Note 7
(RSA)

- Def: Public key is (N, e) where $N = pq$ s.t. p, q are large primes. e (typically small $\mathbb{Z}^+$) such that $e, (p-1)(q-1)$ coprime. Private key is d s.t. $d \equiv e^{-1} \pmod{(p-1)(q-1)}$. Original message x encrypted into $y \equiv E(x) = x^e \pmod{N}$, decryption $D(y) = y^d \equiv x^{ed} \equiv x \pmod{N}$. (Extra: $de - 1 \leq e(p-1)(q-1)$)
- Th 7.1: $D(E(x)) = x \pmod{N}$ $\forall\ x \in \{0, 1, ..., N-1\}$.
- Th 7.2 (FLT): For any prime p and $a \in \{1, 2, ..., p-1\}$, then $a^{p-1} \equiv 1 \pmod{p}$ (Extra: for any x, $x^y = x^{y \pmod{p-1}} \pmod{p}$, $x^p \equiv x \pmod{p}$).
- Th 7.3 (Prime Number Theorem): Let $\pi(n)$ denote # of primes that are $\leq n$. Then $\forall n \geq 17$, we have $\pi(n) \geq \frac{n}{\ln n}$ (in fact $\lim_{n\to\infty} \frac{\pi(n)}{n/\ln n} = 1$).
- HW: $a^{p(p-1)} \equiv 1 \pmod{p^2}$ where p is prime and a, p are coprime
- HW: If $a \equiv b \pmod{m_1}, a \equiv b \pmod{m_2}$ and $\gcd(m_1, m_2) = 1$, then $a \equiv b \pmod{m_1 m_2}$.
- HW (composite FLT): If $n = p_1 p_2 \cdots p_k$ where p_i are distinct primes and $(p_i - 1) \mid (n-1)$ $\forall i$, then $a^{n-1} \equiv 1 \pmod{n}$ $\forall a \in \{i \mid 1 \leq i \leq n \wedge \gcd(n, i) = 1\}$.
- Extra: If $ax + bm = d$, $\gcd(x, m) = d, xu \equiv v \pmod{m}$, then it has a solution $\iff d \mid v$; if so, one such solution is $u \equiv \frac{va}{d} \pmod{\frac{m}{d}}$, (in essence, $a = (\frac{x}{d})^{-1} \pmod{\frac{m}{d}}$,) and there are exactly d-many solutions (of the form $u = \frac{va}{d} + i \cdot \frac{m}{d} \pmod{m}$).
- Extra: $(p-1)! \equiv (p-1) \pmod{p}$: Proof by the fact that $2, ..., p-2$ will pair up with their own inverse, and only ones that map back to themselves are $1, -1$.
- Extra: The existence of a^{-1} allows us to conclude that $ar \pmod{n}$ to $r \pmod{n}$ is a bijection.
- Extra: If there are k numbers that are relatively prime to N in $\{0, \ldots, N-1\}$, then $a^k \equiv 1 \pmod{N}$ if $\gcd(a, N) = 1$ (Proof $\sim$FLT).

Note 8
(Polynomials)

- Fundamental properties (1) A non-zero polynomial of deg d has at most d roots; (2): $d+1$ distinct pairs $(x_1, y_1), \ldots, (x_{d+1}, y_{d+1})$ uniquely decide $p(x)$ of degree (at most) d such that $p(x_i) = y_i$ for $1 \leq i \leq d+1$.
- Lagrange interpolation (Construct the unique $p(x)$ of deg $\leq d$ from $d+1$ pairs)
  1. Def: Let $\Delta_i(x)$ be the degree d polynomial that goes through these $d+1$ points, construct $\Delta_i(x) = \frac{\prod_{j\neq i}(x - x_j)}{\prod_{j\neq i}(x_i - x_j)}$ for $i \in \{1, \ldots, d+1\}$ with the given pairs.
  2. Construct $p(x) = \sum_{i=1}^{d+1} y_i \Delta_i(x)$
- Finite field: $GF(p)$ or $F(p)$ when we work with numbers modulo a prime p
- Note that the reason that we can count the number of polynomials in this setting is because we are working over a finite field.

- Secret Sharing utilizes Lagrange interpolation and works under $GF(p)$. N.B., two people have no more information about the secret than one.
- HW: Min/Max number of roots for non-zero real polynomials f and g
  - $f+g$: min = 0 if sum is even degree, 1 if odd degree; max = $\max(\deg(f), \deg(g))$ if sum is non-zero, and infinity if $f = -g$
  - $f \cdot g$: min = 0; max = $\deg(f) + \deg(g)$
  - f/g (assume polynomial): min = 0; max = $\deg(f)$ - $\deg(g)$
- HW: Consider $f = x^{p-1} - 1, g = x$ be polynomials over $GF(p)$. Both are non-zero polynomials, but their product have a zero on all points in $GF(p)$.
- HW: If f is under $GF(p)$ such that $\deg(f) \geq p$, then there exists a polynomial h with $\deg(h) < p$ such that $f(x) = h(x)$ for all $x \in \{0, 1, \ldots, p-1\}$.
- HW: There are $(p-1) \cdot p^{d-1}$ many f of degree *exactly* $d < p$ over $GF(p)$ such that $f(0) = a$ for some fixed $a \in \{0, 1, \ldots, p-1\}$
- Extra: There are p^p polynomials in $GF(p)$ for some prime p.
- Extra: $(x-1)(x-2)\cdots(x-p+1) \equiv x^{p-1} - 1 \pmod{p}$

Note 9
(Error Correction)

- In mod, $\frac{a}{x} \equiv ax^{-1} \pmod{p}$
- If we want to send x packets, work under $GF(p)$ where $p > x$ and p is prime.
- Consider sharing an n-bit secret, where the secret is encoded as $P(0)$ for a polynomial of degree k modulo p where s shares will be handed out, then $p \geq \max(2^n, k+1, s+1)$.
- Erasure Errors: At least $(n+k)$ packets for k erasure errors. We can uniquely reconstruct $P(x)$ from its values at *any* n distinct points via Lagrange interpolation, since it has degree $n-1$.
- General Errors: At least $(n+2k)$ packets for k general/corruption errors (Extra: The distance between any two codewords must be $2k+1$).
  - Error-locator polynomial (deg k): $E(x) = (x-e_1)(x-e_2)\ldots(x-e_k)$, so k of $k+1$ coefficients unknown.
  - $P(i)E(i) = r_i E(i)$ for $1 \leq i \leq n+2k$
  - Define $Q(x) := P(x)E(x)$ $(= r_x E(x))$, which is a polynomial of degree $(n+k-1)$, and is therefore described by $n+k$ coefficients $a_0, a_1, \ldots, a_{n+k-1}$
  - We have $(n+2k)$ packets (linear equations), one for each value of i, and $(n+2k)$ unknowns. Solve for $E(x)$ and $Q(x)$. Compute $\frac{Q(x)}{E(x)}$ to obtain $P(x)$.
- Error Detection (HW): At least $(n+k)$ packets to detect any errors, for maximum of k errors.
- Extra: To send n-bit message through a channel that (1) loses fraction f of packets, need $\frac{n}{1-f}$ packets; (2) corrupts fraction f of packets, need $\frac{n}{1-2f}$ packets.

Note 10
(Infinity and Countability)

- Def: A set S is countable if there is a bijection between S and $\mathbb{N}$ or some subset of $\mathbb{N}$. (Extra: if there's an injection from S to $\mathbb{N}$, then latter part holds.)
- The Cantor-Bernstein theorem (?): $|B| \leq |A| \leq |B| \iff A, B$ have the same cardinality. (one-to-one + onto (reversed 1-to-1) $\rightarrow$ bijection)
- Notes: The set of all (finite-degree) polynomials with $\mathbb{N}$ coefficients, $N(x)$, is countable.
- Th: The real interval $\mathbb{R}[0,1]$ (and hence also $\mathbb{R}$) is uncountable. (Proof by Diagonalization)
- Th: $|\mathscr{P}(\mathbb{N})| > |\mathbb{N}|$ (Extra: The power set of any infinite set is uncountable).
- Disc: The countable union of countable sets is countable.
- Disc: The subset of any countable set is countable, so the intersection of (a countable or uncountable number of) countable sets is countable.
- Disc: The functions from $\mathbb{N}$ to $\mathbb{N}$ is uncountable. (Diag.)
- Disc: The total number of programs is countably infinite (each is a string of characters), so halting programs $\subseteq$ all programs is countable. Since there are ∞ halting programs (e.g. for each $i \in \mathbb{N}$ the program that prints i), so the total number of halting programs is countably infinite. (Not every function from $\mathbb{N}$ to $\mathbb{N}$ can be written as a program.)
- Disc: Given A is countable, non-empty. $\forall i \in A$, S_i is uncountable; B is uncountable. $\forall i \in B$, Q_i is countable.
  - $A \cap B$ countable; $A \cup B$ uncountable
  - $\cup_{i \in A} S_i$ uncountable; $\cap_{i \in A} S_i$ both could happen (disjoint vs. identical)
  - $\cup_{i \in B} Q_i$ both could happen (identical vs. $B = \mathbb{R}, Q_i = \{i\}$); $\cap_{i \in B} Q_i$ countable
- Disc: Not possible to determine if we ever execute a specific line (depends on the logic of the program), but the number of steps/instructions can be counted.
- Extra: Finite graphs (V, E) is countably infinite (defined isomorphically).

Note 11
(Self-Reference and Computability)

- Th: The Halting Problem is uncomputable; i.e., *TestHalt* that can determine, on all inputs (P, x), whether the program P halts on input x DNE.
- EastHalt is uncomputable (testing whether P halts on input 0). (Proof by Reduction)
- Disc: We can determine whether a program halts/outputs sth in k steps, but not before k^{th} line.
- HW
  - $f(x) = p \pmod{x}$, where $p > 2$ is a prime, then $f : \{\frac{p+1}{2}, \ldots, p\} \rightarrow \{0, \ldots, \frac{p-1}{2}\}$ is a bijection.
  - Countable: $A \times B$, where A and B are both countable.
  - Countable: $\cup_{i \in A} B_i$ where A, B_i are all countable
  - Uncountable: The set of all functions $f : \mathbb{N} \rightarrow \mathbb{N}$ such that f is non-decreasing, i.e. $f(x) \leq f(y)$ whenever $x \leq y$. (Proof by Diag.)
  - Countable: The set of all functions $f : \mathbb{N} \rightarrow \mathbb{N}$ such that f is non-increasing, i.e. $f(x) \geq f(y)$ whenever $x \geq y$.

- Uncountable: The set of all bijective functions from $\mathbb{N}$ to $\mathbb{N}$.

- Extra: We can determine whether a program contains a loop, given finite memory or limited time/steps (since we can list all possible k states and run the program $k+1$ steps). We can't check an arbitrary program as a whole w/o any restriction.

- Extra: There is a computer program that prints all programs and the inputs where they halt.

Note 12
(Counting)

- First Rule of Counting: $n_1 n_2 \cdots n_k$ (order matters)

- Second Rule of Counting: $\binom{n}{k}$ (order doesn't matter)

	Replacement	No Replacement
Order	n^k	$\frac{n!}{(n-k)!}$
No Order	$\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$	? $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

- HW: If there are exactly 20 students currently enrolled in a class, then there are $19 \cdot 17 \cdots 3 \cdot 1 = \frac{20!}{10! \cdot 2^{10}} = \frac{\binom{20}{2}\binom{18}{2}\cdots\binom{2}{2}}{10!}$ different ways to pair up the 20 students.

- Disc: There are $\sum_{k=3}^{n} \frac{n!}{(n-k)! \cdot 2k}$ distinct cycles in a complete graph with n vertices. Assuming no duplicated edges, two cycles are the same by rotation or inversion.

Note 13
(Discrete Probability)

- Def: The outcome of a random experiment is called a sample point, and the sample space Ω is the set of all possible outcomes of the experiment.

- Def: A **probability space** is a **sample space** Ω with $\mathbb{P}[\omega]$ ($\Pr[\omega]$) for each sample point ω, such that

  - (Non-negativity): $0 \leq \mathbb{P}[\omega] \leq 1$ for all $\omega \in \Omega$
  - (Total one): $\sum_{\omega \in \Omega} \mathbb{P}[\omega] = 1$, i.e., the sum of the probabilities over all outcomes is 1.

- If the coin has $\mathbb{P}(H) = p$. Consider any sequence of n flips with exactly r H's, then $\mathbb{P}[\text{this sequence}] = p^r(1-p)^{n-r}$. Now consider the event C that we get exactly r H's when we flip the coin n times, so $\mathbb{P}[C] = \binom{n}{r} p^r (1-p)^{n-r}$

- If we throw m labeled balls into n labeled bins, we have a sample space of size n^m (generalized).

- Throw m (identical or distinct) balls into n distinct bins with uniform probability, then $\mathbb{P}[\text{the first bin has } \textbf{exactly } k \text{ balls}] = \binom{m}{k} \cdot (\frac{1}{n})^k \cdot (1 - \frac{1}{n})^{m-k}$

Note 14
(Conditional Probability)

- Def 14.1 (Conditional Probability): For events $A, B \subseteq \Omega$ in the same Pr space such that $\mathbb{P}[B] > 0$, the conditional Pr of A *given* B is $\mathbb{P}[A|B] = \sum_{\omega \in A \cap B} \mathbb{P}[\omega|B] = \sum_{\omega \in A \cap B} \frac{\mathbb{P}[\omega]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$

- $\mathbb{P}[A \cap B] = \mathbb{P}[A|B] \cdot \mathbb{P}[B] = \mathbb{P}[B|A] \cdot \mathbb{P}[A] \longrightarrow \mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]\mathbb{P}[A]}{\mathbb{P}[B]}$ Bayes' Rule

- $\mathbb{P}[B] = \mathbb{P}[A \cap B] + \mathbb{P}[\overline{A} \cap B] = \mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[B|\overline{A}]\mathbb{P}[\overline{A}] = \mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[B|\overline{A}](1 - \mathbb{P}[A])$ Total Probability Rule
- Thus, combining the two equations above gives: $\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]\mathbb{P}[A]}{\mathbb{P}[B]} = \frac{\mathbb{P}[B|A]\mathbb{P}[A]}{\mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[B|\overline{A}](1-\mathbb{P}[A])}$
- Def 14.2 (Partition of an event): We say that an event A is partitioned into n events $A_1, \ldots, A_n$ if
  1. $A = A_1 \cup A_2 \cup \cdots \cup A_n$, and
  2. $A_i \cap A_j = \emptyset$ for all $i \neq j$ (mutually exclusive)
- Let $A_1, \ldots, A_n$ be a partition of the sample space Ω. Then, the Total Probability Rule for any event B is $\mathbb{P}[B] = \sum_{i=1}^{n} \mathbb{P}[B \cap A_i] = \sum_{i=1}^{n} \mathbb{P}[B|A_i]\mathbb{P}[A_i]$, and the Bayes' Rule (assuming $\mathbb{P}[B] \neq 0$) is $\mathbb{P}[A_i|B] = \frac{\mathbb{P}[B|A_i]\mathbb{P}[A_i]}{\mathbb{P}[B]} = \frac{\mathbb{P}[B|A_i]\mathbb{P}[A_i]}{\sum_{j=1}^{n} \mathbb{P}[B|A_j]\mathbb{P}[A_j]}$
- Def 14.3 (Independence): Two events $A, B \in \Omega$ are independent $\iff \mathbb{P}[A \cap B] = \mathbb{P}[A] \times \mathbb{P}[B]$. For events A, B such that $\mathbb{P}[B] > 0$, the condition $\mathbb{P}[A|B] = \mathbb{P}[A] \equiv$ independence.
- If X, Y are independent, then $\mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY]$, but the converse isn't always true.
- Def 14.4/5 (Mutual independence): Events $A_1, \ldots, A_n$ are said to be mutually independent if (two equivalent definitions):
  - for EVERY subset $I \subseteq \{1, \ldots, n\}$ with size $|I| \geq 2$, we have $\mathbb{P}[\cap_{i \in I} A_i] = \prod_{i \in I} \mathbb{P}[A_i]$
  - for all $B_i \in \{A_i, \overline{A_i}\}, i = 1, \ldots, n$, we have $\mathbb{P}[B_1 \cap \cdots \cap B_n] = \prod_{i=1}^{n} \mathbb{P}[B_i]$
- The independence of every pair of events (so-called pairwise independence) does not necessarily imply mutual independence.
- Theorem 14.1 (Product Rule): $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B|A]$. More generally, for any events $A_1, \ldots, A_n$, we have $\mathbb{P}[\cap_{i=1}^{n} A_i] = \mathbb{P}[A_1] \cdot \mathbb{P}[A_2|A_1] \cdots \mathbb{P}[A_3|A_1 \cap A_2] \cdots \mathbb{P}[A_n| \cap_{i=1}^{n-1} A_i]$ (Proof by Induction)
- Theorem 14.2 (Inclusion-Exclusion): Let $A_1, \ldots, A_n$ be events in some probability space, where $n \geq 2$. Then, we have $\mathbb{P}[\cup_{i=1}^{n} A_i] = \sum_{i=1}^{n} \mathbb{P}[A_i] - \sum_{i<j} \mathbb{P}[A_i \cap A_j] + \sum_{i<j<k} \mathbb{P}[A_i \cap A_j \cap A_k] - \cdots + (-1)^{n-1}\mathbb{P}[A_1 \cap A_2 \cap \cdots \cap A_n]$ (Proof by Induction)
- (Mutually exclusive events) If the events $A_1, \ldots, A_n$ are mutually exclusive, then $\mathbb{P}[\cup_{i=1}^{n} A_i] = \sum_{i=1}^{n} \mathbb{P}[A_i]$
- (Union bound) Upper bound always due to overestimating: $\mathbb{P}[\cup_{i=1}^{n} A_i] \leq \sum_{i=1}^{n} \mathbb{P}[A_i]$

Note 15
(Random Variables, Expectation)

- Def 15.1 (Random Variable): A random variable X on a sample space Ω is a function $X : \Omega \to \mathbb{R}$ that assigns to each sample point $\omega \in \Omega$ a real number $X(\omega)$.

- Def 15.2 (Distribution): The distribution of a discrete random variable X is the collection of values $\{(a, \mathbb{P}[X = a]) : a \in \mathscr{A}\}$, where $\mathscr{A}$ is the set of all possible values taken by X (Note that the collection of events form a partition of the sample space Ω).
- Bernoulli Distribution: $X \sim \text{Bernoulli}(p)$ which takes value in $\{0, 1\}$ such that $p = \mathbb{P}[X = 1] = 1 - \mathbb{P}[X = 0]$ where $0 \leq p \leq 1$.
- Binomial Distribution: $X \sim \text{Bin}(n, p)$ such that $\mathbb{P}[X = i] = \binom{n}{i} \cdot p^i(1-p)^{n-i}$ for $i = 0, 1, \ldots, n$.
- If $X \sim Bin(n, p)$, then $\mathbb{E}[X] = np$. For independent $X, Y \sim Bin(n, p)$, we have $X + Y \sim Bin(2n, p)$.
- The $\text{Bin}(n, p)$ gives probabilistic proof for the Binomial Theorem since by properties of X we have $\sum_{i=0}^{n} \mathbb{P}[X = i] = 1 \longrightarrow \sum_{i=0}^{n} \binom{n}{i} \cdot p^i(1-p)^{n-i} = 1$
- Relation to Error Correction: If we model each packet getting lost with probability p and the losses are independent, then if we transmit $n + k$ packets, the number of packets received is a random variable X with binomial distribution: $X \sim \text{Bin}(n + k, 1 - p)$, so the probability of successfully decoding the original data is: $\mathbb{P}[X \geq n] = \sum_{i=n}^{n+k} \mathbb{P}[X = i] = \sum_{i=n}^{n+k} \binom{n+k}{i} \cdot (1-p)^i p^{n+k-i}$
- Def 15.3: The joint distribution for two discrete random variables X and Y is the collection of values $\{((a, b), P[X = a, Y = b]) : a \in \mathscr{A}, b \in \mathscr{B}\}$, where $\mathscr{A}$ is the set of all possible values taken by X and $\mathscr{B}$ is the set of all possible values taken by Y.
- Def 15.4 (Independence): Random variables X and Y on the same probability space are said to be independent if the events $X = a$ and $Y = b$ are independent for all values a, b. Equivalently, the joint distribution of independent r.v.'s decomposes as $\mathbb{P}[X = a, Y = b] = \mathbb{P}[X = a]\mathbb{P}[Y = b]$, $\forall a, b$.
- Def 15.5 (Expectation): The expectation of a discrete random variable X is defined as $\mathbb{E}[X] = \sum_{a \in \mathscr{A}} a \cdot \mathbb{P}[X = a]$ where the sum is over all possible values taken by the r.v. (Expectation is well defined provided that the sum on the RHS is absolutely convergent, and there are random variables for which expectations do not exist (haven't encountered yet but keep in mind)).
- The expectation can be seen in some sense as a "typical" value of the r.v. (though note that $\mathbb{E}[X]$ may not actually be a value that X can take, like a regular die).
- Theorem 15.1 (Linearity of Expectation): For any two random variables X, Y on the same probability space and any constant c, we have $\mathbb{E}[X+Y] = \mathbb{E}[X]+\mathbb{E}[Y], \mathbb{E}[cX] = c\mathbb{E}[X]$ regardless of whether or not X and Y are independent. If X, Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.
- Indicator variable: $X_n = I_1 + \cdots + I_n$ where $I_i = 1$ if student i gets their own HW and 0 otherwise. So, $\mathbb{E}$[number of students who get their own homeworks] in a class of size n is 1. That is, the expected number of fixed points in a random permutation of n items is always 1, regardless of n.
- HW: If A and B are integer-valued random variables such that for every integer i, $P(A = i) = P(B = i)$, then $P(A = B) > 0$ is not necessarily true. (Counterexample: Let A be 0 with probability $\frac{1}{2}$ and 1 with probability $\frac{1}{2}$. Let $B = 1 - A$. Then A and B are never equal but B takes the values 0, 1 with probability $\frac{1}{2}$ each as well.)
- HW: If $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$, then A is independent with itself; If A, B are independent, then $\overline{A}, \overline{B}$ are independent.

- Extra: $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$ since $\mathbb{E}[Z^2] - \mathbb{E}[Z]^2 = \mathbb{E}\big[(Z - \mathbb{E}[Z])^2\big] \geq 0$ with $Z \cdot \mathbb{E}[Z] = \mathbb{E}[Z^2]$

Extra Sanity Check

- For RSA, $x^{ed} \equiv x \pmod{N}$
- All polynomials that Lagrange Interpolation can output $\neq$ all polynomial solutions (e.g. polynomial with the coordinates $(-1, 1), (0, 0), (2, 4)$ exist in $GF(8)$).
- If two degree d polynomials intersect on $d + 1$ points, they must be the same polynomial.
- f is bijection $\iff$ it is both one-to-one and onto.
- Prove strategy for uncountable: Diagonalization; for uncomputable: Use self-referencing algorithms OR Reduction to the Halting Problem:
  1. Take input from H, and convert it into input for A.
  2. Take output from A, and show how it is correct and valid as output for H.
- There are two ways to show undecidability: Use your program as a subroutine to solve a problem we know is undecidable OR Proof by Diag..
- Care whether problem needs counting or probability.
- When in doubt, always calculate probability by considering the elements as distinct.
- Calculating probability with these steps:
  - What is the sample space (i.e., the experiment and its set of possible outcomes)?
  - What is the probability of each outcome (sample point)?
  - What is the event we are interested in (i.e., which subset of the sample space)?
  - Finally, compute the probability of the event by adding up the probabilities of the sample points contained in it.
- When stuck on probability, consider symmetry.
- The independence of every pair of events (so-called pairwise independence) does not necessarily imply mutual independence (counterexample: flip first coin, flip second coin, both flips are the same).
- Two disjoint events A and B with $\mathbb{P}[A] > 0$ and $\mathbb{P}[B] > 0$ cannot be independent.
- When solving r.v. problems of intuition, always double-check the edge cases of $\mathbb{P}[A] = 0$ or 1.
- A random experiment must define the sample space (the set of possible outcomes) AND a set of probabilities.
- Combinatorial Proofs (Addition is OR, multiplication is AND)
  - $\binom{n}{k+1} = \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k}{k}$
  - $2^n = \sum_{i=0}^{n} \binom{n}{i}$
  - $(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$
  - $\binom{n}{k} \leq n^k$

- $\sum_{i=1}^{n} i\binom{n}{i} = n2^{n-1}$
- $\sum_{i=1}^{n} i\binom{n}{i}^2 = n\binom{2n-1}{n-1}$
- $\binom{2n}{n} = 2\binom{2n-1}{n-1} = \binom{2n-1}{n-1} + \binom{2n-1}{n}$
- $\sum_{i=0}^{n} \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} = 3^n = \sum_{k=0}^{n} \binom{n}{k} 2^k$
- $\sum_{k=b}^{a} \binom{k}{b} = \binom{a+1}{b+1}$ (16fa Q3 (4))
- $\sum_{k=0}^{n} k^2\binom{n}{k} = n(n-1)2^{n-2} + n2^{n-1}$
- $\binom{k+n-1}{n-1} = \sum_{i=0}^{k} \binom{k-i+n-2}{n-2}$

- Extra: In $GF(p)$, $\exists$ $p(x)$ of degree d and $q(x)$ of degree $d-1$ such that a degree 1 polynomial $y(x) = p(x)$ satisfies $p(-y(0)) = 0$, where $d < p-1$ ($\equiv p, q$ share $d-1$ roots).
- Godel's Incompleteness Theorem: Any formal system that is sufficiently rich to formalize arithmetic is either inconsistent (there are false statements that can be proved) or incomplete (there are true statements that cannot be proved).