50.034 - Introduction to Probability and Statistics

Week 1 – Lecture 2

January–May Term, 2019

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Lecture 2 Assigned Readings: 1.3, 1.4, 1.5

Outline of Lecture

- Experiment, sample space, event
- Set theory
- Axioms of Probability
- Properties of probability

Experiment, Sample Space

An experiment is any process (whether real or hypothetical), in which the possible outcomes can be identified ahead of time.

Examples of experiments:

- Rolling a die. (singular: die, plural: dice)
- Checking the age of a student in this class.
- Recording if it rains at SUTD tomorrow at 10am.

The sample space of an experiment, usually denoted by Ω, is the set of **ALL** possible outcomes of that experiment.

Question: What are the sample spaces of the above experiments?

Properties of a sample space

A sample space must be collectively exhaustive.

- This means the sample space contains all possible outcomes.
- For the experiment of rolling a die, $\{1, 2, 4, 5, 6\}$ is not a sample space because the outcome 3 is missing, and hence $\{1, 2, 4, 5, 6\}$ is not collectively exhaustive.

The outcomes in a sample space must be mutually exclusive.

- This means that two different outcomes cannot both occur at the same time.
- For example, for the experiment of rolling a die, you cannot roll and get both 1 and 2 at the same time.
- This also means that you cannot have any repeated values. For example, $\{1, 1, 2, 3, 4, 5, 6\}$ is NOT a sample space.

Event

An event is a subset of outcomes contained in a sample space Ω.

An event is simple if it consists of exactly one outcome, and is compound if it consists of more than one outcome.

An event A is said to occur if the resulting experimental outcome is contained in A.

Event

Consider the experiment of rolling a die, there are six simple events: $\{1\}$, $\{2\}$, ... $\{6\}$.

Example of compound events:
1. Event that the outcome is odd: $\{1, 3, 5\}$.
2. Event that the outcome is greater than 4: $\{5, 6\}$.

In general, exactly one simple event will occur, but many compound events may occur simultaneously.

Example 1

Consider the experiment of tossing a coin that has two faces. What is the sample space? What are all simple events and compound events?

Example 1

Solution:

Let the two possible outcomes be denoted by H and T.

1. The sample space is $\{H, T\}$.
2. There are two simple events, $\{H\}$ and $\{T\}$.
3. There is one compound event $\{H, T\}$.

Sample space vs Population

Sample space versus population—how are they different?
Can you give an example?

Main Difference: A population could possibly have repeated values, but a sample space cannot have any repeated values.

- For example, consider a population of numbers representing the ages of all students in SUTD. Among all students, you may find many students with a common age of 19. Then the population contains many repeated 19's.
- The sample space of the ages of all SUTD students would contain 19, but this age appears only once as an outcome in the sample space.

In other words, a sample space must be a **set**.
Note: A set is a collection with no repeated objects.

Set Theory

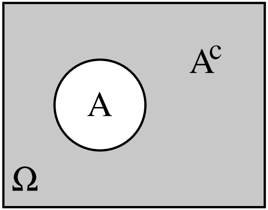
Operations in Set Theory:

The complement of an event A, denoted by A^c, is the set of all outcomes in Ω that are not contained in A.

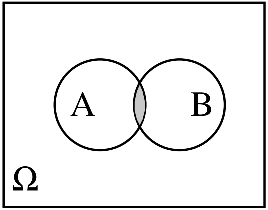
The intersection of two events A and B, denoted by $A \cap B$, is the event consisting of all outcomes that are in both A and B.

The union of two events A and B, denoted by $A \cup B$, is the event consisting of all outcomes that are either in A or in B.

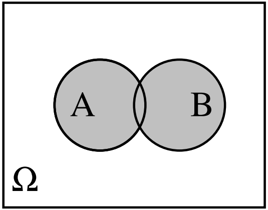
Venn diagrams



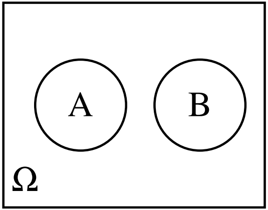
A^c shaded



$A \cap B$ shaded



$A \cup B$ shaded



Disjoint events

Null event

The null event, denoted by $\varnothing$, is the event that consists of no outcomes. The symbol "$\varnothing$" is also usually called the "empty set" or the "null set".

Events A and B are said to be mutually exclusive or disjoint events if $A \cap B = \varnothing$. The symbol "$\cap$" is usually read aloud as "intersect".

Events $A_1, A_2, A_3, \ldots$ are mutually exclusive (or pairwise disjoint) if no two events have any outcome in common.

Example 2: De Morgan's Laws

Use Venn diagrams to verify the following two relationships for any events A and B.

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

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

Probability

Given an experiment and a sample space Ω, the probability of an event A is defined to be a number that represents some measure of the chance that A will occur. We denote this number by $\Pr(A)$.

(In some other textbooks, the notation $\mathrm{P}(A)$ is used instead. For consistency, we follow the notation used in the course textbook and write $\Pr(A)$.)

Example: Let X denote the outcome variable from tossing a coin. Let H represent the outcome of tossing heads.

We can write $\Pr(H) = 0.5$ or $\Pr(X = H) = 0.5$, to mean that the event H has a probability of 0.5.

Axioms of probability

There are three probability axioms:

- **Axiom 1:** For any event A, $\Pr(A) \geq 0$.
- **Axiom 2:** $\Pr(\Omega) = 1$.
- **Axiom 3:** Any infinite sequence of mutually exclusive (disjoint) events $A_1, A_2, A_3, \ldots$ satisfies

$$\Pr(A_1 \cup A_2 \cup A_3 \cup \ldots) = \sum_{i=1}^{\infty} \Pr(A_i).$$

Example 3

Let X be a random variable representing the number of planes arriving at Changi Airport from 2pm to 5pm. We assume X can be modeled by the **Poisson distribution** with parameter λ:

$$\Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, 3, \ldots$$

1. What are the sample space Ω and simple events?
2. What is probability that four planes arrive?
3. Show that $\Pr(\Omega) = 1$. (Hint: $\displaystyle\sum_{i=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$.)

Solution

1. The sample space is $\Omega = \{0, 1, 2, \dots\}$.
   The simple events are $\{0\}, \{1\}, \{2\}, \dots$

2. The probability that four planes arrive is $\Pr(X = 4) = \dfrac{\lambda^4 e^{-\lambda}}{4!}$.

3. Aim: Show that $\Pr(\Omega) = 1$.

$$
\begin{aligned}
\Pr(\Omega) &= \Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2) + \dots \\
&= \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \dots \\
&= e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^k}{k!} \\
&= e^{-\lambda} e^{\lambda} = 1
\end{aligned}
$$

Null events have zero probability

Question: How do we know that $\Pr(\varnothing) = 0$?

Answer: Axiom 3 implies it!
Let $A_1, A_2, \ldots$ be an infinite sequence of events, where $A_i = \varnothing$ for every i. We make the following observations:

- The events are mutually exclusive (disjoint), since $\varnothing \cap \varnothing = \varnothing$.
- $A_1 \cup A_2 \cup \cdots = \varnothing \cup \varnothing \cup \cdots = \varnothing$.

So by Axiom 3, we must have

$$\Pr(\varnothing) = \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i) = \sum_{i=1}^{\infty} \Pr(\varnothing).$$

$\Pr(\varnothing)$ is a number, and the only possible number is $\Pr(\varnothing) = 0$.

Finite sequence of mutually exclusive events

Theorem:
For any finite sequence of mutually exclusive events $A_1, A_2, \ldots, A_n$,

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \Pr(A_1) + \Pr(A_2) + \cdots + \Pr(A_n).$$

Proof: We can extend this finite sequence to an infinite sequence

$$A_1, A_2, \ldots, A_n, A_{n+1}, \ldots,$$

such that $A_k = \varnothing$ for all $k > n$.
The events in this infinite sequence are mutually exclusive, since $\varnothing \cap A_i = \varnothing$ for any event A_i, and since $\varnothing \cap \varnothing = \varnothing$.

- $A_i \cap A_j = \varnothing$ for all $i \neq j$ in $\{1, \ldots, n\}$, since we are given that $A_1, \ldots, A_n$ are mutually exclusive.
- $\Pr(\varnothing) = 0$. (previous slide)
- $A_1 \cup A_2 \cup \cdots \cup A_n = A_1 \cup A_2 \cup \cdots \cup A_n \cup A_{n+1} \cup \cdots$

Thus by Axiom 3, $\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$.

Probability of complements of events

Theorem: For any event A, $\Pr(A^c) = 1 - \Pr(A)$.

Proof: By the definition of complements, $\Omega = A \cup A^c$; hence

$$\Pr(\Omega) = \Pr(A \cup A^c) = \Pr(A) + \Pr(A^c),$$

since A and A^c are disjoint.

(Here, we used $\Pr(A \cup A^c) = \Pr(A) + \Pr(A^c)$ from the previous slide.)

By Axiom 2, $\Pr(\Omega) = 1$; hence $\Pr(A) + \Pr(A^c) = 1$, which we can rewrite equivalently as $\Pr(A^c) = 1 - \Pr(A)$.

Range of probability

Theorem: For any event A, $0 \leq \Pr(A) \leq 1$.

Proof:
Axiom 1 already gives $\Pr(A) \geq 0$, so we are left to show $\Pr(A) \leq 1$.

A^c is an event, so using Axiom 1 again, we know that $\Pr(A^c) \geq 0$.
Thus:

$$\Pr(A) \leq \Pr(A) + \Pr(A^c) = 1.$$

(Here, we used $\Pr(A^c) = 1 - \Pr(A)$ from the previous slide.)

This is interesting!

- We intuitively "know" that any probability cannot exceed 1.
- This property $\Pr(A) \leq 1$ actually follows from $\Pr(A) \geq 0$ (Axiom 1), combined with Axiom 2 and Axiom 3!

Properties of Probability

Here is a summary of *some* properties of probability we covered:
(Note: Each property follows from the three probability axioms.)

- $\Pr(\varnothing) = 0$.
- For mutually exclusive events $A_1, \ldots, A_n$,
  $\Pr(A_1 \cup \cdots \cup A_n) = \Pr(A_1) + \cdots + \Pr(A_n)$.
- For any event A, $\Pr(A^c) = 1 - \Pr(A)$.
- For any event A, $0 \leq \Pr(A) \leq 1$.

There are **more** useful properties of probability covered in Chapter 1.5 of the course textbook. This is part of the assigned readings for this lecture, so please remember to read up!

Probability zero does NOT mean impossible!!

For a null event $\varnothing$, we know that $\Pr(\varnothing) = 0$.

However, $\Pr(A) = 0$ does NOT mean we must have $A = \varnothing$.

- **Example:** Randomly pick a number n in the range $[1, \infty)$. The probability that $n = 500$ is 0, but the event $\{n = 500\}$ is not a null event.
- **Strange Example from Physics:** There is zero probability that a black hole will spontaneously appear in front of you. (But it does not mean this event will never happen.[1])

Similarly, probability 1 does NOT mean "always occurs"!

- $\Pr(A) = 1$ does NOT imply that $A = \Omega$.
- If you randomly pick a number n as above, the probability that $n \neq 500$ is 1, but it does not mean that every number picked in every experiment must be not 500.

[1] Don't worry! Because of Hawking radiation, such a black hole will vanish almost instantaneously. Such black holes are called "micro black holes".

Equally likely outcomes

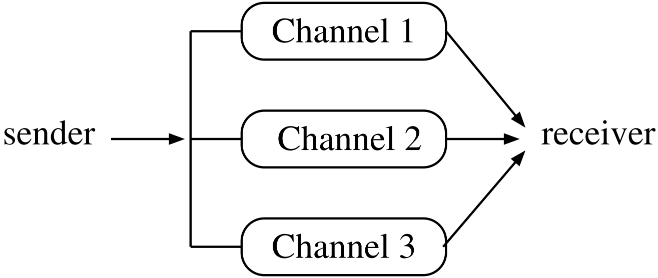
It is common to have an experiment that consists of finitely many outcomes, where each outcome occurs with an **equal probability**.

- Suppose there are n outcomes. This means there are n simple events. Let the n simple events be denoted by $A_1, \ldots, A_n$.
- Then the probability of *every* **simple event** is $\Pr(A_i) = \frac{1}{n}$.
- To figure out the probability of a **compount event**, we need to know how many outcomes there are in the event:

$$\Pr\begin{pmatrix}\text{compound}\\ \text{event}\end{pmatrix} = \frac{\text{number of outcomes in compound event}}{\text{total number of outcomes in the sample space}}.$$

Example 4: Communication Channels

A message is sent via three independent channels in parallel. It is received if any of the channel transmits the message successfully. The success rate (probability) of each channel is 0.5. Find the success rate of the three-channel system.



Example 4

The sample space of the three-channel system has 8 simple outcomes:

1. Ch_1: good, Ch_2: good, Ch_3: good.
2. Ch_1: good, Ch_2: good, Ch_3: bad.
3. Ch_1: good, Ch_2: bad, Ch_3: good.
4. Ch_1: good, Ch_2: bad, Ch_3: bad.
5. Ch_1: bad, Ch_2: good, Ch_3: good.
6. Ch_1: bad, Ch_2: good, Ch_3: bad.
7. Ch_1: bad, Ch_2: bad, Ch_3: good.
8. Ch_1: bad, Ch_2: bad, Ch_3: bad.

Example 4

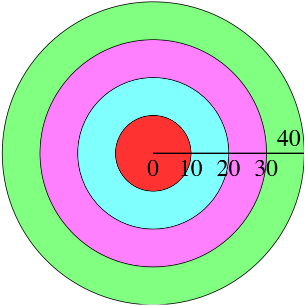
The three-channel system fails only in the case when outcome 8 occurs, so the failure rate is $\frac{1}{8}$.

Let A be the event representing successful transmission. Then:

$$\Pr(A) = 1 - \Pr(A^c) = 1 - \frac{1}{8} = \frac{7}{8}.$$

Example 5

A dart falls on any point of a target board (shown below) with equal probability.



What is the probability of the dart falling within the outermost ring? How about within the center red circle? What is the probability of the dart falling **exactly on** the circle of radius 20?

Example 5

What is the probability of the dart falling within the outermost ring?

Answer: $\dfrac{(\pi \cdot 40^2 - \pi \cdot 30^2)}{\pi \cdot 40^2} = \dfrac{7}{16}$

What is the probability of the dart falling within the center red circle?

Answer: $\dfrac{\pi \cdot 10^2}{\pi \cdot 40^2} = \dfrac{1}{16}$

What is the probability of the dart falling exactly on the circle of radius 20?

Answer: 0

Example 6 (Challenge of the day)

1. What is the probability that a randomly selected student, in a class of 40 students, is born on 1st April?
(Assume that no student has 29th February as his or her birthday.)

Answer: $\frac{1}{365}$ (The class size 40 is irrelevant.)

2. What is the probability that none of the students, in a class of 40 students, are born on 1st April?
(Again assume that no student has 29th February as his or her birthday.)

Answer: $\left(\frac{364}{365}\right)^{40} \approx 0.896$ (The class size 40 is crucial.)

Example 6 (continued)

3. What is the probability that no two students, in a class of 40 students, have the same birthday?
(Assume that no student has 29th February as his or her birthday.)

Solution:

- Student 1 can be born on any day; probability is 1.
- Student 2 can be born on any day other than Student 1's birthday; probability is $\frac{364}{365}$.
- Student 3 can be born on any day other than on the birthdays of Students 1 and 2; probability is $\frac{363}{365}$.
- Student 4 ...

Therefore, Pr(No two students have the same birthday) is

$$1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdots \frac{365 - 39}{365}.$$

Summary

- Experiment, sample space, event
- Set theory
- Axioms of Probability
- Properties of probability

For your **first Cohort Class**, you will learn more about counting methods and combinatorial methods, as well as learn how to solve probability questions similar to the birthday challenge questions.

Assigned readings for Week 1 Cohort Class: 1.6, 1.7, 1.8, 1.10