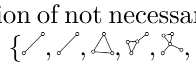
This concept of bijection we subtly utilized previously is actually a fundamental result combinatorialists love to exploit. Before we get there, some basics are necessary.

Definition 2.1.1 (Set).

A *set* is a collection of distinct objects. E.g. The set of integers between 28 and 56 $= \{28, 29, 30, \dots, 56\}$, a set of graphs = {, . . . }, etc.

Note in our definition of a set we require *distinctness*. As such, the collection $\{1, 1, 2, 3, \dots, 29\}$ is not a set. Likewise with any collection with repeating elements, but this permits another important definition:

Definition 2.1.2 (Multiset).

A *multiset* is a collection of not necessarily distinct objects. E.g. The multiset of integers $= \{28, 28, 28, 29, 30, \dots, 56\}$, the multiset of graphs = {, . . . }, etc.

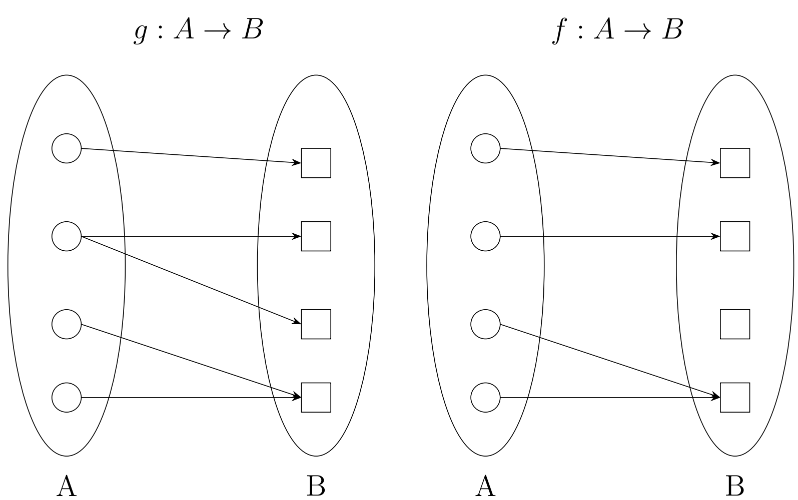
However, we won't touch multisets until later on. For now, let us discuss relations between sets.

Definition 2.1.3 (Function).

A function $f : A \to B$ is a relation between sets A and B such that for each $a \in A$ (a in A), there is a unique $b \in B$ such that $f(a) = b$.

Example 2.1.4.

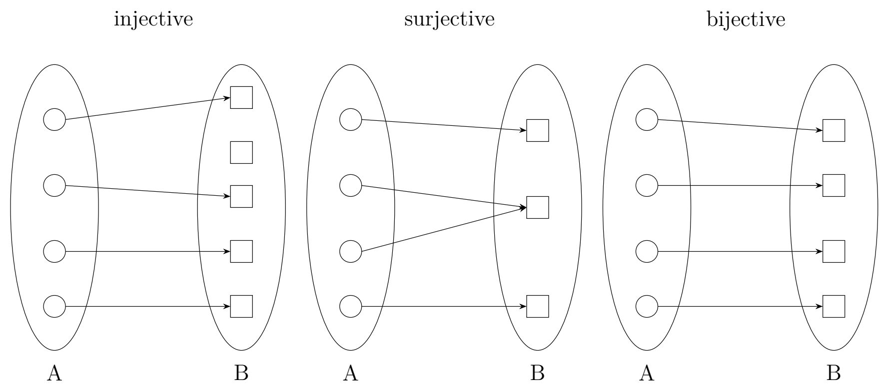
As an example, on the left, we have $g : A \to B$ that assigns the second element in A to two elements in B. This means g is not a function. However, on the right, we have $f : A \to B$ that assigns each element in A only one element in B. This means f is a function.



△

Definition 2.1.5 (Injection, Surjection, and Bijection).

Firstly, a function f is *injective* if distinct elements in A are mapped to distinct elements in B. Secondly, f is *surjective* if every element in B has at least one element in A that gets mapped to it. Lastly, if f is both injective and surjective, then f is *bijective*.



For combinatorial purposes, one nice way to think of a function $f : A \to B$ as the placement of a set of objects A into a set of boxes B:

- If no two objects are placed in the same box, f is injective.
- If no box is left empty, f is surjective.
- If both occur, f is bijective.

We can use this thinking to phrase the following statements:
Let $f : A \to B$ be a function that assigns a set of objects A to be placed into a set of boxes B.

- f is injective means f ensures no two objects are placed in the same box.
- f is surjective means f leaves no box empty.
- f is bijective means f leaves no box empty and ensures no two objects are placed in the same box.

But, why do we care about this as combinatorialists?

Theorem 2.1.6.
Again, let $f : A \to B$ as defined above.

1. *If f is injective, $|A| \leq |B|$.*
2. *If f is surjective, $|A| \geq |B|$.*
3. *If f is bijective, $|A| = |B|$*

where $|\cdot|$ denotes the number of elements in a set.

First, an outline of a proof.

Proof.
Let f be a function such that $f : A \to B$ for sets A and B.

1. If f is injective, it ensures no two objects are in the same box. So, there must be at least enough boxes to house these distinctly assigned objects ($|A| \leq |B|$).
2. If f is surjective, it leaves no box empty. So, there must be at least enough objects to fill each box ($|A| \geq |B|$).
3. If f is bijective, it ensures no two objects are in the same box *and* no box is left empty. As such, we know $|A| \leq |B|$ *and* $|A| \geq |B|$ (i.e. that $|A| = |B|$).

□

This simple statement is incredibly powerful. What this theorem firstly tells us is that there is a way to infer information about the *sizes* of the sets when we have a specific property about the *relationship* between those sets. Most importantly, the theorem also states that if there exists a bijection between two sets, they must be the exact same size. We can exploit this fact to enumerate easily! In fact, we already have.
Think back to our second problem: "how many integers are there between 28 and 352 inclusive?". Perhaps we write out "28 is 1^{st}, 29 is 2^{nd},… " and so on until we reach 352 or count on our 325 fingers one by one. Either way we know this would be annoying and tedious to count explicitly. By defining a bijective function between sets, we eased the trouble, $f(a) = a - 27$ for $a \in \{28, 29, 30, \ldots, 325\}$.
From this final "proof" in 2.1.6, we can infer a lot of information. For clarity, if f is bijective, each object has a distinct box *and* each box has a distinct object. If we wish to reverse the process of assignment f has committed, we could easily do so: simply reverse each assignment!
This thinking leads to the following theorem:

Theorem 2.1.7.
The function $f : A \to B$ is a bijection if and only if f^{-1} exists.

This allows us to formalize our results:

- Let $A = \{28, 29, \ldots, 352\}$, $B = \{1, 2, \ldots, 325\}$, and $f : A \to B$ be defined as $f(a) = a - 27 = b$ for $a \in A$ and $b \in B$.
- Since f subtracts 27, f^{-1} might add 27 ($f(a) = a - 27 = b \implies a = f(a) + 27 \implies f^{-1}(b) = b + 27$).
- Thus, f must be a bijection since there exists an f^{-1}. So, we know $|A| = |B|$.
- Therefore, $|\{1, 2, \ldots, 325\}| = |\{28, 29, \ldots, 352\}| = 325$.

We can use the same process with problems 3 and 4 of 2.0.3 :

- Let $A = \{28, 30, 32, \ldots, 352\}$ and define $f(a) = \frac{1}{2}a - 13 = b$ for $a \in A$. What is f^{-1}? What is B? Conclude $|A| = |B|$.
- Let $A = \{30, 33, 36, \ldots, 348, 351\}$ and define $f(a) = \frac{1}{3}a = b$ for $a \in A$. What is f^{-1}? What is B? Conclude $|A| = |B|$.
- (Extra!) Let $A = \{m, m+1, m+2, \ldots, n-1, n\}$ and define $f(a) = \frac{1}{k}a - \ell$ for $m < k < n$ and $\ell \geq 0$. What is f^{-1}? What is B? Conclude $|A| = |B|$.

2.2 Summary

What we've done so far is introduce important techniques we often use in combinatorics:

- Let's simplify our current problem to a problem we already know how to solve
- Now that we have found a method to solve a problem, let us generalize

These often guide thought in mathematical problem solving and we will utilize both in combinatorics!

Chapter 3

Introduction to counting techniques

3.0.1 Warmup

Consider the following problem:

Question 3.0.1.
Physics 533 and Math 325 occur at the same time (12:10-1PM MWF). Phys 533 has 5 people enrolled while Math 325 has 21.

1. How many people are in Math 325 or Physics 533 from 12:10-1 MWF?

2. How many people are in Math 325 and Physics 533 from 12:10-1 MWF?

This may seem immediately obvious to some, but this highlights key counting principles used in combinatorics.

Answer 3.0.2.

1. *There are 21 people in Math 325. There are 5 in Phys 533. So, there are 21+5 in either during the class period.*

2. *Nobody can be in two places at once, so there must be zero people in both classrooms during the class period.*

This is a nice set up. We know the set of students in 325 is completely separate from the set of students in 533. If we use set theoretic language, the set of students in 325 is *disjoint* from the set of students in 533. More on this later.
For now, consider yet another problem:

Question 3.0.3.
How many ways are there to flip a coin then a six sided dice? More generally, rolling an m sided dice then an n sided dice?

You might note that my first roll *does not affect* my second roll. They are independent! Intuitively, if I roll heads, I have the opportunity to roll 1-6 on the dice. Likewise if roll tails. As such:

Answer 3.0.4.
There are 2 options for my first roll. Since each roll is independent, there are 6 options for my second roll. So, there must be $2 \cdot 6 = 12$ *ways to do this. More generally, since each roll is independent, there are* $m \cdot n$ *ways to do this with an* m *sided dice and an* n *sided dice.*

3.1 More Set Theory

Definition 3.1.1 (Subset)**.**
Let B be some set. Another set A is a *subset* of B if every element of A is an element of B (for every $a \in A$, $a \in B$). We denote this relationship by $A \subseteq B$ if A could equal B itself. If A cannot equal B, we write $A \subset B$.

Definition 3.1.2 (Set Complement)**.**
Let $A \subseteq S$. The *complement* of A with respect to S, $A^C = S \setminus A$, is the set of elements in S that are not in A ($A^C = \{s \in S | s \notin A\}$). The complement of S with respect to itself, $S^C = S \setminus S$, is the set with zero elements, $\emptyset = \{\}$.

We can think of the set complement $S \setminus A$ as considering all the elements of S that are not in A. If there are no such elements (i.e. $A \nsubseteq S$), $S \setminus A$ is $\emptyset$.

Example 3.1.3.
Let $A = \{\text{integers between 28 and 352 divisible by 3}\}$ and $S = \{\text{integers between 28 and 352 inclusive}\}$

- $A \subset S$
- $A^C = S \setminus A = \{\text{integers between 28 and 352 not divisible by 3}\}$

△

Remark 3.1.4.
The choice of S is important. If $S = \mathbb{Z}$ (the set of all integers), then A^C in the previous example would become the set of *all* integers excluding only integers between 28 and 352 that are divisible by 3.

Question 3.1.5.
Let $A \subseteq S$. Is the function $f(A) = A^C$ a bijection? If so, what is f^{-1}?

Answer 3.1.6.
f assigns subset A to its complement $A^C = S \setminus A$.

- *Let us say $f(A) = A^C = B$.*
- *By the definition of a set's complement, B contains every element of S that is not in A. Likewise, A contains every element of S that is not in B.*
- *In other words, $A^C = B$ and $A = B^C$.*
- *So, $f^{-1}(B) = B^C$. Hence, f^{-1} exists and f is a bijection!*

This is a nice property. Consider the problem of enumerating k-element subsets of $S = \{1, 2, 3, 4, \ldots, n\}$ $(k < n)$.

- Let A be the set of all k-elements subsets $A_i \subset S$. Moreover, let B be the set of all the complements of elements in A.
- Then for each A_i, there is some set B_i such that $f(A_i) = A_i^C = B_i \subset S$ of size $n - k$.
- Since $f(A_i) = A_i^C$ is a bijection, the set of all k-element subsets, A, is the same size as the set of all $(n - k)$-element subsets, B. That is, $|A| = |B|$.

We will talk about what the actual sizes $|A|$ and $|B|$ later.

Definition 3.1.7 (Set Union)**.**
The *union* of two sets A and B, denoted $A \cup B$, is the set of all elements in A or B ($A \cup B = \{x | x \in A$ or $x \in B\}$).

Definition 3.1.8 (Set Intersection)**.**
The *intersection* of two sets A and B, denoted $A \cap B$, is the set of all elements in A and B ($A \cap B = \{x | x \in A$ and $x \in B\}$).

Definition 3.1.9 (Disjointness)**.**
If the intersection between sets A and B is empty, denoted $A \cap B = \emptyset$, we say A and B are *disjoint*.

Example 3.1.10.
Let $A = \{$the set of students in Math 325$\}$ and $B = \{$the set of students in Physics 533$\}$.

- $A \cup B = \{$the set of students in either couse$\}$
- $A \cap B = \emptyset$ since now student is in both classes

△

3.2 Addition & Multiplication Principle

Thinking back to the initial problem of this section (3.0.1), we could simply add the number of people in 325 and 533 to get the total number *since no student is in both classes.* This principle of complete separation allowing simple addition can be formalized:

Theorem 3.2.1 (Addition Principle)**.**
If $A \cap B = \emptyset$, $|A \cup B| = |A| + |B|$.
More generally, if $A_i \cap A_j = \emptyset$ for all $i \neq j$ and $1 \leq i, j \leq n$, then $|A_1 \cup A_2 \cup \cdots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n|$.

Definition 3.2.2 (Cartesian Product)**.**
Let $A \times B$ be the set of ordered pairs (a, b) where $a \in A$ and $b \in B$.

Theorem 3.2.3 (Multiplication Principle)**.**
Let $S = A \times B$. Then, $|S| = |A \times B| = |A| \cdot |B|$.
More generally, let $S = A_1 \times A_2 \times \cdots \times A_n$. Then, $|S| = |A_1 \times A_2 \times \cdots \times A_n| = |A_1||A_2| \cdots |A_n|$

Example 3.2.4.
From warmup problem 2, let $A = \{H, T\}$ be the results from flipping a coin and $B = \{1, 2, 3, 4, 5, 6\}$ be the results from rolling the six-sided dice. Then,

$$A \times B = \{(H, 1), (H, 2), (H, 3), \ldots, (T, 5), (T, 6)\}$$

and $|S| = |A \times B| = |A||B| = 2 \cdot 6 = 12$. △

3.2.1 (In)dependence

Note that the *independence* of warmup problem 2 is important. Flipping a coin *has no impact* on the rolling of the dice. Consider a different problem:

Question 3.2.5.
There are four orbs labelled A, B, C, D in an urn. You will draw from the urn four times. For each draw, you take a single orb out, write down its label, and replace it in the urn. How many possible label sequences could you write down?

Answer 3.2.6.

- *Let A_1, A_2, A_3, and A_4 be the sets of possible draws on draw 1, 2, 3, and 4 respectively and S be the set of possible label sequences you could write down.*
- *Since all four balls in the urn before you draw first, $A_1 = \{A, B, C, D\}$.*
- *Moreover, since you placed the orb you drew back into the urn, you have a chance to draw any of the four orbs again: $A_2 = \{A, B, C, D\}$.*
- *Likewise with A_3 and A_4.*

- *Thus, the total possible label sequences there are is:* $|S| = |A_1 \times A_2 \times A_3 \times A_4| = |A_1||A_2||A_3||A_4| = 4^4 = 256$.

Consider a small change to your process:

Question 3.2.7.
You really like writing down labels, so you continue with your setup from the previous problem. However, this time instead of replacing the orb after drawing it, you set it to the side. You then draw another orb and set it next to the previously drawn orb. You repeat this process until there are no orbs left. Then, you write down the sequence of labels you see. How many possible label sequences could you see?

Answer 3.2.8.

- *Let* B_1, B_2, B_3, *and* B_4 *be the sets of possible draws on draw 1, 2, 3, and 4 respectively and* T *be the set of possible label sequences you could write down.*
- *Since all four balls in the urn before you draw first,* $|B_1| = 4$.
- *Whatever you drew, it stays outside of the urn. Since there are 3 orbs left, you have 3 options for your next draw:* $|B_2| = 3$.
- *Continuing on, you have taken yet another option for your third draw away:* $|B_3| = 2$.
- *Lastly, since you drew one of the two remaining orbs, that leaves one left for your final draw* $|B_4| = 1$.
- *Thus,* $|T| = |B_1 \times B_2 \times B_3 \times B_4| = |B_1||B_2||B_3||B_4| = 4 \cdot 3 \cdot 2 \cdot 1 = 24$.

So, replacing the orbs every time gave us a count of 256 while leaving them out every time gave us a count of 24. What happened? One way to think of this is that by replacing the orbs, you have allowed for *repetitions* to exist in your set S (e.g. by drawing A then A then C then A). Contrarily, by leaving the orbs out after drawing, you remove the ability to have repeats. As such, you remove every element in S that has any repeat labels to obtain a new set T, the set of label sequences with *distinct* labels.

Question 3.2.9 (To Ponder)**.**
Let S and T be defined as in the above problem. What is $T^C = S \setminus T$? How big is T^C in this case?

Question 3.2.10 (To Ponder)**.**
Let S be defined as in the above problem. How many label sequences have only 2 repeats? i.e. sequences like (A, A, B, C) or (D, A, C, D).

Question 3.2.11 (To Ponder)**.**
Consider drawing n labelled orbs from an urn with replacement. How many ways can you do this? Without replacement?

3.3 Summary

What we've covered here:

- A function $f : A \to B$ having an inverse function $f^{-1} : B \to A$ means $|A| = |B|$.
- Set complements, unions, and intersections.
- Addition principle: the size of a union of disjoint sets is the sum of their sizes.
- Multiplication principle: the size of a product of two sets is the product of their sizes.
- The difference between counting with dependence versus without.

3.4 "Counting"

3.4.1 Week's Warmups

Question 3.4.1.

How many total subsets of $S = \{1, 2, 3, \ldots, 50\}$ are there?

Question 3.4.2.

Calculate $1 + 2 + 3 + \cdots + 49 + 50$. Could you form a general formula for $1 + 2 + 3 + 4 + \cdots + n$?

Question 3.4.3.

Determine the number of subsets of $S = \{1, 2, 3, \ldots, 50\}$ whose sum of elements is larger than or equal to 638.

Answer 3.4.4.

1. *Consider a subset $\mathcal{S}$ of S.*

   - *For each $i = 1, 2, 3, \ldots, 50$, either i is in $\mathcal{S}$ or not.*
   - *For each i, define $\mathcal{S}_i = \{0, 1\}$ where 0 is the case where i is in $\mathcal{S}$ and 0 where i is not.*
   - *Then, the total number of ways we can create subsets is given by $|\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3 \times \cdots \times \mathcal{S}_{50}|$.*
   - *By the multiplication principle, this tell us that there are*

     $$\begin{aligned} |\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3 \times \cdots \times \mathcal{S}_{50}| &= |\mathcal{S}_1| \cdot |\mathcal{S}_2| \cdot |\mathcal{S}_3| \cdots |\mathcal{S}_{50}| \\ &= 2 \cdot 2 \cdot 2 \cdots 2 = 2^{50} \end{aligned}$$

     possible subsets of S.

   This is a totally valid solution.
   As you may know, there are multiple ways to solve a problem. So, let's try and solve this problem by finding a bijection. Consider taking a subset $A \subseteq S$.

   - *For each element s in S, s is either in A or not.*
   - *We could think of assigning each element s a 1 or 0 depending on if it is in A (say, 1 if $s \in A$, 0 if $s \notin A$).*
   - *Then, $f(A)$ could be a vector of length n with entries 1 or 0. For example, $f(\{1, 3, 4, 6\}) = [1, 0, 1, 1, 0, 1, 0, \ldots, 0]^T = X_A$.*
   - *What f^{-1} be then? If the i-th entry of X_A is 1, put i in a set A. If the i-th entry of X_A is 0, exclude i from A.*
   - *Thus, f^{-1} exists and f is a bijection. So, the number of subsets A of S equals the number of possible 0,1-vectors X_A of size $|S|$.*
   - *How many vectors are there? There's 2 options for each $|S|$ entries of a vector X_A. So, again, by the multiplication principle, there are $2^{|S|}$ possible such vectors. And so, the total number of subsets is equal to $2^{|S|}$.*
   - *In our case, this means there are 2^{50} subsets.*

   There are multiple ways to solve a problem. The benefit of this is that we may learn different things about the objects in question. For example, in the bijection proof we learn about the relationship between 0,1-vectors of size n and total subsets of a set of size n. They are equivalent!

2. *You might try the following:*

$$\begin{aligned} L &= 1+2+3+\cdots+49+50 \\ \implies 2L &= 1+2+3+\cdots+49+50 \\ &\quad +50+49+48+\cdots+2+1 \\ \implies 2L &= (1+50)+(2+49)+(3+48) \\ &\quad +\cdots+(49+2)+(50+1) \\ &= 51+51+51+\cdots+51+51 \\ &= 50(51) \\ \implies 2L &= 50(51) \implies L = \frac{50(51)}{2} = 1275 \end{aligned}$$

This generalizes to $\frac{n(n+1)}{2}$ *when we sum 1 through* n.

3. *Again, let us try and find a bijection. Let* T *be the collection of all possible subsets of* S. *From warmup problem 1, we know* $|T| = 2^{50}$. *From warmup problem 2, the sum of elements of* S *is 1275.*

   - *Let* A *be the set of all subsets of* S *whose sum is* ≥ 638 *and* B *be the set of all subsets of* S *whose sum is* < 638. *Since either a subset sums to* ≥ 638 *or* < 638 *and not both,* $A \cup B = T$ *and* $A \cap B = \emptyset$. *Thus,* $|T| = |A \cup B| = |A| + |B|$.
   - *Moreover, if* $A_i \in A$, *then the sum of elements in* A_i *is* ≥ 638. *Since the sum of elements in* S *is 1275, the sum of elements in* A_i^C *must be less than or equal* $1275 - 638 = 637$.
   - *From Friday's class, we know* $f(A_i) = A_i^C$ *is a bijection. So,* $|A| = |B|$.
   - *So,* $|T| = |A| + |B| = |A| + |A| = 2|A|$.
   - *Since* $|T| = 2^{50} = 2|A|$, *we know* $|A| = 2^{49}$.

   You might try and generalize this to $S = \{1, 2, 3, \cdots, n\}$ *with a specific choice of sum for the subsets, say 325.*

3.4.2 Permutations, Combinations, & Double Counting

Combinatorics differs itself from many fields by its proof techniques. We have been talking about bijections the last few classes. What bijections allow us to do is say the following:

- We want to find a formula to count something
- We can transform the problem into an equivalent problem with a bijection
- Finding a formula in this equivalent problem tells us the count of the original problem

But, what if we found formulas for both problems? These are equivalent problems, so they should have equivalent formulas!
This is what is referred to as "double counting" or "counting with equivalence" or "counting in two ways". The backbone of double counting can be summarized as follows:

Proposition 3.4.5.
If $f(n)$ *and* $g(n)$ *are functions that count the solutions to a problem involving* n *objects, then* $f(n) = g(n)$ *for every* n.

That is, if there are two equivalent ways to count something, they must be equivalent all the time!

Definition 3.4.6 (Combinatorial Proof & Combinatorial Identities)**.**
A *combinatorial proof* works as follows:

- **Problem**: There is a problem of counting the number of solutions to a problem on n objects.

- **LHS**: Count one way and find a function $f(n)$ counts these solutions.
- **RHS**: Count another way and find a function $g(n)$ counts these solutions.
- **Conclusion**: Then, $f(n) = g(n)$ for every n.

The result, $f(n) = g(n)$, is what we call a *combinatorial identity*.

Remark 3.4.7.
This first step is crucial. Sometimes we are given the problem we are going to solve (e.g. “How many ways are there to permute n objects?”). Then, we find two equivalent formulas. Other times, we are given an identity to prove and we must identify a problem to count (e.g. “Prove $\binom{n}{k} = \binom{n}{n-k}$”). Then, we must show that both the left hand side and right hand side count the solutions to the problem we have identified.

Definition 3.4.8 (Selections & Combinations)**.**
A *selection* is the creation of a subset of a set. The resulting subset is called a *combination*. We do not care about the ordering of the elements. Simply what elements are chosen.

Definition 3.4.9 (Number of Combinations)**.**
The number of ways to select a subset of k objects from a set n distinct objects, $C(n, k)$, is denoted:

$$C(n, k) =_n C_k = \binom{n}{k}$$

The right hand side is what we call a *binomial coefficient*. We read this as “n choose k”.

Example 3.4.10.
How many ways are there to make a sundae with 3 toppings from an ice cream shop with 10 toppings? △

Definition 3.4.11 (Arrangements & Permutations)**.**
An *arrangement* is a selection <u>and an ordering of the subset</u>. Each resulting ordered arrangement is called a *permutation*.

Definition 3.4.12 (Number of Permutations)**.**
The number of ways to arrange k objects from a set of n distinct objects, $P(n, k)$, is denoted:

$$P(n, k) =_n P_k = (n)_k$$

Example 3.4.13.
Out of the 20 karaoke competitors, how many ways are there to schedule the performances of the 4 finalists? △

Example 3.4.14.
To recall a previous example “How many ways can we draw 4 labelled orbs (A,B,C,D) from an urn without replacement?” Using the multiplication principle, we found there was 4 options for the first draw, 3 for the second, and so on giving a count of $4 \cdot 3 \cdot 2 \cdot 1 = 4! = 12$. Equivalently, there are $P(4, 4)$ ways. △

Question 3.4.15.
How many ways are there to arrange k objects from a set of n distinct objects?

Answer 3.4.16.

- ***Identify Problem***: *How many ways are there to arrange k objects from a set of n distinct objects?*
- ***LHS***: *We know by definition $P(n, k)$ counts this.*
- ***RHS***: *We can think of assigning n objects to k boxes. For box 1, there are n possible objects that can be assigned to it. For box 2, there are $n - 1$ possible objects. Continuing on to the kth box, all $k - 1$ boxes before it have an object, so there are $n - (k - 1)$ possible objects that can be assigned box k. Using the multiplication principle, there are $n \cdot (n - 1) \cdot (n - 2) \cdot (n - (k - 1))$ ways to do this.*

- ***Conclusion**: Thus, there are $P(n,k) = n \cdot (n-1) \cdot (n-2) \cdot (n-(k-1))$ ways to arrange k objects from a set of n.*

The proof above allows us to immediately say the following as well:

Theorem 3.4.17 (Number of Permutations)**.**

$$P(n,k) = n \cdot (n-1) \cdot (n-2) \cdot (n-(k-1)) = \frac{n!}{(n-k)!}$$

where $n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$ and referred to as "n factorial". By convention, $0! = 1$.

Coming back to a version of our orbs and urn problem: "You draw from an urn with 4 labeled orbs (A,B,C,D) one by one without replacement. How many ways are there to do this?"

Answer 3.4.18.
We are arranging 4 objects from a set of 4. There are $P(4,4)$ ways to do this. So, there are $P(4,4) = \frac{4!}{(4-4)!} = 4!$ ways to do this.

We also asked the following: "You draw the orbs from the urn with replacement, you write down 4 labels. How many ways are there to do this so their is only a pair of repeats?" (e.g. (A,B,C,A) is okay but not (D,D,C,D)).

Answer 3.4.19.
Consider a set of four slots S in which you will write the labels. We select one of the labels to repeat, there are 4 ways to do this. We select 2 slots from the 4 slots for this label, there are $C(4,2) = \binom{4}{2}$ ways to do this. After this label has been placed, there are 3 labels to be arranged in 2 slots, there are $P(3,2) = \frac{3!}{(3-2)!} = 3!$ to do this. This means there are $4 \cdot C(4,2) \cdot P(3,2) = 4 \cdot \binom{4}{2} \cdot 3!$ ways in total to get exactly one pair in this case. This is a totally valid response to a combinatorics question. For an concise answer, we can simplify to obtain 144 ways.

Question 3.4.20 (To Ponder)**.**
How many ways are there to draw n labeled orbs from an urn with replacement?

Question 3.4.21 (To Ponder)**.**
How many ways are there to draw n labeled orbs from an urn with replacement only m times when $m < n$? What about $m > n$?

Now that we have some practice with combinations and permutations, let's get some practice with combinatorial proofs and show how they are different from non-combinatorial proofs.

Question 3.4.22.
Use a combinatorial proof to show that $P(n,k) = P(k,k)C(n,k)$.

Combinatorial Proof.

- **Identify Problem**: How many ways are there to arrange k objects from a set of n objects?
- **LHS**: By definition, $P(n,k)$ counts this.
- **RHS**: $C(n,k)$ counts the number of ways to select a subset of k objects from a set of n. This ignores the order of the objects. For each k-element selection, there are $P(k,k)$ ways to arrange the elements. Thus, by the multiplication principle, there are $C(n,k)P(k,k)$ ways to arrange k elements from a set of n objects.
- **Conclusion**: Therefore, $P(n,k) = P(k,k)C(n,k)$.

□

From the proof above, we can immediately algebraically deduce:

Theorem 3.4.23 (Number of Combinations)**.**

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

Non-Combinatorial Proof.

$$\begin{aligned} P(n,k) = P(k,k)C(n,k) &\implies \frac{n!}{(n-k)!} = \frac{k!}{(k-k)!}C(n,k) \\ &\implies C(n,k) = \frac{n!}{k!(n-k)!} = \binom{n}{k} \end{aligned}$$

□

What we've done here is simply use algebra and pre-established relationships to reach our result. We have counted nothing. This is an example of a *non-combinatorial proof.*

Using the above theorem, here is another good example of a non-combinatorial proof:

Theorem 3.4.24.

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

(That is, the number of ways to select k objects from n is equal to the number of ways to select $n-k$ objects from n.)

Non-Combinatorial Proof.

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

□

How could we change this into a combinatorial proof?

Theorem 3.4.25.

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

Combinatorial Proof.

- **Identify Problem**: How many ways are there to select k objects from a set of n?
- **LHS**: By definition, $\binom{n}{k}$ counts the number of ways to select k objects from a set of n objects.
- **RHS**: By definition, $\binom{n}{n-k}$ counts the number of ways to select $n-k$ objects from a set of n. By including $n-k$ objects in a set, we are creating a set of k excluded objects. If instead, we select the $n-k$ objects from a set of n to exclude, we are including the other k objects. Thus, $\binom{n}{n-k}$ counts the number of ways to select k from a set of n.
- **Conclusion**: Therefore, $\binom{n}{k} = \binom{n}{n-k}$.

□

Double counting gives us many ways to count the same thing! A few great examples:

Theorem 3.4.26.

$$\begin{aligned}\binom{n}{k} = \binom{n}{n-k} &= \frac{n}{n-k}\binom{n-1}{k} \\ &= \frac{n}{k}\binom{n-1}{k-1} = \binom{n-1}{k} + \binom{n-1}{k-1}\end{aligned}$$

3.5 Pigeonhole Principle

3.5.1 Warmup Problem

Question 3.5.1.
There are n married couples waiting at tables around a dance floor.

1. Pairing the $2n$ people up to dance, how many ways are there to select a pair of dance partners?
2. How many people must be on the dance floor to guarantee a married couple is somewhere on the dance floor?

Answer 3.5.2.

1. *This is a case where we would like to use our new notion of combination:*
   - *We are selecting 2 people from* $2n$ *people.*
   - *There are* $C(2n, 2) = \binom{2n}{2}$ *ways to do this.*
2. *We want to* guarantee *that a married couple is on the dance floor. If we invite* n *people to the floor, there is a chance we could have a couple. But, what if invite only one spouse from each couple? Then there is no couple on the floor. Consider this case.*
   - *Select one spouse from each married couple to the dance floor. There are now* n *people ready to dance.*
   - *If we invite any one else to the dance floor, we are guaranteed their spouse is on the dance floor.*
   - *So, to* guarantee *we have a couple on the dance floor, we need to invite* $n + 1$ *people to the floor.*

This is what we call the *pigeonhole principle*. It is an incredible powerful tool:

Theorem 3.5.3 (Pigeonhole Principle: Simple Form)**.**
If $n + 1$ *objects are distributed into* n *boxes, then at least one box contains two or more objects.*

Proof.
If each of the n boxes has at most one object, then the total number of objects is at most $1+1+1+\cdots+1 = n$. Since there is one object not accounted for, we can say there exists some box that contains at least two objects. □

Example 3.5.4. Out of Garrett, the Dean of the University, and the Chair of the Math & Stats Department, there are at least two people either inside Garrett's office or outside of Garrett's office. △

We can also exploit a more generally case of the pigeonhole principle when we have potentially more than $n + 1$ objects:

Theorem 3.5.5 (Pigeonhole Principle: Strong Form)**.**
Let $q_1, q_2, \dots, q_n$ *be positive integers. If*

$$q_1 + q_2 + q_3 + \cdots + q_n - n + 1$$

objects are distributed into n *boxes, then either the 1st box contains at least* q_1 *objects or the 2nd box contains at least* q_2 *objects,...,* *or the* n*th box contains at least* q_n *objects.*

Proof.

We distribute the $q_1 + q_2 + q_3 + \cdots + q_n - n + 1$ objects into n boxes. If each ith box contains less than q_i objects, then the number of objects in boxes is less than

$$(q_1 - 1) + (q_2 - 1) + \cdots + (q_n - 1) = q_1 + q_2 + \cdots + q_n - n$$

Since there is an object not accounted for, we can say there exists a box i with at least q_i objects. □

Example 3.5.6.

Out of the 21 students in Math 325, there are 3 whose birthday lie on the same day of the week. △

What both the simple form and strong form of the pigeonhole principle say in plain english is: ***distributing too many objects into too few boxes forces overlaps***.

Recall the 4 E's of combinatorics:

- Enumeration: How many ways...?
- Existence: Does there exist...?
- Extremal: What is the largest/smallest possible...?
- Expectation: What is the expected number of objects such that...?

We have been covering how to answer questions of enumeration. The pigeonhole principle is our first technique to answer questions of existence!

For example, we can answer the following existence question using the pigeonhole principle:

Question 3.5.7.

Consider choosing 101 integers from $S = \{1, 2, 3, 4, \ldots, 200\}$. Show that there are two chosen integers such that one divides the other. Hint: try factoring out as many 2's as possible from each term in S.

Answer 3.5.8.

Each number in $S = \{1, 2, 3, 4, \ldots, 200\}$ is either even or odd. Consider writing each integer $s \in S$ as $s = 2^k \cdot t$ where t is odd and $k \geq 0$. Each t comes from the set $T = \{1, 3, 5, \ldots, 199\}$ of size 100. By picking 101 integers from S, we are also picking 101 odd factors from T. Since 101=100+1 factors are chosen from 100 factors, by the pigeonhole principle, there are two integers chosen from S that share a factor from T. That is, from the chosen integers of S, there are two, say s_1 and s_2, such that $s_1 = 2^m \cdot t$ and $s_2 = 2^n \cdot t$. Either $m > n$ and s_2 divides s_1 or $n > m$ and s_1 divides s_2. As such, there are two integers such that one divides the other.

This problem exemplifies why the pigeonhole principle is so useful:

- How many ways are there to select 101 integers from $S = \{1, 2, 3, \ldots, 200\}$? $C(200, 101) = \binom{200}{101} > 10^{57}$.
- This is an insanely large number. More than the number of atoms in 1 million Earths (Earth is about $\approx 10^{51}$ atoms).
- But, without going through every single possible subset choice, we know something about *every single one of these choices*. Namely, that there must be two numbers in each such that one divides the other.

To end this section, let's work through a few problems:

Question 3.5.9.

Pretend we are going to group the 21 Math 325 students into 4 study groups. Each group will have a leader.

- How many ways can we select the leaders?
- Pretend we number the groups 1,2,3, and 4. How many ways can we assign leaders to groups?
- Use the pigeonhole principle to show that there is a group with at least 2 people.

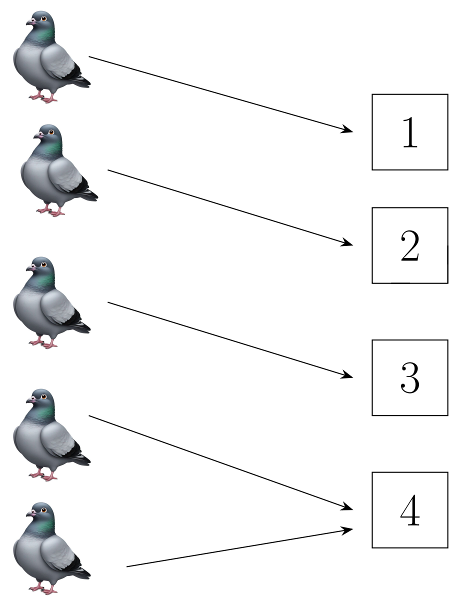


Figure 3.1: Simple example of the pigeonhole principle (using pigeons!)

- Use the pigeonhole principle to show that there is a group with at least 6 people.

Answer 3.5.10.

1. *We are selecting 4 people from 20. There are* $C(20, 4) = \binom{20}{4}$ *ways to do this.*
2. *We are arranging 4 people into 4 groups. There are* $P(4, 4) = (4)_4 = 4!$ *ways to do this.*
3. *We are distributing 21 students into 4 groups. Considering the first 5, we are placing 4+1 students into 4 groups. So, by the pigeonhole principle, there must be a group with at least two people.*
4. *Again, we are distributing 21 students into 4 groups. In other words, we are placing 21 = 6+6+6+6-4+1 people into 4 groups. By the pigeonhole principle, either group 1 has at least 6, group 2 has at least 6, and so on. So, there exists at least one group with at least 6 people in it.*

Chapter 4

Combinatorial Thinking Practicum

In this section, we'll go through a few key problems to develop our intuition for combinatorial problems.

4.1 Trailing Zeroes in a Product

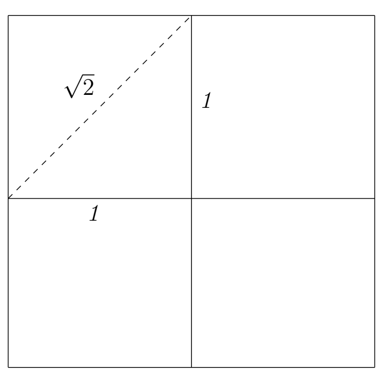
Question 4.1.1.
Consider the set $S = \{1, 2, 3, \ldots, 325\}$. How many zeroes are at the end of the product $1 \cdot 2 \cdot 3 \cdots \cdot 324 \cdot 325$? E.g. in $1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$ there is one zero at the end, but in $1 \cdot 2 \cdot 3 \cdot 4 = 24$ there are none.

Answer 4.1.2.
Potential Answer *: Note that there are more powers of two than 5 in our set (i.e. the powers of 5 are* $\{5, 25, 125\}$ *while powers of 2 are* $\{2, 4, 8, 16, 32, \ldots, 256\}$*). We want to pair up 2's and 5's until we run out of 5's. Try for yourself to verify there are 65 multiples of 5, 13 multiples of 25, and 2 multiples of 125 in* S*. Then, there are at least 80 zeroes. Try and find why this implies there are* exactly *80 zeroes.*

4.2 Points in a Square

Question 4.2.1.
Given five points inside a 2 by 2 square, show that there are two points whose distance is at most $\sqrt{2}$.

Answer 4.2.2.
Potential Answer *: Consider partitioning the square into 4 equally sized squares.*



The maximum distance between any two points is between the corners themselves. This distance is $\sqrt{2}$*. Since we have 5 points in 4 squares, by the pigeonhole principle, we have at least 2 points in one square. These two points could be on the corners, which means the distance between them is at most* $\sqrt{2}$*.*

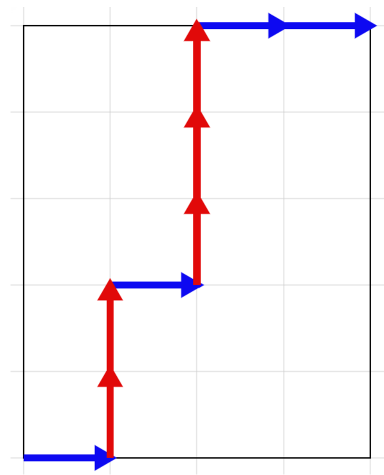
4.3 Lattice Paths

Definition 4.3.1.

Consider an $m \times n$ grid with corners $(0,0), (0,n), (m,0)$ & (m,n). A *lattice path* is a path starting at (0,0) and ending at (m,n) only moving right and up.

Example 4.3.2.

Below is an example of a lattice path on the 5×4 grid.



△

Question 4.3.3.

How many lattices paths are there an $m \times n$ grid?

Answer 4.3.4.

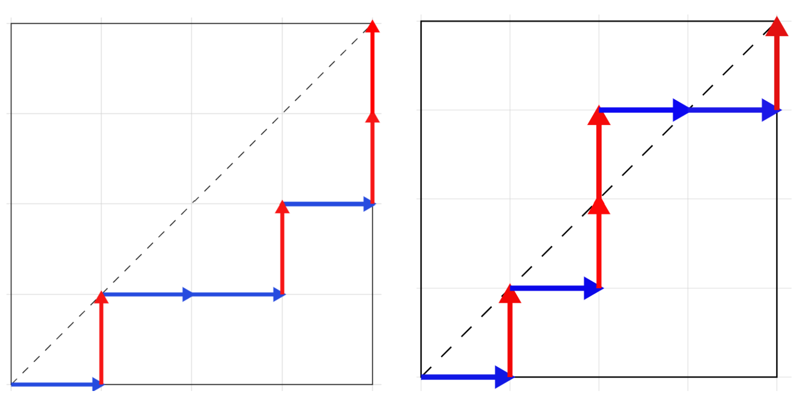
Potential Answer : *Consider the steps we must make as a list* $L = \{\_, \_, \_, \dots, \_\}$. *To get from* $(0,0)$ *to* (m,n), *we must move right* m *steps and we must move up* n *steps. In total, we must make* $m+n$ *steps (i.e.* $|L| = m+n$*). Note that if we choose when to go up/right, this determines when we must go right/up. There are* $\binom{m+n}{n}$ *ways to select which steps should be up's. We then fill the rest of the steps with right's. Likewise, we could have also counted the ways to place right's,* $\binom{m+n}{m}$, *and then place the up's. As such, there are* $\binom{m+n}{n} = \binom{m+n}{m}$ *lattice paths in this case.*

Question 4.3.5.

How many lattice paths on an $m \times n$ grid do not cross the line $y = x$? i.e. They can meet the points (i,i) but not $(i+1,i)$.

Example 4.3.6.

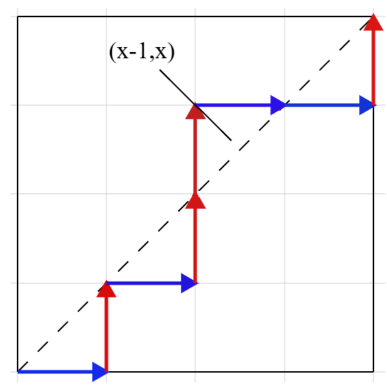
Below is an example of paths we want to consider. On the left, we have a path that does not cross $y = x$ while on the right we do.



△

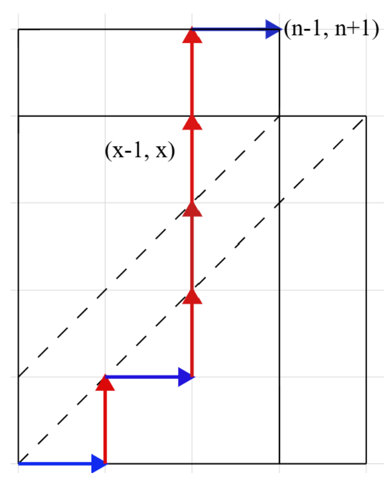
Answer 4.3.7.

Potential Answer : *Consider the point after a path crosses the diagonal. Using the example above, we have something like*



Assume we cross for some path. This point, $(x-1, x)$, we have gone up one more time than we'd like. This means we will have $n-x$ more ups and $n-x+1$ more rights to get to (n, n).

Keeping the path up to $(x-1, x)$ the same, let's reverse the direction of everything afterward. Before, we had one more right than up ($n-x+1$ rights and $n-x$ ups). Now, we will now have one more up than right ($n-x$ rights and $n-x+1$ ups). Since every step before $(x-1, x)$ is this same, we end up making $n-x$ more rights to end up at $n-1$ to the right and making $n-x+1$ ups to end up at $n+1$ up. So, we have a lattice path of the $(n-1)\times(n+1)$ grid.



This operation defines our function f taking paths that cross the diagonal in the $n \times n$ grid to lattice paths of the $(n-1)\times(n+1)$ grid. f^{-1} is simply taking a lattice path on the $(n-1)\times(n+1)$ grid and reversing steps after crossing the diagonal (try formalizing yourself!).

Since we have a bijection, there are the same number of paths that cross on the $n \times n$ grid as lattice paths on the $(n-1)\times(n+1)$ grid. There are $\binom{n-1+n+1}{n-1} = \binom{2n}{n-1}$ such lattice paths on the $(n-1)\times(n+1)$ grid. Removing this from the total $\binom{2n}{n}$ lattices path on the $n \times n$ grid, we get the number that do not cross the diagonal:

$$\binom{2n}{n} - \binom{2n}{n-1}$$

This expression forms an exceptionally useful sequence of numbers, the *Catalan Numbers*. They count many, many things including rooted binary trees, polygon triangulations, and more. For a reference list, see https://oeis.org/A000108.

4.4 Tilings

We are going to place objects on a $1 \times n$ board. "Tiling" the board means we cover the n spots exactly.

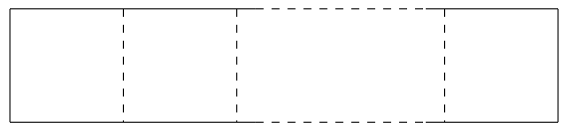


Figure 4.1: A $1 \times n$ board we will tile

Question 4.4.1.
Let t_n be the number of ways there are to tile the $1 \times n$ board. Find an expression for t_n using t_{n-1} and t_{n-2} (i.e. some function $t_n = f(t_{n-1}, t_{n-2})$).

Answer 4.4.2.
Potential Answer *: Consider a tiling of a $1 \times n - 1$ board. How can we extend this tiling to a tiling of $1 \times n$? By adding a 1×1 square. Likewise, we can extend a tiling of a $1 \times n - 2$ board to a tiling of $1 \times n$ by adding a domino. The extension of these tilings never produce the same tiling of the $1 \times n$ board. Moreover, any tiling must end in either a domino or a square. So, a tiling of the $1 \times n$ board is counted by these two cases. As such, $t_n = t_{n-1} + t_{n-2}$.*

Consider then, t_1. There is only 1 way to tile a 1×1 board, with a square. t_2 gives 2: two squares or one domino. You might try finding higher and higher values:

$$\begin{aligned} t_1 &= 1 \\ t_2 &= 2 \\ t_3 &= 3 \\ t_4 &= 5 \\ t_5 &= 8 \\ &\vdots \end{aligned}$$

You might immediately recognize the sequence, the *Fibonacci Numbers*! This is arguably the most ubiquitous sequence in combinatorics. The list of problems with Fibonacci numbers as the solution is immense. For a brief list, refer here https://oeis.org/A000045.

Question 4.4.3.
Recall we defined t_n to be the number of ways to tile the $1 \times n$ board with 1×1 squares and 1×2 dominoes. How many ways are there to tile a $2 \times n$ board with 1×2 dominoes?

Answer 4.4.4.
Potential Answer *: Bijection! (t_n = the number of tilings of the $2 \times n$ board with dominoes)*

- *f: Take a $2 \times n$ tiling with 1×2 dominoes. Split the board down the middle (horizontally) perhaps with some domino-cutting-strength scissors. We now have two copies of a tiling of the $1 \times n$ board with 1×1 squares and 1×2 dominoes.*
- *f^{-1}: Take two copies of the same $1 \times n$ tiling with 1×1 square and 1×2 dominoes. Stack them on top of one other. Take the squares that lay on top of one another and replace them with a vertical domino. We now have the tiling of the $2 \times n$ board with dominoes.*

4.5 {1,2}-sums

Question 4.5.1.

Using just 1 and 2, how many ways are there to form a sequence that sum to n? E.g. 1+1+2=1+2+1=4 are two examples of sequences that sum to 4.

Answer 4.5.2.

Potential Answer *: We have already solved this!*

- *f: Take a sum $q_1 + q_2 + q_3 + \cdots + q_k = n$ where each q_i is either 1 or 2. We can create a tiling of the $1 \times n$ board using k tiles as follows:*

  - *If $q_i = 1$, the i-th tile will be a square.*
  - *If $q_i = 2$, the i-th tile will be a domino.*

  Since the sum of the q_i terms must be n, we have a tiling of the $1 \times n$ board.

- *f^{-1}: Take a tiling of the $1 \times n$ board using k tiles. If tile i is a square, let $q_i = 1$. If tile i is a domino, let $q_i = 2$. Since the sums of the lengths of the tiles must be n, the sum of q_i terms must be n.*

We have a bijection, so there are the same number of ways to do each. Since there are t_n ways to tile the $1 \times n$ board, we have t_n ways to form the $\{1, 2\}$-sums.

Chapter 5

Combinatorial Models

5.1 Week's Warmups

Question 5.1.1.

How many ways are there to select 4 mathematicians and 10 statisticians from a group of 50 mathematicians and 60 statisticians?

Question 5.1.2.

How many ways are there to select 4 mathematicians or 10 statisticians from a group of 50 mathematicians and 60 statisticians?

Question 5.1.3.

I need to downsize. I select 3 scarves and 2 jackets to donate to the local thrift store. If I have 10 scarves and 6 jackets, how many ways can I do this?

Answer 5.1.4.

Selecting 3 scarves from 10 gives $\binom{10}{3}$. Selecting 2 jackets from 6 gives $\binom{6}{2}$. So, there are $\binom{10}{3}\binom{6}{2}$ ways to do this.

5.2 What is a Combinatorial Model?

Consider this final warmup problem. Notice, I could have reversed this process as well, selecting jackets then scarves. So, $\binom{10}{3}\binom{6}{2} = \binom{6}{2}\binom{10}{3}$. But, what if I instead asked the following question? (To motivate this section's topic)

Question 5.2.1.

Show $\binom{10}{3}\binom{6}{2} = \binom{6}{2}\binom{10}{3}$.

Extremely easy to show algebraically, but we can double count!

Proof.

First, we must identify a problem that these formulae count. Perhaps, we can hypothetically select scarves and jackets!

- **Problem**: How many ways are there to select 3 scarves and 2 jackets from 10 scarves and 6 jackets?
- **LHS**: I don't care about my scarves as much. So, we can select the 3 scarves first in $\binom{10}{3}$ ways. We can select the 2 jackets second in $\binom{6}{2}$ ways. So, in total we can select 3 scarves and 2 jackets from 10 scarves and 6 jackets in $\binom{10}{3}\binom{6}{2}$ ways.
- **RHS**: I need time to admire my scarf collection. So, we can select the 2 jackets first in $\binom{6}{2}$ ways then the 3 scarves in $\binom{10}{3}$ ways. So, in total we can select 3 scarves and 2 jackets from 10 scarves and 6 jackets in $\binom{6}{2}\binom{10}{3}$ ways.

- **Conclusion**: Therefore, $\binom{10}{3}\binom{6}{2} = \binom{6}{2}\binom{10}{3}$.

□

The trick is that I created a scenario to help me prove the combinatorial identity $\binom{10}{3}\binom{6}{2} = \binom{6}{2}\binom{10}{3}$. This is the idea behind a *combinatorial model*. For clarity, recall the process of double counting:

Definition 5.2.2. Double counting is the proof technique commonly used in combinatorics to prove identities. Consider a question of the form "Show ______=______". Double counting would proceed as follows:

- **Identify Problem**: Find a problem that the LHS and RHS count solutions to.
- **LHS**: Count once to obtain a formula, LHS.
- **RHS**: Count the same problem a different way to obtain a formula, RHS.
- **Conclusion**: Therefore, LHS=RHS.

Often, the hard part is identifying a problem to count. This is where *combinatorial models* come in. Much like the jacket and scarf scenario, we can use a hypothetical scenario that is counted by the identity we are given (LHS=RHS).

5.3 Tiling Model

We have already discussed an important combinatorial model: the ***tiling model***. Recall, we proved the following:

Theorem 5.3.1. *The number of tilings of a* $1 \times n$ *board with squares and dominoes,* t_n*, is given by* F_n*, the* (n)*-th Fibonacci number.*

Where the Fibonacci number F_i is defined by $F_i = F_{i-1} + F_{i-2}$ for $i \geq 2$ with $F_0 = 1$ and $F_1 = 1$.

We then used the tiling model to prove the following result:

Theorem 5.3.2. *Using just 1 and 2, there are* F_n *ways to form a sequence that sums to* n*. E.g. 1+1+2=1+2+1=4 are two examples of sequences that sum to 4.*

What we did was find a correspondence between a $1 \times n$ tiling using squares and dominoes with sums of 1's and 2's.

This is the key: ***we took an identity we wish to solve and used an interpretable scenario to prove it***. Combinatorial models are scenarios that allow us to prove results. Just like with bijections, combinatorial models allows us to take a problem we do not know how to solve and equate it with one we do.

Question 5.3.3.
Using a tiling model, show that $F_0 + F_1 + F_2 + \cdots + F_n = F_{n+2} - 1$

Answer 5.3.4.
Problem*: How many ways are there to tile a* $1 \times (n+2)$ *board using at least one domino?*

RHS*: Taking all the tilings of the board, we have* F_{n+2}*. Removing the tiling of all squares, we have the number of tilings with at least one domino,* $F_{n+2} - 1$*.*

LHS*: Consider the rightmost domino in a tiling of the* $1 \times (n+2)$ *board. To the right of this domino, there will be all squares. Assuming there are* k *squares to the right of this rightmost domino, we have a board of length* $n-k$ *to the left of the domino. This board can be tiled in* F_{n-k} *ways. Since we can have* $k = 0$ *squares to the right all the way to* $k = n$ *squares to the right, we consider tilings of each* $1 \times (n-k)$ *board* *for all $k = 0, 1, 2, \ldots, n$. We can't have a shared tiling of differently sized boards, so we add F_{n-k} for each size of board. Thus, we have $F_0 + F_1 + \cdots + F_n$ total ways to tile the $1 \times (n+2)$ board.*

***Conclusion**: Since these expressions count the number of ways to tile a $1 \times (n+2)$ board using at least one domino, we conclude $F_0 + F_1 + F_2 + \cdots + F_n = F_{n+2} - 1$.*

5.4 Lattice Path Model

We also discussed another important model: the ***lattice path model***. Recall, we found the following:

Theorem 5.4.1. *There are $\binom{m+n}{m}$ lattice paths on a $m \times n$ grid.*

Theorem 5.4.2. *There are $\binom{2n}{n} - \binom{2n}{n+1}$ lattice paths on the $n \times n$ grid that do not cross the $y = x$ line.*

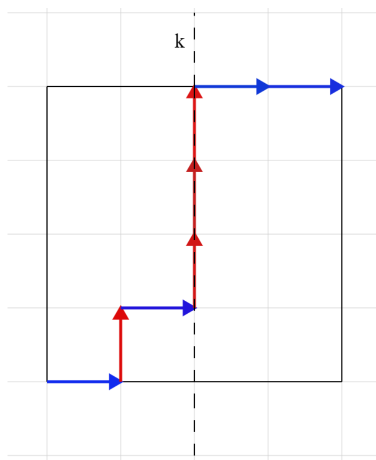
Again, we can use this model to prove other results:

Question 5.4.3.
Use a lattice path model to show $\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2$ (or equivalently $\binom{2n}{n} = \sum\limits_{k=0}^{n} \binom{n}{k}^2$).

Answer 5.4.4.
***Problem**: How many lattice paths are there on the $n \times n$ grid?*

***LHS**: We know from before, that there $\binom{2n}{n}$ lattice paths.*

***RHS**: Consider some column k on our grid:*



At some point, every lattice path must meet each column $k = 0, 1, 2, \ldots, n$ by definition of a lattice path. Using the same intuition when proving the general formula for lattice paths, consider the path we took to meet the column k. We must have gone right k times. There are $\binom{n}{k}$ ways to select when to go right. After meeting this column, there are $n-k$ more right moves to make. We can do this in $\binom{n}{n-k}$ ways. So, there are $\binom{n}{k}\binom{n}{n-k} = \binom{n}{k}\binom{n}{k} = \binom{n}{k}^2$ ways for each k. Since selecting when to move right determines when we move up, we have covered every lattice path. Therefore, there are $\sum_{k=0}^{n} \binom{n}{k}^2$ total lattice paths.

***Conclusion**: Since both formulas count how many lattices paths there are, $\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$.*

5.5 Committee Selection Model

The last model we'll discuss is the model of **committee selection**. We consider the following simple scenario:

- There are some number of people in a club.
- They select committee members.
- Perhaps they select a chair or some number of chairs for the committee.

We then count the number of ways they could make this selection. Apparently simple, but very useful!

Question 5.5.1.
Use a committee selection model to show

$$n\binom{2n}{n} = (n+1)\binom{2n}{n+1}$$

Answer 5.5.2.
Problem: *How many ways are there to select a committee of size n with a chair from a club of $2n$ people?*

LHS: *The chair could be considered a committee member. We could select n committee members in $\binom{2n}{n}$ ways. Then, select a chair from the committee members in $\binom{n}{1} = n$ ways. So, there are $n\binom{2n}{n}$ ways to create a committee of size n with a chair.*

RHS: *The chair could not be considered a committee member. We could select $n+1$ people who could be a committee member or the chair in $\binom{2n}{n+1}$ ways. We could then select the chair from the $n+1$ people in $\binom{n+1}{1} = n+1$ ways. So, there are $(n+1)\binom{2n}{n}$ ways to create a committee of size n with a chair.*

Conclusion: *Since the LHS and RHS count the number of ways to form a committee in this case, $n\binom{2n}{n} = (n+1)\binom{2n}{n}$.*

5.6 Committee Selection Model

Question 5.6.1.
Use a committee selection model to show that

$$k\binom{n}{k} = (n-k+1)\binom{n}{k-1}$$

Answer 5.6.2.
Problem: *How many ways are there to select a committee of size k with a chair from a club of n people?*

LHS: *Selecting the committee can be done in $\binom{n}{k}$ ways. From that committee, we can select a chair in $\binom{k}{1} = k$ ways. So, there are $k\binom{n}{k}$ total ways we can select a committee with chair in this case.*

RHS: *Selecting the $k-1$ committee members who are* not *the chair can be done in $\binom{n}{k-1}$ ways. We can then select the chair from the leftover $n-(k-1)$ people in $\binom{n-k+1}{1} = n-k+1$ ways. So, there are $(n-k+1)\binom{n}{k-1}$ total ways we can select a committee in this case.*

Conclusion: *Since the LHS and RHS count the number of ways to form a committee in this case, $k\binom{n}{k} = (n-k+1)\binom{n}{k-1}$.*

5.7 Combinatorial Model Practice

Question 5.7.1. For each of the following, use the given model to describe what the LHS and RHS of the given identities is counting in terms of the model.

- Committee Selection: $\binom{m+n}{3} = \binom{m}{3} + \binom{n}{3} + \binom{m}{2}n + \binom{n}{2}m$

- Tiling: $\sum_{k=0}^{n} \binom{n-k}{k} = F_n$
- Lattice Path: $\frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n+1}$

Question 5.7.2. Show $F_{m+n} = F_m F_n + F_{m-1} F_{n-1}$ using a tiling model.

5.8 Week's Warmup

Question 5.8.1.

Given 5 points on the surface of a sphere, show that there is a closed hemisphere that contains 4 points. (A closed hemisphere is a sphere cut in half including the boundary that was cut)

Question 5.8.2.

1. How many ways are there to select k objects from n?
2. After selecting, how many ways are there to permute those k objects?
3. How many ways are there permute k objects from n?

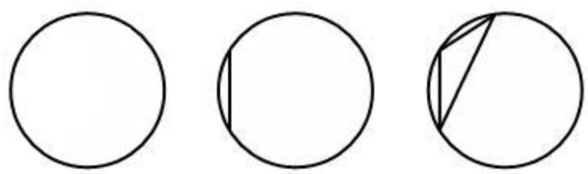
Courtesy of Zvezdelina Stankova, a combinatorial thinking builder:

Question 5.8.3.

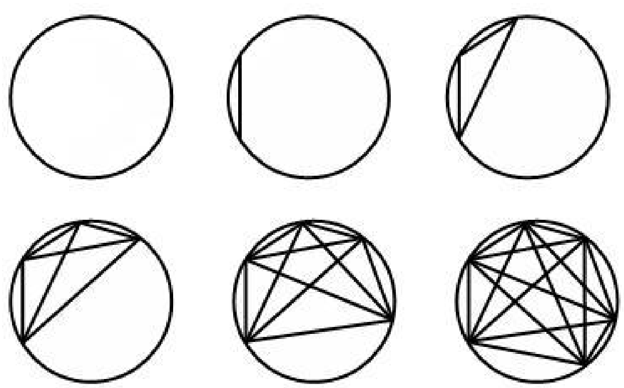
Given n points on the boundary of a circle, connect each pair of points such that only two lines intersect at any point. How many regions does this create?

Answer 5.8.4.

For 1 point, this is an easy answer: 1 region, the circle itself. For 2, also fairly straightforward: 2. For 3, we get 4. Powers of 2! Pictures from: ThreeSixty360.



We can continue this process and find the answer for higher numbers of points:

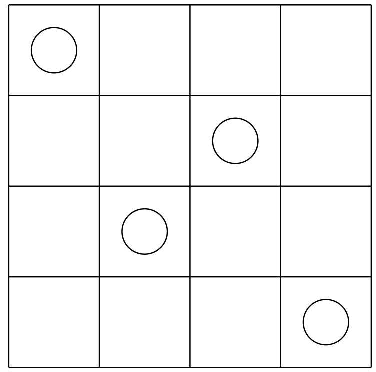


With a fair bit of zooming in, you may count 1,2,4,8,16, and then... 31. This strange sequence exemplifies the need for combinatorial reasoning. Although we may conjecture a relationship between a sequence and a process like this, without proof, we know close to nothing! So, be careful when conjecturing.

Question 5.8.5.

Given an $n \times n$ board, how many ways can we place n points so that no point shares a column or row with any other point?

This question formulated differently, is what we call the "non-attacking rook position" problem. Given an $n \times n$ chess board, we want to places n rooks so that no rook can attack any others. As an example, consider the following placement.



Each rook has been placed in a spot such that it is the only rook in its respective row and column.

Answer 5.8.6.

Going column by column, consider placing a rook in the first column. No other have been placed, so we can place this rook in n ways. Now that this rook has been placed, the rook we will place in second column can be placed anywhere so long as it doesn't share the row from the first rook. There are $n-1$ ways we can do this. Likewise, we can place the rook to be in column 3 in any row except the rows of the 2 previous. There are $n-2$ ways. Continuing on, we have $n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = n!$ total ways. (One could equally think of this as a permutation problem!)

5.9 Graphs

One fundamental combinatorial object we study in combinatorics is the *graph*. You may have already seen these. For example, see Figure 5.1. We will formalize the idea in this section.

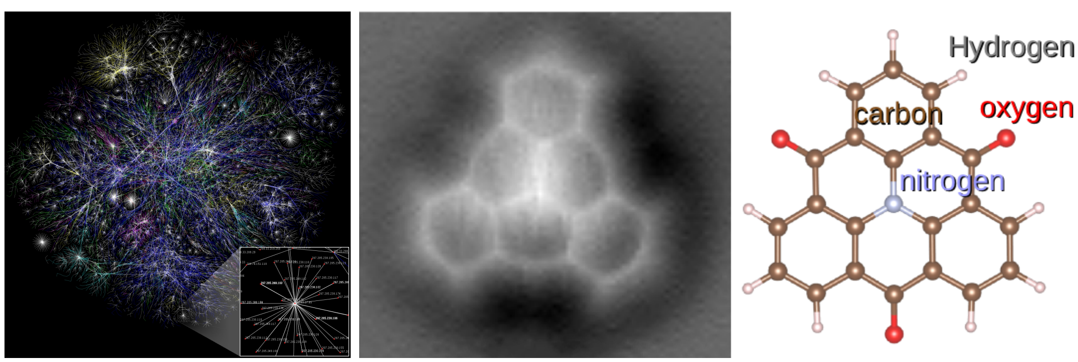


Figure 5.1: Examples of graphs. On the left, a network of the internet. On the right, a graph corresponding to a molecule. Courtesy of https://en.wikipedia.org/wiki/Computernetwork and https://en.wikipedia.org/wiki/Molecule respectively.

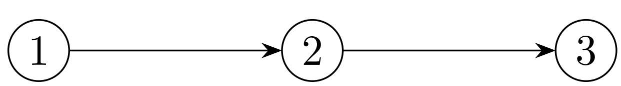
Definition 5.9.1.

A *graph* $G = (V, E)$ is a collection of vertices V and edges E where E is a set of pairs of vertices.

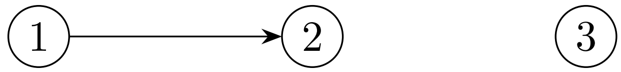
We can synonymously call vertices and edges nodes and arcs, but I will try to remain consistent with the former. We typically assume E is a set of edges defined over V. Otherwise, this object would indeed be hard to interpret. Consider the following graphs.

Example 5.9.2.

Let $G = (\{1,2,3\}, \{(1,2),(2,3)\})$. Then, we can visualize G as



Notice, if we remove $(2,3)$, we end up with a different, yet similar graph $H = (\{1,2,3\}, \{(1,2)\}$

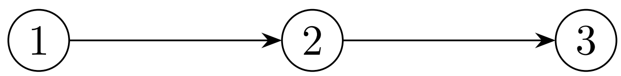


△

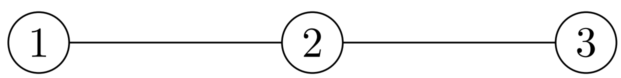
Note that in drawing these graphs, we have an ordered pair (x, y) telling us our edge begins at x and ends at y. As such, we have an arrow indicating the starting node and ending node for each edge. A graph where this *direction* is important we call a *directed graph*. If instead, order does not matter, we have an *undirected graph*.

Definition 5.9.3.
A graph $G = (V, E)$ is *directed* when E consists of *ordered pairs* of vertices (e.g. $(2,3)$). Conversely, a graph is *undirected* when E consists of *unordered pairs* of vertices (e.g. $\{2,3\}$).

Example 5.9.4. Consider the graph from before, $G = (\{1,2,3\}, \{(1,2),(2,3)\}$.

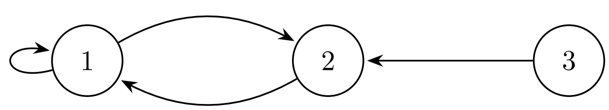


We can form an undirected analogue, $H = (\{1,2,3\}, \{\{1,2\},\{2,3\}\}$.

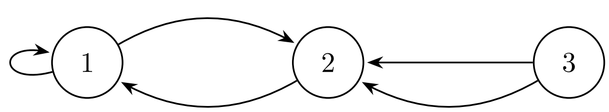


△

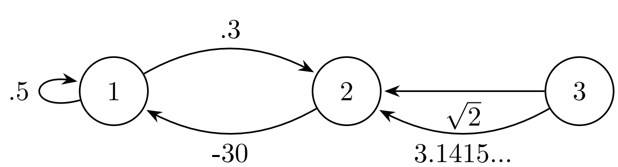
Note, we do not require that the vertices in an edge be distinct. That is, we could have a *self-loop* (x, x) for some $x \in V$.



We also do not require that there is only one edge between a pair of nodes.



We can even weight edges!



These objects can be very complicated indeed. As such, we can further restrict our graphs to what we call *simple*.

Definition 5.9.5.
A *simple* graph is an undirected, unweighted graph without self-loops and multiple edges.

A simple graph is arguably the easiest graph to think of. But, simple and not simple graphs have the pros and cons when dealing with them in practice. Consider the following question:

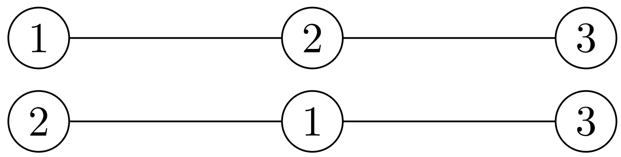
Question 5.9.6.

How many simple graphs in general are there on n nodes?

Answer 5.9.7.

For an edge to exist, we have to pair up two nodes. How many ways can we do this? $\binom{n}{2}$ *ways. For each of those potential edges, there is 2 possibilities: the edge exists or it doesn't. So, by multiplication principle, there are* $2^{\binom{n}{2}}$ *possible graphs in general.*

Okay, $2^{\binom{n}{2}}$ in general. But, consider the two following graphs.



Intuitively, these are representatives of the same structure. One could say these are the same graph with different labels. This idea is what we call *isomorphic graphs*. The question of "How many simple *non-isomorphic* graphs are there on n nodes?" is a much more difficult question to answer. We will come back to this.

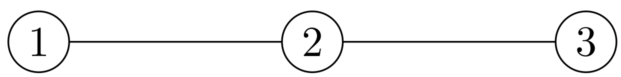
There are many features of graphs we care about. One particular example is the number of edges each node is represented in.

Definition 5.9.8.

The number of edges incident with a node v is the *degree* of v denoted $d(v)$.

Example 5.9.9.

For the following graph, the degree of each node is: $d(1) = 1$, $d(2) = 2$, and $d(3) = 1$.



△

Theorem 5.9.10 (Handshaking Lemma)**.**

Let $G = (V, E)$*. Then,*

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

Proof.

For each node v, $d(v)$ counts the number of edges v appears in (i.e. the number of edges (v, u) or $\{v, u\}$). The sum over all v will count (v, u) and (u, v) for every edge in the graph. As such, we count twice the number of edges in the graph. □

Corollary 5.9.11.

In a graph G, there are an even number of odd degree nodes.

Theorem 5.9.12.

Let $x, y \in$*. Then,*

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Proof.

On the left, we have n factors of $(x + y)$, $(x + y)^n = (x + y)(x + y) \cdots (x + y)$. The coefficient of x^k will be the number of ways to choose x from the total factors. Since we have n, we also choose y^{n-k} as a consequence. This corresponds to selecting k terms to contribute to the coefficient of $x^k y^{n-k}$ from n factors. Since we can do this in $\binom{n}{k}$ ways, the coefficient of $x^k y^{n-k}$ is $\binom{n}{k}$. Adding from $k = 0$ to $k = n$ gives *all* the ways we can do this. □

Question 5.9.13.

Using the binomial theorem, how many *directed* graphs are there on n nodes in general?

Chapter 6

Recurrence Relations & Generating Functions

6.1 Week's Warmup

Question 6.1.1.
Given the following expressions, we'll assume $n \geq 0$ is an integer. Using the given values, we can plug in to the expression to find the next term in the sequence. Conjecture a function $a_n = f(n)$ that does not depend on any terms before it in the sequence.

- $a_n = 2a_{n-1} + 1$ where $a_0 = 0$
- $a_{n+1} = a_n + a_{n-1}$ where $a_0 = 0$ and $a_1 = 1$
- $a_n = (a_{n-1})^{\frac{1}{n}}$ where $a_0 = 1$

Answer 6.1.2.

- *Plugging in* $a_0 = 0$ *to obtain* a_1*, we get* $a_1 = 2(0) + 1 = 1$*. We then use this to obtain* a_2*. We get* $a_2 = 2(1)+1 = 3$*. Repeating this process we get* 7 *then* 15 *then* 31 *and so on. The sequence we get looks something* $\{1, 3, 7, 15, 31, 63, 127, \dots\}$*. You may notice these are almost powers of 2... any conjectures for* $f(n)$*?*
- *Plugging in similar to before, we get* $\{0, 1, 1, 2, 3, 5, 8, \dots\}$ *the famous sequence we've been calling the "Fibonacci Sequence". What would* $f(n)$ *look like?*
- *A rather boring sequence, we get* $\{1, 1, 1, 1, \dots\}$*. A simple formula for* $f(n)$*:* $f(n) = 1$*.*

But, how can we find $f(n)$ *in general? That is the topic of this section.*

The mathematical object we have just interacted with is what we call a *recurrence relation*. Firstly, note the following example.

Example 6.1.3.
$a_{n+1} = 3a_n - 1$ △

This by itself does not generate a sequence. Notice in our warmup, we were given some known values. These are necessary to determine a sequence (otherwise why do we use them to count??).

Definition 6.1.4.
An equation relating the n-th term in a sequence as a combination of the previous terms is called a *recurrence relation*. The given first element (multiple elements) is called the initial value (values).

Once we have some initial values, the rest of the sequence is generated uniquely. We will typically deal with integer valued recurrence relations, but this need not be the case.

Example 6.1.5.

$a_{n+1} = 3a_n - 1$ where $a_0 = 2$ △

So, how do we find the function $a_n = f(n)$? Though we may able to guess at what this function may be, we need a mathematically rigorous way to determine $f(n)$. What we will do is use a new type of machinery: *generating functions*. Recall the following:

Theorem 6.1.6.

The geometric series $\sum_{k=0}^{\infty} x^k$ *converges to* $\frac{1}{1-x}$ *when* $|x| < 1$*. That is, if* $|x| < 1$*, then*

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

Under this assumption $|x| < 1$, we can find a simple, closed formula for an infinite sum. We will implicitly assume this is satisfied for the series we write. If we pretend we had $(cx)^k$ instead of x^k for some constant c, we will implicitly assume $|cx| < 1$ so that we get a nice formula for the infinite series.

Definition 6.1.7.

The *ordinary generating function* (OGF) of some sequence $\{a_0, a_1, a_2, a_3, a_4, \ldots\}$ is the *power series* of the form

$$A(x) = \sum_{k=0}^{\infty} a_k x^k$$

Theorem 6.1.8.

Let the OGFs of $\{a_0, a_1, a_2, a_3, a_4, \ldots\}$ *and* $\{b_0, b_1, b_2, b_3, b_4, \ldots\}$ *be* $A(x)$ *and* $B(x)$ *respectively. Also, let* c *be some scalar. Then,*

- $A(x) + B(x)$ *generates* $\{a_0 + b_0, a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4, \ldots\}$
- $cA(x)$ *generates* $\{c \cdot a_0, c \cdot a_1, c \cdot a_2, c \cdot a_3, c \cdot a_4, \ldots\}$
- $x^k A(x)$ *generates* $\{\underbrace{0, 0, \ldots, 0}_{k\ zeroes}, a_0, a_1, a_2, a_3, a_4, \ldots\}$
- $\frac{d}{dx}(A(x)) = A'(x)$ *generates* $\{a_1, 2a_2, 3a_3, 4a_4, 5a_5, \ldots\}$
- $A(x) \cdot B(x)$ *generates* $\{c_0, c_1, c_2, c_3, c_4, \ldots\}$ *is where* $c_n = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0$
- $\frac{A(x)}{1-x}$ *generates* $\{s_0, s_1, s_2, s_3, s_4, \ldots\}$ *where* $s_n = \sum_{k=0}^{n} a_k$

Theorem 6.1.9.

Consider $ax^2 + bx + c$*. Let*

$$q_1(x) = \frac{-b + \sqrt{b^2 - 4ac}}{2x}$$
$$q_2(x) = \frac{-b - \sqrt{b^2 - 4ac}}{2x}$$

Then, we can factor $ax^2 + bx + c$ *in the following ways*

$$\begin{aligned} ax^2 + bx + c &= a(x - q_1(a))(x - q_2(a)) \\ &= c(1 - q_1(c)x)(1 - q_2(c)x) \end{aligned}$$

Formula	Series	Sequence
$\frac{1}{1-x}$	$\sum_{k=0}^{\infty} x^k$	$\{1, 1, 1, \dots\}$
$\frac{c}{1-bx}$	$\sum_{k=0}^{\infty} cb^k x^k$	$\{c \cdot b^0, c \cdot b^1, c \cdot b^2, \dots\}$
$\frac{1-x^n}{1-x}$	$\sum_{k=0}^{n-1} x^k$	$\{\underbrace{1, 1, 1, \dots, 1}_{n\ ones}, 0, 0, \dots\}$
$\frac{x^n}{1-x}$	$\sum_{k=n}^{\infty} x^k$	$\{\underbrace{0, 0, \dots, 0}_{n\ zeroes}, 1, 1, 1, \dots\}$
$\frac{d}{dx}\left(\frac{1}{1-x}\right)$	$\sum_{k=1}^{\infty} kx^{k-1}$	$\{1, 2, 3, 4, \dots\}$

Chapter 7

Principle of Inclusion-Exclusion (P.I.E.) & the D.I.E. Method

7.1 Principle of Inclusion-Exclusion

Thus far, we've covered a very simple counting problem using set theory: For two sets A, B, if they are disjoint ($|A \cap B| = 0$), we know that $|A \cup B| = |A| + |B|$. For example, recall the problem of disjoint sets of students: "How many people are in Math 325 or Physics 533 if they occur at the same time?". That is, to count the number of things in the set $A \cup B$ we could just count the number of things in A once, count the number of things in B once, and sum them.

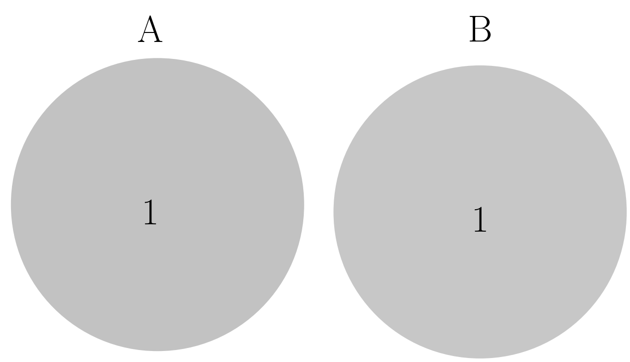


Figure 7.1: In the disjoint case, $|A \cup B| = |A| + |B|$.

However, in the case that A and B are not disjoint, we run into a problem. If we count the number of elements in A, then count the number of elements in B, we end up counting their shared elements twice!

So, to count $|A \cup B|$, we must correct this over counting by *excluding* a copy of the set we over-counted. Counting A, we get $|A|$. Counting B, we get $|B|$. Since we count $|A \cap B|$ twice, we can count $A \cup B$ in a straightforward way: $|A \cup B| = |A| + |B| - |A \cap B|$. That is, by counting A and B, we over-count $A \cap B$, so we correct this to find $|A \cup B|$ (see Figure 7.2 for a visual aid).

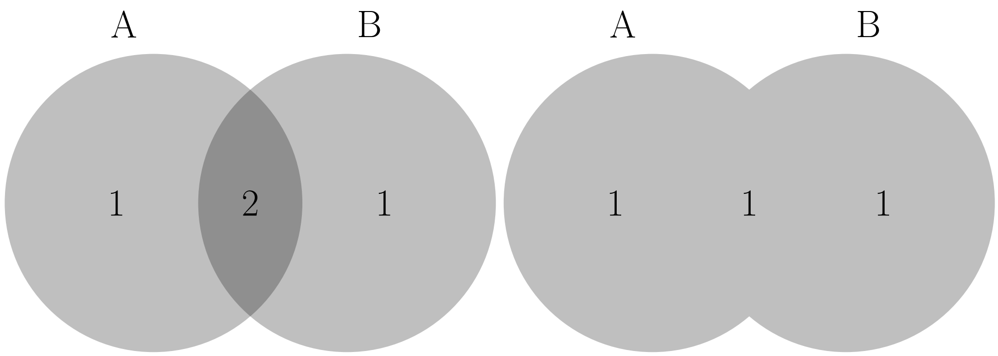


Figure 7.2: On the left, we have $|A| + |B|$ where we count $A \cap B$ twice. To fix this, on the right, we remove a copy of $|A \cap B|$ to get the true count $|A \cup B| = |A| + |B| - |A \cap B|$

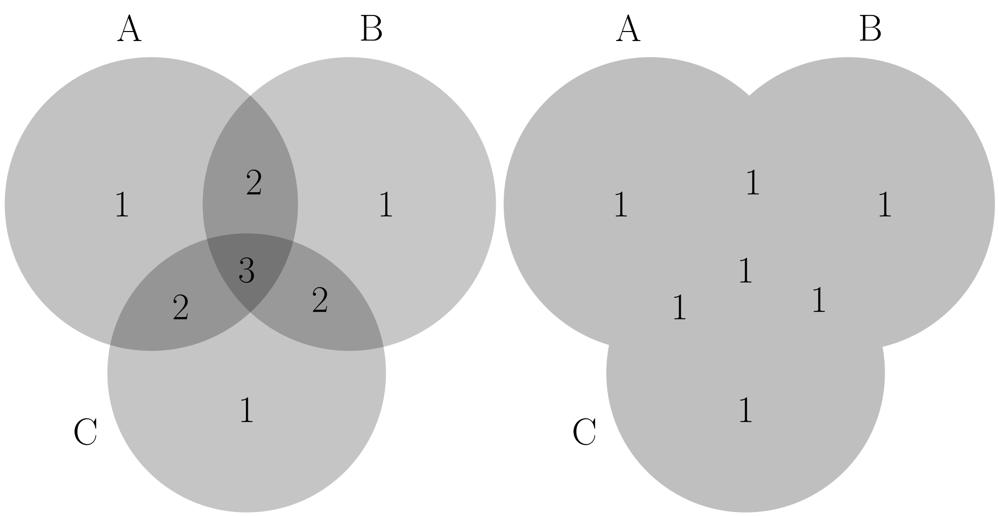


Figure 7.3: On the left, we count A then B then C to get $|A| + |B| + |C|$. We need to remove the pairwise intersections once. This removes the triple-wise intersection entirely so we include it back: $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$

This is the idea of the *Principle of Inclusion-Exclusion (P.I.E.)*: We want to count the union of sets so...

- Include count of the individuals
- Exclude count of the pairwise intersections Include count of the triple-wise intersections
- Exclude count of the quadruple-wise intersections
- Include count of the quintuple-wise intersections

$$\vdots$$

$$\vdots$$

For those who are curious, there is a formula:

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1} |A_i| - \sum_{1\le i<j\le n} |A_i \cap A_j| + \sum_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1}|A_1 \cap \cdots \cap A_n|$$
$$= \sum_{\emptyset \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|+1} \left|\bigcap_{i\in I} A_i\right|$$

though this is algebraic formulation is not often used in practice.

7.2 D.I.E. Method

Oftentimes we are faced with a sum of the following form:

$$\sum_{k=1}^{n} (-1)^k f_k = (-1)^n f_{n-1}$$

where f_k is the k-th Fibonacci number. Luckily, we have been equipped with experience with these numbers. We may even have an idea of rephrasing this sum combinatorially. For example, we know that f_k counts the number of ways to tile a $1 \times k$ board with squares and dominos. This alternating seems to define a relationship between the tilings. How might we prove such a sum?

Thinking this specific example through, we see for even k, we have a coefficient 1. For odd k, we have a coefficient -1. It is almost like the even tilings are naturally assigned 1 and odd tilings are naturally assigned -1. This sum seems to tell us that the set of tilings of every size has been split in two. But, for some reason they haven't been paired up all the way (i.e. why is there a $(-1)^n f_{n-1}$ on the right hand side?).

This vein of thinking led to the useful cousin of inclusion and exclusion called the D.I.E. method. It proceeds as follows:

1. *Desciption*: Describe what the sum counts *ignoring* the alternating
2. *Involution*: Define a sign-reversing involution between the sets of opposite parity
3. *Exception*: Count how many exclusions occur

Without defining involution yet, we'll discuss how a result might be proven. We already have a combinatorial model set up, so let's try the D.I.E. method out:

Question 7.2.1.
Prove that

$$\sum_{k=1}^{n} (-1)^k f_k = (-1)^n f_{n-1}$$

Answer 7.2.2.

1. *Desciption: Ignoring the alternating, $\sum_{k=1}^{n} f_k$ counts the number of ways to tile $1 \times k$ boards of size 1 through n using squares and dominos.*
2. *Involution: Recall that even tilings are assigned 1 by our sum and odd tilings assigned -1. Consider "toggling" the last tile of a tiling by replacing it with the opposite tile. That is, if the tiling ends in a square, remove it and place a domino. Likewise, if the tiling ends in a domino, remove it and place a square. Since removing a square and replacing it with a domino increases the size by 1, the tiling length changes from even to odd or odd to even. Likewise, replacing a domino with a square changes the length of the tiling from even to odd or odd to even. So, no matter what, we change "parity" (even/odd switch) and are "sign-reversing" ($1/{-1}$ switching). Also, if we toggle the last tile twice, we result in the same tiling and are an "involution".*

3. Exception: *What could go wrong? Well, if we have a tiling of the n board that ends in a square, we will toggle it and end up with a tiling of the $n+1$ board. But, we don't care about these tilings because our sum ends at n-sized boards. So, how many are there? Well, if our n board tiling ends in a square and we remove it, what do we have? A tiling of the $n-1$ board. There are f_{n-1} such tilings. So, whatever n is assigned, we have not paired up f_{n-1} tilings in that assignment. That is, we have $(-1)^n f_{n-1}$ as a leftover term.*

4. Conclusion: *We have paired up each odd tilings with an even tilings via a sign reversing involution* except f_{n-1} *tilings assigned $(-1)^n$. So,*

$$\sum_{k=1}^{n}(-1)^k f_k = (-1)^n f_{n-1}$$

Voilà! We have cleverly paired up our tilings to prove an otherwise hard to intuit sum. This is the idea of D.I.E. method. Another example for clarity:

Question 7.2.3.
Prove that

$$\sum_{k=0}^{n}(-1)^k \binom{n}{k} = 0$$

Answer 7.2.4.

1. Desciption: *Ignoring the alternating, $\sum\limits_{k=0}^{n}\binom{n}{k}$ counts the number of subsets of $\{1, 2, \ldots, n\}$ of sizes 1 through n.*

2. Involution: *Note that for even k, we have a coefficient of 1. Meanwhile, odd k gives a coefficient of -1. So, even sized subsets are assigned 1. Meanwhile, odd subsets are assigned -1. Consider "toggling" the element 1 in a subset. That is, if 1 is in the subset, remove it. If not, add it. Then, the parity of the subset will change (even to odd or odd to even) and so the assignment (1 to -1 or -1 to 1). So, this operation is sign-reversing. Removing 1 then including 1 will result in the same subset. Likewise with including 1 and then removing it again. So, this operation is an involution.*

3. Exception: *If a subset has 1 in it and we remove it, we obtain a subset of opposite parity. Otherwise, the subset does not have 1 and we add it to obtain a subset of opposite parity. So, each subset is paired up and we have no exceptions.*

4. Conclusion: *We have paired up each odd sized subsets of $\{1, 2, \ldots, n\}$ with even sized subsets of $\{1, 2, \ldots, n\}$ via a sign reversing involution with no exceptions. So,*

$$\sum_{k=0}^{n}(-1)^k \binom{n}{k} = 0$$

Pretty clever! Hopefully by now you have some idea for what a "sign-reversing involution" does. Formally, this is what we are after:

Definition 7.2.5. A function $f : A \to B$ is an *involution* if $f(f(a)) = a$ for all $a \in A$. The involution f is then *sign-reversing* if $f(a)$ and a differ in sign for all $a \in A$.

Chapter 8

Probability, Graphs, & Ramsey Theory

8.1 Probability Basics

You probably (pun-intended) have an idea of what probability is. We experience this most every day of our lives. Thoughts like "I figure the bus will be late today" or "I think I will pass this class" are both simple examples of probabilistic thinking. *We are assigning an event some chance of occurring.* We will now formalize this idea mathematically.

Definition 8.1.1.
A *sample space* $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ is a set of events (outcomes, occurences, etc.). A *probability* $\mathbb{P} : \Omega \to [0, 1]$ is a function assigning elements of Ω a number such that the following hold

1. $\mathbb{P}[\omega_i] \geq 0$ for all i
2. $\sum\limits_{i=1}^{n} \mathbb{P}[\omega_i] = 1$

Intuitively, a sample space is simply a set and a probability is just a function with special restrictions. Even more intuitively, a sample space is a set of events and a probability is a likelihood those events occur.

Enough formalism, examples below.

Example 8.1.2.
Let $\Omega = \{\text{heads, tails}\}$ and $\mathbb{P}[\text{heads}] = \mathbb{P}[\text{tails}] = \frac{1}{2}$. Then, we have the famed "fair coin toss". If you have not flipped a coin yet, please do. Note that $\mathbb{P}[\omega_i] = \frac{1}{2} \geq 0$ for all i. Also, $\sum\limits_{i=1}^{n} \mathbb{P}[\omega_i] = \mathbb{P}[\text{heads}] + \mathbb{P}[\text{tails}] = \frac{1}{2} + \frac{1}{2} = 1$. So, the $\mathbb{P}[\cdot]$ we have chosen is indeed a *probability.* △

Example 8.1.3.
Let $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\mathbb{P}[i] = \frac{1}{6}$ for all $i \in \Omega$. Then, we have the "dice roll". Note that $\mathbb{P}[\omega_i] = \frac{1}{6} \geq 0$ for all i. Also, $\sum\limits_{i=1}^{n} \mathbb{P}[\omega_i] = \mathbb{P}[1] + \mathbb{P}[2] + \mathbb{P}[3] + \mathbb{P}[4] + \mathbb{P}[5] + \mathbb{P}[6] = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$. So, the $\mathbb{P}[\cdot]$ we have chosen is again a *probability.* △

A lot of the rules of combinatorics have probabilistic counterparts. For example, the addition rule says that if sets A and B are disjoint ($|A \cap B| = 0$), then $|A \cup B| = |A| + |B|$. Likewise, the following holds.

Theorem 8.1.4.

- *If A and B are* mutually exclusive *events (they cannot occur at the same time), then* $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$.

- *If $A_1, A_2, \dots, A_n$ are all mutually exclusive, then* $\mathbb{P}\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \mathbb{P}[A_i]$.

Likewise, there is a product rule for probability:

Theorem 8.1.5.

- *If A and B are* independent *events (A has no impact on B), then* $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$.
- *If $A_1, A_2, \dots, A_n$ are all independent, then* $\mathbb{P}\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n} \mathbb{P}[A_i]$.

Example 8.1.6.

Consider an urn with n orbs labeled $\{1, 2, \dots, n\}$. We will draw an orb without looking and then putting it back.

1. Since there are n orbs, the probability we draw any specific orb is $\frac{1}{n}$ ($\mathbb{P}[i \in \{1, 2, \dots, n\}] = \frac{1}{n}$).
2. We are only drawing one orb at a time. So, drawing orb j or orb i are mutually exclusive (we cannot draw both at once). For example, seeing 2 or 3 are cannot happen at the same time. So, $\mathbb{P}[2 \cup 3] = \mathbb{P}[2] + \mathbb{P}[3] = \frac{1}{n} + \frac{1}{n} = \frac{2}{n}$.
3. If we draw one orb, put it back, and draw again, then the event of seeing orb j first has no impact on seeing orb i second. So, the events are independent. For example, drawing 1 first has no impact on whether we draw 5 second. So, $\mathbb{P}[1 \cap 5] = \mathbb{P}[1]\mathbb{P}[5] = \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2}$

$\triangle$

This is the best case scenario. We do not have to care about overlap of events. But, you might wonder what happens if events are not independent or mutually exclusive? In general, we have the following:

Theorem 8.1.7.

For two events A and B, the probability of A and B happening ($\mathbb{P}[A \cap B]$) is the probability of B happening ($\mathbb{P}[B]$) times the probability of A happening given that B has already happened *($\mathbb{P}[A|B]$). That is,*

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

We call $\mathbb{P}[A|B]$ a conditional probability *and we say "probability of A given B".*

Hopefully this seems rather intuitive: "A *and* B happening means B must have happened. Now given that B has happened, consider A". Note that the choice of A and B here is arbitrary. That is, we can also say $\mathbb{P}[A \cap B] = \mathbb{P}[B|A]\mathbb{P}[A]$.

Example 8.1.8.

Consider again the orb and urn problem with n numbered orbs. We will draw without looking. This time, we will not put the drawn orb back in the urn. Consider drawing 2 and setting it to the side. Now, there are only $n-1$ orbs in the urn. This affects the probability of drawing any of the remaining orbs. Originally, $\mathbb{P}[i] = \frac{1}{n}$ for all i in the urn. Without 2, we have $\mathbb{P}[i] = \frac{1}{n-1}$ for all i in the urn. $\triangle$

Consider the second part of the definition of a probability: $\sum_{i=1}^{n} \mathbb{P}[\omega_i] = 1$. What this is saying is that the probability that *anything* happens is 1. That is, there is a 100% chance of something happening. Formally, we would say $\mathbb{P}[\Omega] = 1$ (the probability of our sample space is 1). This leads to the following thm.

Theorem 8.1.9.

- $\mathbb{P}[A] = 1 - \mathbb{P}[A^C]$ *for all* $A \in \Omega$
- $\mathbb{P}\left[\bigcup_{i=1}^{n} A_i\right] = 1 - \mathbb{P}\left[\bigcap_{i=1}^{n} A_i^C\right]$ *for* $A_i \in \Omega$

Proof.
$\mathbb{P}[A] = \mathbb{P}[\Omega \backslash A^C] = 1 - \mathbb{P}[A^C]$ and then DeMorgan's laws to obtain second point. □

We are now ready to define what is called a random variable. Much like a variable in the traditional sense, a random variable is a placeholder that takes on values from some set. A random variable differs from a traditional variable since we also require that it takes *random* values from Ω according to $\mathbb{P}$. Formally,

Definition 8.1.10.
A random variable $X : \Omega \to \mathbb{R}$ is a function that maps the events in Ω equipped with the probability $\mathbb{P}[\cdot]$.

Example 8.1.11.
The above definition may be scary but think about the following:

- Let $\Omega = \{\text{heads, tails}\}$ and $X : \Omega \to \Omega$ be the outcome of a coin toss.
- Let $\Omega = \{\text{heads, tails}\}$ and $X : \Omega \to \{0, 1\}$ be the outcome of a coin toss where heads will be assigned 0 and tails will be assigned 1.
- Let $\Omega = \{\text{tilings of the } 1 \times n \text{ board with squares and dominoes}\}$ and $X : \Omega \to \{1, 2, \ldots, n\}$ be the length of a randomly draw board.

△

Definition 8.1.12.
The *expectation* (expected value) of a random variable is defined as

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

Example 8.1.13.
Let X be the outcome of a fair coin toss where 0 is heads and 1 is tails. The expected value of X is the average: $\frac{1}{2}$.

$$\mathbb{E}[X] = 0 \cdot \mathbb{P}[X = 0] + 1 \cdot \mathbb{P}[X = 1] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

△

We can think of the expected value in the following way: "Given that we have assigned events in Ω the probabilities $\mathbb{P}$, what will be the average of events we see?" In the coin toss example, we should see heads half the time and tails half the time. Assigning these 0 and 1 respectively, we should see 0 and 1 half the time. So, the average of things we should see is $\frac{1}{2}$.

A couple useful properties of expectation:

Theorem 8.1.14.

- $\mathbb{E}[f(X)] = \sum_x f(x) \cdot \mathbb{P}[X = x]$ *for random variable X and function f*
- $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ *for scalars a, b and random variables X and Y*
- $\mathbb{E}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i \mathbb{E}[X_i]$ *for scalars a, b and random variables X_i*

Coming soon:

- Indicator variables
- Erdös-Renyi model
- Probabilistic method