Administrative Announcements

- **Homework 1**: Due Thursday, February 11 at 11:59pm. *Be sure to follow the collaboration policy (see syllabus).*
- **Recitation 1**: Rescheduled for Tuesday (today) from 5:30-6:30pm in CIT 368. *Matlab tutorial (HW1 problem 5).*
Axioms of Probability for Infinite Spaces

Conditional Probability and Independence

1.2 PROBABILISTIC MODELS

A probabilistic model is a mathematical description of an uncertain situation. It must be in accordance with a fundamental framework that we discuss in this section. Its two main ingredients are listed below and are visualized in Fig. 1.2.

Elements of a Probabilistic Model

• The sample space \( \Omega \), which is the set of all possible outcomes of an experiment.
• The probability law, which assigns to a set \( A \) of possible outcomes (also called an event) a nonnegative number \( P(A) \) (called the probability of \( A \)) that encodes our knowledge or belief about the collective "likelihood" of the elements of \( A \). The probability law must satisfy certain properties to be introduced shortly.

Sample Spaces and Events

Every probabilistic model involves an underlying process, called the experiment, that will produce exactly one out of several possible outcomes. The set of all possible outcomes is called the sample space of the experiment, and is denoted by \( \Omega \). A subset of the sample space, that is, a collection of possible
Sample Spaces & Probability Laws

Ω = a set, or unordered “list”, of all possible distinct outcomes from some random experiment

Probabilities of events (sets of outcomes) must satisfy axioms:

1. Non-negativity: \( P(A) \geq 0 \)
2. Normalization: \( P(\Omega) = 1 \)
3. Countable additivity: For disjoint events \( A_1, A_2, A_3, \ldots \)
   \[
P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots
   \]
Repeatedly flip a coin with probability of Heads $p$, count the number of tosses $X$ until the first Head is observed:

$$P(X = 1) = p, \quad P(X = 2) = (1 - p)p, \quad P(X = 3) = (1 - p)^2 p, \ldots$$

$$P(X = k) = (1 - p)^{k-1} p \text{ for } k = 1, 2, 3, \ldots$$

The number of possible outcomes is infinite: there is no $k$ after which the next toss is guaranteed to be Heads.

Example:

Your laptop hard drive independently fails on each day with (hopefully small) probability $p$. What is the distribution of the number of days until failure?
Geometric Probabilities

- Repeatedly flip a coin with probability of Heads $p$, count the number of tosses $X$ until the first Head is observed:

  \[
P(X = 1) = p, \quad P(X = 2) = (1-p)p, \quad P(X = 3) = (1-p)^2p, \ldots
  \]

  \[
P(X = k) = (1-p)^{k-1}p \text{ for } k = 1, 2, 3, \ldots
  \]

- Recall the geometric series:

  \[
  \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \quad 0 < q < 1.
  \]

- Verify that geometric probabilities are normalized:

  \[
  \sum_{k=1}^{\infty} (1-p)^{k-1}p = p \sum_{k=0}^{\infty} (1-p)^k = \frac{p}{1 - (1-p)} = 1
  \]
Repeatedly flip a coin with probability of Heads $p$, count the number of tosses $X$ until the first Head is observed:

$$P(X = 1) = p, \ P(X = 2) = (1 - p)p, \ P(X = 3) = (1 - p)^2 p, \ldots$$

$$P(X = k) = (1 - p)^{k-1} p \text{ for } k = 1, 2, 3, \ldots$$

What is the probability that the number of tosses $X$ is odd?

$$P(X \text{ odd}) = \sum_{k=1}^{\infty} (1 - p)^{2(k-1)} p = \frac{1}{2 - p}$$

For a fair coin, this equals

$$P(X \text{ odd}) = \frac{2}{3} \text{ if } p = \frac{1}{2}$$
Another Birthday Problem

Suppose there are $m$ students in a class. What is the probability that at least one student in the class has the same birthday as the instructor?

- The probability that all $m$ students have distinct birthdays is

$$\left(1 - \frac{1}{N}\right)^m = \left(\frac{N - 1}{N}\right)^m$$

- Interpretation: Check all $m$ birthdates one at time

**Assumptions:**
- Birthdays are equally likely to occur on any of $N=365$ days
- No dependence between birthdays of different people

*Not completely true, but fairly accurate approximations.*
Suppose there are $m$ students in a class. What is the probability that at least one student in the class has the same birthday as the instructor?

Probability that $m$ students do not share instructor’s birthday:

$$
\left(1 - \frac{1}{N}\right)^m = \left(\frac{N-1}{N}\right)^m
$$

In a class of 70 students, the probability that at least one has the same birthday as the instructor is

$$
1 - \left(\frac{364}{365}\right)^{70} \approx 0.175
$$
The Birthday “Paradox”

The probability that \( m \) pairs of birthdays on \( N \) days are distinct:

\[
P(D_m) = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{m-1}{N}\right) = \prod_{i=0}^{m-1} \left(1 - \frac{i}{N}\right)
\]

In a class of 70 students, the probability that at least two share a birthday is approximately 0.9987

\[
P(D_m) \leq \prod_{i=0}^{m-1} e^{-\frac{i}{N}} = e^{-\sum_{i=0}^{m-1} \frac{i}{N}} = e^{-\frac{m(m-1)}{2N}}
\]
CS145: Lecture 3 Outline

- Axioms of Probability for Infinite Spaces
- Conditional Probability and Independence
Conditional Probability

- \( P(A | B) = \) probability of \( A \), given that \( B \) occurred
  - \( B \) is our new universe

- **Definition:** Assuming \( P(B) \neq 0 \),
  \[
P(A | B) = \frac{P(A \cap B)}{P(B)}
  \]

  \( P(A | B) \) undefined if \( P(B) = 0 \)

- Under discrete uniform law, where all outcomes equally likely:
  \[
P(A | B) = \frac{\text{number of elements of } A \cap B}{\text{number of elements of } B} = \frac{|A \cap B|}{|B|}
  \]
Example: Two-Sided Cards

A hat contains three cards.
One card is black on both sides.
One card is white on both sides.
One card is black on one side and white on the other.
The cards are mixed up in the hat. Then a single card is drawn and placed on a table. If the visible side of the card is black, what is the chance that the other side is white?

Label the faces of the cards:
- $b_1$ and $b_2$ for the black–black card;
- $w_1$ and $w_2$ for the white–white card;
- $b_3$ and $w_3$ for the black–white card.

$\{\text{black on top}\} = \{b_1, b_2, b_3\}$
$\{\text{white on bottom}\} = \{b_3, w_1, w_2\}$

$$P(\text{white on bottom}|\text{black on top}) = \frac{\#(\text{white on bottom and black on top})}{\#(\text{black on top})} = \frac{1}{3}$$

Observing the top face of the card provides information about the color of the bottom face.
Example: Roll of a 6-Sided Die

- The probability of "the outcome of a die roll is even" is $\frac{3}{6}$.
- The probability of the event "the outcome is $\leq 4$" is $\frac{4}{6}$.

$$P(\text{die even} \mid \text{die } \leq 4) = \frac{2}{4} = \frac{1}{2} = P(\text{die even})$$

Observing that the die roll was at most 4 does not provide information about whether it was even.
Independence of Two Events

- "Defn:" \( P(B \mid A) = P(B) \)
  - "occurrence of \( A \) provides no information about \( B \)'s occurrence"

- Recall that \( P(A \cap B) = P(A) \cdot P(B \mid A) \)

- Defn: \( P(A \cap B) = P(A) \cdot P(B) \)

- Symmetric with respect to \( A \) and \( B \)
  - applies even if \( P(A) = 0 \)
  - implies \( P(A \mid B) = P(A) \)
Sources of Independence

- Event associated with "independent" physical processes are independent.
- But independent events do not have to be related to independent physical processes.
- Example:
  - The probability of "the outcome of a die roll is even" is $\frac{3}{6}$.
  - The probability of the event "the outcome is $\leq 4$" is $\frac{4}{6}$.
  - The probability of "an even outcome $\leq 4$" is
    \[
    \frac{2}{6} = \frac{12}{36} = \frac{3}{6} \cdot \frac{4}{6}
    \]
  \(\Rightarrow\) the two events are independent.
- The "intuition" here is that there are the same number of odd and even outcomes that are $\leq 4$. Thus, the "information" that the outcome is $\leq 4$ does not "help" in deciding if it is odd or even.
Events that are NOT Independent

- Assume events are non-degenerate: \(0 < P(A) < 1, 0 < P(B) < 1\)
- Nested events are not independent:
  \(\text{If } A \subset B, \quad P(B \mid A) = 1 \neq P(B).\)
- Mutually exclusive events are not independent:
  \(\text{If } A \cap B = \emptyset, \quad P(A \cap B) = 0 \neq P(A)P(B).\)
Serial versus Parallel Systems

Let \( W_i \) be the event that component \( C_i \) works without failure

\[
\{ \text{system works} \} = W_1 \cap W_2
\]

Assume component failures are independent events, and that

\[
P(\tilde{W}_1) = 0.9 \text{ and } P(W_2) = 0.8
\]

\[
P(\text{system works}) = P(W_1 W_2) = P(W_1)P(W_2) = 0.9 \times 0.8 = 0.72
\]

\[
\{ \text{system works} \} = W_1 \cup W_2
\]

\[
P(\text{system works}) = 1 - (0.1)(0.2) = 0.98
\]
Conditioning may Affect Independence

- Conditional independence, given \( C \), is defined as independence under probability law \( P(\cdot | C) \)

- Assume \( A \) and \( B \) are independent

- If we are told that \( C \) occurred, are \( A \) and \( B \) independent?

\[
P(A \cap B | C) = 0 \neq P(A | C)P(B | C)
\]

Definition of conditional independence:
\[
P(A \cap B | C) = P(A | C)P(B | C)
\]

For this example:
\[
P(A \cap B) = P(A)P(B)
\]
Example: Conditioning & Independence

- Two unfair coins, $A$ and $B$:
  $P(H \mid \text{coin } A) = 0.9$, $P(H \mid \text{coin } B) = 0.1$
  choose either coin with equal probability

- Once we know it is coin $A$, are tosses independent?

  *Yes, by definition.*

- If we do not know which coin it is, are tosses independent?

  *No, consider probability that second coin is heads given first.*

- Compare:
  $P(\text{toss 11} = H)$
  $P(\text{toss 11} = H \mid \text{first 10 tosses are heads})$
Example: Monty Hall Problem

- A prize is equally likely to be behind one of three doors.
- You choose a door and before that door is opened another door without the price is opened.
- You are offered to switch your choice to the other closed door.
- Should you do it?
Example: Monty Hall Problem

- **Strategy 1: don’t switch:** With probability $\frac{1}{3}$ the prize is behind the door you had chosen.

- **Strategy 2: switch:**
  - If the prize is behind your first choice you loose - probability $\frac{1}{3}$.
  - Else, with probability $\frac{2}{3}$, the prize was behind one of the other two doors. One is open and has no prize, the prize is behind the second door.

- Opening one of the doors increases your knowledge about the other door.

- Switching increases wining probability from $\frac{1}{3}$ to $\frac{2}{3}$. 