Introduction to Probability
Gambling at its core

- **16th century**
  - Cardano: *Books on Games of Chance*
  - First systematic treatment of probability

- **17th century**
  - Chevalier de Mere posed a problem to his friend Pascal. Which is likelier:
    - At least one 1 in four rolls of a single die?
    - At least one double 1 in 24 rolls of two dice?
  - Pascal corresponded with Fermat for several years, in attempt to resolve this puzzle

- **Early 18th century**
  - Bernoulli and De Moivre (*Doctrine of Chances*)
  - Early mathematical treatments of probability
Basic Definitions (grounded in set theory)

- An **random experiment** is the process of observing something uncertain.
  - A coin flip, a roll of a die, etc.
- An **elementary outcome** is a result of a random experiment.
- The **sample space** $\Omega$ is the set of all possible elementary outcomes.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Sample Space $\Omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toss a fair coin</td>
<td>$\Omega = {H, T}$</td>
</tr>
<tr>
<td>Toss two fair coins</td>
<td>$\Omega = {HH, HT, TH, TT}$</td>
</tr>
<tr>
<td>Roll a fair die</td>
<td>$\Omega = {1, 2, 3, 4, 5, 6}$</td>
</tr>
<tr>
<td>Roll a loaded die (25% chance of 1)</td>
<td>$\Omega = {1, 2, 3, 4, 5, 6}$</td>
</tr>
<tr>
<td>Roll two dice</td>
<td>$\Omega = {(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (2,6), \ldots, (6,1), (6,2), \ldots, (6,6)}$</td>
</tr>
</tbody>
</table>
Event Probabilities

- An **event** is a subset of the sample space (i.e., a set of elementary outcomes).
- The **probability** of an event is its likelihood of occurring.
- How do you compute probabilities?
  - First find the number of times the relevant elementary outcomes occur in the sample space.
  - Then, divide by the total number of elementary outcomes: i.e., the size of the sample space.

| Event Description                           | \( \Omega \)       | \( |\{H\}| / |\{H, T\}| = 1/2 \) | \( |\{HH\}| / |\{HH, HT, TH, TT\}| = 1/4 \) | \( |\{1\}| / |\{1, 2, 3, 4, 5, 6\}| = 1/6 \) | \( p(1) = .25 \) & \( p(\{2, 3, 4, 5, 6\}) = .75 \) | \( |\{2\}| / |\{2, 3, 4, 5, 6\}| (.75) = .15 \) |
|--------------------------------------------|---------------------|-----------------------------------|-------------------------------------------------|-------------------------------------------------|--------------------------------------------|--------------------------------------------|
| Toss a fair coin                           | \( \Omega = \{H, T\} \) |                                   |                                                 |                                                 |                                            |                                            |
| Toss two fair coins                        | \( \Omega = \{HH, HT, TH, TT\} \) |                                   |                                                 |                                                 |                                            |                                            |
| Toss a fair die                            | \( \Omega = \{1, 2, 3, 4, 5, 6\} \) |                                   |                                                 |                                                 |                                            |                                            |
| Toss a loaded die (25% chance of 1)        | \( \Omega = \{1, 2, 3, 4, 5, 6\} \) |                                   |                                                 |                                                 |                                            |                                            |
Fair Coin Tosses

- What is the probability of seeing a head when we flip a fair coin?
  - There are two elementary outcomes in the sample space, heads and tails ({H, T}).
  - The (sole) elementary outcome we are interested in is heads.
  - Thus, the probability is $\frac{1}{2}$.

- What is the probability of seeing a head when we flip two fair coins?
  - There are four elementary outcomes in the sample space: {HH, HT, TH, TT}.
  - The elementary outcomes (plural!) we are interested in are HH, HT, and TH.
  - Thus, the probability is $\frac{3}{4}$.

| Toss a fair coin, See heads | $\Omega = \{H, T\}$ | $|\{H\}| / |\{H, T\}| = 1/2$ |
|-----------------------------|---------------------|---------------------------------|
| Toss two fair coins See heads | $\Omega = \{HH, HT, TH, TT\}$ | $|\{HH, HT, TH\}| / |\{HH, HT, TH, TT\}| = 3/4$ |
Rolling the dice

- What is the probability of seeing a 1 when you roll a fair die?
  - There are six elementary outcomes in the sample space: \( \{1, 2, 3, 4, 5, 6\} \).
  - The elementary outcome we are interested in is 1.
  - Thus, the probability is \( \frac{1}{6} \).

- What is the probability of seeing a 2 when you roll a loaded die that favors 1?
  - There are six elementary outcomes in the sample space: \( \{1, 2, 3, 4, 5, 6\} \).
  - The probability of seeing a 1 is .25. The probability of seeing something else is .75.
  - The elementary outcome we are interested in is 2.
  - Thus, the probability is .15.

| Toss a fair die  | \( \Omega = \{1, 2, 3, 4, 5, 6\} \) | \(|\{1\}| / |\{1, 2, 3, 4, 5, 6\}| = \frac{1}{6} \) |
|-----------------|----------------------------------|------------------------------------------------|
| Toss a loaded die (25% chance of 1) | \( \Omega = \{1, 2, 3, 4, 5, 6\} \) | \( p(1) = .25 \) & \( p(\{2, 3, 4, 5, 6\}) = .75 \) \(|\{2\}| / |\{2, 3, 4, 5, 6\}| \cdot (.75) = .15 \) |
ICA

- In the game of Monopoly, a player rolls two dice. The player then adds their rolls together, and moves that many spaces.
- There is a Free Parking space, and if a player lands on it, they get a payout of some amount of (Monopoly) money.
- It is your turn, and you are 8 spaces away from the Free Parking space.
- What is the probability you land on Free Parking on your next turn?
## Enumerating the Sample Space

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
## Finding the Event of Interest

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
Solution

- There are 36 elementary outcomes in all.
- We can roll two dice that sum to 8 in five different ways.
  - The event of interest is \{(6,2), (5,3), (4,4), (3,5), (2,6)\}.
- Hence, the probability that you will land on Free Parking is 5/36!
The Axioms of Probability

- Probability is a branch of mathematics. As such, it is an axiomatic science.
- An axiom is a statement that is regarded as self-evidently true.
- In probability, there are 3 basic axioms:
  - Probabilities cannot be negative.
  - Probabilities cannot exceed 1.
    - Axioms 1 and 2 imply that probabilities are bounded between 0 and 1, inclusive.
  - The probability of the union of two disjoint events, \( A \) and \( B \), is the sum of their respective probabilities: i.e., \( P(A \text{ or } B) = P(A) + P(B) \)
    - In particular, the sum of the probabilities of all elementary outcomes is 1.
- We will check that our definition of probabilities satisfies these basic axioms: \( P(A) = \frac{|A|}{|\Omega|} \).
Axiom 1

Probabilities cannot be negative:

• Recall how we calculate probabilities: we count how many times an outcome of interest occurs, and then we divide that number by the sample size (i.e., the total number of possible outcomes).
• At worst, an event occurs 0 times (e.g., rolling a 7 on a six-sided die)!
• This implies that probabilities are bounded below by 0.
Axiom 2

Probabilities cannot exceed 1:

- The sample space includes all possible outcomes.
- The probability of something in the sample space occurring is 1!
- Impossible outcomes (outside the sample space) cannot occur, so no event has greater probability than the sample space.
Axiom 3

The probability of the union of two disjoint events, A and B, is the sum of their respective probabilities: i.e., \( P(A \text{ or } B) = P(A) + P(B) \):

- \( P(A \text{ or } B) = \frac{|A \cup B|}{|\Omega|} \)
- \( P(A) + P(B) = \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} \)
- Must show: \( \frac{|A \cup B|}{|\Omega|} = \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} \)
- But \( \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} = \left(\frac{|A| + |B|}{|\Omega|}\right) = \frac{|A \cup B|}{|\Omega|} \), because A and B are disjoint
Example of Axiom 3

- If A and B are mutually exclusive, then
  - $P(A \text{ or } B) = P(A) + P(B)$

- Example: Roll two fair coins
  - Event A: “two heads”
  - Event B: “two tails”
  - What is the probability of A or B?
  - $P(A \text{ or } B) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

| Event A       | \{HH\}         | $|\{HH\}| / |\{HH, HT, TH, TT\}| = 1/4$ |
|---------------|---------------|-------------------------------------|
| Event B       | \{TT\}        | $|\{TT\}| / |\{HH, HT, TH, TT\}| = 1/4$ |
Corollary of Axiom 3

The sum of the probabilities of all elementary outcomes must be 1:

- The sample space is the union of all elementary outcomes.
- Elementary outcomes do not overlap.
  - For example, you cannot roll both a 1 and a 2 at the same time!
- So, if we add up the probabilities of all the elementary outcomes, we should get the probability of the entire sample space.
- But we already know the probability of something in the sample space occurring is 1.
Laws of Probability

Definition of probability:

- \( P(A) = \frac{|A|}{|\Omega|} \)

Laws implies by the axioms:

- \( P(\emptyset) = 0 \)
- \( P(\Omega \setminus A) = 1 - P(A) \)
- If \( A \subseteq B \), then \( P(A) \leq P(B) \)
- Inclusion-exclusion principle
Inclusion-Exclusion Principle

- If A and B are not mutually exclusive, then
  - \( P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \)

- Example: Flip two fair coins
  - Event A: “Heads on the first coin”
  - Event B: “Tails on the second”
  - What is the probability of A or B?
  - \( P(A \text{ or } B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} \)

| Event A       | \{HH, HT\} | \(|\{HH, HT\}| / |\{HH, HT, TH, TT\}| = 1/2\) |
|---------------|------------|-----------------------------------|
| Event B       | \{HT, TT\} | \(|\{HT, TT\}| / |\{HH, HT, TH, TT\}| = 1/2\) |
| Event A and B | \{HT\}     | \(|\{HT\}| / |\{HH, HT, TH, TT\}| = 1/4\) |
ICA

Assume the following:
- There is a 60% chance of rain today
- There is a 75% chance of rain tomorrow
- There is a 20% chance of no rain either day

What is the probability it will rain today or tomorrow?
What is the probability it will rain today and tomorrow?
Define the relevant events

A: Rain today
B: Rain tomorrow
A or B: Rain today or tomorrow
A and B: Rain today and tomorrow
DeMorgan’s Law

¬A: No rain today
¬B: No rain tomorrow

¬A and ¬B: No rain today and no rain tomorrow
¬(¬A and ¬B) = A or B: Rain today or rain tomorrow
The probability of rain today or tomorrow

\[ P(A \text{ or } B) \]
\[ = P(\text{Rain today or rain tomorrow}) \]
\[ = 1 - P(\text{No rain today and no rain tomorrow}) \]
\[ = 1 - .2 \]
\[ = .8 \]
The probability of rain today and tomorrow

\[ P(A \text{ and } B) \]
\[ = P(A) + P(B) - P(A \text{ or } B) \]
\[ = P(\text{Rain today}) + P(\text{Rain tomorrow}) - P(\text{Rain today or tomorrow}) \]
\[ = .6 + .75 - .8 \]
\[ = .55 \]
Conditional Probability

- In 2016, Brown had an acceptance rate of 9%.
  - Does this mean that all applicants had an identical 9% chance of getting in?
  - Probably not! If you won the Nobel Peace Prize in your senior year of high school, then you probably had a higher likelihood of getting in.
- The probability of gaining weight given that you don’t exercise is probably higher than the probability of gaining weight if you do.
- Conditional probabilities are probabilities, given a (set of) conditions.
- We write $P(A|B)$, and say “the probability of A given B”
A Worked Example

- You roll two die, and record the sum of the rolls.
- Given that the first roll is even, what is the probability the sum is greater than or equal to 8?
Enumerating the Sample Space

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
# Conditioning the Sample Space

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
Finding the Event of Interest

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
Computing the Conditional Probability

- After conditioning on the first roll being even, there are 18 elementary outcomes.
- Among these outcomes only, we can roll two dice whose sum is at least 8 in nine different ways.
  - The event of interest is \{(2,6), (4,4), (4,5), (4,6), (6,2), (6,3), (6,4), (6,5), (6,6)\}.
- Hence, the conditional probability is $9/18 = ½!$
Formula for Conditional Probability

- \( P(A \mid B) = \frac{|A \text{ and } B|}{|B|} = \frac{P(A \text{ and } B)}{P(B)} \)
- Why? Because \( \frac{|A \text{ and } B|}{|B|} = \frac{|A \cap B|}{|\Omega|} / \frac{|B|}{|\Omega|} \)
- Notice, this implies the probability of \( P(A \text{ and } B) = P(A \mid B)P(B) \)
- Moreover, \( P(A \text{ and } B) = P(A \mid B)P(B) = P(B \mid A)P(A) \)
Independent vs. Dependent Events

Example 1

- Let A be the event that you get an A in this class
- Let B be the event that you take another computer science course
- Do these sound like independent events?
- Not so much

Example 2

- Let A be the event that you get an A in this class
- Let B be the event that I go to East Side Pockets for lunch
- These sound like independent events
Independent Events

- Let A be the event that you get an A in this class
  - Suppose this event has probability .8
- Let B be the event that I go to East Side Pockets for lunch
  - Suppose this event has probability .4
- What is the probability that you get an A in this class, given that I go to East Side Pockets for lunch?
- What I eat for lunch does not affect your grade, so the probability of you getting an A remains .8, regardless of my decision.
Independent Events (cont’d)

- Recall the following, from the definition of conditional probability:
  - $P(A \text{ and } B) = P(A | B)P(B)$
  - $P(A \text{ and } B) = P(B | A)P(A)$
- Mathematically, two events are independent if:
  - $P(A | B) = P(A)$
  - $P(B | A) = P(B)$
- Consequently, two events are independent if $P(A \text{ and } B) = P(A)P(B)$. 