

# Introducing Probability

*Definitions, axioms, and examples*

Dr Andrew Mitchell 

*a.j.mitchell@ucl.ac.uk*

*Lecturer in AI and Machine Learning for Sustainable Construction*

2025-01-30

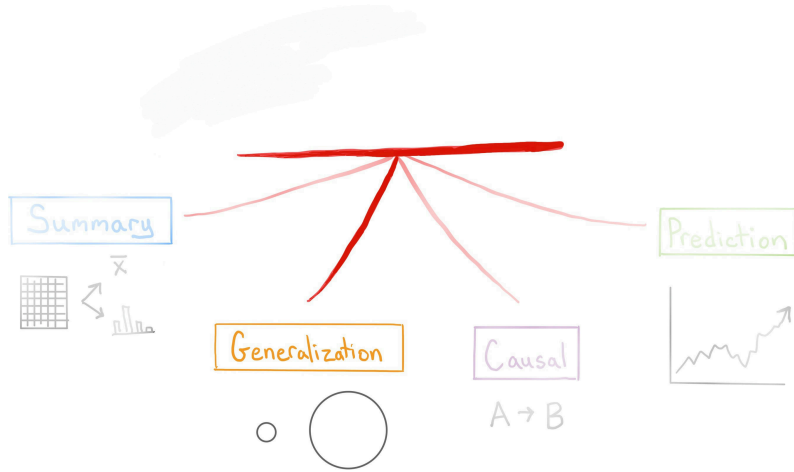
# Part 1: Introducing Probability

# Introducing Probability

## Why Study Probability?

- ▶ Quantify uncertainty
- ▶ Make valid generalizations
- ▶ Understand sampling variation
- ▶ Evaluate statistical claims

# Introducing Probability



## Real-World Applications:

- ▶ Election polling
- ▶ Clinical trials
- ▶ Sports analytics

# Introducing Probability

## ► Dating probabilities!

In an enormously entertaining paper written about a decade ago, the economist Peter Backus estimated his chance of finding a girlfriend on any given night in London at about 1 in 285,000 or 0.0000034%. As he writes, this is either depressing or cheering news for a person, depending on what you had estimated your chance to be before reading the paper and doing a similar computation for yourself. The interesting point in the paper was using a probabilistic argument (originally developed by the astronomer and astrophysicist Frank Drake to estimate the probability of extra-terrestrial civilizations) to think about his dating problems. Anyone can follow the arguments put forward by Backus, including his statements that use probability.

We all have some notion of chance or probability, and can ask questions like: - What is the chance you will get an A in Stat 20? (About 32%, based on last fall.) - What is the chance the 49ers will win the Super Bowl this year? (They are the favorites, with an implied probability

# Introducing Probability

of about 54.5%.) - What is the chance you will roll a double on your next turn to get out of jail while playing Monopoly? (One in six.) - What is the chance that Donald Trump will win the Presidential election? (About 47%.)

# De Méré's Paradox

## *The Games*

### **Game 1:**

- ▶ Roll a fair die 4 times
- ▶ Bet on at least one six
- ▶ De Méré's calculation:  $4 \cdot (1/6) = 2/3$

### **Game 2:**

- ▶ Roll two dice 24 times
- ▶ Bet on at least one double six
- ▶ De Méré's calculation:  $24 \cdot (1/36) = 2/3$

## *Simulation Results*

```
set.seed(123)
# Game 1: 1000 trials of 4 rolls
```

# De Méré's Paradox

```
die <- 1:6
game1 <- replicate(1000, any(sample(die, 4, replace = TRUE) == 6))
cat("Game 1 probability:", mean(game1), "\n")

# Game 2: 1000 trials of 24 rolls of pair
game2 <- replicate(1000, any(sample(1:36, 24, replace = TRUE) == 36))
cat("Game 2 probability:", mean(game2))
```

## *Key Insights*

1. De Méré's calculations were wrong
2. Simple multiplication doesn't work
3. Need proper probability theory
4. Simulation helps verify results



# De Méré's Paradox

In 17th century France, gamblers would bet on anything. In particular, they would bet on a fair six-sided die landing 6 at least once in four rolls. Antoine Gombaud, aka the Chevalier de Méré, was a gambler who also considered himself something of a mathematician. He computed the chance of a getting at least one six in four rolls as  $2/3(4 \cdot (1/6) = 4/6)$ . He won quite often by betting on this event, and was convinced his computation was correct. Was it?

The next popular dice game was betting on at least one double six in twenty-four rolls of a pair of dice. De Méré knew that there were 36 possible outcomes when rolling a pair of dice, and therefore the chance of a double six was  $1/36$ . Using this he concluded that the chance of at least one double six in 24 rolls was the same as that of at least one six in four rolls, that is,  $2/3 (24 \cdot 1/36)$ . He happily bet on this event (at least one double six in 24 rolls) but to his shock, lost more often than he won! What was going on?

# De Méré's Paradox

The simulation results show that De Méré's calculations were incorrect. By the end of this unit, you'll understand why and be able to conduct simulations like these yourself in R!

# Basic Probability Concepts

## 1. Experiment:

# Basic Probability Concepts

# Basic Probability Concepts

# Basic Probability Concepts

## 1. Experiment:

- ▶ Action involving chance
- ▶ Finite possible outcomes
- ▶ Example: coin toss, die roll

## 2. Outcome Space ():

# Basic Probability Concepts

# Basic Probability Concepts



# Basic Probability Concepts

## 1. Experiment:

- ▶ Action involving chance
- ▶ Finite possible outcomes
- ▶ Example: coin toss, die roll

## 2. Outcome Space ():

- ▶ Set of all possible outcomes
- ▶ Example for coin: = {Heads, Tails}
- ▶ Example for die: = {1, 2, 3, 4, 5, 6}

## 3. Event:

# Basic Probability Concepts

# Basic Probability Concepts

# Basic Probability Concepts

## 1. Experiment:

- ▶ Action involving chance
- ▶ Finite possible outcomes
- ▶ Example: coin toss, die roll

## 2. Outcome Space ():

- ▶ Set of all possible outcomes
- ▶ Example for coin: = {Heads, Tails}
- ▶ Example for die: = {1, 2, 3, 4, 5, 6}

## 3. Event:

- ▶ Collection of outcomes
- ▶ Subset of outcome space
- ▶ Example: “rolling an even number”
- ▶ Denoted by capital letters: A, B, C

## 4. Probability $P(A)$ :

# Basic Probability Concepts

- ▶ Measure of likelihood
- ▶ Between 0 and 1
- ▶ For equally likely outcomes:  $P(A) = k/n$
- ▶  $k$  = outcomes in event,  $n$  = total outcomes

First, let's establish some terminology:

**Experiment:** An action, involving chance, that can result in a finite number of possible outcomes (results of the experiment). For example, a coin toss is an experiment, and the possible outcomes are the coin landing heads or tails.

**Outcome space:** This is just a set. It is the collection of all the possible outcomes of an experiment is called an outcome space or sample space, and we denote it by the upper case Greek letter  $\Omega$  ("Omega"). For example, if we toss a coin, then the corresponding outcome space is  $\Omega = \{\text{Heads}, \text{Tails}\}$ . If we roll a die, then the corresponding outcome space  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . We will denote a set by enclosing the elements of the set in braces:  $\{ \}$ .

# Basic Probability Concepts

Event: A collection of outcomes as a result of the experiment being performed, perhaps more than once. For example, we could toss a coin twice, and consider the event of both tosses landing heads. We usually denote events by upper case letters from the beginning of the alphabet: A, B, C, .... An event is a subset of the outcome space, and we denote this by writing  $A \subset \Omega$ .

For any event A, we write the probability of A as  $P(A)$ .

# Equally Likely Outcomes

## Definition:

When all outcomes have same probability:  $P(\text{each outcome}) = \frac{1}{n}$

For event  $A$  with  $k$  outcomes:  $P(A) = \frac{k}{n}$

## Examples:

- ▶ Fair coin:  $P(H) = P(T) = 1/2$
- ▶ Fair die:  $P(\text{each face}) = 1/6$

## Visualization:

When all the possible outcomes in a finite outcome space of size  $n$  happen with the same probability, which is  $1/n$ .

# Equally Likely Outcomes

Let's say that there are  $n$  possible outcomes in the outcome space  $\Omega$ , and an event  $A$  has  $k$  possible outcomes out of those  $n$ . If all the outcomes are equally likely to happen (as in a die roll or coin toss), then we say that the probability of  $A$  occurring is  $k/n$ .

For example, suppose we toss a fair coin, and I ask you what is the chance of the coin landing heads. Like most people, you reply 50%. Why? Well... (you reply) there are two possible things that can happen, and if the coin is fair, then they are both equally likely, so the probability of heads is  $1/2$  or 50%.

Here, we have thought about an event (the coin landing heads), seen that there is one outcome in that event, and two outcomes in the outcome space, so we say the probability of the event,  $P(\text{Heads})$ , is  $1/2$ .



# Axioms of Probability

## The Three Axioms:

1. Non-negativity:
  - ▶  $P(A) \geq 0$  for any event  $A$
2. Total Probability:
  - ▶  $P(\Omega) = 1$  for sample space  $\Omega$
3. Addition Rule:
  - ▶ For mutually exclusive events:
  - ▶  $P(A \cup B) = P(A) + P(B)$

## Important Consequences:

1. Complement Rule:
  - ▶  $P(A) + P(A^c) = 1$
2. Impossible Event:
  - ▶  $P(\emptyset) = 0$

# Axioms of Probability

## 3. Probability Range:

►  $0 \leq P(A) \leq 1$

In order to compute the probabilities of events, we need to set some basic mathematical rules called axioms (which are intuitively clear if you think of the probability of an event as the proportion of the outcomes that are in it). There are three basic rules that will help us compute probabilities:

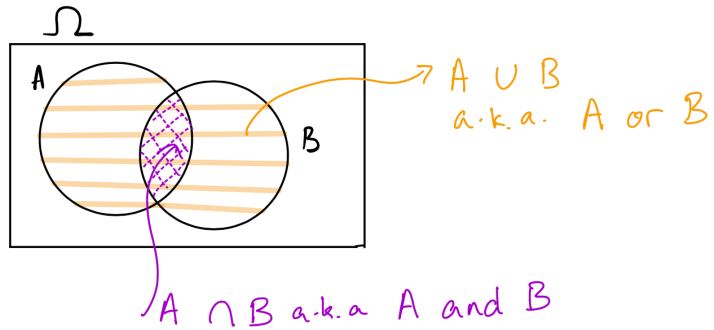
Axiom 1: The chance of any event is at least 0:  $P(A) \geq 0$  for any event  $A$ .

Axiom 2: The chance of an outcome being in  $\Omega$  is 1:  $P(\Omega) = 1$ . This is true because we can consider that the probability of  $\Omega$  is the number of outcomes in  $\Omega$  divided by  $n$ , which is  $n/n = 1$ .

# Axioms of Probability

Axiom 3: If A and B are mutually exclusive ( $A \cap B = \emptyset$ ), then  $P(A \cup B) = P(A) + P(B)$ . That is, for two mutually exclusive events, the probability that either of the two events might occur is the sum of their probabilities. This is called the addition rule.

# Set Operations & Venn Diagrams

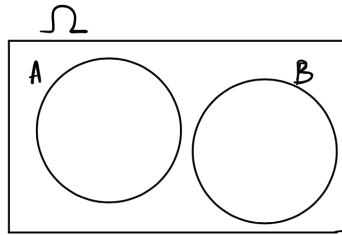


## Key Operations:

- ▶ Union ( $A \cup B$ ): “or”
- ▶ Intersection ( $A \cap B$ ): “and”
- ▶ Complement ( $A^c$ ): “not  $A$ ”

## Mutually Exclusive Events:

# Set Operations & Venn Diagrams



$$A \cap B = \emptyset$$

$$P(A \cap B) = 0$$

$$P(A \cup B) = P(A) + P(B)$$

- ▶ No outcomes in common
- ▶  $A \cap B = \emptyset$
- ▶  $P(A \cap B) = 0$
- ▶ Example: rolling a 6 AND a 1

Union of events: Given events A, B, we can define a new event called A or B, which consists of all the outcomes that are either in A or in B or in both. This is also written as  $A \cup B$ , read as “A union B”.

# Set Operations & Venn Diagrams

Intersection of events: Given events  $A$ ,  $B$ , we can define a new event called  $A$  and  $B$ , which consists of all the outcomes that are both in  $A$  and in  $B$ . This is also written as  $A \cap B$ , read as “ $A$  intersect  $B$ ”.

Mutually exclusive events: If two events  $A$  and  $B$  do not overlap, that is, they have no outcomes in common, we say that the events are mutually exclusive. If  $A$  and  $B$  are mutually exclusive, then we know that if one of them happens, the other one cannot. We denote this by writing  $A \cap B = \emptyset$  and read this as  $A$  intersect  $B$  is empty.

For example, if we are playing De Méré’s second game, the event  $A$  that we roll a pair of sixes and the event  $B$  that we roll a pair of twos cannot happen on the same roll. These events  $A$  and  $B$  are mutually exclusive.

# Examples: Fair Die

## *Setup*

### **Fair six-sided die:**

- ▶  $= \{1, 2, 3, 4, 5, 6\}$
- ▶ Each outcome:  $P = 1/6$

Let  $A = \text{“rolling even”}$

- ▶  $A = \{2, 4, 6\}$
- ▶  $P(A) = 3/6 = 1/2$

## *Probability Table*

Outcome	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

## *Simulation*

# Examples: Fair Die

```
set.seed(123)
die <- 1:6
rolls <- sample(die, 1000, replace = TRUE)
table(rolls)/1000
```

Consider rolling a fair six-sided die: six outcomes are possible so  $= \{1, 2, 3, 4, 5, 6\}$ . Since the die is fair, each outcome is equally likely, with probability  $= 1/6$ .

Let  $A$  be the event that an even number is rolled. Then the set  $A$  can be written  $\{2,4,6\}$ . Since all of these outcomes are equally likely:

$$P(A) = 1/6 + 1/6 + 1/6 = 3/6$$

The simulation shows how the empirical probabilities approach the theoretical probabilities with a large number of trials.



# R Implementation: Key Functions

## *sample()*

```
# Draw from box of tickets
die <- 1:6
# Two draws with replacement
sample(die, size = 2, replace = TRUE)
```

## *set.seed()*

```
# Control random number generation
set.seed(123)
sample(die, size = 2, replace = TRUE)
# Same result with same seed
set.seed(123)
sample(die, size = 2, replace = TRUE)
```

## *seq()*

# R Implementation: Key Functions

```
# Create sequence of numbers
# By value
seq(from = 1, by = 2, to = 9)
# By length
seq(from = 1, by = 2, length = 5)
```

Three useful functions for working with probability:

1. `sample()`: randomly picks out elements (items) from a vector
  - ▶ Arguments: `x` (vector to sample from), `size` (number of items), `replace` (whether to replace drawn items)
2. `set.seed()`: returns the random number generator to the point given by the seed number
  - ▶ Arguments: `n` (seed number to use)
  - ▶ Makes random results reproducible
3. `seq()`: creates a sequence of numbers

# R Implementation: Key Functions

- ▶ Arguments: from (start), by (increment), to/length (end/length)
- ▶ Useful for creating vectors of outcomes

# Simulating Probability: Examples

## Die Rolling Example:

```
set.seed(123)
# 1000 rolls
rolls <- replicate(1000,
  sample(1:6, 1, replace = TRUE))

# Plot distribution
data.frame(rolls) |>
  ggplot(aes(x = factor(rolls))) +
  geom_bar(aes(y = after_stat(prop)),
    fill = "blue", width = 0.98) +
  labs(x = "Die Face",
    y = "Proportion",
    title = "1000 Die Rolls")
```

# Simulating Probability: Examples

## Key Points:

1. Simulation Process:
  - ▶ Define outcomes
  - ▶ Set sample size
  - ▶ Use appropriate sampling
  - ▶ Calculate proportions
2. Long-run Behavior:
  - ▶ Empirical  $\approx$  Theoretical
  - ▶ Law of Large Numbers
  - ▶ More trials = Better estimate

Let's simulate rolling a die and counting how many times we see each face. If we roll it 6 times, we don't really expect to see each face exactly once, and as you can see in the simulation, the results vary.

# Simulating Probability: Examples

What about if we roll the die many more times? We should see each face about equally often. The important takeaway here is that we have a theoretical probability distribution of the outcomes, and we have what actually happens when we perform the experiment over and over. Eventually, the empirical distribution begins to look like the theoretical distribution.

# Summary

## 1. Foundations:

# Summary



# Summary

# Summary

## 1. Foundations:

- ▶ Experiments and outcomes
- ▶ Sample spaces and events
- ▶ Probability axioms
- ▶ Set operations

## 2. Key Concepts:

# Summary

# Summary

# Summary

## 1. Foundations:

- ▶ Experiments and outcomes
- ▶ Sample spaces and events
- ▶ Probability axioms
- ▶ Set operations

## 2. Key Concepts:

- ▶ Equally likely outcomes
- ▶ Mutually exclusive events
- ▶ Addition rule
- ▶ Complement rule

## 3. Implementation:

# Summary

# Summary

# Summary

## 1. Foundations:

- ▶ Experiments and outcomes
- ▶ Sample spaces and events
- ▶ Probability axioms
- ▶ Set operations

## 2. Key Concepts:

- ▶ Equally likely outcomes
- ▶ Mutually exclusive events
- ▶ Addition rule
- ▶ Complement rule

## 3. Implementation:

- ▶ `sample()` for random draws
- ▶ `set.seed()` for reproducibility
- ▶ `seq()` for creating sequences



# Summary

- ▶ Simulation for verification

## 4. Applications:

- ▶ Games of chance
- ▶ Real-world probabilities
- ▶ Statistical inference
- ▶ Decision making

In this lecture, we:

- ▶ Introduced equally likely outcomes and defined the outcome space of an experiment
- ▶ Using equally likely outcomes, defined the probability of an event as the ratio of the number of outcomes in the event to the number of total outcomes in the outcome space
- ▶ Wrote down the axioms (fundamental rules) of probability, after defining unions, intersections, and mutually exclusive events and Venn diagrams

# Summary

- ▶ In the “Ideas in Code” section, explored how to simulate probabilities using `sample()` and `replicate()`, and learned another useful function `seq()`