Note 1 (Propositional Logic)

- $P \implies Q \equiv \neg P \vee Q$
- $P \implies Q \equiv \neg Q \implies \neg P$
- $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$
- $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$
- $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$
- $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$
- $\neg(\forall x P(x)) \equiv \exists x \neg P(x)$
- $\neg(\exists x P(x)) \equiv \forall x \neg P(x)$

Note 2/3 (Proofs)

- Direct proof
- Proof by contraposition
- Proof by cases
- Proof by induction
  - Base case (prove smallest case is true)
  - Inductive hypothesis (assume $n = k$ true for weak induction, assume $n \leq k$ true for strong induction)
  - Inductive step (prove $n = k + 1$ is true)
- Pigeonhole principle
  - Putting $n + m$ balls in n bins $\implies \geq 1$ bin has ≥ 2 balls

Note 4 (Sets)

- $\mathscr{P}(S)$ = powerset/set of all subsets; if $|S| = k$, $|\mathscr{P}(S)| = 2^k$
- One to one (injection); $f(x) = f(y) \implies x = y$
- Onto (surjection); $(\forall y \exists x)(f(x) = y)$; “hits" all of range
- Bijection: both injective and surjective

Note 5 (Countability & Computability)

- Countable if bijection to $\mathbb{N}$
- Cantor-Schroder-Bernstein Theorem: bijection between A and B if there exists injections $f : A \to B$ and $g : B \to A$
- Cantor diagonalization: to prove uncountability, list out possibilities, construct new possibility different from each listed at one place (ex. reals $\in (0, 1)$, infinite binary strings, etc)
- $A \subseteq B$ and B is countable $\implies$ A is countable
- $A \supseteq B$ and A is uncountable $\implies$ B is uncountable
- Infinite cartesian product sometimes countable ($\emptyset \times \emptyset \times \cdots$), sometimes uncountable ($\{0,1\}^\infty$)
- Halting Problem: can't determine for every program whether it halts (uncomputable)
- Reduction of `TestHalt(P, x)` to some task (here, `TestTask`)
  - define inner function that does the task if and only if `P(x)` halts
  - call `TestTask` on the inner function and return the result in `TestHalt`

Note 6 (Graph Theory)

- K_n has $\frac{n(n-1)}{2}$ edges
- Handshaking lemma: total degree = $2e$
- Trees: (all must be true)
  - connected & no cycles
  - connected & has $n - 1$ edges ($n = |V|$)
  - connected & removing an edge disconnects the graph
  - acyclic & adding an edge makes a cycle
- Hypercubes:
  - n-length bit strings, connected by an edge if differs by exactly 1 bit
  - n-dimensional hypercube has $n2^{n-1}$ edges, and is bipartite (even vs odd parity bitstring)
- Eulerian walk: visits each edge once; only possible if connected and all even degree or exactly 2 odd degree
- Eulerian tour: Eulerian walk but starts & ends at the same vertex; only possible if all even degree and connected
- Planar graphs
  - $v + f = e + 2$
  - $\sum_{i=1}^{f} s_i = 2e$ where s_i = number of sides of face i
  - $e \leq 3v - 6$ if planar (because $s_i \geq 3$)
  - $e \leq 2v - 4$ if planar for bipartite graphs (because $s_i \geq 4$)
  - nonplanar if and only if the graph contains K_5 or $K_{3,3}$
  - all planar graphs can be colored with ≤ 4 colors

Note 7 (Modular Artithmetic)

- x^{-1} (modular inverse) exists mod m if and only if $\gcd(x, m) = 1$
- Extended Euclidean Algorithm:

x	y	$\lfloor x/y \rfloor$	a	b	
35	12	2	$\boxed{-1}$	$\boxed{3}$	answer
12	11	1	1	-1	
11	1	11	$\boxed{0}$	$\boxed{1}$	
$\boxed{1}$	0				start
gcd					

  - new a = old b
  - new $b = a - b\left\lfloor \frac{x}{y} \right\rfloor$
  - if $\gcd(x, y) = 1$, then $a = x^{-1} \bmod y$, $b = y^{-1} \bmod x$
- Chinese Remainder Theorem:
  - find bases b_i that are $\equiv 1 \bmod m_i$ and $\equiv 0 \bmod m_j$ for $j \neq i$
    $\to$ $b_i = c_i(c_i^{-1} \bmod m_i)$ where $c_i = \prod_{i \neq j} m_j$
  - $x \equiv \sum a_i b_i \pmod{\prod m_i}$
  - solution is unique mod $\prod m_i$
  - m_i must be pairwise relatively prime in order to use CRT

Note 8 (RSA)

- Scheme: for primes p, q, find e coprime to $(p - 1)(q - 1)$
  - public key: $N = pq$ and e
  - private key: $d = e^{-1} \bmod (p - 1)(q - 1)$
  - encryption of message m: $m^e \pmod N = y$
  - decryption of encrypted message y: $y^d \pmod N = m$
- Fermat's Little Theorem (FLT): $x^p \equiv x \pmod p$, or $x^{p-1} \equiv 1 \pmod p$ if x coprime to p
- Prime Number Theorem: $\pi(n) \geq \frac{n}{\ln n}$ for $n \geq 17$, where $\pi(n)$ = # of primes $\leq n$
- Breaking RSA if we know d:
  - we know $de - 1 = k(p - 1)(q - 1)$, where $k \leq e$ because $d < (p - 1)(q - 1)$
  - so $\frac{de-1}{k} = pq - p - q + 1$; $pq = N$, so we can find p, q because we know d, e, k

Note 9 (Polynomials)

- Property 1: nonzero polynomial of degree d has at most d roots
- Property 2: $d + 1$ pairs of points (x_i distinct) uniquely defines a polynomial of degree at most d
- Lagrange Interpolation:
  - $\Delta_i(x) = \prod_{i \neq j} \frac{x - x_j}{x_i - x_j}$
  - $P(x) = \sum_i y_i \Delta_i(x)$
- Secret Sharing (normally under $GF(p)$):
  - $P(0)$ = secret, $P(1), \ldots, P(n)$ given to all people
  - $P(x)$ = polynomial of degree $k - 1$, where k people are needed to get the secret
- Rational Root Theorem: for $P(x) = a_n x^n + \cdots + a_0$, the roots of $P(x)$ that are of the form $\frac{p}{q}$ must have $p \mid a_0$, $q \mid a_n$

Note 10 (Error Correcting Codes)

- Erasure Errors: k packets lost, message length n; need to send $n + k$ packets because $P(x)$ of degree $n - 1$ needs n points to define it
- General Errors: k packets corrupted, message length n; send $n + 2k$ packets
- Berlekamp Welch:
  - $P(x)$ encodes message (degree $n - 1$)
  - $E(x)$ constructed so that roots are where the errors are (degree k); coefficients unknown
  - $Q(x) = P(x)E(x)$ (degree $n + k - 1$)
  - substitute all (x_i, r_i) into $Q(x_i) = r_i E(x_i)$, make system of equations
  - solve for coefficients; $P(x) = \frac{Q(x)}{E(x)}$

Note 11 (Counting)

- 1st rule of counting: multiply # of ways for each choice
- 2nd rule of counting: count ordered arrangements, divide by # of ways to order to get unordered
- $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ = # ways to select k from n
- Stars and bars: n objects, k groups $\to$ n stars, $k-1$ bars $\to \binom{n+k-1}{k-1} = \binom{n+k-1}{n}$
- Zeroth rule of counting: if bijection between A and B, then $|A| = |B|$
- Binomial theorem: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$
- Hockey-stick identity: $\binom{n}{k+1} = \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k}{k}$
- Derangements: $D_n = (n-1)(D_{n-1} + D_{n-2}) = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$
- Principle of Inclusion-Exclusion: $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$ More generally, alternate add/subtract all combinations
- Stirling's approximation: $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

Note 12 (Probability Theory)

- Sample points = outcomes
- Sample space = Ω = all possible outcomes
- Probability space: $(\Omega, \mathbb{P}(\omega))$; (sample space, probability function)
- $0 \le \mathbb{P}(\omega) \le 1, \forall \omega \in \Omega$; $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$
- Uniform probability: $\mathbb{P}(\omega) = \frac{1}{|\Omega|}, \forall \omega \in \Omega$
- $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$ where A is an event
- If uniform: $\mathbb{P}(A) = \frac{\#\text{ sample points in } A}{\#\text{ sample points in } \Omega} = \frac{|A|}{|\Omega|}$
- $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$

Note 13 (Conditional Probability)

- $\mathbb{P}(\omega \mid B) = \frac{\mathbb{P}(\omega)}{P(B)}$ for $\omega \in B$
- $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \to \mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\,\mathbb{P}(B)$
- Bayes' Rule:
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B \mid A)\,\mathbb{P}(A) + \mathbb{P}(B \mid \overline{A})\,\mathbb{P}(\overline{A})}.$$
- Total Probability Rule (denom of Bayes' Rule):
$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbb{P}(B \mid A_i)\,\mathbb{P}(A_i)$$
for A_i partitioning Ω
- Independence: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$ or $\mathbb{P}(A \mid B) = \mathbb{P}(A)$
- Union bound: $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} \mathbb{P}(A_i)$

Note 14 (Random Variables)

- Bernoulli distribution: used as an indicator RV
- Binomial distribution: $\mathbb{P}(X = i) = i$ successes in n trials, success probability p
  - If $X \sim \text{Bin}(n,p)$, $Y \sim \text{Bin}(m,p)$ independent, $X+Y \sim \text{Bin}(n+m,p)$
- Hypergeometric distribution: $\mathbb{P}(X = k) = k$ successes in N draws w/o replacement from size N population with B objects (as successes)
- Joint distribution: $\mathbb{P}(X = a, Y = b)$
- Marginal distribution: $\mathbb{P}(X = a) = \sum_{b \in B} \mathbb{P}(X = a, Y = b)$
- Independence: $\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\,\mathbb{P}(Y = b)$
- Expectation: $\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x\,\mathbb{P}(X = x)$
- LOTUS: $\mathbb{E}[g(X)] = \sum_{x \in \mathcal{X}} g(x)\,\mathbb{P}(X = x)$
- Linearity of expectation: $\mathbb{E}[aX + bY] = a\,\mathbb{E}[X] + b\,\mathbb{E}[Y]$
- X, Y independent: $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$

Note 15 (Variance/Covariance)

- Variance: $\text{Var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$
  - $\text{Var}(cX) = c^2\,\text{Var}(X)$, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X,Y)$
  - if indep: $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$
- Covariance: $\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$
- Correlation: $\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$, always in $[-1, 1]$
- Indep. implies uncorrelated (Cov = 0), but not other way around, ex.
$$X = \begin{cases} 1 & p = 0.5 \\ -1 & p = -0.5 \end{cases}, \quad Y = \begin{cases} 1 & X = -1, p = 0.5 \\ -1 & X = -1, p = 0.5 \\ 0 & X = 1 \end{cases}$$

Note 16 (Geometric/Poisson Distributions)

- Geometric distribution: $\mathbb{P}(X = i)$ = exactly i trials until success with probability p; use $X - 1$ for failures until success
  - Memoryless Property: $\mathbb{P}(X > a + b \mid X > a) = \mathbb{P}(X > b)$; i.e. waiting $> b$ units has same probability, no matter where we start
- Poisson distribution: λ = average # of successes in a unit of time
  - $X \sim \text{Pois}(\lambda)$, $Y \sim \text{Pois}(\mu)$ independent: $X + Y \sim \text{Pois}(\lambda + \mu)$
  - $X \sim \text{Bin}(n, \frac{\lambda}{n})$ where $\lambda > 0$ is constant, as $n \to \infty$, $X \to \text{Pois}(\lambda)$

Note 20 (Continuous Distributions)

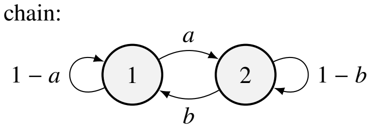
- Probability density function:
  - $f_X(x) \ge 0$ for $x \in \mathbb{R}$
  - $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$
- Cumulative density function: $F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$, $f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x)$
- Expectation: $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x$
- LOTUS: $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x$
- Joint distribution: $\mathbb{P}(a \le X \le b, c \le Y \le d)$
  - $f_{XY}(x,y) \ge 0, \forall x, y \in \mathbb{R}$
  - $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$
- Marginal distribution: $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}y$; integrate over all y
- Independence: $f_{XY}(x,y) = f_X(x) f_Y(y)$
- Conditional probability: $f_{X|A}(x) = \frac{f_X(x)}{\mathbb{P}(A)}$, $f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}$
- Exponential distribution: continuous analog to geometric distribution
  - Memoryless property: $\mathbb{P}(X > t + s \mid X > t) = \mathbb{P}(X > s)$
  - Additionally, $\mathbb{P}(X < Y \mid \min(X,Y) > t) = \mathbb{P}(X < Y)$
  - If $X \sim \text{Exp}(\lambda_X)$, $Y \sim \text{Exp}(\lambda_Y)$ independent, then $\min(X,Y) \sim \text{Exp}(\lambda_X + \lambda_Y)$ and $\mathbb{P}(X \le Y) = \frac{\lambda_X}{\lambda_X + \lambda_Y}$
- Normal distribution (Gaussian distribution)
  - If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ independent: $Z = aX + bY \sim \mathcal{N}(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$
- Central Limit Theorem: if $S_n = \sum_{i=1}^{n} X_i$, all X_i iid with mean μ, variance σ^2,
$$\frac{S_n}{\sigma} \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right); \quad \frac{S_n - n\mu}{\sigma\sqrt{n}} \approx \mathcal{N}(0,1).$$

Note 17 (Concentration Inequalities, LLN)

- Markov's Inequality: $\mathbb{P}(X \ge c) \le \frac{\mathbb{E}[X]}{c}$, if X nonnegative, $c > 0$
- Generalized Markov: $\mathbb{P}(|Y| \ge c) \le \frac{\mathbb{E}[|Y|^r]}{c^r}$ for $c, r > 0$
- Chebyshev's Inequality: $\mathbb{P}(|X - \mu| \ge c) \le \frac{\text{Var}(X)}{c^2}$ for $\mu = \mathbb{E}[X]$, $c > 0$
  - Corollary: $\mathbb{P}(|X - \mu| \ge k\sigma) \le \frac{1}{k^2}$ for $\sigma = \sqrt{\text{Var}(X)}$, $k > 0$
- Confidence intervals:
  - For proportions, $\mathbb{P}(|\hat{p} - p| \ge \varepsilon) \le \frac{\text{Var}(\hat{p})}{\varepsilon^2} \le \delta$, where δ is the confidence level (95% interval $\to \delta = 0.05$)
  - $\hat{p}$ = proportion of successes in n trials, $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$
  - $\implies n \ge \frac{1}{4\varepsilon^2\delta}$
  - For means, $\mathbb{P}\left(\left|\frac{1}{n}S_n - \mu\right| \ge \varepsilon\right) \le \frac{\sigma^2}{n\varepsilon^2} = \delta$
  - $S_n = \sum_{i=1}^{n} X_i$, all X_i's iid mean μ, variance σ^2
  - $\implies \varepsilon = \frac{\sigma}{\sqrt{n\delta}}$, interval $= S_n \pm \frac{\sigma}{\sqrt{n\delta}}$
  - With CLT, $\mathbb{P}(|A_n - \mu| \le \varepsilon) = \mathbb{P}\left(\left|\frac{(A_n - \mu)\sqrt{n}}{\sigma}\right| \le \frac{\varepsilon\sqrt{n}}{\sigma}\right) \approx 1 - 2\Phi\left(-\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1 - \delta.$ Here, $A_n = \frac{1}{n}S_n$ and CLT gives $A_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$; use inverse cdf to get ε
- Law of large numbers: as $n \to \infty$, sample average of iid $X_1, \ldots X_n$ tends to population mean

Note 24 (Markov Chains)

- Markov chain = sequence of states
- X_n = state at time step n
- $\mathcal{X}$ = state space (finite)
- π_n = distribution of states at time step n
- Markov property (memoryless): $\mathbb{P}(X_{n+1} = i \mid X_n, X_{n-1}, \dots, X_0) = \mathbb{P}(X_{n+1} = i \mid X_n)$
- Transition matrix ($\mathbf{P}$): transition probabilities, from row to column
- $\pi_n = \pi_0 \mathbf{P}^n$
- Invariant distribution: $\pi = \pi\mathbf{P}$; solve balance equations in addition to $\sum \pi(i) = 1$
- Irreducible Markov chain: any state can be reached from any other
- All irreducible Markov chains have a unique invariant distribution:
$$\pi(i) = \lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} \mathbf{1}\{X_m = i\}.$$
That is, average time spent at state $i = \pi(i)$
- Periodicity: let $d(i) = \gcd\{n > 0 : \mathbb{P}(X_n = i \mid X_0 = i) > 0\}$
  - If $d(i) > 1$, periodic with period d
  - If $d(i) = 1$, aperiodic
  - If irreducible, all states have the same $d(i)$
- Aperiodic irreducible Markov chains will always converge to the invariant distribution
- Hitting time: # of steps before first reaching state i
  - Let $\beta(j)$ denote this value, starting at state j; set up system of equations and solve
- $\mathbb{P}(A \text{ before } B)$: similarly, let $\alpha(j)$ denote this probability, starting at state j; set up system of equations and solve (use TPR)
- Similar problems: let $f(i)$ denote # of steps or probability, starting at state i; set up equations and solve
- 2-state Markov chain:



$\mathbf{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$ and $\pi = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$

Other

- In CS70, naturals and whole numbers both include 0
- Sum of finite geometric series: $\frac{1-r^n}{1-r}a$ where r is ratio, a is first term, n is number of terms
- Memoryless property independence: if X, Y both memoryless (i.e. geometric or exponential), then $\min(X,Y)$ and $\max(X,Y) - \min(X,Y)$ are independent
- If $S_n = \sum_{i=1}^{n} X_i$, then (useful for variance of indicators)
$$\mathbb{E}\left[S_n^2\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + \sum_{i\neq j} \mathbb{E}\left[X_i X_j\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + 2\sum_{i<j} \mathbb{E}\left[X_i X_j\right]$$
- Coupon Collector Problem:
  - n distinct items, each with equal probability; X_i = # of tries before ith new item, given $i-1$ already
  - $S_n = \sum X_i$ = total tries before getting all items
  - We have $X_i \sim \text{Geom}(\frac{n-i+1}{n})$ because $i-1$ old items, so $n-i+1$ new items
  - Hence,
$$\mathbb{E}[S_n] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{n}{n-i+1} = n\sum_{i=1}^{n} \frac{1}{i} \approx n(\ln n + 0.5772)$$

Note: Do not blindly copy down the content on this cheatsheet (this is one of the reasons why I typed this in LaTeX, as you cannot just print it out and use it). My goal in making my cheatsheet publicly available is to ease your studying and to provide a general guideline for the kinds of things that you could put on your cheatsheet. The most helpful part of a cheatsheet is the process of making it—you synthesize and review the course material yourself and pick out what *you* think is important.

I guarantee that there are concepts in the notes/lectures/discussions not on my cheatsheet that you should know, and I also guarantee that there are things on my cheatsheet that may be of no use to you. This is a collection of items that helped *me* when I took the class, and only you know what benefits you the most. My advice is to make your own cheatsheet first, without referencing mine, and then going over my cheatsheet afterward to add things that you may have missed.

Good luck on your exams!

Table 1: Common Discrete Distributions

Distribution	Parameters	PMF ($\mathbb{P}(X = k)$)	CMF ($\mathbb{P}(X \leq k)$)	Expectation ($\mathbb{E}[X]$)	Variance ($\mathrm{Var}(X)$)	Support
Uniform	$\mathrm{Uniform}(a, b)$	$\frac{1}{b - a + 1}$	$\frac{k - a + 1}{b - a + 1}$	$\frac{a + b}{2}$	$\frac{(b - a + 1)^2 - 1}{12}$	$X \in [a, b]$
Bernoulli	$\mathrm{Bernoulli}(p)$	$\begin{cases} 1 & p \\ 0 & 1 - p \end{cases}$	—	p	$p(1 - p)$	$X \in \{0, 1\}$
Binomial	$\mathrm{Bin}(n, p)$	$\binom{n}{k} p^k (1 - p)^{n-k}$	—	np	$np(1 - p)$	$X \in \mathbb{N}$
Geometric	$\mathrm{Geom}(p)$	$p(1 - p)^{k-1}$	$1 - (1 - p)^k$	$\frac{1}{p}$	$\frac{1 - p}{p^2}$	$X \in \mathbb{N}$
Poisson	$\mathrm{Pois}(\lambda)$	$\frac{\lambda^k e^{-\lambda}}{k!}$	—	λ	λ	$X \in \mathbb{N}$
Hypergeometric	$\mathrm{Hypergeometric}(N, K, n)$	$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$	—	$n\frac{K}{N}$	$n\frac{K(N - K)(N - n)}{N^2(N - 1)}$	$X \in \mathbb{N}$

Table 2: Common Continuous Distributions

Distribution	Parameters	PDF ($f_X(x)$)	CDF ($F_X(x) = \mathbb{P}(X \leq x)$)	Expectation ($\mathbb{E}[X]$)	Variance ($\mathrm{Var}(X)$)	Support
Uniform	$\mathrm{Uniform}(a, b)$	$\frac{1}{b - a}$	$\frac{x - a}{b - a}$	$\frac{a + b}{2}$	$\frac{(b - a)^2}{12}$	$X \in [a, b]$
Exponential	$\mathrm{Exp}(\lambda)$	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$X \in [0, \infty)$
Normal/Gaussian	$\mathcal{N}(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$	$\Phi(x)$	μ	σ^2	$X \in \mathbb{R}$