The next result is best understood through Pascal's Triangle.

Theorem 1.3.4 *For $n \in \mathbb{N}$ and $r = 0, 1, \ldots, n-1$*

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}.$$

Pascal's triangle:

Proof 1.3.5 (proof: $\left(\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}\right)$) *Consider,*

$$(1+x)^n = (1+x)(1+x)^{n-1} = (1+x)^{n-1} + x(1+x)^{n-1}.$$

Both sides are polynomials in x, and two polynomials are equal if and only they have the same coefficients, so we may equate coefficients of x^r for any $r = 0, 1, \ldots, n$.

On the left, the coefficient of x^r is $\binom{n}{r}$ (by binomial theorem).

On the right, the coefficient of x^r in $(1+x)^{n-1}$ is $\binom{n-1}{r}$, and in $x(1+x)^{n-1}$, the coefficient on x^r is $\binom{n-1}{r-1}$.

Thus $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

Theorem 1.3.6 ***(Multinomial Coefficients)*** *Let* $r_1 + r_2 + \cdots + r_k = n$*. The coefficient of* $x_1^{r_1} x_2^{r_2} \cdots x_k^{r_k}$ *in* $(x_1 + x_2 + \cdots + x_k)^n$ *(such coefficients are called* ***multinomial coefficients****) is*

$$\binom{n}{r_1, r_2, \ldots, r_k} = \frac{n!}{r_1! \cdot r_2! \cdot \ldots \cdot r_k!}.$$

Example 1.3.7 *What is the coefficient of* $x_1^3 x_3^4 x_4^2$ *in the expansion of* $(x_1 + x_2 + x_3 + x_4 + x_5)^9$*?*

Proof 1.3.8 *Hint: Similar to permutations with repeated elements.*

Chapter 2

Probability

2.1 Probability Concepts and Rules

Mathematics is used to model real world phenomena.

Deterministic model (ideal situation): Predicts the outcome of an experiment with certainty based on given initial conditions. e.g. velocity of a falling object

$$v = gt.$$

Probabilistic, or stochastic, model (randomness): When the same initial conditions can lead to a variety of outcomes, these models provide a value (probability) to the possible outcomes. e.g. rolling a die results in one of six numbers facing up.

Assign each outcome the value $\frac{1}{6}$.

2.1.1 Classical Probability Concept:

When there are N possible (equally likely) outcomes of which k are considered successful, then the *probability* of a success is the ratio $\frac{k}{N}$.

Example 2.1.1 *Probability of...*

- *tossing tails with a balanced coin:*

- *drawing an ace from deck of cards:*

- *rolling either 3 or 5 with a (fair, six-sided) die:*

- *rolling a total of 5 with a (fair) pair of (six-sided) dice:*

These values describe the frequency of the successful outcome; the proportion of time the event occurs in the **long run**.

LONG RUN

LONG RUUUUUUUNN

2.1.2 Sample Spaces

The set of all possible outcomes of an experiment is called the **sample space**.

Example:

Experiment	Sample space
single coin toss	$\{H, T\}$
roll of two dice	$\{(d_1, d_2) \mid d_1, d_2 \in \{1, 2, 3, 4, 5, 6\}\}$
sum of two dice	
drawing a card	all (n, s) with $n \in \{2, \ldots, 10, J, Q, K, A\}$ and $s \in \{\spadesuit, \heartsuit, \diamondsuit, \clubsuit\}$
three coin tosses	

All previous examples were *finite* sample spaces. The following experiments have infinite sample spaces.

Example 2.1.2 • *Tossing a coin until heads is reached:*

$$\{H, TH, TTH, TTTH, TTTH, \ldots\}$$

- *Playtime for two AA alkaline batteries in a Wii remote:*

$$\{t \text{ hours} \mid t \in [0, 50]\}$$

Continuous and Discrete Sample Spaces:

There is an important distinction between the sample spaces in the previous example; the outcomes of the first example (coin toss) may be listed, whereas the outcomes in the other two belong to a continuum of values.

A **discrete sample space** has only finitely many, or a countably infinite number of elements.

A **continuous sample space** is an interval in $\mathbb{R}$, or a product of intervals lying in $\mathbb{R}^n$.

The important distinction is how probabilities are assigned.

Events: While individual elements of a sample space are called **outcomes**, subsets of a sample space are called **events**. If the outcome of an experiment lies in an event, we say that event has occurred.

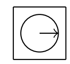
Example 2.1.3 *Experiment: Tossing a coin three times.*

Sample space: $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

Event A: Getting at least two heads: $\{HHH, HHT, HTH, THH\}$

Event B: Getting exactly two tails

Event C: Getting two consecutive heads

Example 2.1.4 *Experiment: Spinning a probability spinner.* 

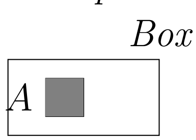
Sample space: $\{\theta \text{ degrees} | \theta \in [0, 360)\}$

Event A: Landing between 90 and 180 degrees, $[90, 180]$

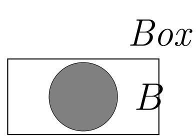
Event B: Landing either between 45 and 90 degrees or between 270 and 315 degrees, $[45, 90] \cup [270, 315]$

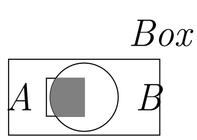
Event C: Landing precisely on 180 degrees.

Example 2.1.5 *Experiment: Dropping a pencil head first into a rectangular box.*
Sample space: All points on the bottom of the box.

Event A: 

(pencil lands in shaded region)

Event B: 

Event $A \cap B$*:* 

2.1.3 Union, Intersection, Complement

Let A and B be events in sample space S; i.e. A and B are subsets of a set S.

The **union** of A and B is the set of outcomes that are in either A or B or both.

$$A \cup B = \{x \in S | x \in A \text{ or } x \in B\}.$$

The **intersection** of A and B is the set of outcomes that are in both A and B.

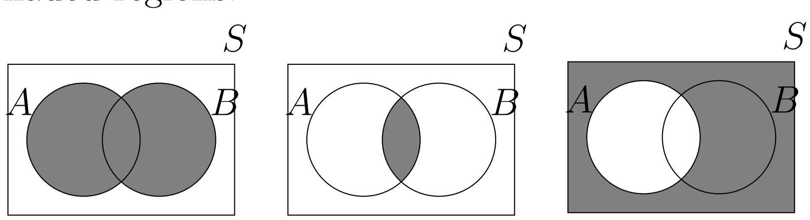
$$A \cap B = \{x \in S | x \in A \text{ and } x \in B\}.$$

The **complement** of A in S is the set of outcomes in S that are not in A.

$$A' = \{x \in S | x \notin A\} = S \setminus A.$$

Venn Diagrams:

A **Venn diagram** is a visual depiction of subsets of some "universal" set. Subsets are represented (usually by disks) lying within a rectangle representing the universal set. Sets of interest are represented by shaded regions.



$A \cup B$ $\qquad$ $A \cap B$ $\qquad$ A'

In the pencil dropping example, Venn diagrams gave a literal representation of the events in that experiment, but these representations can be used in more general situations to help visualize relationships between different subsets of a sample space.

Mutually Exclusive Events

A set which has no elements is called the **empty set**, denoted $\emptyset$.

Example 2.1.6 *Experiment: Rolling two dice.*

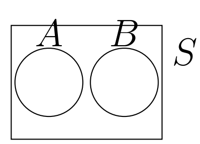
Event A: Rolling at least one six.

$$A = \{(d_1, 6), (6, d_2) | d_1, d_2 \in \{1, 2, 3, 4, 5, 6\}\}$$

Event B: Sum of dice equals 4.

Event C: Rolling at least one six and having a sum of 4.

$$C = A \cap B = \emptyset.$$



$A \cap B = \emptyset$

Sets with empty intersection are called disjoint, and the events in this case are called **mutually exclusive**.

2.1.4 Algebra of Sets

Let A, B and C be subsets of a universal set S.

- Idempotent laws:

  $$A \cup A = A, \qquad A \cap A = A$$

- Associative laws:

  $$(A \cup B) \cup C = A \cup (B \cup C), \qquad (A \cap B) \cap C = A \cap (B \cap C)$$

- Commutative laws:

  $$A \cup B = B \cup A, \qquad A \cap B = B \cap A$$

- Distributive laws:

  $$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

- Identity laws:

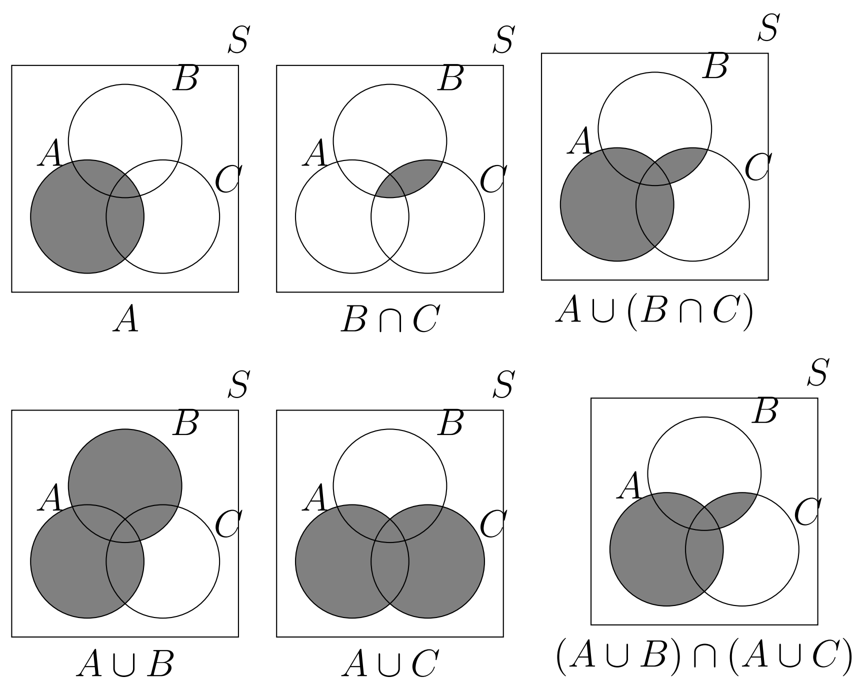
  $$A \cup \emptyset = A, \quad A \cup S = S, \quad A \cap S = A, \quad A \cap \emptyset = \emptyset$$

- Complement laws:
  $(A')' = A, \quad A \cup A' = S, \quad A \cap A' = \emptyset, \quad S' = \emptyset, \quad \emptyset' = S$

- DeMorgan's Laws:
  $(A \cup B)' = A' \cap B', \quad (A \cap B)' = A' \cup B'$

Set Operations:

Example 2.1.7 *Use Venn diagrams to verify the distributive law* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.



Proof: (of $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$)
[How does one even "prove" such a statement mathematically?]

Exercise: Use Venn diagrams to verify DeMorgan's Laws.

$$(A \cup B)' = A' \cap B', \quad (A \cap B)' = A' \cup B'$$

The Probability of an Event:

A **probability**, or **probability measure**, is a function P which maps events in the sample space S to real numbers.

In order to assign probabilities in a meaningful way, P must satisfy the following called the **postulates** (or **axioms**) **of probability**.

P1: The probability of any event A in S is a non-negative real number, i.e. $P(A) \geq 0$.

P2: $P(S) = 1$.

P3: If $A_1, A_2, A_3, \ldots,$ is a finite or infinite sequence of (pairwise) mutually exclusive events in S then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

(P is **countably additive**)

2.1.5 Postulates of Probability

- Interpreting a probability as a frequency, or a proportion of time, it makes sense that $P(A) \geq 0$; in fact we will show that $0 \leq P(A) \leq 1$ for any event A.

- P2 says that the probability that outcome of the experiment lies in S must be assigned value 1. Since this is certain to happen, we interpret $P(A) = 1$ as "A happens 100 percent of the time."

- P3 is for consistency. For example, if events A_1 and A_2 share no common outcomes, then the probability that either event occurs, $P(A_1 \cup A_2)$, is the sum of their individual probabilities.

A technical detail has been overlooked in the postulates of probability presented above. In a discrete sample space S, an "event" can be any subset of S, however in the continuous case one has to be more careful about which subsets of S are allowed as events. A precise definition for these allowable events comes in a course on **measure theory**. In this course we won't require that level of detail; i.e. the subsets we assign probabilities to will be allowable events.

Single Die Roll:
Let S be the sample space for rolling a die once.

Example 2.1.8 *Each outcome in S is its own event, call these $A_1, \dots, A_6$.*

Events $A_1, \dots, A_6$ are mutually exclusive, and any event E in S is a union of these, for example let $E = A_2 \cup A_4 \cup A_5$.

By the classical probability concept, $P(E) = \frac{3}{6}$ (successes/number of outcomes), and $P(A_i) = \frac{1}{6}$ for each i.

It follows that this satisfies the postulates of probability:

- $P(B) \geq 0$ *for any* $B \subset S$.
- $P(S) = \frac{6}{6} = 1$.
- *P3 is satisfied: for example* $P(E) = \frac{3}{6} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = P(A_2) + P(A_4) + P(A_5)$.

Example 2.1.9 *Suppose we assigned probabilities in this experiment in a different way. Using the same notation as before say for any event* B *we specify that*

$$P(B) = \sum_{A_i \in B} P(A_i), \text{ and}$$

$$P(A_1) = \tfrac{1}{2}, P(A_2) = \tfrac{1}{4}, P(A_3) = \tfrac{1}{8},$$
$$P(A_4) = 0, P(A_5) = \tfrac{1}{16}, P(A_6) = \tfrac{1}{16}$$

Are the postulates of probability still satisfied?

Example 2.1.10 *An experiment has four possible outcomes* A, B, C, D *that are mutually exclusive. Explain why the following assignments of probabilities are not permissible.*

(a) $P(A) = 0.12, P(B) = 0.63, P(C) = 0.45, P(D) = -0.20$

(b) $P(A) = \frac{9}{120}, P(B) = \frac{45}{120}, P(C) = \frac{27}{120}, P(D) = \frac{46}{120}$