ECE 302: Chapter 02 Probability Model

Fall 2019

Prof Stanley Chan

School of Electrical and Computer Engineering
Purdue University
1. Set Theory
Set

Definition
A **set** is a collection of objects. We denote \( A = \{\xi_1, \xi_2, \ldots, \xi_n\} \) as a set, and \( \xi_i \) be the \( i \)-th element in the set.

Notation:
- \( \xi \in A \): An object \( \xi \) is in set \( A \).
- \( \xi \notin A \): An object \( \xi \) is not in set \( A \).

Finite, Countable, Uncountable.
- Finite: \( A = \{0, 1\} \)
- Countable: \( A = \{2, 4, 6, 8, \ldots\} \)
- Uncountable: \( A = \{x \mid 0 < x < 1\} \)

Open and Closed Intervals.
- \((a, b) = \{x \mid a < x < b\}\)
- \([a, b] = \{x \mid a \leq x \leq b\}\)
Subset

Definition

A **subset** of $A$ is sub-collection of objects in $A$. That is, $B \subseteq A$ if for any $\xi \in B$, this $\xi$ is also in $A$.

**Proper subset**: $B \subset A$.

**Example.** If $A = \{1, 2, 3, 4, 5, 6\}$, then $B = \{1, 2\}$ is a proper subset of $A$.

**Example.** If $A = \{1, 2, 3\}$, then the set $B = \{1, 2, 3\}$ is an improper subset of $A$.

**Theorem**

If $A \subseteq B$ and $B \subseteq A$, then $A = B$.

**Proof.** Suppose $A \subset B$ (which means $A \neq B$. You may also consider $B \subset A$). Then there exist $x \in B$ but $\notin A$. But $B \subseteq A$ requires all $x \in B$ also $\in A$. So we reach a contradiction. The only way to resolve the contradiction is to make $A = B$. 
Empty and Universal Set

Definition (Empty Set)
A set is **empty** if it contains no element. We denote an empty set as $\emptyset$.

An empty set is a subset of any set.

Definition
The **universal set** is the set containing all elements. We denote a universal set as $\Omega$.

Any set is a subset of $\Omega$, including $\Omega$ itself.
Finite Union

Definition (Finite Union)

The **finite union** of two sets $A$ and $B$ contains all elements in $A$ or in $B$. That is,

$$A \cup B = \{\xi \mid \xi \in A \text{ or } \xi \in B\}.$$ 

Example. If $A = \{1, 2, 3, 4\}$, $B = \{1, 5, 6\}$, then $A \cup B =$

Example. If $A = \{t \mid 3 < t \leq 4\}$, $B = \{t \mid t \geq 3.5\}$, then $A \cup B =$
Definition (Infinite Union)

The **infinite union** of $A_1, A_2, \ldots, A_n$ is denoted as

$$A \overset{\text{def}}{=} \bigcup_{n=1}^{\infty} A_n.$$ 

It holds that $x \in A$ if $x$ is in **at least one** of $A_1, A_2, \ldots, A_n$.

**Example.** If $A_n = [0, 1 - \frac{1}{n}]$, then

(a) $A = [0, 1]

(b) $A = [0, 1)$
Finite Intersection

Definition

The **finite intersection** of two sets $A$ and $B$ contains all elements in $A$ and in $B$. That is,

$$A \cap B = \{ \xi \mid \xi \in A \text{ and } \xi \in B \}.$$ 

Example. If $A = \{1, 2, 3, 4\}$, $B = \{1, 5, 6\}$, then $A \cap B =$

Example. If $A = \{t \mid 3 < t \leq 4\}$, $B = \{t \mid t \geq 3.5\}$, then $A \cap B =$
Infinite Intersection

**Definition**

Define the **infinite intersection** of $A_1, A_2, \ldots, A_n$ as

$$A = \bigcap_{n=1}^{\infty} A_n$$

Then, $x \in A$ if $x$ is in all of $A_1, A_2, \ldots, A_n$.

**Example.** If $A_n = [0, 1 + \frac{1}{n})$, then

(a) $A = [0, 1]$

(b) $A = [0, 1)$
Complement

Definition

The complement of a set $A$ is the set containing all elements in $\Omega$ but not in $A$. That is

$$A^c = \{ \xi \mid \xi \in \Omega \text{ and } \xi \notin A \}.$$ 

Example. Let $A = \{\text{even integers}\}$, $\Omega = \{\text{integers}\}$, then $A^c =$
Difference

Definition
The **difference** $A \setminus B$ is the set containing all elements in $A$ but not in $B$. 

$$A \setminus B = \{ \xi \mid \xi \in A \text{ and } \xi \notin B \}.$$
**Disjoint**

**Definition**

Two sets $A$ and $B$ are **disjoint** if $A \cap B = \emptyset$. For a collection of sets \{\(A_1, A_2, \ldots, A_n\}\), we say that the collection is disjoint if $A_i \cap A_j = \emptyset$. 
Partition

Definition

A collection of sets \( \{A_1, \ldots, A_n\} \) is a **partition** to the universal set \( \Omega \) if it satisfies the following conditions:

- (non-overlap) \( \{A_1, \ldots, A_n\} \) is disjoint;
- (decompose) \( A_1 \cup A_2 \cup \ldots \cup A_n = \Omega \).
Set Operations

There are four set operations:

- **Commutative (Order does not matter).**
  \[
  A \cap B = B \cap A
  \]
  \[
  A \cup B = B \cup A
  \]

- **Associative (How to do multiple union and intersection)**
  \[
  A \cup (B \cup C) = (A \cup B) \cup C
  \]
  \[
  A \cap (B \cap C) = (A \cap B) \cap C
  \]
Set Operations

- **Distributive (How to mix union and intersection)**
  \[ A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \]
  \[ A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \]

- **De Morgan’s Law (How to complement over intersection and union)**
  \[ (A \cap B)^c = A^c \cup B^c \]
  \[ (A \cup B)^c = A^c \cap B^c \]
2. Probability Model
What is Probability?

- It is a **number**.

- Always between **0** and **1**.

- Always the probability of **an event**.

**Example.** The probability of getting a Head when tossing a coin:

\[ P(“H”) = \]
Three Elements of a Probability Model

1. Sample Space
2. Event
3. Probability Law
Sample Space

Definition (Sample Space)

A sample space $\Omega$ is the collection of all possible outcomes.

We denote $\omega$ as an element in $\Omega$.

Example.

- Coin flip:
  $\Omega =$

- Throw a dice:
  $\Omega =$

- Waiting time for a bus in West Lafayette:
  $\Omega =$
**Definition (Event)**

An **event** $F$ is a subset in the sample space $\Omega$.

**Outcome VS Event:**

**Example.** Throw a dice. Let $\Omega = \{1, 2, 3, 4, 5, 6\}$.

- $F_1 = \{\text{even numbers}\} = \{2, 4, 6\}$.
- $F_2 = \{\text{less than 3}\} = \{1, 2\}$.

**Example.** Wait a bus. Let $\Omega = \{0 \leq t \leq 30\}$.

- $F_1 = \{0 \leq t < 10\}$
- $F_2 = \{0 \leq t < 5\} \cup \{20 < t \leq 30\}$.
Event Space

Definition (Event Space)
The collection of all possible events is called the

**Event Space or σ-field**
denoted as $\mathcal{F}$. $\mathcal{F}$ satisfies the following two properties:

- If $F \in \mathcal{F}$, then $F^c \in \mathcal{F}$
- If $F_1, F_2, \ldots \in \mathcal{F}$, then $F_i \cap F_j \in \mathcal{F}$ and $F_i \cup F_j \in \mathcal{F}$.

**Example.** $\Omega = \{H, T\}$, the event space is $\{\emptyset, H, T, \Omega\}$. 
Probability Law

Definition

A **probability law** is a function $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ that maps an event $A$ to a real number in $[0, 1]$. The function must satisfy three axioms known as **Probability Axioms**.

I. Non-negativity:

II. Normalization:
III. Additivity:

For any disjoint subsets \( \{ A_1, A_2, \ldots \} \), it holds that

\[
P \left[ \bigcup_{n=1}^{\infty} A_n \right] = \sum_{n=1}^{\infty} P[A_n].
\]
Properties of Probability

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}[A]$. 

2. For any $A \subseteq \Omega$, $\mathbb{P}[A] \leq 1$. 

3. $\mathbb{P}[\emptyset] = 0$. 

For any $A$ and $B$,

$$P[A \cup B] = P[A] + P[B] - P[A \cap B].$$
Properties of Probability

(Union Bound) For any $A$ and $B$,

$$\mathbb{P}[A \cup B] \leq \mathbb{P}[A] + \mathbb{P}[B].$$
Properties of Probability

6. If $A \subseteq B$, then $\mathbb{P}[A] \leq \mathbb{P}[B]$ 

Example. $A = \{ t \leq 5 \}$, and $B = \{ t \leq 10 \}$, then $\mathbb{P}[A] \leq \mathbb{P}[B]$. 
Example

Let the events $A$ and $B$ have $\mathbb{P}[A] = x$, $\mathbb{P}[B] = y$ and $\mathbb{P}[A \cup B] = z$. Find the following probabilities.

(a) $\mathbb{P}[A \cap B]$

(b) $\mathbb{P}[A^c \cap B^c]$
Example

(c) $\mathbb{P}[A^c \cup B^c]$

(d) $\mathbb{P}[A \cap B^c]$
\[ P[\cdot] \text{ is a measure of the size of the event.} \]

- \( \Omega = \{1, 2, 3, 4, 5, 6\} \)
- The measure \( P[\cdot] \) is a counter
- Counts the number of events

- \( \Omega = [0, 1] \times [0, 1] \)
- The measure \( P[\cdot] \) is an integration
- Integrates the area of events
$\mathbb{P}[\cdot]$ is a **measure** of the size of the event.

- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $E = \{1\}$, $\mathbb{P}[E] = 1/6$
- $E = \{1, 3\}$, $\mathbb{P}[E] = 2/6$
- $\mathbb{P}[E_1] \leq \mathbb{P}[E_2]$ if $E_1 \subseteq E_2$.

- $\Omega = [0, 1] \times [0, 1]$
- $E$ = shaded region, $\mathbb{P}[E] = \text{area}$.
- $E = \{(x_0, y_0)\}$, $\mathbb{P}[E] = 0$.
- $\mathbb{P}[E]$ can be $0$ even if $E \neq \emptyset$. 
Measure Zero

Definition
Let $\Omega$ be the sample space. A set $A \in \Omega$ is said to have measure zero if for any given $\epsilon > 0$,

- There exists a countable number of subsets $A_n$ such that $A \subseteq \bigcup_{n=1}^{\infty} A_n$, and
- $\sum_{n=1}^{\infty} P[A_n] < \epsilon$. 

**Measure Zero Example**

**Example.** Let $\Omega = [0, 1]$. Then the set $\{0.5\} \subset \Omega$ has measure zero.

- Fix any $\epsilon$.
- Create a very small interval around 0.5.
- Shrink the radius of the interval to make sure it is less than $\epsilon$.

**Facts.**

- Discrete space: No issue about measure zero.
- **Example.** $\{1\}$ in $\{1, \ldots, 6\}$ has non-zero probability.
- Continuous space: Isolated points have measure zero.
- **Example.** $\{1\}$ in $[1, 6]$ has zero probability.
Almost Surely

Definition
An event $A \in \mathbb{R}$ is said to hold \textbf{almost surely (a.s.)} if

$$\mathbb{P}[A] = 1,$$

except for all measure-zero sets in $\mathbb{R}$.

Example. Let $\Omega = [0, 1]$. Then

$$\mathbb{P}[(0, 1)] = 1 \quad \text{(almost surely)}$$

because the points 0 and 1 have measure zero in $\Omega$.

Note: In this course we will skip “a.s.” if the context is clear.
3. Conditional Probability
Definition (Conditional Probability)

Assume $\mathbb{P}[B] \neq 0$. The **conditional probability** of $A$ given $B$ is

\[
\mathbb{P}[A | B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} \quad \text{and} \quad \mathbb{P}[A \cap B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[\Omega]}.
\]

**Difference:**
Understanding Conditional Probability

Assume 12 equally likely outcomes

\[ P(A) = \frac{5}{12} \quad P(B) = \frac{6}{12} \]

If told \( B \) occurred:

\[ \begin{align*}
P(A | B) &= \quad \quad \quad \quad \quad \quad P(B | B) = \\
&= \frac{3}{12} \quad \quad \quad \quad \quad \quad \frac{2}{12} \quad \quad \quad \quad \quad \quad \frac{4}{12}
\end{align*} \]

https://ocw.mit.edu/resources/res-6-012-introduction-to-probability-spring-2018/part-i-the-fundamentals/
Example: two rolls of a 4-sided die

Let $B$ be the event: $\min(X, Y) = 2$.

Let $M = \max(X, Y)$.

$P(M = 1 \mid B) =$

$P(M = 3 \mid B) =$

https://ocw.mit.edu/resources/res-6-012-introduction-to-probability-spring-2018/part-i-the-fundamentals/
Example

Conditional probability can appear in daily life:

- \( A = \) You are admitted to Harvard
- \( B = \) Your SAT score is 2200 or above

The probabilities are:

- \( P[A] \) very small (4.5% in 2019 \(^1\))
- \( P[B] \) hard, but not too bad
- \( P[A|B] \) may not be very high, because a lot of applicants are outstanding
- \( P[B|A] \) probably higher, because you will not be reviewed if your SAT is low

And do not forget, the 4.5% is a conditional probability conditioned on those who applied only.

\(^1\)https://www.thecrimson.com/article/2019/3/29/2023-admit-numbers/
Bayes Theorem

Theorem (Bayes Theorem)

For any two events $A$ and $B$ such that $\mathbb{P}[A] > 0$ and $\mathbb{P}[B] > 0$, it holds that

$$
\mathbb{P}[A \mid B] = \frac{\mathbb{P}[B \mid A] \mathbb{P}[A]}{\mathbb{P}[B]}.
$$
Law of Total Probability

Theorem (Law of Total Probability)

Let \( \{A_1, A_2, \ldots, A_n\} \) be a partition of \( \Omega \), i.e., \( A_1, \ldots, A_n \) are disjoint and \( \Omega = A_1 \cup A_2 \cup \ldots \cup A_n \). Then, for any \( B \subseteq \Omega \),

\[
P[B] = \sum_{i=1}^{n} P[B \mid A_i] P[A_i].
\]
Law of Total Probability

**Figure**: Law of total probability decomposes the probability $\mathbb{P}[B]$ into multiple conditional probabilities $\mathbb{P}[B | A_i]$. The probability of obtaining each $\mathbb{P}[B | A_i]$ is $\mathbb{P}[A_i]$. 

\[ \mathbb{P}[B | A_1], \mathbb{P}[B | A_2], \mathbb{P}[B | A_3], \mathbb{P}[B | A_4] \]
Example 1: Tennis Tournament

Consider a tennis tournament. Your probability of winning the game is

0.3 against $\frac{1}{2}$ of the players (Event A).

0.4 against $\frac{1}{4}$ of the players (Event B).

0.5 against $\frac{1}{4}$ of the players (Event C).

What is the probability of winning the game?
Example 2: Communication Channel

Consider a communication channel shown below. The probability of sending a 1 is $p$ and the probability of sending a 0 is $1 - p$. Given that 1 is sent, the probability of receiving 1 is $1 - \eta$. Given that 0 is sent, the probability of receiving 0 is $1 - \varepsilon$.

Find $\mathbb{P}[\text{Receive 1}]$

Find $\mathbb{P}[\text{Send 1}]$

Find $\mathbb{P}[\text{Receive 1} | \text{Send 1}]$

Find $\mathbb{P}[\text{Send 1} | \text{Receive 1}]$
Properties of Conditional Probability

**Theorem**

Let \( \mathbb{P}[B] > 0 \). The conditional probability \( \mathbb{P}[A \mid B] \) satisfies Axiom I to Axiom III.

**Proof.**

- **Axiom I:**

- **Axiom II:**

- **Axiom III:**
Independence

Definition

Two events $A$ and $B$ are statistically **independent** if

Disjoint VS Independent.
Examples of Independence

**Example 1.** Throw a dice twice. Let

\[ A = \{ \text{1st dice is 3} \} \quad \text{and} \quad B = \{ \text{2nd dice is 4} \}. \]

Are \( A \) and \( B \) independent?
Examples of Independence

Example 2. Throw a dice twice. Let

\[ A = \{1st \text{ dice is } 1\} \quad \text{and} \quad B = \{\text{sum is 7}\}. \]

Are \( A \) and \( B \) independent?

- Think about \( \mathbb{P}[A|B] \).
- If you know the sum is 7, then the pair has to be (1,6), (2,5), (3,4), (4,3), (5,2), (6,1).
- The chance of getting first dice = 1 is still 1/6. It has been not been changed by \( B \).
- So independent.
- How about you change \( B = \{\text{sum is 8}\} \)?
Examples of Independence

**Example 3.** Throw a dice twice. Let

\[ A = \{ \text{1st dice is 2} \} \quad \text{and} \quad B = \{ \text{sum is 8} \}. \]

Are \( A \) and \( B \) independent?

- Think about \( \mathbb{P}[A|B] \).
- If you know the sum is 8, then the pair has to be (2,6), (3,5), (4,4), (5,3), (6,2).
- The chance of getting first dice = 2 is no longer 1/6. It has been changed by \( B \).
- So dependent.
Example 4. Throw a dice twice. Let

$$A = \{ \text{max is 2} \} \quad \text{and} \quad B = \{ \text{min is 2} \}. $$

Are $A$ and $B$ independent?
Independence Via Conditional Probability

- Recall that $\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.
- If $A$ and $B$ are independent, then $\mathbb{P}[A \cap B] = \mathbb{P}[A] \mathbb{P}[B]$

Therefore,

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A] \mathbb{P}[B]}{\mathbb{P}[B]} = \mathbb{P}[A].$$

Interpretation.
Prisoner’s Dilemma

- Three Prisoners: A, B, C. The King decides to release 2 and kill 1.
- You were A.
- Your chance of release is 2/3.
- Suppose you know the guard well. You can ask him about which of B or C will be released.
- But if you find out B (or C) is released, your chance becomes 1/2.
- How come!!