Discrete Math (Math 55) @UC Berkeley

Adrian Fry

Fall 2024

Contents

This is my attempt at taking notes for my math class[1]: Math 55 @UC Berkeley as taught by Dr. Nikhil Srivastava. LaTeXtemplate thanks to Ian Kerio, Evan Chen. Textbook used is Discrete Mathematics and its applications, by Kenneth Rosen. Any mistakes are my own

---

§1 Logic

§1.1 8.29.2024

§1.1.1 intro

Discrete math

- integers Z
- graphs
- not continuous objects

This class will be built up from scratch (naive set theory)

§1.1.2 Propositional Logic

"Mathematics is the science that draws necessary conclusions." - B. Pierre

Definition 1.1 (Proposition)**.** A **proposition** is a declarative sentence which is True or False, but not both.

Example 1.2

Sentence: Pigs can fly
Proposition? Yes
Truth Value: False

Sentence: $x + 5 = 9$
Proposition? No

[1] we'll see how long this keeps up

> **Remark 1.3.** A proposition must be 100% precisely unambiguous / precisely specified (must agree on the definition)

There are 3 basic logical connectives

- Negation $\neg$

  Definition 1.4 (negation)**.** The **negation** of a proposition p is another proposition $\neg p$ that has the opposite truth value of the proposition p.

p	$\neg p$
T	F
F	T

- Conjunction $\wedge$

  Definition 1.5 (conjunction)**.** Given 2 propositions p and q, their **conjunction** is the proposition "p and q" denoted $p \wedge q$, where their conjunction is true only if both p and q are true.

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

- Disjunction $\vee$

  Definition 1.6 (disjunction)**.** Given 2 propositions, their **disjunction** is "p or q" denoted $p \vee q$. $p \vee q$ is true if at least one of p and q is true, otherwise it is false.

These 3 connectives appear throughout mathematics

Definition 1.7 (conditional)**.** If p and q are propositions, the **conditional** $p \to q$ is the statement "p implies q", or "if p then q"

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Definition 1.8 (converse implication)**.** $q \to p$ is the converse of $p \to q$

Definition 1.9 (biconditional)**.** Given p, q propoistions, the **biconditional** proposition "p IFF q" is denoted $p \leftrightarrow q$. So The truth value of the biconditional is T if both p and q have the same truth value.

Problem 1.10. Is $p \leftrightarrow q$ the same as $(p \to q) \wedge (q \to p)$?

Definition 1.11 (equivalence)**.** 2 compound propositions C_1, C_2 are **equivalent**, denoted $C_1 \equiv C_2$ if they have the same truth value for all truth values of the propositional variables in them.

By drawing out a truth table, we can mechanically determine the solution, which is that the 2 statements are equivalent.

> **Remark 1.12.** You might wonder if $C_1 \equiv C_2$ is a proposition?
> Notice that if we draw the truth table for this, we can see that it is equivalent to $C_1 \leftrightarrow C_2$, or $\equiv\equiv\leftrightarrow$

Definition 1.13 (Tautology)**.** A compound proposition that's always true (for all truth values of its propositional variables) is a called a **tautology**.

Example 1.14 (monopoly tautology)

Consider the statement: "If I rolled the most doubles, then I lost, or if I lost, then I had the least hotels."
This statement is hard to decipher, but we can break it down into a compound proposition, and then determine whether this proposition is a tautology.
Let $d, l,$ and h be the following propositions: d = "I rolled the most doubles."
l = "I lost."
h = "I had the least hotels."
The logical structure of the sentence is: $C = (d \to l) \vee (l \to h)$
The truth value of this proposition is always true.

---

§1.2 9.3.2024

§1.2.1 More on logical equivalence, Rules of Inference (Propositional logic)

> **Remark 1.15** (Review)**.** Last week we Reasoned about the truth values of propositions without knowing them.

Below are some examples of equivalence

Example 1.16 (1)

$p \vee \neg p \equiv T$ - Tautology
$p \wedge \neg p \equiv F$ - Contradiction

Example 1.17 (2)

$p \to q \equiv (\neg p) \vee q$
"If it rains, I get wet." is equivalent to saying: "It does not rain or I get wet"

Example 1.18 (3 - Contrapositive)

$p \to q \equiv \neg q \to \neg p$
This becomes a surprisingly powerful tool when writing proofs

Example 1.19 (4 - de Morgan's Laws)

$\neg(p \wedge q) \equiv \neg p \vee \neg q$
$\neg(p \vee q) \equiv \neg p \wedge \neg q$

Remark 1.20 (Disjunction Rules)**.** If all of the logical connectives in a statement are $\vee$ (or), then the statement is associative - any reordering of parentheses is equivalent. It is also commutative

Rules of inference:

- If $p \wedge (p \to q) \to q \equiv T$ (Modus Ponens)
- $\neg q \wedge (p \to q) \equiv negp$
- $(p \to q) \wedge (q \to r) \equiv p \to r$

§1.2.2 Predicate Logic

Remark 1.21 (Motivation)**.** We want to say statements like "Every integer is even or odd". That is, we want to say things about infinitely many contexts. Attempting to do this in propositional logic gives (1 is odd $\vee$ 1 is even) $\vee$ (2 is odd $\vee$ 2 is even), and so on. Which is infinite

Definition 1.22 (Predicate)**.** A **predicate** (aka propositional function), is a statement containing one or more variables from a domain, which becomes a proposition when each variable is instantiated.

Example 1.23

Letting P be the predicate, and x the variable, we can write
$P(x)$ is the statement "x is even."
The Domain must be specified so we can write:
$P(x)$ is the statement "x is even when $x \in \mathbb{N}$"
And thus we have $P(2)$ = "2 is even," $P(3)$ = "3 is even

Example 1.24

$Q(x, y) = x < y$, with $x, y \in \mathbb{R}$

Definition 1.25 (Quantifiers $\forall$ and $\exists$)**.** The **Universal Quantification** of $P(x)$ is the proposition, "For all x in the domain, P(x)." Denoted $\forall\, xP(x)$ such that $x \in$ domain.
The **Existential Qualification** of $P(x)$ is the proposition, "There exists some x in the domain such that $P(x)$." Denoted $\exists\, xP(x)$ such that $x \in$ domain.

Remark 1.26. The same statement can be written both using logical notation or in English.

Terminology: A variable appearing in a qualifier is **bound**. Otherwise it is **free**. A statement in which all variables are bound is a proposition.
The professor goes on a tangent about $\forall$ and $\exists$ and what they mean.

Remark 1.27. If the domain is **finite**, you can write $\forall\, xP(x) \equiv P(d_1) \wedge P(d_2) \wedge \ldots \wedge P(d_n)$
$\exists\, xP(x) \equiv P(d_1) \vee P(d_2) \vee \vee P(d_n)$

Remark 1.28. Can use $\rightarrow$, $\wedge$ to specify domain:
$\forall x\ (x \in \mathbb{Z} \rightarrow P(x) \vee R(x))$
$\exists x\ (x \in \mathbb{Z} \wedge x^2 = 2)$

Example 1.29 (Negation of Quantifiers)

"Every integer is prime," or $\forall x \in \mathbb{Z}, x \in \mathbb{P}$
"There exists some non-prime integer," or $\exists x \in \mathbb{Z}$ s.t. $x \notin \mathbb{P}$

Definition 1.30 (Logical Equivalence for statements in Predicate Logic)**.** 2 propositions with quantifiers are **logically equivalent**, denoted $C_1 \equiv C_2$ if $\forall$ predicates $\wedge$ $\forall$ domains, they have the same truth value.

Remark 1.31 (De Morgan for Qualifiers)**.** Notice $\neg\forall xP(x) \equiv \exists x \neg P(x)$, and vice versa

Example 1.32 (Nesting of Qualifiers)

Can write expressions like: $\forall x$, $\exists y$ s.t. $(x < y)$ where $x, y \in \mathbb{Z}$
But reversing the order of $\forall$ and $\exists$, we get a statement which is false.

§1.3 9.5.2024

§1.3.1 Rules of Inference for Quantifiers

Definition 1.33 (instantiation)**.** $\forall$ in the domain implies $P(x) \to$ "$P(x)$ is true for an arbitrary element a of the domain."
$\exists$ in the domain s.t. $P(x) \to$ "It is possible to choose an element in the domain s.t. $P(x)$ is true."

The opposite of instantiation is generalization

§1.3.2 Direct Proof

Definition 1.34 (Argument)**.** An argument is a sequence of propositions, each of which is a premise (assumption) or a conclusion (follows from the previous propositions via the Rules of Inference).

An argument is **valid** is the above property holds A **proof** is a valid argument used to establish the truth of a mathematical proposition.

Definition 1.35 (Direct Proof)**.** Used to prove the statement: $\forall\, x \in$ Domain, $P(x) \to Q(x)$

Definition 1.37 (Even)**.** An integer n is even if there exists an integer k s.t. $n = 2k$.

Definition 1.38 (Odd)**.** $\forall\, n \in \mathbb{N}$, n is Odd $\leftrightarrow$ it is not Even, or if there exists an integer k s.t. $n = 2k + 1$

Proposition 1.39 (A)

$\forall\, n \in$ Odd, n^2 is also Odd.

Proof: Let $n \in \mathbb{N}$ and n is odd. By definition, $\exists\, k \in \mathbb{N}$ s.t. $n = 2k + 1$ ($\exists$ instantiation). Observe that

$$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

Let $h = (2k^2 + 2k)$. Since $h \in \mathbb{N}$, we have $n^2 = 2h + 1$ and thus is odd($\exists$ generalization). Thus $\forall\, n$ odd, n^2 is odd. ($\forall$ generalization) □

§1.3.3 Contrapositive

Definition 1.40 (Contrapositive)**.** $\forall\, x,\, P(x) \to Q(x) \equiv \forall\, x, \neg Q(x) \to \neg P(x)$

Proposition 1.42 (B)

For every integer n, if n^2 is odd, then n is also odd

Proof (by contrapositive): Let $n \in \mathbb{N}$ and n is even. By definition, we can choose k s.t. $n = 2k$ ($\exists$ instantiation). Observe that

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

Thus, n^2 is even ($\exists$ generalization), as desired ($\forall$ n even, n^2 is even ($\forall$ generalization)).

§1.3.4 Contradiction

Definition 1.43 (Rationals)**.** $x \in \mathbb{R}$ has the property $x \in \mathbb{Q}$ if *exists* $p, q \in \mathbb{N}$ s.t. $x = \frac{p}{q}$

Definition 1.44 (irrational)**.** A real number x is irrational if it is not rational.

Proposition 1.45 (C)

If $x \in \mathbb{Q}$, we can choose p, q with no common factors s.t. $x = \frac{p}{q}$

Theorem 1.46 ($\sqrt{2}$ is irrational)

Proof: Assume for contradiction that $\sqrt{2}$ is rational. Then $\exists$ $p, q \in \mathbb{N}$ s.t. $\sqrt{2} = \frac{p}{q}$. Thus $\sqrt{2}^2 = 2 = \frac{p^2}{q^2}$. But if $\frac{p^2}{q^2} = 2$, then $p^2 = 2q^2$, which implies p^2 is even, and thus p is even (by **Proposition** B). By definition, we can choose come k s.t. $p = 2k$. Substituting, we have $(2k)^2 = 2q^2$, thus $q^2 = 2k^2$, so q^2 is even, which implies q is even.(by **Proposition** B) But p and q have no common factors, so we have a contradiction. □

---

§1.4 9.10.2024

§1.4.1 More on Proofs

Remark 1.47 (The bare minimum)**.** A good proof is

1. Clear - every word in the proof is well-defined - variables bound with qualifiers - possible to translate into formal logic
2. Correct - valid (each line follows from the previous line) - no mistakes

TIPS:

- Use *keywords* to indicate logical structure:

  - Let x be arbitrary.
  - We can choose y s.t.. . .
  - We conclude that. . .
  - By definition. . .
  - Assume $P(x)$. . .

- A proof has a beginning, middle, and end. Be aware of where you are
  - Beginning: Assume hypotheses
  - Middle: Rules of Inference, Observations
  - End: Conclusion

- Use complete sentences

§1.4.2 Sets

Definition 1.48 (Set)**.** A **Set** is an unordered "collection" of objects[2], which are called its **elements**. A set **contains** its elements. Denoted $x \in A$.

Key Property: For every x, "$x \in A$" is a proposition.
2 **ways to specify**

1. Roster notation:
   - $A = \{1, 2, 5\}$
   - $\mathbb{Z}$ = the set of all integers = $\{0, 1, 2, \ldots\}$
2. Set-Builder notation:
   - Given a predicate $P(x)$, $A = \{x : P(x)\}$ is a set.
   - Even = $\{x \in \mathbb{Z} : x \text{ is even}\}$

Definition 1.49 (The empty set)**.** The set with no elements is called the empty set, denoted $\emptyset$
Formally: $\forall x, (x \notin \emptyset)$

Definition 1.50 (Equality of sets)**.** 2 sets are equal if every element in 1 is in the other and vice versa.

Definition 1.51. Given A, B, A is a subset of B denoted $A \subseteq B$ if $\forall x(x \in A \to x \in B)$

Operations on Sets: $\cup, \cap, -, \mathcal{P}, \times$

Definition 1.52 (Set operations)**.** $A \cup B = \{x : x \in A \vee x \in B\}$ - Union
$A \cap B = \{x : x \in A \wedge x \in B\}$ - Intersection
$A - B = \{x : x \; in A \wedge x \notin B\}$ - Difference
$\overline{A} = \{x : x \notin A\}$ - Complement

[2] We are using naive Set theory and assuming every object is concrete

Definition 1.53 (Power set)**.** The **Power Set** of A, denoted $\mathcal{P}(A)$ is the set of all subsets of A

$$\mathcal{P}(A) = S : S \subseteq A$$

> **Remark 1.54.** If A has n elements, $\mathcal{P}(A)$ has 2^n elements.

Definition 1.55 (Cartesian Product)**.** Given A, B, their **Cartesian Product** is the set of ordered pairs denoted $A \times B = \{(a, b) : a \in A \wedge b \in B\}$

Example 1.56 (are two)

$\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$

---

§2 Sets

§2.1 9.12.2024

§2.1.1 Functions

Example 2.1

$f(x) = x^2$ gives a parabola

Definition 2.2 (Function)**.** If A, B are nonempty sets, a **function** from A to B denoted $f : A \to B$is a rule which assigns exactly one element of B to every element of A. This assignment is denoted $f(a) = b$, where $a \in A$ and $b \in B$. A is the domain and B is the codomain.

> **Remark 2.3.** Every element of A must go to a unique element of B

Example 2.4

$f : \mathbb{R} \to \mathbb{R}$ s.t. $f(x) = x^2$
and a different function
$f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ s.t. $g(x) = x^2$

Example 2.5

A is a set, $p : A \to \mathcal{P}(A)$ where $p(x) = \{x\}$

Definition 2.6 (Surjections)**.** $f : A \to B$ is **onto** or **surjective** if for every $b \in B$, there exists an

$a \in A$ s.t. $f(a) = b$. That is:

$$\forall b \in B, \exists a \in A \text{s.t.} f(a) = b$$

Example 2.7

Question: $A \subset B$. Can $f : A \to B$ be onto?

Answer: When A, B are finite, no, but

Definition 2.8 (Injections)**.** $f : A \to B$ is **one-to-one** or a **injective** if for every $a_1, a_2 \in A$ $(f(a_1) = f(a_2) \to a_1 = a_2)$

Remark 2.9 (Bijections)**.** Notice that our use of one-to-one is sometimes different than other uses, when it often referrs to a bijection, which is injective both ways, or injective and surjective. Notice that this implies bijections are invertible.

§2.1.2 Cardinality

Definition 2.10 (Cardinality)**.** The **cardinality** of A, denoted $|A|$, is the number of elements in the set

Definition 2.11 (finite and infinite sets)**.** A set A is **finite** if it has exactly n elements for some non-negative integer n. n is the cardinality of A

Proposition 2.12

If A, B are finite, and $f : A \to B$ is onto, then $|A| \geq |B|$

Example 2.13

If $|A| < |B|$, then it cannot be onto. Proof by pigeonhole principle.

---

§2.2 9.17.2024

§2.2.1 Cardinality2

Definition 2.14 (Cardinality of not necessarily finite sets)**.** If A, B are sets, they have the same cardinality if there is a bijection $f : A \to B$.

If there is an injective function $f : A \to B$, then $|A| \leq |B|$.

If $|A| \leq |B|$ but $|A| \neq |B|$ then we write $|A| < |B|$

Definition 2.15 (Countability)**.** A set A is countable if there is a bijection to $\mathbb{Z}_+$, or $|A| = |\mathbb{Z}_+|$. A function $f : \mathbb{Z}_+ \to A$ is called a **sequence**, written $a_1, a_2, a_3, \ldots$, where $a_n = f(n)$.
A sequence in which every element of A occurs exactly once is called an **enumeration**.

Example 2.16 (Cardinality of $\mathcal{P}(\mathbb{Z})$)

Is the power set $\mathcal{P}(\mathbb{Z})$ countable? No, since we proved before that the cardinality of a power set is strictly greater than the cardinality of it's original set.

Definition 2.17 ((0,1))**.** $(0, 1) = \{0.d_1d_2d_3d_4 \cdots : d_i \in \{0, \ldots, 9\}$ and d_j are not all $0\}$

Theorem 2.18 ((0, 1) is uncountable)

Proof: Assume $f : \mathbb{Z}_+ \to (0, 1)$. We will show f is not onto (codomain $\neq (0, 1)$). Consider the array:

$$\begin{aligned} f(1) &= 0.d_{11}d_{12}\ldots \\ f(2) &= 0.d_{21}d_{22}\ldots \\ f(3) &= 9.d_{31}d_{32}\ldots \end{aligned}$$

where d_ij is the jth digit of $f(i)$. Look at the following example:

$$\begin{aligned} f(1) &= 0.4362473 \\ f(2) &= 0.8765834 \\ f(3) &= 0.3498738 \\ x &= 0.356\ldots \end{aligned}$$

We will construct an $x \in (0, 1)$ such that $f(n) \neq x$ for all $n \in \mathbb{Z}_+$. Let $x = 0.x_1x_2x_3\ldots$ defined by $x_n \neq d_{nn}$
Observe that for every n, $f(n) \neq x$ bas they differ in the nth digit, by construction. Thus f is not onto. □

Remark 2.19. $|\mathbb{Z}_+| < |(0, 1)| = |\mathcal{P}|$

§2.2.2 "Prove or Disprove 'S'"

1. Write down S and $\neg S$
2. Write down all definitions.
3. Form a belief about S or $\neg S$
4. Use the steps above to try to write a proof
   a) Success - backslash q e d
   b) Failure - articulate what you have learned and go back to 4

§3 Number Theory

§3.1 9.19.2024

§3.1.1 Division and Divisibility

Definition 3.1 (Divisibility)**.** If $a \neq 0$ and b are integers, then $a|b$ (a divides b) if $\exists\ k \in \mathbb{Z}$ s.t. $b = ka$

Example 3.2

$3|9$ because let $k = 3$.
$3|0$ because let $k = 0$
$n|0$ by the same logic.

Proposition 3.3 (Properties of Divisibility)

Below are some properties of divisibiltiy

1. Transitivity
2. If $a|b$ and $a|c$ then $a|b + c$

Theorem 3.4 (Division Algorithm)

If $a \in \mathbb{Z}$ and $d \in \mathbb{Z}_+$, then there exist unique integers q, r s.t. $a = dq + r$ and $0 \leq r < d$

[Well-Ordering Principle] Every nonempty subset of $\mathbb{Z}_{\geq 0}$ contains a least element.
$\forall S \subseteq \mathbb{Z}_{\geq 0}$, $(S \neq \emptyset \rightarrow \exists y \in S$ s.t $(\forall x \in S, x > y))$

§3.1.2 Modular Arithmetic

Example 3.5

Time is $n \mod 12$,
Days are $m \mod 7$

Definition 3.6. If $a, b \in \mathbb{Z}$ and $m \in \mathbb{Z}_+$, then a is congruent to $b \mod m$, denoted $a \equiv b \mod m$ if $a \mod m = b \mod m$ or equivalently, $m|b - a$

Proposition 3.7

Notice that these properties are similar to the divisibility properties

1. If $a \equiv b \mod m$ and $b \equiv c \mod m$, then $a \equiv c \mod m$
2. If $a \equiv b \mod m$ and $c \in \mathbb{Z}$ then $ca \equiv cb \mod m$
3. If $a \equiv b \mod m$ and $c \equiv d \mod m$ then
   a) $a + c \equiv b + d \mod m$
   b) $ac \equiv bd \mod m$

§3.1.3 Representations of Numbers

In decimal, $d_k, d_{k-1}, \ldots, d_1 d_0$ means $d_0 10^0 + d_1 10^1 + \cdots + d_k 10^k$

Fact 3.8 (base b representation)**.** If $b \in \mathbb{Z}_+$, every $a \in \mathbb{Z}_{\geq 0}$ can be written uniquelly as a sum of powers of b, ex. there are unique $a_0, a_1, \ldots, a_k$ s.t. $a = a_1 b^0 + a_1 b^1 + \cdots + a_k b^k$. This is called the base b representation of a, denoted $a = (a_k a_{k-1} \ldots a_0)_b$

---

§3.2 9.24.2024

§3.2.1 Primes

Definition 3.9. An integer $n > 1$ is **prime** is it has no divisors other than 1 and itself.

$$n \text{ is prime} \leftrightarrow \forall a (a|n \rightarrow a = 1 \text{ or } a = n$$

A number that is not prime is composite. i.e., there exists $a, b > 1$ such that $n = ab$

We are now going to prove that there are ∞ primes.

Lemma 3.10

If $c > 1$ is composite, there exists a prime $p < c$ such that $p|c$.

Proof. Assume $c > 1$ is composite. Let $S = \{a \in \mathbb{N} : a|c \wedge a > 1\}$ (the set of all nontrivial divisors). We know this is non-empty because c is composite. So we can choose a least element p (by the well ordering principle). Assume for contradiction that p is composite. By definition, there exists $b > 1$ such that $b|p$. But $b|p$ and $p|c$ implies $b|c$, thus b is in S. But $b > p$ because we chose a smallest element of S. Then we have a contradiction. □

Theorem 3.11

There are infinitely many primes

Proof. Assume for contradiction there are finitely many primes $p_1, p_2, \ldots p_n$. Consider $q = (\prod_{i \in I} p_i) + 1$. Since q is not prime (it is bigger than all the primes), it must be composite. By lemma 3.2.10, we can choose some prime p_i such that $p_i | q$. But p_i also divides $q - 1$. This is impossible, thus we have a contradiction. □

Theorem 3.12 (The Fundamental Theorem of Arithmetic)

Every integer $n > 1$ can be written uniquely as a product of primes arranged in nondecreasing order. (By making it a monotonic sequence (ordered) we have made it unique)

§3.2.2 GCD

Definition 3.13 (Greatest Common Divisor)**.** If $0 < a, b \in \mathbb{Z}$, the greatest common divisor of a and b, denoted $\gcd(a, b)$ is the largest $d \in \mathbb{N}$ such that $d|a$ and $d|b$.

Fact: If $a > b > 1$ and $a = bq + r$, $0 \leq r < b$, then $\gcd(a, b) = \gcd(b, r)$

§3.2.3 Euclidean Algorithm

Euclidean Algorithm: Repeatedly apply the key observation until $r = 0$ - This always occurs as $r < b$ always.

Lemma 3.14

If $a > b > 1$, and $a = bq + r$, $0 \leq r < b$, then for all $d \in \mathbb{Z}_+$, ($d|a$ and $d|b \rightarrow d|b \wedge d|r$).

Proof. $\rightarrow$ *direction*
*Assume*a¿b¿1,a = bq + r,0≤ $r < b$. Assume $d \in \mathbb{N}$ with $d|a$ and $d|b$. Observe that $d|bq$, so $d|a - bq$. Thus $d|r$.
The left direction is similar. □

Theorem 3.15 (Bézouts Lemma)

If $a, b \in \mathbb{Z}$, not both 0, then there exists some $s, t \in mathbbZ$ (can be negative) such that $\gcd(a, b) = sa + tb$ (professor accidentally called this Bézouts Theorem)

Proof. Follows by substituting remainders in the euclidean algorithm □

We will try another way.

Proof. Consider $S = \{sa+tb : s,t \in \mathbb{Z} \wedge sa = tb \geq 1\} \subseteq \mathbb{N}$. Observe that $S \neq \emptyset$ since $a^2+b^2 \in S$. Then we can choose a least element g (by the well ordering principle), such that $g = sa + tb$. We will show that $g = \gcd(a, b)$

Proposition 3.16

$g|a$ and $g|b$ Assume towards contradiction the $g \nmid a$. Choose r such that $0 < r < g$ and $a = gq + r$. Observe $r = a - gq = a - (sa + tb)q = (1 - q)a + (-tq)b \in S$ (since $r > 0$). But this contradicts the minimality of g □

Proposition 3.17 (If for $d \in \mathbb{N}$, $d|a$ and $d|b$, then $d|g$)

This proof is left as an excercise for the reader

□

Corollary 3.18 (Strong Bézout)

If $a, b \in \mathbb{Z}$ not both 0, then $\gcd(a, b) = \min\{sa + tb : s,t \in \mathbb{Z}, sa + tb \geq 1\}$

Remark 3.19 (For the exam)**.**

1. Proofs aren't unique.
2. Complete sentences, clearly specifying final proof (can use existential qualifiers)
3. Use definitions (know them)

---

§4 Induction

§4.1 10.8.2024

§4.1.1 Induction

Induction is a new **Inference Rule** for proving statements of the form: $\forall n \in \mathbb{N} P(n)$

Example 4.1

$\forall n \in \mathbb{N}(n! < 2^n)$.

Definition 4.2 (Principle of Mathematical Induction)**.** Let $P(n)$ be a predicate. Induction tells us that from

1. A basis step $P(1)$
2. An inductive step $\forall k \in \mathbb{N}(P(k) \to P(k + 1))$

We get: $\forall n \in \mathbb{N} P(n)$

Proposition 4.3

$\forall n \in \mathbb{N}(\sum_{i=1}^{n} i = \frac{n(n+1)}{2})$

Proof. We proceed by induction.
Basis step: For $n = 1$ we have $1 = \frac{1*2}{2} = 1$.
Inductive step: Suppose $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Then $\sum_{i=1}^{n+1} i = \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1)+2n+2}{2} = \frac{(n+2)(n+1)}{2}$.
Then for all n, the identity holds. □

Remark 4.4. The validity of induction follows from the well-ordering principle.

Let $S_k = \sum_{i=1}^{k}(2i+1)$ (the sum of odd numbers to k

Proposition 4.5

$\forall n \in \mathbb{N}\ (\exists m \in \mathbb{N} | S_n = n^2)$

Proof. We proceed by induction:
Basis step: For $n = 1$, there exists $m = 1$ s.t. $S_1 = 1 = 1^2$
Inductive Step: Assume $k \in \mathbb{N}$. Assume $P(k)$. We will show $P(k+1)$. By $P(k)$, we know that $S_k = k^2$. Observe that $S_{k+1} = \sum_{i=1}^{k+1}(2i+1) = \sum_{i=1}^{k+1}(2i+1) + 2(k+1) - 1 = \sum_{i=1}^{k+1} = k^2 + 2k + 1 = (k+1)^2$, establishing $P(k+1)$, as desired □

Remark 4.6. Sometimes, it is easier to prove a stronger statement by induction ex. its more difficult to prove for some arbitrary m rather that $m = k$

Strong induction next time due to technical difficulties.

---

§4.2 10.10.2024

§4.2.1 Tilings

Proposition 4.7

For every $n \in \mathbb{N}$, for every $2^n \times 2^n$ chessboard with an arbitrary square removed, the board can be perfectly tiled with triominos.

Proof. We proceed by induction:

- Basis step: If $n = 1$, the board must be one of 4 cases, each of which can be tiled by a single triomino.
- Induction Step: Suppose we can tile a $2^k \times 2^k$ board with one piece removed. Then we can tile 3 other boards with one extra tile, which we can have in any place we want. Then we can organize four of the $2^k \times 2^k$ boards together so that there is a missing triomino in the center, which we can fill in with another triomino.

□

§4.2.2 Strong Induction + Recursive Definition

Definition 4.8 (Strong Induction)**.** Suppose you know $P(1)$. Then $\forall k \geq 1$, $P(1) \wedge P(2) \wedge \cdots \wedge P(k) \to P(k+1)$.

We are looking for a way to define $f : \mathbb{Z}_+ \to \mathbb{Z}$ or $f : \mathbb{Z}_{\geq 0} \to \mathbb{Z}$
So far we have three ways:

1. Explicit formula
2. $a_1, a_2, a_3, \ldots$ and editing

New Way: to define $f : \mathbb{Z}_+ \to A$:

- Basis Step: For some $b \geq 1$, define $f(1), f(2), \ldots, f(b)$.
- Recursive Step: For every $k \geq b$, define $f(k+1)$ in terms of $f(1), f(2), \ldots, f(k)$.

Example 4.9

Fibonacci

Example 4.10

$f : \mathbb{Z}_+ \to \mathbb{R}$

- Base: $f(1) = 1$
- Recursive $f(k+1) = \sqrt{1 + f(k)}$

Definition 4.11 (Golden Ratio)**.** $\frac{1+\sqrt{5}}{2}$

Theorem 4.12

For $n \geq 2$, $f(n) \geq \phi^{n-1}$

Proof. We proceed by **STRONG** induction

- Base Case(s): For $n = 2$, $f(2) = 2 \geq \phi^1$. For $n = 3$, $f(3) = 3 \geq \phi^2$
- Inductive Step: Assume $k \geq 3$. Assume $P(2), P(3), \ldots, P(k)$. We will show $P(k+1)$. Observe that $f(k+1) = f(k) + f(k-1) \geq \phi^{k-1} + \phi^{k-2} = \phi^k$

□

---

§5 Graph Theory

§5.1 10.15.2024

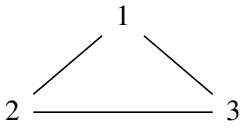
§5.1.1 Graphs

> **Remark 5.1.** A **graph** is a mathematical abstraction of **pairwise** relationships.

Definition 5.2. A **simple graph** $G = (V, E)$ is a pair of sets:
V, the set of vertices
E, the set of edges
Such that E is a set of unordered pairs of elements of V i.s., $E \subseteq \{S \subseteq V : |S| = 2\}$

Example 5.3

$V = \{1, 2, 3\}$, $E = \{\{1, 2\}, \{2, 3\}, \{1, 3\}\}$



Definition 5.4. A **multigraph** $G = (V, E)$ is a pair of sets:
V, the set of vertices
E, the multiset of edges
Such that E is a set of unordered pairs of elements of V i.s., $E \subseteq \{S \subseteq V : |S| = 2\}$

> **Remark 5.5.**
> - We will only be working with undirected graphs
> - V will be finite
> - If E has repetitions, G is a multigraph, otherwise it is a simple graph

Motivation:

1. Internet: $V = \{\text{servers}\}$, $E = \{xy : x \text{ is connected to } y \text{ by a wire}\}$
2. Social Network: $V\{\text{people}\}$, $E = \{\text{a relationship between two people}\}$

§5.1.2 Degree and Handshaking

Terminology: Given $G = (V, E)$, if $e = xy \in E$, then x is **adjacent** to y in G. e is **incident** with x or y. x, y are **endpoints** of e.

Definition 5.6. The **degree** of a vertex is the number of edges incident with it. $\deg(v) = |\{e \in E : e \text{ is incident with } v\}|$

Theorem 5.7 (Handshaking)

If $G = (V, E)$ is a graph,

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

Proof. Assume $G = (V, E)$ is a graph. Consider $I = \{(e, v) : e \in E, v \in V, e \text{ is incident with } v\}$ We will count the number of incidences $|I|$ in two ways

1. Each edge $e \in E$ participates in exactly two indices. No incidence participates in more than one edge. Thus, $|I| = 2|E|$
2. Each vertex $v \in V$ participates in exactly $\deg(v)$ incidences. And no incidence participates in more than one vertex. Thus $|I| = \sum_{v \in V} \deg(v) = 2|E|$.

□

Corollary 5.8

In any graph, there are an even number of odd degree vertices.

Proof. Suppose G has n_1 even degree vertices, and n_2 odd degree vertices. Observe

$$\sum_{v \in V} \deg(V) = \sum_{v \in V: \text{ v is odd}} \deg(v) + \sum_{v \in V: \text{ v is even}} \deg(v) \equiv 0 + n_2 (\mod 2) \equiv 2|E| (\mod 2) \equiv 0 (\mod 2)$$

□

§5.1.3 Ramsey Theory

Theorem 5.9

In any group of 6 people, there are 3 acquaintances or 3 strangers.

Proof. Consider the graph $G = (V, E)$
$V = \{1, 2, 3, 4, 5, 6\}$, $E = \{xy : x \text{ and } y \text{ are acquaintances.}$
Choose an arbitrary vertex x.
Case 1: $\deg(x) \leq 2$: Choose $a, b, c \in V$ such that $xa, xb, xc \notin E$. Observe that either ab, bc, ca

are all edges, in which case we are done, **or** without loss of generality, $bc \notin E$, which implies $xa, xb, bc \notin E$, so $\{x, b, c\}$ are strangers, as desired
Case 2: Homework □

Theorem 5.10

For all groups of 18 there exists 4 who are acquaintances or strangers

Theorem 5.11

For all $k \geq 2$, there exists N such that for all groups of $\geq N$, people there exists subgroups of k strangers or acquaintances

§5.1.4 Connected Components

Definition 5.12. A **path** of length k in a graph $G = (V, E)$ is an alternating sequence of vertices $x_i \in V$ and edges $e_i \in E : x_0, e_1, x_2, e_2, \cdots, e_k, x_k$ such that for all $i = 1, \cdots, k$, $e_i = \{x_{i-1}, x_i\}$. A path in which no vertex is repealed is called **simple**.

Definition 5.13. A graph $G = (V, E)$ is **connected** if for all $x, y \in V$, there is a path in G between x and y.

> **Remark 5.14.** If there is a path from x to y, then there is a path from y to x, by reversing the sequence. If there is a path from x to y, and from y to z, then there is a path from x to z.

Example 5.15 (Cycle)

$V = \{0, \ldots, n-1\}$
$E = \{xy : x - 1 \equiv 1(\mod n)\}$

Example 5.16 (Complete)

$V = \{1, \ldots, n\}$
$E = \{S : S \subseteq V, |S| = 2\}$

Example 5.17 (Independent Set)

$V = \{1, 2, \ldots n\}$
$E = \emptyset$

Definition 5.18 (Subgraph)**.** A graph $H = (W, F)$ is a **subgraph** of $G = (V, E)$ if $W \subseteq V$ and $F \subseteq E$.

Example 5.19

Independent $\subseteq$ Cycle $\subseteq$ Complete (on $\{1, \ldots, n\}$)

Definition 5.20. A **connected component** of $G = (V, E)$ is a subgraph $H \subseteq G$ which is

1. Connected
2. Maximal, i.e., for every subgraph $H' \subseteq G$ such that $H \subseteq H'$, H' is not connected

Theorem 5.21

Every graph is a disjoint union of its connected components. (Proved in HW using strong induction)

---

§5.2 10.17.2024

§5.2.1 Connected Components

Definition 5.22. A graph $H = (F, W)$ is a connected component of $G = (V, E)$ if it is a maximal connected subgraph of G. i.e., $H \subseteq G$, H is connected, $\forall\ H' \subseteq G$, $(H' \not\equiv H \rightarrow H$ is not connected)

Theorem 5.23

Every graph is a disjoint union of its connected components.

Lemma 5.24

If $G = (V, E)$ is a nonempty graph, then it has at least 1 connected component.

Proof. Choose a vertex $x \in V$. Define the set $W = \{y \in V :$ there is a path from x to y in $G \cup \{x\}\}$. Consider the subgraph $H = (W, F)$ where $F = \{xy \in E : x, y \in W\}$ (This is called an **induced subgraph**).
Let z, y be distinct vertices in W. Then there exists a path $xz \in W$ and a path $yx \in W$. By transitivity, we can concatenate these paths, thus there exists a path $yz : y \rightarrow z \in G$. Observe that the concatenation is a path in H, so H is connected.
To see that H is maximal, assume $H \subset H'$. This implies that there is some vertex $y \in H'$ such that $y \notin W$. By construction of W, there is no path from y' to x in G, so there is no path from y' to x in H', so H' is connected. □

5.23. Use strong induction. □

§5.2.2 k-coloring

Definition 5.25. A **k-coloring** of $G = (V, E)$ is a function $F : V \to \{1, 2, \ldots, k\}$ such that for all $xy \in E$, $(f(x)f(y))$

Example 5.26

The complete graph with 3 vertices is 3-colorable

The smallest k such that G is k-colorable is called its **chromatic number** denoted $\chi(G)$.

Theorem 5.27

IF G is a graph, with maximum degree D then $\chi(G) \leq D + 1$

Remark 5.28. For all $i \leq n$, $S_i = \{x \in V : f(x) = i\}$ is an independent set.

Proof. By induction. Assume $D \geq 0$. Let $P(n)$ be the statement: "For all graphs G with $\leq n$ vertices, if $\max \deg(G) \leq D$, then $\chi(G) \leq D + 1$."

- Base Case: Every graph with 1 vertex is 1-colorable, so it must also be 2-colorable.
- Inductive Step: Assume $k \in \mathbb{N}$. Assume $P(k)$. We will show $P(k + 1)$. Assume G has $k + 1$ vertices and maximum degree D. Choose vertex $x \in V$.
  Let $y_1, \ldots, y_m$ be the vertices adjacent to $x \in G$.
  Let $H = (W, F)$ be the subgraph obtained by deleting $x : W = V/\{x\}$, $F = \{yz \in E : y, z \in W\}$. Since H has maximum degree $\leq D$ and k vertices, $P(k)$ implies that there is a $D + 1$-coloring $f' : W \to \{1, \ldots, D + 1\}$ of H.
  Observe that $\{1, \ldots, D + 1\} - \{f(y_1), \ldots, f(y_k)\} \neq \emptyset$ since $m \leq D$. Let j be any color in the set.
  Define $f : V \to \{1, \ldots, D + 1\}$ be $\begin{cases} f(y) = f'(y) \text{ for } y \neq x \\ f(x) = j \end{cases}$ f is a valid coloring of H and also of the edges $xy_1, xy_2, \ldots, xy_m \in G$ by construction, so f is a valid coloring of G and G is $D + 1$ colorable, as desired.

□

§5.2.3 2-colorings

Definition 5.29. A path $x_0, e_1, x_1, \ldots, e_k, x_k$ is called a **circuit** if $x_0 = x_k$

Remark 5.30. For coloring, graphs are simple.

Theorem 5.31

A graph G is 2-colorable IFF it does not contain a circuit of odd length.

Proof. ($\rightarrow$) Assume G is 2-colorable. Choose a 2-coloring $f : V \rightarrow \{1, 2\}$. Assume G contains a circuit $C = x_0, x_1, \ldots, x_k$ with $x_k = x_0$. Consider the sequence $f(x_0), f(x_1), \ldots, f(x_k)$. Observe that $f(x_i) \neq f(x_{i+1}$ for all $i \leq k - 1$, so $f(x_i) = f(x_0)$ for all even i and $f(x_i) \neq f(x_0)$ for all odd i, i.e., the colors alternate as $i = 0, \ldots, k$. In particular, $f(x_k) \neq f(x_0)$, a contradiction since $x_k = x_0$. □ □

---

§5.3 10.24.2024

§5.3.1 Leonhard Euler

[3]

Example 5.32

THE $\mathbf{S}^{\mathbf{ev}}_{\mathbf{en}}\mathbf{Br}^{\mathbf{id}}_{\mathbf{ges}}$ OF **Könisburg**

Definition 5.33. An **Eulerian circuit** in a graph $G = (V, E)$ is a circuit which traverses every edge exactly once.

Theorem 5.34

If a graph $G = (V, E)$ with $|V| \geq 2$ has an eulerian circuit, then the degree of every vertex v has to be even.

Proof. Assume $G = (V, E)$ is connected and $|V| \geq 2$. Let $C = x_0, e_1, \ldots, e_m, x_0$ be an eulerian circuit of G. Assume $v \neq x_0$ is a vertex of G. Suppose C visits V exactly k times. Observe that in each visit, C traverses exactly 2 edges incident with V. Since C is eulerian, these pairs of edges must be disjoint, and every edge incident with v appears in a pair. Thus the edges incident with v can be partitioned into k pairs, so $\deg(v) = 2k$, as desired. By handshaking, $\sum \deg(v) = 2|E|$, so x_0 is even, completing the proof. □

Theorem 5.35

If a connected graph $G = (V, E)$ has $\deg(v)$ even for all $v \in V$ and $|V| \geq 2$ then G has an eulerian circuit.

[3] len-hard yule-her

Lemma 5.36

If $G = (V, E)$ is a graph with $\deg(V) \geq 2$, for all $v \in V$, then G contains a cycle.

Proof. We will prove by contrapositive. Assume $G = (V, E)$ is acyclic. We will show G has a vertex v with $\deg(v) \geq 2$. Let $G_1, \ldots, G_k$ be the connected components of G. Observe that each G is connected and acyclic since $G_i \subseteq G$ and G is acyclic, so G_i is a tree for $i = 1, \ldots, k$. If some G_i has at least 2 vertices, it must have a leaf l, with $\deg(l) = 1$. But then l has degree 1 in G as well, since G_i is a connected component. If all G_i have exactly one vertex, then we have an independent set and thus $\deg(V) = 0$ □

Proof. of Theorem 5.35: Proceed by induction on the number of edges m.

- Our induction hypothesis is: Let $P(n)$ = "If G has m edges, ≥ 2 vertices, is connected, and has all even degrees, then G has an eulerian circuit." We will show $P(m)$ for all $m \geq 2$.
- Base Case: The only connected (multi) graphs with 2 edges and all even degrees (and no self loops) has an eulerian circuit.
- Inductive Step: Assume $m \geq 2$, $P(2), P(3), \ldots, P(m)$. Assume $G = (V, E)$ is a connected graph with $m + 1$ edges, all even degrees, at least 2 vertices.
  By the Lemma, G contians a cyc;e $C = (V_c, E_c)$. If $C = G$, then traversing C yields an eulerian circuit of G.
  If $C \neq G$, let $H = G - C = (V, E - E_c)$ be the graph obtained by removing the edges of C from G.
  Let $H_1, \ldots, H_k$ be the connected components of H with at least 2 vertices. Observes that for all $i = 1, \ldots, k$,
  1. H_i is connected and has ≥ 2 vertices.
  2. Each H_i has at most $|E| - |E_c| \leq m + 1 - 2 \leq m - 1$
  3. The degree of every vertex $v \in H_i$ is $\deg_{H_i}(v) = \deg_G(v) - \deg_C(v)$.

  Since $\deg_G(v)$ is even and $\deg_C(v) \in \{0, 2\}$, $\deg_{H_i}(v)$ is even.
  Claim: For every $i = 1, \ldots, k$, there exists a vertex $s_i \in H_i$ such that C visits s_i
  HW
  By induction, each H_i has an eulerian circuit C_i which WLOG starts and ends at s_i. In a connected graph, if there is an eulerian circuit, then for all $v \in V$ there is one starting at v.
  Let C' be equal to C with the first occurence of s_i in C replaced by C_i, for $i = 1, \ldots, k$
  Observe that
  1. C' is a circuit
  2. C' traverses every edge of $C \cup H_1 \cup H_2 \cup \cdots \cup H_k = G$ exactly, once, as desired.

□

Remark 5.37.
- Naming objects is important
- Proofs are not unique
- Many proofs give algorithms.

---

§6 Counting

§6.1 10.29.2024

§6.1.1 Prototypical Examples

Goal: Count the number of objects satisfying a given property. i.e., find the cardinality of $S = \{x : P(x)\}$.

Example 6.1

Bit strings of length 10, $S = \{(b_1, \ldots, b_{10}) : b_i \in \{0, 1\}\}$

Example 6.2

Bit strings of length 5 ending in 00 or beginning with 1. $S = \{(b_1, \ldots, b_5) : b_i \in \{0, 1\}, b_4 = b_5 = 0 \vee b_1 = 1\}$

Example 6.3

Number of rankings of {Cal, Stanford, UCLA, USC}.

Example 6.4

Ordered sequences of 5 distinct cards from a deck of 52. $\{(c_1, c_2, \ldots, c_5) : c_i \in \{\text{cards}\}, c_i\text{distinct}\}$

Example 6.5

Unordered sets of 5 distinct cards from a deck of 52 Ordered sequences of 5 distinct cards from a deck of 52. $\{\{c_1, c_2, \ldots, c_5\} : c_i \in \{\text{cards}\}, c_i\text{distinct}\}$

Example 6.6

Simple graphs with vertex set $V = \{1, 2, \ldots, n\}$

§6.1.2 Principles of Counting

1. If $f : A \to B$ is a bijection, then $|A| = |B|$.
2. $|A \times B| = |A||B| \xrightarrow[\text{induction}]{} |A_1 \times \cdots \times A_k = |A_1| \times |A_2| \times \cdots \times |A_k|$
3. $|A \cup B| = |A| + |B| - |A \cap B|$ This is called the principle of inclusion exclusion **PIE**
4. Suppose an object from a set S is uniquely specified by a sequence of k choices $C_1, C_2, \ldots, C_k$ and
   a) The number of ways to make $C_1 = n_1$
   b) Given C_1, number of ways to make $C_2 = n_2$
   c) Given $C_1, C_2, \ldots, C_{k-1}$ Number $C_k = n_k$

   Then $|S| = n_1 n_2 \ldots n_k$ = the number of ways to make $(C_1, C_2, \ldots, C_k)$. (**Product Rule**)

Example 6.1:
Observe: $S = \{0,1\} \times \cdots \times \{0,1\}10$ times $= \{0,1\}^{10}$
$2^{10} = 1024$

Example 6.2:
Observe: $S = \{A \cup B\}$, $A = \{\underline{b} : b_4 = b_5 = 0\}$, $B = \{\underline{b} : b_1 = 1\}$ PIE: $|S| = |A| + |B| - |A \cap B|$
$f : A \to \{0,1\}^3$ is a bijection, so $f(\underline{b}) = (b_1, b_2, b_3)$.
$|A| = |\{0,1\}|^3 = 2^3 = 8$.
Similarly, we can show that $|B| = 2^4 = 16$.
Then, we see that $|A \cap B| = 2^2 = 4$. Thus we have

$$|S| = 8 + 16 - 4 = 20$$

Example 6.3:
Product Rule:
Observe: We want to decompose our "big" choice into a sequence of small choices. Process:

- Choose school number one
- Choose school number two given school number one
- Choose school number three given schools two and one
- Given the first 3, choose our fourth school.

Our first choice has 4 options, our second choice has only 3 options our third one 2, and our last one we only have 1 school. Then we have:

$$4 * 3 * 2 * 1 = 4!$$

We can consider a decision tree of our options

> **Remark 6.7.** Each $s \in S$ is uniquely specified by a sequence of choices i.e., there is a bijection from the set of orderings to the ways of ordering it

Definition 6.8. An ordering of n distinct objects is called a **permutation**.

The number of permutations of n distinct objects is $n!$, proven by the product rule.
Example 6.4:
Observe: By the product rule, first let us describe our process.

- Choose c_1
- Given c_1, choose c_2
- Given c_1, c_2, choose c_3
- et cetera

Next, we can see that $|S| = \frac{52!}{47!}$

Definition 6.9. An ordered sequence of r distinct objects chosen from n distinct objects is called an r-permutation. The number is $\frac{n!}{(n-r)!}$

Example 6.5:
Observe: We will proceed by division rule: Attempt: Process:

- Choose a card c_1
- Choose $c_2 \neq c_1$
- et cetera

The problem with this is that it doesn't uniquely specify an unordered set of cards

Definition 6.10. $f : A \to B$ is m-to-1 if it is onto and $\forall\, b \in B$, $|f^{-1}(\{b\})| = m$,

Notice that our mapping from choices to unordered hands is an m-to-1 function. So we have the same thing as before but we are dividing by the number of orderings of our sequence by 5!. So we have $\binom{52}{5}$ or $\frac{52!}{47!5!}$

§6.2 10.31.2024

§6.2.1 Binomial Theorem

Theorem 6.11

If $n \geq 1$, then

$$(x+y)^n = \Sigma_0^n \binom{n}{k} x^k y^{n-k}$$

Proof. Observe that the expansion of $(x+y)^n$ consists of terms corresponding to xy-strings of length n, since each term is defined by choosing a veriable x or y from each paranthesized factor of $(x+y)\ldots(x+y)$. Grouping monomials, the coefficient of $x^k y^{n-k}$ is the number of xy-strings of length n with exactly k x's, which is equal to the number of k-subsets of $\{1,\ldots,n\}$, which is $\binom{n}{k}$ □

§6.2.2 Combinatorial Identities

§6.2.3 Permutations and Combinations with Repitition

stars and bars or something

---

§6.3 11.5.2024

§6.3.1 Ball and Urns

no stars and bars are here

§6.3.2 Recurrence Relations

---

§7 Probability

§7.1 11.14.2024

§7.1.1 Probability

joke: what does Math stand for?

A: Math Ath Th H

Axioms of Classical Probability

1. An **experiment** is a well-defined procedure with a set of outcomes. An outcome is a complete specification of (all the relevant details of) an experiment.

2. The set of possible outcomes of an experiment is called a **sample space** S.
3. An **Event** is a subset of the sample space $E \subseteq S$
4. Given that,
   a) S is finite
   b) All outcomes in S are equally likely

   the probability of an event E denoted $\mathbb{P}(E)$ is $(E) = \frac{|E|}{|S|}$

What does (E) mean?

- Frequentist Interpretation: If you repeat the experiment many times, the outcome will lie in E $\approx (E)$ fraction of the time.
-

---

§7.2 12.3.2024

§7.2.1 Coupon Collector

Example 7.1

Roll a fair die until I see two distinct numbers; i.e., $(1, 1, 1, 1, 1, 1, 4)$. What is the expected number of rolls?
We can view this experiment in multiple different steps:

1. Roll the die once
2. Roll the die repeatedly until a new number pops up

X = total number of rolls $X_1 + X_2$, where X_1 = number of rolls in step one, and $X_2 = \sim$ [a] Geometric $(\frac{5}{6})$ number of rolls in step two. Then $\mathbb{E}X = 1 + \frac{6}{5} = \frac{11}{5}$

[a] $\sim$ means "has distribution"

Coupon Collector:
Suppose we have n coupons total. $\{1, 2, 3, \ldots, n\}$. In each trial, get one uniformly at random.
X = number of trials until you get at least one of each. What is the expected value of X
Let
$X_0 = 1$ = number of trials to see the first new coupon
X_1 = number of trials between seeing the first and second new coupons
...
X_i = number of trials after X_{i-1} = (number of trials to see i new coupons) to see the $i + 1$st new coupon.
...

So our formula is

$$\sum_{i=0}^{n-1} \frac{n-i}{n} \approx n \int_1^n \frac{1}{x} dx = n \ln n$$

§7.2.2 Algebra with random variables

Given random variables $X, Y : S \to \mathbb{R}$, we defined

1. $(X+Y)(s) = X(s) + Y(s)$ - SUM
2. $(cX)(s) = x \cdot X(s)$ where $c \in \mathbb{R}$ - scalar mult
3. $XY : S \to \mathbb{R}$ by $(XY(s) = X(s) + Y(s))$ - Product

Example 7.2

Suppose we have two independent dice rolls, with:
X = number on first
Y = number on second
XY = product of the two numbers

$\mathbb{E}1_E 1_F = E1_{E\cap F} = \mathbb{P}(E \cap F)$
$\mathbb{E}1_E \mathbb{E}1_F = \mathbb{P}(E)\mathbb{P}(F)$
These two are not equal, except for when E, F are independent.

Definition 7.3. Random variables $X, Y : S \to R$ are **independent** if $\forall\, r_1, r_2 \in \mathbb{R}$ the events $\{X = r_1\}, \{Y = r_2\}$ are independent (i.e., $\mathbb{P}(x = r_1 \cap Y = r_2) = \mathbb{P}(X = r_1 \mathbb{P}(Y = r_2))$

Theorem 7.4

If X, Y are independent random variables, then $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$

Remark 7.5. Linearity of expectation does not require independence - with sums you don't need independence, but with products you do.

Polynomials: Given $X, Y : S \to \mathbb{R}$, we can define expressions such as $(X+Y)^2 = X^2 + 2XY + Y^2$

§7.2.3 Variance

Motivation: Given X, we want to understand how close X "typically" is to $\mathbb{E}X$.

Definition 7.6. Variance is defined as the square root of the standard deviation

Definition 7.7. If a random variable X has $\mathbb{E}X = \mu_x$, the **variance** of X denoted $V(X)$ is

$$\mathbb{E}(X - \mu_X)^2$$

Remark 7.8. The standard deviation $\text{stdev}(x) = \sqrt{V(X)}$ means the "typical" deviation from μ_X

Remark 7.9. Why don't we do this? $\mathbb{E}(X - \mu_X)$

Answer: because $\mathbb{E}(X - \mu_X) = 0$
Follow up question: Why don't we just compute the absolute values
Answer: Pain to calculate, but also, $\mathbb{E}(X - \mu_X)^2 = X - \mu_X, X - \mu_X$ in an inner product space.
This is the real reason why we square it. (Hidden euclidean geometry)

Theorem 7.10

If X and Y are independent random variables, $V(X + Y) = V(X) + V(Y)$

Corollary 7.11

The same follows for n pairwise independent random variables

The above follow computationally

---

§7.3 12.5.2024

§7.3.1 Markov's Inequality

Given $S, p : S \to \mathbb{R}$, $X : S \to \mathbb{R}$. How close is X to $\mathbb{E}X$, typically?

Theorem 7.12

If Y is a nonnegative random variable, then for all $t > 0$, $\mathbb{P}(Y \geq t) \leq \frac{\mathbb{E}Y}{t}$

This tells us that A nonnegative random variable is unlikely to be much larger than its Expected Value.

Example 7.13

Y = random midterm 2 score. $\mathbb{E}Y = 30$.
Let $t = 40$. Markov's inequality tells us that $\mathbb{P}(Y \geq 40) \leq \frac{30}{40} = \frac{3}{4}$.

Proof. Let $Y \geq 0, t > 0$.

$$\begin{aligned}\mathbb{E}(Y) &:= \sum_{r \geq 0} r\mathbb{P}(Y = r) \\ &= \sum_{r \geq 0, r \geq t} r\mathbb{P}(Y = r) + \sum_{r \geq 0, r < t} rP(Y = r)\end{aligned}$$

$$\geq \sum_{r\geq 0, r\geq t} r\mathbb{P}(Y=r)$$
$$\geq \sum_{r\geq 0, e\geq t} t\mathbb{P}(Y=r) = t\sum_{r\geq 0, e\geq t}\mathbb{P}(Y=r)$$
$$= t\mathbb{P}(Y\geq t)$$

□

§7.3.2 Chebyshev's Inequality

Theorem 7.14

If Y is a random variable, $\forall t > 0$, $\mathbb{P}(|Y - \mathbb{E}Y| \leq \frac{V(Y)}{t^2}$

Proof. Given Y, $t > 0$, let $\mu_x = \mathbb{E}Y$ and define $Z = (Y - \mu_Y)^2$, which is nonnegative. Notice $\mathbb{E}Z = \mathbb{E}(Y - \mu_Y)^2 = V(Y)$.
By Markov, $\mathbb{P}(Z \geq t^2) \leq \frac{\mathbb{E}(Z)}{t^2} = \frac{V(Y)}{Z}$
But $\mathbb{P}(Z \geq t^2) = \mathbb{P}((Y - \mu_Y)^2 \geq t^2) = \mathbb{P}(|Y - \mu_Y| \geq t)$, which is Chebyshev □

Example 7.15

Prototypical Example throughout statistics: Given a coin with unknown bias $q \in [0,1]$, how can you find q?

Idea: Flip a coin n times independently.

Let $X_i = \begin{cases} 1 \text{ if ith flip is H where i} = 1, \ldots, \text{n} \\ 0 \text{ otherwise} \end{cases}$ Define $\hat{q} = \frac{X_1+X_2+\cdots+X_n}{n}$ = fraction of flips which were H.

- $\mathbb{E}\hat{q} = \frac{\mathbb{E}X_1+\mathbb{E}X_2+\cdots+\mathbb{E}X_n}{n} = \frac{q+\cdots+q}{n} = q$
-

$$\begin{aligned} V(\hat{q}) &= \mathbb{E}(\frac{X_1+\cdots+X_n}{n})^2 - (\mathbb{E}(X_1+\cdots+X_n))^2 \\ &= \frac{1}{n^2}[\mathbb{E}(X_1+\cdots+X_n)^2 - (\mathbb{E}(X_1+\cdots+X_n))^2] \\ &= \frac{1}{n^2}V(X_1+\cdots+X_n) \\ &= \frac{1}{n^2}[V(X_1)+\cdots+V(X_n)] \\ &= \frac{nq(1-q)}{n^2} \\ &= \frac{q(1-q)}{n} \leq \frac{1}{4n} \text{ for all } q \in [0,1] \end{aligned}$$

By Chebyshev:

$$\mathbb{P}(|\hat{q}-q| \geq t) \leq \frac{V}{(}\hat{q})t^2 \leq \frac{1}{4nt^2}$$

For $t = \frac{1}{20}$, $n = 1000$, $\mathbb{P}(\hat{q}-q| \geq \frac{1}{20}) \leq \frac{1}{4 \cdot 1000\frac{1}{400}} = \frac{1}{10}$. Therefore, with 90% probability, $\hat{q}$ is within $\frac{1}{20}$ of the bias.

Remark 7.16. • In general, $V(xC) = c^2V(X)$ for $c \in \mathbb{R}$.

§7.3.3 Law of Large Numbers

Theorem 7.17

If $X_1, \ldots, X_n$ are independent random variables with the same distribution, $\mathbb{E}X < \infty$, $V(X) < \infty$, for all $t > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|\frac{X_1+\cdots+X_n}{n} - \frac{\mathbb{E}X_1+\ldots X_n}{n}| \geq t) = 0$$