Axioms of Probability. *Axioms for probability:*

1. $P(E) \geq 0$ *for any event* E
2. $P(S) = 1$
3. *If* $E_1, E_2, E_3, \ldots$ *are disjoint events, then* $P(E_1 \cup E_2 \cup E_3 \cup \ldots) = \sum_{i=1}^{\infty} P(E_i)$

Thoerem (Basic theorems of probability)**.** *Let* A *and* B *be events.*

1. $P(\emptyset) = 0$
2. *If* $A \subseteq B$*, then* $P(A) \leq P(B)$
3. $P(A) = 1 - P(A^C)$
4. $P(A) = P(A \cap B) + P(A \cap B^C)$
5. $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$

Method. The number of ways to select k elements from an n-element set is...

	Order matters	Order doesn't matter
With replacement	n^k	$\binom{n+k-1}{k}$
Without replacement	$\frac{n!}{(n-k)!}$	$\binom{n}{k} = \frac{n!}{(n-k)!k!}$

Definition. Let A and B be events with $P(B) \neq 0$. The **conditional probability of** A **given** B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Definition. Events A and B are **independent** if and only if $P(A \cap B) = P(A)P(B)$.

Thoerem (Multiplication rule for probabilities)**.** *If* $P(B) \neq 0$*, then* $P(A \cap B) = P(A|B)P(B)$*.*

Thoerem (The Law of Total Probability)**.** *If event* B *has probability strictly between* 0 *and* 1*, then for any event* A*,* $P(A) = P(A|B)P(B) + P(A|B^C)P(B^C)$*.*

Thoerem (Bayes' Law)**.** *If* A *and* B *are events with positive probability, then* $P(B|A) = \dfrac{P(A|B)P(B)}{P(A)}$

Definition. A **random variable** X assigns a number to each outcome in the sample space S.

1. All random variables have a **cumulative distribution function (CDF)**: $F(x) = P(X \leq x)$.
2. A discrete random variable has a **probability mass function (PMF)**: $m(x) = P(X = x)$.

Definition. The total number of successes in n independent, identically distributed (iid) Bernoulli trials with parameter p is a random variable with a **binomial distribution**. The PMF of a random variable X having a binomial distribution with parameters n and p is $\boxed{b(x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \ldots, n}$

Proposition. *The mean of a binomial distribution is* $\mu = np$*.*

Definition. Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed (iid) Bernoulli trials, all with probability of success p. Let N be the trial on which the first success occurs. The random variable N is said to have a **geometric distribution** with parameter p and its PMF is $\boxed{g(n) = p(1-p)^{n-1} \text{ for } n = 1, 2, 3, \ldots}$

Proposition. *The mean of a geometric distribution is* $\mu = \dfrac{1}{p}$*.*

Definition. Suppose n elements are to be selected without replacement from a population of size N of which k are successes. The number of successes selected is a **hypergeometric** random variable and its PMF is

$\boxed{h(x) = \dfrac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}}$ for those integers x which make all terms positive.

Proposition. *The mean of a hypergeometric distribution is* $\mu = \dfrac{nk}{N}$*.*