Basic Probability Review

Preliminaries: Set Theory

Definitions:

(i) a **set** is a collection of objects.

(ii) An object contained in a set is called an **element** of that set.

(iii) A **subset** of a given set is a collection of some, but not necessarily all, elements of the given set.

Notation: sets are often denoted by capital letters; elements by lower-case letters.

$s \in S$ indicates that s is an element of set S.
$x \notin S$ indicates that x is not an element of set S.
$A \subset S$ indicates that A is a subset of S.

A set is specified by specifying all elements contained in that set. Standard notation is to specify the elements between curly brackets, such as:
$S = \{1, 2, 3, 4, 5, 6\}$

Note that sets are unordered and specifying all elements of a set uniquely defines that set. For example, $S = \{1, 2, 3, 4, 5, 6\} = \{2, 4, 6, 1, 3, 5\} = \{1, 1, 2, 2, 2, 3, 4, 5, 6, 5, 6\}$

Let $A = \{s \in S : \text{ s is prime }\}$. The ":" is read as "such that". Thus, A is a set that contains $s \in S$ such that s is also prime. In this case, $A = \{2, 3, 5\}$

Infinite sets (sets containing infinitely many elements) are important in mathematics and probability. Consider, for example, the set of natural numbers (i.e., the counting numbers $1, 2, 3, 4, 5, \ldots$),

$$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\}$$

Examples of sets and subsets:

(a) $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ (the set of integers)

(b) $\mathbb{Q} = \{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{Z}, n \neq 0\}$ (the set of rational numbers)

(c) $\mathbb{R}$ = the real numbers

Note that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$

(d) $E = \{n \in \mathbb{N} : \text{ n is even }\}$. Note that $E \subset \mathbb{N}$

(e) Let S be the set of outcomes of an experiment in which a six-sided die is rolled once.

(f) Let W be the set of all people with Systemic Lupus Erythematosus.

(g) let V be a subset of 30 randomly selected individuals from W, above.

Set Operations

Let S be a set, and let $A \subset S$ and $B \subset S$.

Definitions:

(i) $A \cup B := \{x : x \in A \text{ or } x \in B\}$ is the **union** of A and B (note that $A \cup B$ includes any element that is both in A and in B).

(ii) $A \cap B := \{x : x \in A \text{ and } x \in B\}$ is the **intersection** of A and B.

(iii) The **complement** A with respect to S, denoted A^c or $\bar{A}$, is the set $\{x \in S : x \notin A\}$.

(iv) The **empty set**, denoted $\emptyset$, is the set containing no elements. That is, $\emptyset = \{\}$.

A **Venn Diagram** is a graphical way to represent the relationships between sets.

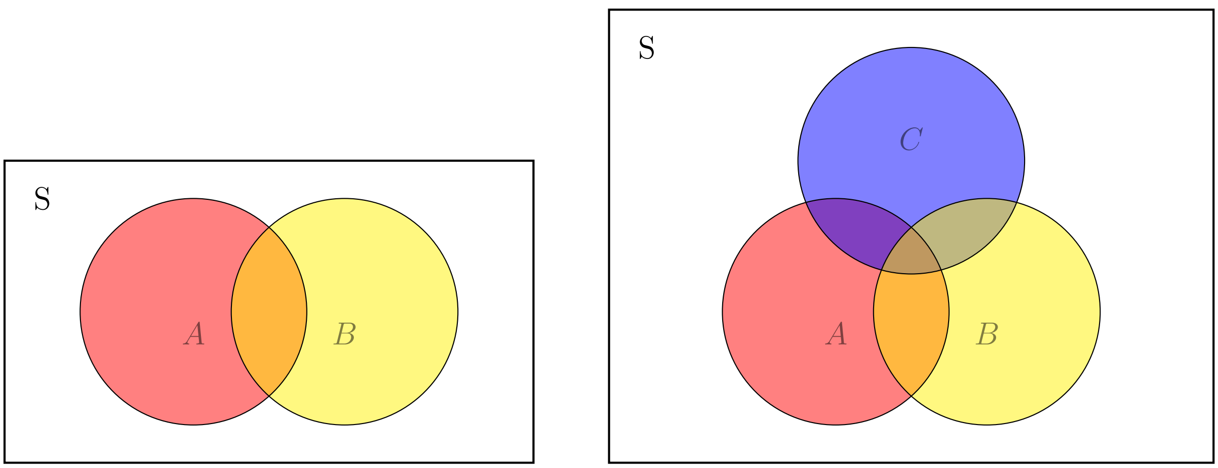

Figure 1: Venn diagrams representing sets $A, B, C \subset S$.

Exercises:

1. Draw a Venn Diagram with sets A and B and shade the following sets:

   (a) $A \cup B$

   (b) $A \cap B$

   (c) A^c

   (d) B^c

   (e) $(A \cup B)^c$

(f) $(A \cap B)^c$

(g) $A \cap B^c$ (also known as the set difference of A and B, denoted $A - B$ or $A \backslash B$)

(h) $A^c \cup B^c$

(i) $A^c \cap B^c$

(j) $\emptyset^c$

2. Draw a Venn Diagram with sets A, B, and C and shade the following sets:

   (a) $A \cup B$

   (b) $C \cap B$

   (c) A^c

   (d) C^c

   (e) $(A \cup B \cup C)^c$

   (f) $(A \cap B \cap C)^c$

   (g) $A^c \cup B^c \cup C^c$

   (h) $A^c \cap B^c \cap C^c$

   (i) $(A \cap B^c) \cup C^c$

3. Use a Venn Diagram to determine which of the following are true and which are false

   (a) $A \cap B \subset A \cup B$

   (b) $A \cap B^c \subset A$

   (c) $A \cap B^c \subset B$

   (d) $(A \cup B)^c = A^c \cup B^c$ (two sets are equal if they contain the same elements, or equivalently, each set is a subset of the other set)

   (e) $(A \cup B)^c = A^c \cap B^c$

   (f) $A \cap B \cap C \subset A \cup B \cup C$

Introduction to Probability

Consider an experiment to be performed.

Definitions:

(i) the **sample space** is the set of all possible outcomes of the experiment.

(ii) An **event** is a subset of the sample space.

We will typically use S to denote the sample space.

"Definition" of Probability: The **probability** of an event is a number between 0 and 1, inclusive, that represents the "likelihood" of the occurrence of that event. The probability of event A is denoted as $P(A)$.

The precise definition of probability relies on Kolmogorov's axioms, from which it follows that $P(S) = 1$ and $P(\emptyset) = 0$. In general, an event with probability 1 is certain to occur, and an event with probability 0 cannot occur. When S contains infinitely many elements, there are generally many events with probability 1 that are not equal to S and many events with probability 0 that are not empty.

To enhance intuition further, one may multiply probabilities by 100, thus converting them to percentages. However, when thinking of probabilities as percentages, many probability formulas do not apply.

When all outcomes in S are equally likely,

$$P(A) = \frac{|A|}{|S|} = \frac{\text{number of elements (outcomes) in } A}{\text{number of elements (outcomes) in } S}$$

Exercises:

1. What is the probability of choosing a red marble from a bag containing 5 blue marbles, 6 green marbles, and 8 red marbles?
2. What is the probability of choosing a red marble from a bag containing 5 blue marbles, 6 green marbles, and 8 yellow marbles?
3. What is the probability of choosing a red or a blue marble from a bag containing 5 blue marbles, 6 green marbles, and 8 red marbles?
4. What is the probability of choosing a red or a yellow marble from a bag containing 5 blue marbles, 6 green marbles, and 8 red marbles?
5. What is the probability of choosing a marble that is not red from a bag containing 5 blue marbles, 6 green marbles, and 8 red marbles?

Useful Formulas:

(i) $P(A) = 1 - P(A^c)$

(ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iii) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Exercises: Suppose that a fair 20-sided die is rolled

1. What is the probability that an even number is rolled?
2. What is the probability that a multiple of 5 is rolled?
3. What is the probability that a multiple of 5 is not rolled?
4. What is the probability that an even number or a multiple of 5 is rolled?
5. What is the probability that an even number or an odd number is rolled?

Definition: Let A and B be events. If $A \cap B = \emptyset$, A and B are said to be **disjoint** or **mutually exclusive**.

If A and B are disjoint, $P(A \cup B) = P(A) + P(B)$.

Exercises: Suppose that a fair 20-sided die is rolled. Let A be the event that an even number is rolled, B be the event that a multiple of 6 is rolled, C be the event that a multiple of 7 is rolled.

1. Which pairs of events defined above are disjoint?
2. Which pairs of events defined above are not disjoint?
3. Find $P(A \cup B)$, $P(A \cup C)$, and $P(B \cup C)$.

Conditional Probability and Independence

Definition: Let A and B be events with $P(B) > 0$. The **conditional probability** of A given B is denoted $P(A|B)$ and defined as

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

Multiplication Formula:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Note: in general, $P(A|B) \neq P(B|A)$.

Exercises:

1. Suppose that two fair 4-sided dice are rolled.
   (a) What is the probability that the sum on the faces is 7?
   (b) Given that the first roll results in a 4, what is the probability that the sum on the faces is 7?
   (c) Given that the first roll results in a 2, what is the probability that the sum on the faces is 7?

2. A factory has an old and a new machine. The old machine produces 40% of the products, and the new machine produces 60% of the products. There is a 3% chance that a product produced by the old machine is defective and a 1% chance that a product produced by the new machine is defective.

   (a) What is the probability that a randomly selected product is produced by the old machine and is defective?

   (b) What is the probability that a randomly selected product is produced by the new machine and is defective?

   (c) What is the probability that a randomly selected product produced by the company is defective?

   (d) Given that a randomly selected product produced by the company is defective, what is the probability that it was produced by the old machine? The new machine?

Definition: Events A and B are **independent** if

$$P(A \cap B) = P(A)P(B).$$

Note: this definition implies that if A and B are independent and $P(A), P(B) > 0$, $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

Exercises:

1. Suppose that two fair 6-sided dice are rolled.

   (a) What is the probability that both faces show an even number?

   (b) What is the probability that both faces show a prime number?

   (c) What is the probability that both faces show a multiple of 3?

   (d) What is the probability that at least one of the faces shows a multiple of 3?

2. Suppose that three fair 6-sided dice are rolled.

   (a) What is the probability that all three faces show an even number?

   (b) What is the probability that all three faces show a prime number?

   (c) What is the probability that all three faces show a multiple of 3?

   (d) What is the probability that at least one of the faces shows a multiple of 3?

   (e) What is the probability that the sum of the faces is 3?

   (f) Provide several examples of independent events associated with this experiment.

   (g) How does the multiplication formula for three or more independent events generalize (note that what it means for three or more events to be independent has not been formally defined here)?

Counting Techniques

Recall that when all outcomes in the sample space are equally likely, $P(A) = \frac{|A|}{|S|}$. Counting the number of elements (outcomes) in a given set is not always simple. This section contains some counting rules.

Exercise:

1. Provide an example of an experiment and an associated sample space in which all outcomes are not equally likely.

The Multiplication Rule:

Suppose that events $A_1, A_2, \ldots, A_k$ occur in succession and that there are n_i ways that A_i can occur, $i = 1, \ldots, k$. Then the total number of different ways that $A_1, A_2, \ldots, A_k$ occur in succession is $n_1 \times n_2 \times \ldots \times n_k$.

Exercise:

1. A menu lists three kinds of soups and two main courses. If a meal consists of a soup and a main course, how many different meals are possible?

Note: when the number of choices is small, a *tree diagram* may be helpful.

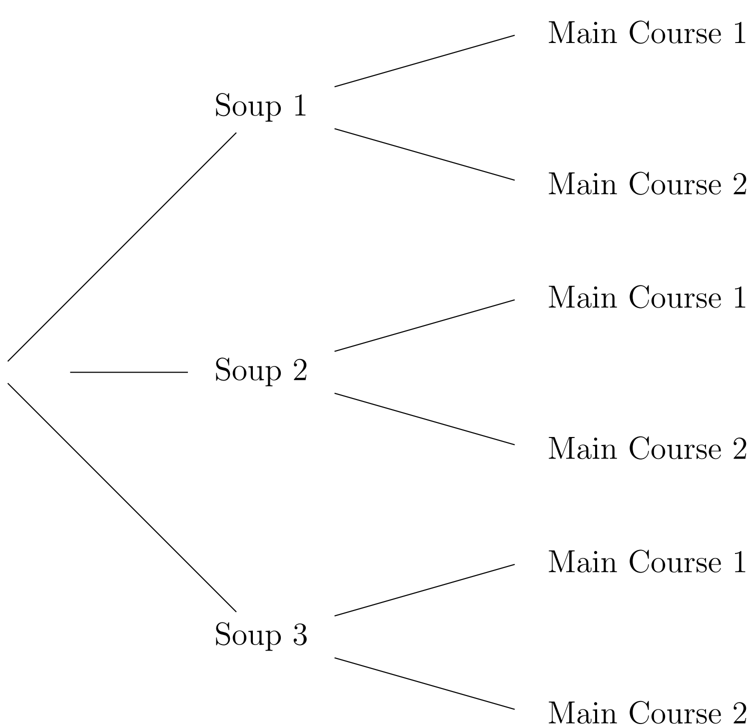

Figure 2: A tree diagram to represent the different menu choices.

Exercises:

1. A menu lists three kinds of soups, four main courses, and two desserts. If a meal consists of a soup, a main course, and a dessert, how many different meals are possible?

2. How many different answer keys are possible for a five-question true/false quiz.

3. If a student guesses each answer with a 1/2 probability of guessing correctly, what is the probability that the student will answer each question of a five-question true/false quiz correctly?

4. An exam contains 10 multiple-choice questions with four choices for each question. How many different answer keys are possible?

5. An exam contains 5 multiple-choice questions with four choices for each question and five true/false questions. How many different answer keys are possible?

6. An exam contains 5 multiple-choice questions with four choices for each question and five true/false questions. What is the probability that a student who guesses the answer the each question (equally likely to guess any available answer) does not guess any answer correctly?

Definition: A **permutation** is an ordered arrangement of objects. The number of permutations of k different objects, selected from n ($n \geq k$) different objects is

$$_nP_k = \frac{n!}{(n-k)!},$$

where "n factorial", $n! = n \times (n-1) \times \ldots \times 1$ for $n \in \mathbb{N}$, and $0! = 1$ by definition.

Exercises:

1. How many different orderings are there of a math book, a chemistry book, a physics book, and a biology book?

2. Ten runners compete in a race. How many different ways are there to assign first, second, and third place?

3. There are eight students in a class. How many different ways are there to select a President, Vice President, and Secretary?

4. There are eight students in a class. Suppose you are in the class. Suppose that a President, Vice President, and Secretary are selected. What is the probability that you will be selected for one of the positions?

5. There are eight students in a class. Suppose you are in the class. Suppose that a President, Vice President, and Secretary are selected. What is the probability that you will be selected for President?

6. How many different rearrangements are there of the letters in NYSTCE?

Definition: A **combination** is a particular selection of k out of n $(n \geq k)$ different objects for which order does not matter.The number of different combinations of k from n $(n \geq k)$ different objects is

$$_nC_k = \binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Exercises:

1. There are eight students in a class. How many different ways are there to select three students to serve on a committee?
2. There are eight students in a class. Suppose you are in the class. What is the probability that you will be selected for the committee?
3. There are eight students in a class. Suppose you are in the class. What is the probability that you will not be selected for the committee?
4. There are eight students in a class. Suppose that five are male and three are female. A committee of three is selected at random. What is the probability that the committee contains all females?
5. There are eight students in a class. Suppose that five are male and three are female. A committee of three is selected at random. What is the probability that the committee contains all males?
6. There are eight students in a class. Suppose that five are male and three are female. A committee of three is selected at random. What is the probability that the committee contains one female and two males?
7. An experiment is repeated five times. Each outcome is success or failure and the outcomes are independent. Suppose that each success probability is 1/3. What is the probability that all five outcomes are successes?
8. An experiment is repeated five times. Each outcome is success or failure and the outcomes are independent. Suppose that each success probability is 1/3. What is the probability that exactly two of the five outcomes are successes?
9. An experiment is repeated five times. Each outcome is success or failure and the outcomes are independent. Suppose that each success probability is 1/3. What is the probability that no more than two of the five outcomes are successes?

Binomial Probabilities

The same experiment is repeated n times, where $n \in \mathbb{N}$, and each repetition of the experiment results in a success or failure. Each repetition of the experiment is referred to as a *trial*.

Suppose trials are independent and the success probability of each trial is p, where $0 \leq p \leq 1$. Thus, the failure probability of each trial is $1 - p := q$. Let

$$X = \text{ the number of successes in } n \text{ trials}$$

Definition: X described above is called a **binomial random variable**, and $P(X = k)$ for $k = 0, 1, \ldots, n$ are the *binomial probabilities*:

$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

Definition: If $n = 1$, X described above is called a **Bernoulli random variable**, and

$$P(X = 1) = p \qquad \text{and } P(X = 0) = 1 - p = q$$

Exercises:

1. Five fair 6-sided dice are rolled.
   (a) What is the probability that all five faces show a 1?
   (b) What is the probability that exactly two of the faces show a 1?
   (c) What is the probability that at most two of the faces show a 1?
   (d) What is the probability that all five faces show a multiple of three?
   (e) What is the probability that exactly two of the faces show a multiple of three?
   (f) What is the probability that at most two of the faces show a multiple of three?
2. You take an eight-question multiple choice quiz with four choices for each question and one correct choice per question. Suppose you guess the answer to each question with equal probability of picking any of the four choices.
   (a) What is the probability that you guess the answers to all of the questions correctly?
   (b) What is the probability that you guess the answers to exactly three of the questions correctly?
   (c) What is the probability that you guess the answers to at most three of the questions correctly?
   (d) What is the probability that you guess the answers to more than three questions correctly?

Pascal's Triangle Consider the triangle below, in which each row begins and ends with 1 and the other numbers are obtained by adding the numbers above it to the left and right. This is called *Pascal's Triangle*.

$$
\begin{array}{ccccccccccccccccc}
 & & & & & & & & 1 & & & & & & & & \\
 & & & & & & & 1 & & 1 & & & & & & & \\
 & & & & & & 1 & & 2 & & 1 & & & & & & \\
 & & & & & 1 & & 3 & & 3 & & 1 & & & & & \\
 & & & & 1 & & 4 & & 6 & & 4 & & 1 & & & & \\
 & & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & \\
 & & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & & \\
 & 1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1 & \\
1 & & 8 & & 28 & & 56 & & 70 & & 56 & & 28 & & 8 & & 1
\end{array}
$$

Compare with

$$
\begin{array}{ccccccccccccccccc}
 & & & & & & & & \binom{0}{0} & & & & & & & & \\
 & & & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & & & \\
 & & & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & & & \\
 & & & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & & & \\
 & & & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & & & \\
 & & & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & & & \\
 & & \binom{6}{0} & & \binom{6}{1} & & \binom{6}{2} & & \binom{6}{3} & & \binom{6}{4} & & \binom{6}{5} & & \binom{6}{6} & & \\
 & \binom{7}{0} & & \binom{7}{1} & & \binom{7}{2} & & \binom{7}{3} & & \binom{7}{4} & & \binom{7}{5} & & \binom{7}{6} & & \binom{7}{7} & \\
\binom{8}{0} & & \binom{8}{1} & & \binom{8}{2} & & \binom{8}{3} & & \binom{8}{4} & & \binom{8}{5} & & \binom{8}{6} & & \binom{8}{7} & & \binom{8}{8}
\end{array}
$$

Exercise:

1. Show that $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, $n \in \mathbb{N}$, and $k = 1, \ldots, n$.

Theorem 0.1 (The Binomial Theorem).

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Exercises:

1. Find the coefficient of x^9y^{2007} in the expansion of $(x+y)^{2016}$.
2. Find the coefficient of x^9 in the expansion of $(x+1)^{2016}$.
3. Find the coefficient of x^9 in the expansion of $(x-1)^{2016}$.
4. Show that $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n$.

Geometric Probability

The probability of selecting a point at random in a region R that is a part of region T, assuming that any point of the region is "equally likely" to be selected, is

$$\frac{\text{area (length) of region } R}{\text{area (length) of region } T}$$

Exercises:

1. A dart is thrown at a circular target inside of a 4 by 4 square. Assume that the dart has probability 1 of hitting the square. Assume each point in the square is "equally likely" to be hit.
   (a) Find the probability that the dart hits a circular target of radius 1 that lies entirely within the square.
   (b) Find the probability that the dart hits a circular target of radius 2 that lies entirely within the square.
   (c) Find the probability that the dart hits the upper half of the square.
   (d) Find the probability that the dart hits the upper half of a circular target of radius 2 that lies entirely within the square.
   (e) Suppose that 1000 darts are thrown at the circular target. How many would be expected to hit the target of radius 1 lying entirely within the square?
   (f) Suppose that 1000 darts are thrown at the circular target. How many would be expected to hit the target of radius 2 lying entirely within the square?
2. A point is selected at random on a line segment of length 3. What is the probability that the point is selected from the middle third of the line segment?