Lecture 1

- **Read:** Chapter 1.1-1.4.

- Set Theory
- Elements of Probability
- Representing Sample Spaces

Probability

- Life and our environment are uncertain
- Common analysis method for uncertain situations
  - Use "long-term averages," i.e., probabilities
- Common approach for decision under uncertainty
  - Optimizing the "average value" of the result of the decision
- Probability theory deals with phenomena whose outcome is not fully predictable
  - but exhibit some regularity when observed many times

History

- Cardano, Galileo (16th Century)
  - Emergence - gambling in Italy...
- Pascal, Fermat (17th Century)
  - Rational thought
- Bernoulli, Laplace, Poisson, Gauss (18th-19th Century)
  - Mathematical organization
- Kolmogorov (20th Century)
  - Axiomatization

Course Objective

1. Develop your ability to describe uncertain events in terms of probabilistic models
2. Develop your skill in probabilistic reasoning

Motivation - Applications

- Engineering
  - Communications, information theory
  - Signal processing and systems control
  - Queuing theory and modeling computer systems
  - Decision and resource allocation under uncertainty
  - Reliability, estimation, and detection
- Statistics: collection and organization of data so that useful inferences can be drawn from them
- Physics, statistical mechanics, thermodynamics
- Computer science: randomized algorithms, random search
- Economics and finance: investment/insurance risk assessment

Two Interpretations for Probabilities

- **Frequency of occurrence:** probability $=$ % of successes in a moderately large number of situations (*Reality may or may not involve repetition!*)
  - When is this appropriate? For example, 50% probability that a coin comes up heads versus 90% probability that Homer wrote the Odyssey and the Iliad?
- **Subjective belief:** probability $=$ an expert's opinion that an event occurred/will occur
  - For example, likelihood that a medication will work when it is used for the first time

Role of Math

- “Probability is common sense reduced to calculation.” (Laplace)
- “The book of the universe is written in the language of mathematics.” (Galileo)
- Probabilistic analysis is mathematical, but intuition dominates and guides the math. (Our goal!)
- Problem formulation in terms of probabilities is typically more challenging than the calculations. (Need to work lots of problems!)

Getting Started

- But, thinking probabilistically is fairly unnatural, unless you are used to it! So, let us get to work.
- Basic idea is to assign probabilities to collections (sets) of possible outcomes, so we start by briefly reviewing set theory.

Set Theory Preliminaries

- Venn Diagrams
- Universal Set/Empty Set
- Union/Intersection
- Complement
- Mutually Exclusive/Collectively Exhaustive

Set Theory Review: Sets

A set A is a collection of objects which are *elements* of the set.

- If x is an element of A, we write $x \in A$
- If x is *not* an element of A, we write $x \notin A$
- A set with no elements is the **empty set** $\emptyset$
- The set with *all* the elements relevant to a particular context is called the **universal set**, say S

Set Theory Review: Describing Sets

"Make a list" versus "describe its elements"

- **list of elements:** $A = \{x_1, x_2, ..., x_n\}$, e.g., possible outcomes of the roll of a die, $\{1, 2, 3, 4, 5, 6\}$ or coin toss, $\{H, T\}$
- **properties of elements:** $A = \{x|x \text{ satisfies } P\}$, e.g., $\{x|x \text{ is an even integer}\}$ or $\{x : 0 \leq x \leq 1\}$. Note that "|" and ":" both mean "such that"

Set Theory Review: Describing Sets

- **countable vs. uncountable:** A set is **countable** if it can be written down as a list, otherwise it is **uncountable.**
- **ordered pair** of two objects (x,y): e.g., set of scalars
  - Note that order is indicated by the use of (x,y) versus {x,y}
- **subset:** $A \subset B$ if every element in A is in B
- **equality of sets:** $A = B$ if and only if $A \subset B$ and $B \subset A$

Set Theory Review: Set Operations and Venn Diagrams

- **union:** (logical OR) of two sets $A \cup B$, i.e., in A or B **or both** (e.g., round and/or blue elements)
  - One can also define the union of a finite or even infinite number of sets, e.g.:

$$\cup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \ldots = \{x : x \in A_i \text{ for some } i\}$$
$$\cup_{\alpha \in R}^{\infty} B_\alpha = \{x : x \in B_\alpha \text{ for some } \alpha\}$$

- **intersection**: (logical AND) of two sets $A \cap B$ (e.g., round and blue elements)
  - Similarly, one can also define intersections of a finite or infinite number of sets, e.g.:

$$\cap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \ldots = \{x : x \in A_i \text{ for all } i\}$$

Set Theory Review: Set Operations and Venn Diagrams

- **complement**: of A in S is A^c, the set of elements which are not in A, e.g., the complement of $\emptyset$ is S
- **difference**: of two sets $A \backslash B = A \cap B^c$
- **disjoint or mutually exclusive sets**: have no common elements, i.e., $A \cap B = \emptyset$ iff A and B are disjoint
- **collectively exhaustive**: a (possibly infinite) collection of sets $A_1, ..., A_n$ is said to be **collectively exhaustive** iff $\cup_{i=1}^{n} A_i = S$
- **partition**: a (possibly infinite) collection of sets $A_1, ..., A_n$ is said to be a **partition** of S iff $\cup_{i=1}^{n} A_i = S$ (i.e., they are collectively exhaustive) and the sets are disjoint (i.e., mutually exclusive), e.g., A and A^c are a partition of S

Set Theory Review: Algebra of Sets

Elementary properties follow from the definitions:

- **Associative**: $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$
- **Distributive**: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- **de Morgan's Law**: $(A \cup B)^c = A^c \cap B^c$, similarly $(A \cap B)^c = A^c \cup B^c$
  - **Proof:** We can show that $(A \cup B)^c \subset A^c \cap B^c$ as follows. If $x \in (A \cup B)^c$, then x is not in A and not in B, thus x must be in both A^c and B^c. Similarly, one can establish that $A^c \cap B^c = (A \cup B)^c$

What is Probability?

- a number between 0 and 1.
- a physical property (like mass or volume) that can be measured?
- measure of our knowledge?

Probabilistic Models

Going from *experiments* in the physical world to *probabilistic models*

- Experiment= Procedure + Observation, e.g., flip a coin and see if it landed heads or tails or transmit a waveform over a channel and observe what is received
- Real Experiments are TOO complicated
- Instead we analyze/develop models of experiments
  - A coin flip is equally likely to be H or T
- Probabilistic model is (usually) a simplified mathematical description used to *study* the situation

Example 1.1

An experiment consists of the following procedure, observation, and model:

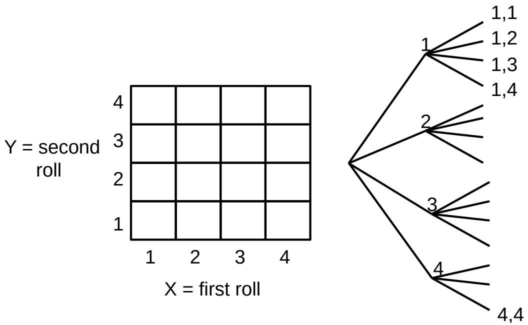
- Procedure: Flip a coin and let it land on a table.
- Observation: Observe which side (head or tail) faces you after the coin lands.
- Model: Heads and tails are equally likely. The result of each flip is unrelated to the results of previous flips.

Components of Probabilistic Models

- An *outcome* s, s_1, or x of an experiment is one of the possible observations of that experiment.
- The *sample space* S of an experiment is the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes.
- An *event* A is a set of outcomes of an experiment, i.e., $A \in S$, e.g. $\{s_1\}$.
- $B_1, B_2, ..., B_n$ make up an *event space* or *partition* S iff
  - $B_i \cap B_j = \emptyset, i \neq j$
  - $\cup_{i=1}^{n} B_i = S$

Representing Sample Spaces

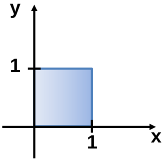
- Sequential models: sample space vs. tree-based sequential description, e.g., two rolls of a tetrahedral (four-sided) die.



- Let us look at the event that the second roll is 4.
- Let us exhibit a partition, e.g., the sets where the first roll is 1, 2, 3, or 4.

Representing Sample Spaces

- A continuous sample space: throw a dart at a square target with area 1, e.g., $S = \{(x, y) | 0 \leq x, y \leq 1\}$



Correspondences

Set Algebra	Probability
set	event
universal set	sample space
element	outcome

Example 1.9

Flip four coins, a penny, a nickel, a dime, and a quarter. Examine the coins in order (penny, then nickel, then quarter) and observe whether each coin shows a head (h) or a tail (t). What is the sample space? How many elements are in the sample space?

. . . .

The sample space consists of 16 four-letter words:

$$\{\text{tttt, ttth, ttht, ..., hhhh}\}$$

Event Spaces

- An *event space* is a collectively exhaustive, mutually exclusive set of events.
- **Example 1.10:** For $i = 0, 1, 2, 3, 4$,

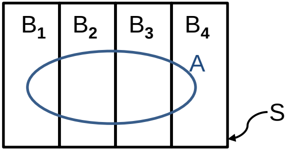
$$B_i = \{\text{outcomes with } i \text{ heads}\}$$

  - Each B_i is an event containing one or more outcomes.
  - The set $B = \{B_0, B_1, B_2, B_3, B_4\}$ is an event space.

Theorem 1.2

- For an event space $B = \{B_1, B_2, ...\}$ and any event A in the sample space, let $C_i = A \cap B_i$.
- For $i \neq j$, the events C_i and C_j are mutually exclusive $(C_i \cap C_j = \emptyset)$ and

$$A = C_1 \cup C_2 \cup \cdots$$



Probability Measure (or Law) and Axioms of Probability

A **probability measure (or law)** $P[\cdot]$ is a function that maps events in the sample space to real numbers ($P[A] \mapsto [0, 1]$) such that

- **Axiom 1** (nonnegativity) For any event A, $P[A] \geq 0$.
- **Axiom 2** (normalization) $P[S] = 1$.
- **Axiom 3** (additivity) For any countable collection $A_1, A_2, \ldots$ of mutually exclusive (i.e., disjoint) events
$$P[A_1 \cup A_2 \cup \ldots] = P[A_1] + P[A_2] + \ldots.$$

Consequences of the Axioms

Theorem 1.7: The probability measure $P[\cdot]$ satisfies

- $P[\emptyset] = 0$.
- $P[A^c] = 1 - P[A]$.
- For any A and B,

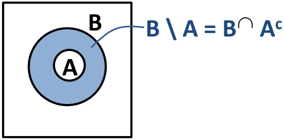
$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

- If $A \subset B$, then $P[A] \leq P[B]$.

Consequences of the Axioms (cont.)

- $P[A^c] = 1 - P[A]$
  Proof:
  $$A \cap A^c = \emptyset$$
  $$A \cup A^c = S$$
  $$P[A \cup A^c] = P[A] + P[A^c] = 1$$

- If $A \subset B$, then $P[A] \leq P[B]$.
  Proof: $B = A \cup [B \backslash A]$, A and $B \backslash A$ are disjoint
  $$P[B] = P[A] + P[B \backslash A]$$
  $$\Rightarrow P[B] \geq P[A]$$



- $P[A \cup B] = P[A] + P[B] - P[A \cap B]$

Consequences of the Axioms (cont.)

- **Theorem 1.4:** If

$$B = B_1 \cup B_2 \cup \ldots \cup B_m$$

and for $i \neq j$,

$$B_i \cap B_j = \emptyset$$

then

$$P[B] = \sum_{i=1}^{m} P[B_i]$$

Problem 1.3.5

A student's score on a 10-point quiz is equally likely to be any integer between 0 and 10. What is the probability of an A, which requires the student to get a score of 9 or more? What is the probability the student gets an F by getting less than 4?

Problem 1.4.5-modified

Before the completion of a phone conversation, a cellphone user is equally likely to get zero drops (D_0), one drop (D_1), or more than one drop (D_2). Also, a caller is on foot (F) with probability $5/12$ or in a vehicle (V).

- Find three ways to fill in the following probability table:

	D_0	D_1	D_2
F			
V			

Problem 1.4.5-modified (cont.)

- If $1/4$ of all callers are on foot making calls with no drops and $1/6$ of all callers are vehicle users making calls with a single drop, what is the table?

	D_0	D_1	D_2
F	$1/4$	$1/6$	0
V	$1/12$	$1/6$	$1/3$

Discrete Models

Building models means

1. defining sample/event space and
2. specifying a suitable probability law, i.e., consistent with the axioms

Examples:

1. fair coin toss $S = \{H, T\}$, P[H]= P[T]= 1/2
2. three fair coin tosses $S = \{HHH, HHT, HTH, ...\}$, each outcome with probability 1/8, suppose $A = \{\text{exactly 2 heads occur}\} = \{HHT, HTH, THH\}$

$$
\begin{aligned}
P[A] &= P[\{HHT, HTH, THH\}] \\
&= P[\{HHT\}] + P[\{HTH\}] + P[\{THH\}]
\end{aligned}
$$

Discrete Probability Law

If S consists of a countable set of outcomes, then for any event $A = \{s_1, s_2, ..., s_n\}$,

$$P[A] = P[\{s_1\}] + P[\{s_2\}] + ... + P[\{s_n\}]$$

Equally Likely Outcomes: Discrete Uniform Probability Law

If $S = \{s_1, s_2, ..., s_n\}$, i.e., consists of n possible outcomes, and they are *equally likely*, then for any event A, we have

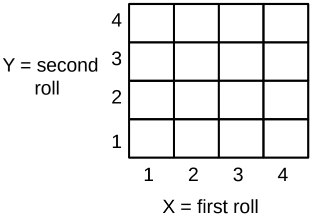
$$P[A] = \frac{(\#\text{ of elements in A})}{n}$$

$$\text{e.g.,} P[\{s_i\}] = \frac{1}{n}$$

Note: For such laws, computing probabilities boils down to counting events! Here arises a basic link between combinatorics and probability. Later, we will go over counting methods, e.g., permutations, combinations, etc.

Equally Likely Outcomes: Example

- Two rolls of a fair four-sided die
- 16 outcomes with the same probability $1/16$



- Find probability that at least one roll is 4, or probability that first roll is equal to the second