# Probability

## CS 220 — Applied Discrete Mathematics

{April 30, May 5}, 2025

# Terminology

Everything you have learned about **counting** constitutes the basis for computing the probability of events to happen.

## Definitions

An **experiment** is a task, procedure, generative process, etc.

Each time the experiment is performed, it yields one **outcome**. If it is performed multiple times, it may yield a different outcome each time.

The set of possible outcomes is called the **sample space**.

An **event** is a subset of the sample space.

## Example

**Experiment:** Roll two dice (one red, one green).
**Sample space:** $\{(1,1),(1,2),(2,1),(1,3),(2,2),(3,1),\ldots,(6,6)\}$

"The dice sum to 7" corresponds to an event.
"The red result is larger than the green result" is another event.
"Both dice show 6s" is another event (containing a single outcome).

# Probability of an Event

If all outcomes in the sample space are equally likely, the following definition of probability applies:

**Definition (Probability)**

Suppose $S$ is a finite sample space of **equally likely** outcomes.

The **probability** of an event $E$, where $E \subseteq S$, is

$$p(E) = \frac{|E|}{|S|}$$

Probability values range from $0$ to $1$.

- A $0$ probability means the event will never happen.
- A $1$ probability means the event will always happen.

# Examples

$$p(E) = \frac{|E|}{|S|}$$

- An urn contains 4 green balls and 5 red balls.
  What is the probability that a ball chosen from the urn is green?

- What is the probability of winning the lottery “6/49” — that is, picking the correct set of 6 numbers out of 49?

# Examples

$$p(E) = \frac{|E|}{|S|}$$

- An urn contains 4 green balls and 5 red balls.
  What is the probability that a ball chosen from the urn is green?

  **Solution:** The sample space has $9$ possible outcomes.
  The event "chosen ball is green" contains $4$ of these outcomes.

  Therefore, the probability of this event is $4/9$ or approximately $44.44\%$.

- What is the probability of winning the lottery "6/49" — that is, picking the correct set of $6$ numbers out of $49$?

# Examples

$$p(E) = \frac{|E|}{|S|}$$

- An urn contains 4 green balls and 5 red balls.
  What is the probability that a ball chosen from the urn is green?

  **Solution:** The sample space has $9$ possible outcomes.
  The event "chosen ball is green" contains $4$ of these outcomes.

  Therefore, the probability of this event is $4/9$ or approximately $44.44\%$.

- What is the probability of winning the lottery "6/49" — that is, picking the correct set of $6$ numbers out of $49$?

  **Solution:** The sample space has $C(49, 6)$ possible outcomes.
  Only $1$ of these outcomes actually wins the lottery.

$$p(Win) = \frac{1}{C(49, 6)} = \frac{1}{13\,983\,816}$$

# Complementary Event

## Definition (Complementary Event)

Let $E$ be an event in a sample space $S$.

Its **complementary event**, written $\overline{E}$, is defined as $S - E$, meaning "$E$ does not happen". Its probability is the following:

$$p(\overline{E}) = 1 - p(E)$$

The probability of $\overline{E}$ can be calculated from the definition:

$$p(\overline{E}) = p(S - E) = \frac{|S - E|}{|S|} = \frac{|S| - |E|}{|S|} = 1 - \frac{|E|}{|S|} = 1 - p(E)$$

This rule is useful if it is easier to determine the probability of the complementary event than the probability of the event itself.

# Examples

$$p(E) = 1 - p(\overline{E})$$

- A sequence of 10 bits is randomly generated.
  What is the probability that at least one of these bits is zero?

- In a group of 36 people, what is the probability that at least 2 of them have the same birthday? (Assume all birthdays are equally likely.)

# Examples

$p(E) = 1 - p(\overline{E})$

- A sequence of 10 bits is randomly generated.
  What is the probability that at least one of these bits is zero?

  **Solution:** The sample space $S$ has $2^{10} = 1024$ outcomes.

  Let $E$ refer to the event "at least one of the bits is zero".

  The complementary event $\overline{E}$ means "none of the bits is zero".
  It has only one outcome, namely 11111 11111. Thus $p(\overline{E}) = 1/1024$.

  $$p(E) = 1 - p(\overline{E}) = 1 - \frac{1}{1024} = \frac{1023}{1024}$$

- In a group of 36 people, what is the probability that at least 2 of them have the same birthday? (Assume all birthdays are equally likely.)

# Examples

$p(E) = 1 - p(\overline{E})$

- A sequence of 10 bits is randomly generated.
  What is the probability that at least one of these bits is zero?

  **Solution:** The sample space $S$ has $2^{10} = 1024$ outcomes.

  Let $E$ refer to the event "at least one of the bits is zero".

  The complementary event $\overline{E}$ means "none of the bits is zero".
  It has only one outcome, namely 11111 11111. Thus $p(\overline{E}) = 1/1024$.

  $$p(E) = 1 - p(\overline{E}) = 1 - \frac{1}{1024} = \frac{1023}{1024}$$

- In a group of 36 people, what is the probability that at least 2 of them have the same birthday? (Assume all birthdays are equally likely.)

  **Solution:** The sample space $S$ contains all possibilities for the birthdays of the 36 people, so $|S| = 365^{36}$. Event $E$ is "at least 2 people share a birthday".

  Consider $\overline{E}$: "all 36 people have a different birthday".
  The event $\overline{E}$ contains $P(365, 36)$ outcomes ($365$ possibilities for the first person's birthday, $364$ for the second, and so on).

  $$p(E) = 1 - p(\overline{E}) = 1 - \frac{P(365,36)}{365^{36}} \approx 1 - 0.168 \approx 0.832, \text{ or } 83.2\%$$

# Unions of Events

Let $E_1$ and $E_2$ be events in the sample space $S$.

If $E_1$ and $E_2$ are disjoint, then we have

$$p(E_1 \cup E_2) = p(E_1) + p(E_2)$$

Otherwise, we have

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

Does this remind you of something?

# Unions of Events

Let $E_1$ and $E_2$ be events in the sample space $S$.

If $E_1$ and $E_2$ are disjoint, then we have

$$p(E_1 \cup E_2) = p(E_1) + p(E_2)$$

Otherwise, we have

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

Does this remind you of something?

Of course, the sum rule and inclusion-exclusion.

# Example: Unions of Events

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

What is the probability that a positive integer selected at random from the set of integers $\{1, \ldots, 100\}$ is divisible by 2 or 5?

# Example: Unions of Events

$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$

What is the probability that a positive integer selected at random from the set of integers $\{1, \dots, 100\}$ is divisible by 2 or 5?

**Solution:**

Let $E_2$ be "divisible by 2"; that is, $E_2 = \{2, 4, 6, \dots, 100\}$, and so $|E_2| = 50$
Let $E_5$ be "divisible by 5"; that is, $E_5 = \{5, 10, \dots, 100\}$, and so $|E_5| = 20$.

Then $p(E_2) = 0.5$ and $p(E_5) = 0.2$. But these events overlap:

What is $E_2 \cap E_5$?

# Example: Unions of Events

$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$

What is the probability that a positive integer selected at random from the set of integers $\{1, \ldots, 100\}$ is divisible by 2 or 5?

**Solution:**

Let $E_2$ be "divisible by 2"; that is, $E_2 = \{2, 4, 6, \ldots, 100\}$, and so $|E_2| = 50$
Let $E_5$ be "divisible by 5"; that is, $E_5 = \{5, 10, \ldots, 100\}$, and so $|E_5| = 20$.

Then $p(E_2) = 0.5$ and $p(E_5) = 0.2$. But these events overlap:

What is $E_2 \cap E_5$? It is "divisible by both 2 and 5" which is just "divisible by 10".
So $E_2 \cap E_5 = \{10, 20, 30, \ldots, 100\}$, and so $|E_2 \cap E_5| = 10$, so $p(E_2 \cap E_5) = 0.1$

Now we can calculate using inclusion-exclusion:

$$p(E_2 \cup E_5) \;=\; p(E_2) + p(E_5) - p(E_2 \cap E_5) \;=\; 0.5 + 0.2 - 0.1 \;=\; 0.6$$

# Example: Unions of Events

$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$

What is the probability that a positive integer selected at random from the set of integers $\{1, \ldots, 100\}$ is divisible by 2 or 5?

**Solution:**

Let $E_2$ be "divisible by 2"; that is, $E_2 = \{2, 4, 6, \ldots, 100\}$, and so $|E_2| = 50$
Let $E_5$ be "divisible by 5"; that is, $E_5 = \{5, 10, \ldots, 100\}$, and so $|E_5| = 20$.

Then $p(E_2) = 0.5$ and $p(E_5) = 0.2$. But these events overlap:

What is $E_2 \cap E_5$? It is "divisible by both 2 and 5" which is just "divisible by 10".
So $E_2 \cap E_5 = \{10, 20, 30, \ldots, 100\}$, and so $|E_2 \cap E_5| = 10$, so $p(E_2 \cap E_5) = 0.1$

Now we can calculate using inclusion-exclusion:

$$p(E_2 \cup E_5) \;=\; p(E_2) + p(E_5) - p(E_2 \cap E_5) \;=\; 0.5 + 0.2 - 0.1 \;=\; 0.6$$

We get the same result if we use inclusion-exclusion to calculate $|E_2 \cup E_5|$:

$$p(E_2 \cup E_5) \;=\; \frac{|E_2 \cup E_5|}{|S|} \;=\; \frac{50 + 20 - 10}{100} \;=\; 0.6$$

# Discrete Probability Distributions

What if the outcomes of an experiment are not equally likely?

**Definition (Discrete Probability Distribution)**

Let $S$ be a finite sample space.

A **discrete probability distribution** directly assigns a probability to every outcome in $S$. It must satisfy two conditions:

1. $0 \leq p(s) \leq 1$ for every $s \in S$, and
2. $\sum_{s \in S} p(s) = 1$.

If we consider the distribution as the function $p : S \to [0, 1]$, then that function is called a **probability mass function**.

If an experiment on $S$ has the discrete probability distribution $p$, then the probability of an event $E \subseteq S$ is defined as follows:

$$p(E) = \sum_{s \in E} p(s)$$

# Notation

In this course, we'll generally talk about only one experiment at a time. Each experiment has

- a sample space $S$, and
- a (discrete) probability distribution on $S$

We will typically "overload" the notation $p(\ )$:

- $p(s)$ where $s \in S$ — the probability of the single outcome $s$
- $p(E)$ where $E \subseteq S$ — the probability of the event $E$
- $p(E \mid F)$ where $E, F \subseteq S$ — the conditional probability of $E$ given $F$

All implicitly depend on the experiment and its probability distribution.

Other courses and literature might use slightly different notations, especially to talk about multiple experiments at a time.

# Uniform Distribution

## Definition (Uniform Distribution)

Let $S$ be a finite sample space. The **uniform distribution** on $S$ is defined by

$$p(s) = \frac{1}{|S|} \qquad \text{for all } s \in S$$

(Does this satisfy the conditions on discrete probability distributions?)

# Uniform Distribution

## Definition (Uniform Distribution)

Let $S$ be a finite sample space. The **uniform distribution** on $S$ is defined by

$$p(s) = \frac{1}{|S|} \qquad \text{for all } s \in S$$

(Does this satisfy the conditions on discrete probability distributions?)

If an experiment on $S$ has a uniform distribution $p$,
then the probability of an event $E \subseteq S$ is

$$p(E) = \sum_{s \in E} p(s) = \sum_{s \in E} \frac{1}{|S|} = \frac{|E|}{|S|}$$

In other words, all of our examples so far have been assuming that the sample space is uniformly distributed ("outcomes are equally likely").

# Example: A Non-Uniform Distribution (1)

A die is biased so that the number 3 appears twice as often as every other number. What are the probabilities of all possible outcomes?

# Example: A Non-Uniform Distribution (1)

A die is biased so that the number 3 appears twice as often as every other number. What are the probabilities of all possible outcomes?

**Solution:** There are 6 possible outcomes: $S = \{s_1, \dots, s_6\}$.
Let $p$ be the discrete probability distribution for this experiment.

$$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6) = x \qquad p(s_3) = 2x$$

# Example: A Non-Uniform Distribution (1)

A die is biased so that the number 3 appears twice as often as every other number. What are the probabilities of all possible outcomes?

**Solution:** There are 6 possible outcomes: $S = \{s_1, \ldots, s_6\}$.
Let $p$ be the discrete probability distribution for this experiment.

$$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6) = x \qquad p(s_3) = 2x$$

Since the probabilities must add up to 1, we have:

$$1 = \sum_{s \in S} p(s) = 7x$$

And therefore:

$$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6) = \frac{1}{7} \qquad p(s_3) = \frac{2}{7}$$

# Example: A Non-Uniform Distribution (2)

$$p(E) = \sum_{s \in E} p(s)$$

For the biased die from the previous example, what is the probability that an odd number appears when we roll the die?

# Example: A Non-Uniform Distribution (2)

$p(E) = \sum_{s \in E} p(s)$

For the biased die from the previous example, what is the probability that an odd number appears when we roll the die?

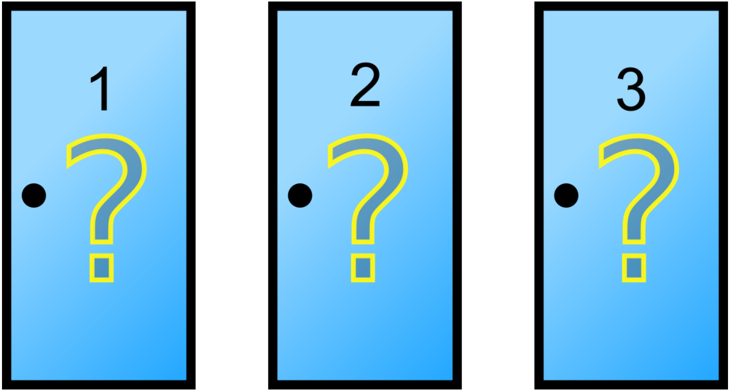
**Solution:** Let $E_{odd} = \{s_1, s_3, s_5\}$. We want to calculate $p(E_{odd})$:

$$p(E_{odd}) = \sum_{s \in E_{odd}} p(s) = \frac{1}{7} + \frac{2}{7} + \frac{1}{7} = \frac{4}{7} \approx 57.14\%$$

# Conditional Probability

# The Monty Hall Problem

You are on a TV show.[1] There are three doors, and you must pick one. Behind one door is a brand new **car**. Behind the other two are **goats**.




[1] https://en.wikipedia.org/wiki/Monty_Hall_problem

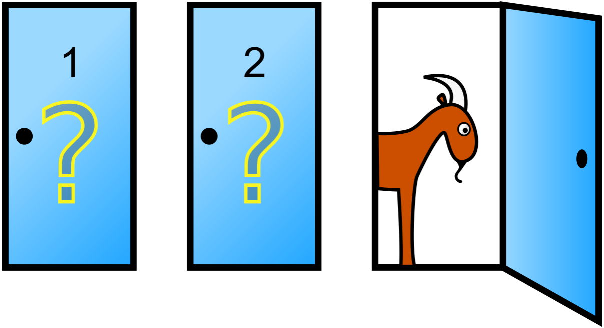
# The Monty Hall Problem — The Twist

You pick a door; let's say it's Door #1.

The host, who knows which doors have which prizes, must open one of the other doors, revealing a goat; let's say it's Door #3.

The host then says, "Do you want to stay with #1 or switch to #2?"

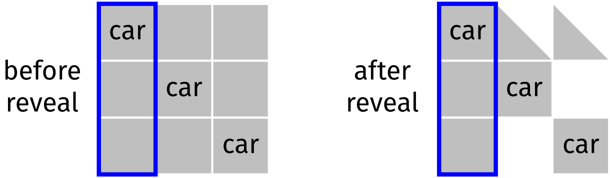
Should you stay or should you switch? Does it matter?



# The Monty Hall Problem — Solution

The answer is, surprisingly, yes you should switch!

Here's how possibilities change, **given** that you picked Door #1 and then Monty revealed a door with a goat:




The **effective** sample space has shrunk!

| Door #1 | Door #2 | Door #3 | Revealed | Stay | Switch |
|---|---|---|---|---|---|
| Car | Goat | Goat | 2 or 3 | Car | Goat |
| Goat | Car | Goat | 3 | Goat | Car |
| Goat | Goat | Car | 2 | Goat | Car |

# Conditional Probability

Suppose that we toss a coin 3 times.
What is the probability that we get tails at least twice?

Solution:

- There are $2^3 = 8$ different outcomes.
- Tails at least twice: $\{\text{HTT}, \text{THT}, \text{TTH}, \text{TTT}\}$
- So $\frac{4}{8} = \frac{1}{2}$. Equivalently, $\frac{C(3,2)+C(3,3)}{2^3} = \frac{3+1}{8} = \frac{1}{2}$.

What is the probability that we get tails at least twice
*if the first toss landed heads*?

Solution:

- If the first toss is H, the possible outcomes are $\{\text{HHH}, \text{HHT}, \text{HTH}, \text{HTT}\}$.
- Tails occurs at least twice in only one of those: $\{\text{HTT}\}$.
- So the conditional probability is $\frac{1}{4}$.

# Conditional Probability

**Definition (Conditional Probability)**

Let $S$ be a sample space and let $E$ and $F$ be events in $S$.

The **conditional probability** of $E$ given $F$, written $p(E \mid F)$, is defined as

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

It is the probability that $E$ occurs assuming that $F$ occurs.

That is, to measure the conditional probability of $E$ given $F$:

- we use $F$ as the effective sample space (it must be possible!)
- we must “trim” $E$ down to the parts that fit in $F$: that is $E \cap F$

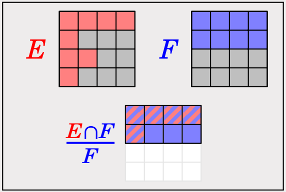
# Example: Calculating Conditional Probability

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

What is the probability that a random string of 4 bits contains at least 2 consecutive zeroes, given that its first bit is a zero?

**Solution:**

- $S$ is "bit strings of length 4"; there are $2^4 = 16$ of them
- $E$ is "bit string contains at least two consecutive zeroes"
- $F$ is "first bit of the string is a zero"
- $E \cap F = \{0000, 0001, 0010, 0011, 0100\}$
- $p(E \cap F) = \frac{|E \cap F|}{|S|} = \frac{5}{16}$
- $p(F) = \frac{|F|}{|S|} = \frac{8}{16}$
- $p(E|F) = \frac{5/16}{8/16} = \frac{5}{8} = 0.625$



# Using Conditional Probability

From the definition of conditional probability:

$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$

we can rewrite it in the following form:

$$p(A \cap B) = p(A) \cdot p(B|A)$$

which we can interpret as follows:

> *The probability that both $A$ and $B$ both happen is*
> *the probability that, first, $A$ happens,*
> *times the conditional probability that $B$ happens given $A$.*

It is not necessary that $A$ causes $B$.
It is not even necessary that $A$ happens before $B$ in time!

# Example: Using Conditional Probability

$$p(A \cap B) = p(A) \cdot p(B \mid A)$$

A gambler has a fair die and a biased die in her pocket.
(Recall, the biased die rolls 3 twice as often as any other number.)

She takes out a random die for a game. We can't tell which one.

How likely is it that she rolls a 3?

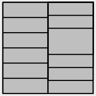
# Example: Using Conditional Probability

$p(A \cap B) = p(A) \cdot p(B \mid A)$

A gambler has a fair die and a biased die in her pocket.
(Recall, the biased die rolls 3 twice as often as any other number.)

She takes out a random die for a game. We can't tell which one.

How likely is it that she rolls a 3?

$S$ is 

**Solution:**

- The sample space consists of outcomes like "used fair die, rolled a 1", "used biased die, rolled a 6", etc. But its distribution is non-uniform!

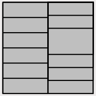
# Example: Using Conditional Probability

$p(A \cap B) = p(A) \cdot p(B \mid A)$

A gambler has a fair die and a biased die in her pocket.
(Recall, the biased die rolls 3 twice as often as any other number.)

She takes out a random die for a game. We can't tell which one.

How likely is it that she rolls a 3?

$S$ is 

**Solution:**

- The sample space consists of outcomes like "used fair die, rolled a 1", "used biased die, rolled a 6", etc. But its distribution is non-uniform!
- Let $F$ be "used fair die" and $\overline{F}$ be "used biased die". So $p(F) = \frac{1}{2}, p(\overline{F}) = \frac{1}{2}$.

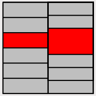
# Example: Using Conditional Probability

$p(A \cap B) = p(A) \cdot p(B \mid A)$

A gambler has a fair die and a biased die in her pocket.
(Recall, the biased die rolls 3 twice as often as any other number.)

She takes out a random die for a game. We can't tell which one.

How likely is it that she rolls a 3?

$E \subseteq S$ is 

**Solution:**

- The sample space consists of outcomes like "used fair die, rolled a 1", "used biased die, rolled a 6", etc. But its distribution is non-uniform!
- Let $F$ be "used fair die" and $\overline{F}$ be "used biased die". So $p(F) = \frac{1}{2}, p(\overline{F}) = \frac{1}{2}$.
- Let $E$ be "rolled a 3".

# Example: Using Conditional Probability

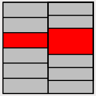
$p(A \cap B) = p(A) \cdot p(B \mid A)$

A gambler has a fair die and a biased die in her pocket.
(Recall, the biased die rolls 3 twice as often as any other number.)

She takes out a random die for a game. We can't tell which one.

How likely is it that she rolls a 3?

$E \subseteq S$ is



**Solution:**

- The sample space consists of outcomes like "used fair die, rolled a 1", "used biased die, rolled a 6", etc. But its distribution is non-uniform!
- Let $F$ be "used fair die" and $\overline{F}$ be "used biased die". So $p(F) = \frac{1}{2}, p(\overline{F}) = \frac{1}{2}$.
- Let $E$ be "rolled a 3". It can be split into two parts: $E = (E \cap F) \cup (E \cap \overline{F})$.
  That is, "rolled a 3" could be "rolled a 3 with fair die" or "rolled a 3 with biased die".

$$p(E \cap F) = p(F) \cdot p(E \mid F) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$$

$$p(E \cap \overline{F}) = p(\overline{F}) \cdot p(E \mid \overline{F}) = \frac{1}{2} \cdot \frac{2}{7} = \frac{1}{7}$$

- Then $p(E) = p(E \cap F) + p(E \cap \overline{F}) = \frac{1}{12} + \frac{1}{7} = \frac{19}{84} \approx 0.226$.
  This is the **marginal probability** of $E$ — the **total** over all the possible "causes".

# Inference

# Inference

You are standing in a dark kitchen, in front of the sink.

You can barely see two switches on the wall. One is for the lights, and one is for the garbage disposal, but you don't know which is which.

You pick a switch at random.

What is the probability that you chose the light switch?

# Inference

You are standing in a dark kitchen, in front of the sink.

You can barely see two switches on the wall. One is for the lights, and one is for the garbage disposal, but you don't know which is which.

You pick a switch at random.

What is the probability that you chose the light switch?

---

You flip the switch that you chose and the lights turn on.

What is the probability that you chose the light switch?

# Inference

You are standing in a dark kitchen, in front of the sink.

You can barely see two switches on the wall. One is for the lights, and one is for the garbage disposal, but you don't know which is which.

You pick a switch at random.

What is the probability that you chose the light switch? **prior** probability

---

You flip the switch that you chose and the lights turn on. **evidence**

What is the probability that you chose the light switch? **posterior** prob.

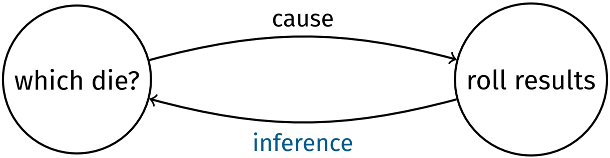
In this example, the evidence is conclusive. What if it isn't?

# Probabilistic Inference

A gambler has both a fair die and a biased die in her pocket.
(Recall, the biased die rolls 3 twice as often as any other number.)

She selects a random die for a game. We can't tell which one.
She rolls the die, and the result is a 3.

How likely is it that she is using the fair die?



# Probabilistic Inference

- Let $F$ be the event that the gambler selects the fair die.
  Then $\overline{F}$ is the event that she selects the biased die.

$$p(F) = \frac{1}{2} \qquad p(\overline{F}) = \frac{1}{2}$$

- Let $X$ be the event that the die comes up 3.
  We know the conditional probabilities of $X$ given $F$ and $\overline{F}$:

$$p(X|F) = \frac{1}{6} \qquad p(X|\overline{F}) = \frac{2}{7}$$

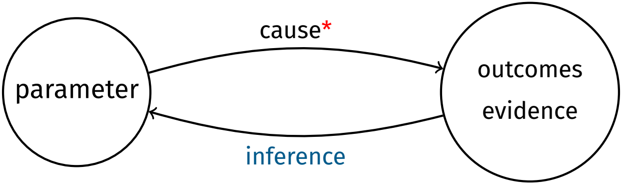
- We **want** the conditional probability of $F$ given $X$ — that is, $p(F|X)$!
  That is, we have **observed** a roll result of 3 (event $X$).
  Given that **evidence**, we want to **infer** what die she is using.

(Note: The probability $p(F)$ doesn't *change*; it refers to the probability at an earlier point in time, before the die was actually chosen.)

# Bayes' Theorem

**Bayes' Theorem** provides a way to reason about the likelihood that the die is biased based on the evidence of the outcomes.

Bayes' Theorem is the cornerstone of many algorithms in machine learning where the goal is to determine the likelihood of some event based on data obtained from observations.



# Bayes' Theorem

## Bayes' Theorem

Let $A$ and $B$ be events. Then

$$p(A \mid B) = \frac{p(A) \cdot p(B \mid A)}{p(B)}$$

Bayes' Theorem is usually applied when $A$ is a "hidden cause" and $B$ is an "observable effect". Then we say

- $p(A)$ is the **prior probability** of $A$.
- $p(B \mid A)$ is the **conditional probability** of $B$ given $A$.
- $p(B)$ is the **marginal probability** of $B$, sometimes* calculated by

$$p(B) = p(B \cap A) + p(B \cap \overline{A}) = p(A) \cdot p(B \mid A) + p(\overline{A}) \cdot p(B \mid \overline{A}))$$

- $p(A \mid B)$ is the **posterior probability** of $A$ given $B$.

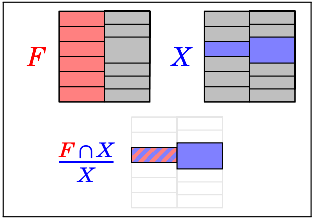
# Using Bayes' Theorem

The prior probability of $F$ ("fair die"):

$$p(F) = \frac{1}{2}$$

The conditional probability of $X$ ("rolled 3") given $F$:

$$p(X|F) = \frac{1}{6}$$




The marginal probability of $X$:

$$p(X) = p(F) \cdot p(X|F) + p(\overline{F}) \cdot p(X|\overline{F}) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{2}{7} = \frac{19}{84}$$

Then by Bayes' Theorem the posterior probability of $F$ given $X$ is

$$p(F|X) = \frac{p(F) \cdot p(X|F)}{p(X)} = \frac{1/2 \cdot 1/6}{19/84} = \frac{7}{19} \approx 0.37$$

# Exercise: Bayes' Theorem

$$p(A \mid B) = \frac{p(B \mid A) \cdot p(A)}{p(B)}$$

Suppose an elven paladin has one 6-sided die, one 10-sided die, and one 20-sided die. The paladin selects a die at random, rolls it out of our view, and reports a 2.

How likely is it that the 6-sided die was selected?

**Hints:**

- There are *three* choices for the die this time. Let the events be called $D_6, D_{10}$, and $D_{20}$. They are disjoint, and $D_6 \cup D_{10} \cup D_{20} = S$.
- Let $X$ be the event "rolled a 2".

# Independent Events

# Independent Events

## Example

A die is rolled.

- What is the probability that its result is a multiple of 3?

- What is the probability that its result is a multiple of 3, given that the result is even?

# Independent Events

## Example

A die is rolled.

- What is the probability that its result is a multiple of 3?
  Solution: $\frac{2}{6} = \frac{1}{3}$.
- What is the probability that its result is a multiple of 3, given that the result is even?
  Solution: $\frac{1}{3}$.

Even vs odd does not affect whether the result is a multiple of 3.

We say these events are independent — despite being based on the same "physical action"!

# Independent Events

## Definition (Independent)

Let $E$ and $F$ be events in $S$. Then $E$ and $F$ are **independent** if and only if

$$p(E \cap F) = p(E) \cdot p(F)$$

Equivalently:

$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{p(E) \cdot p(F)}{p(F)} = p(E)$$

$$p(F|E) = \frac{p(F \cap E)}{p(E)} = \frac{p(F) \cdot p(E)}{p(E)} = p(F)$$

# Independent Bernoulli Trials

## Definition (Bernoulli Trial)

A **Bernoulli trial** is an experiment whose set of outcomes is partitioned into an event labeled “success” and an event labeled “failure”.

## Examples

- Experiment: Flipping a coin.
  Success = $\{\text{H}\}$, failure = $\{\text{T}\}$.
- Experiment: Rolling a die.
  Success = $\{2, 4, 6\}$, failure = $\{1, 3, 5\}$.
- Experiment: Rolling a die.
  Success = $\{1, 2, 3, 4, 5\}$, failure = $\{6\}$.
- Experiment: Randomly picking a number from $\{1, \ldots, 100\}$.
  Success = “prime”, failure = “not prime”.

# Example: Bernoulli Trials

When we roll a die, we consider a result of 1–5 a success, and 6 is a failure.

- If we roll the die 3 times, what is the probability we get a success, a failure, and then another success?

# Example: Bernoulli Trials

When we roll a die, we consider a result of 1–5 a success, and 6 is a failure.

- If we roll the die 3 times, what is the probability we get a success, a failure, and then another success?

  **Solution:** Each roll is independent, so

$$p(sfs) = p(s) \cdot p(f) \cdot p(s) = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{216}$$

# Example: Bernoulli Trials

When we roll a die, we consider a result of 1–5 a success, and 6 is a failure.

- If we roll the die 3 times, what is the probability we get a success, a failure, and then another success?

  **Solution:** Each roll is independent, so

$$p(sfs) = p(s) \cdot p(f) \cdot p(s) = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{216}$$

- What is the probability of 2 successes and 1 failure, in any order?

# Example: Bernoulli Trials

When we roll a die, we consider a result of 1–5 a success, and 6 is a failure.

- If we roll the die 3 times, what is the probability we get a success, a failure, and then another success?

  **Solution:** Each roll is independent, so

$$p(sfs) = p(s) \cdot p(f) \cdot p(s) = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{216}$$

- What is the probability of 2 successes and 1 failure, in any order?

  **Solution:** Generalizing from the previous solution, the probability of any sequence of 2 successes and 1 failure is $\frac{25}{216}$. But there are $3$ ways to choose which 2 of the 3 trials are successful.

$$p(2s, 1f) = p(ssf) + p(sfs) + p(fss) = 3 \cdot \frac{25}{216} = \frac{25}{72}$$

# Bernoulli Trials

## Bernoulli Trials

The probability of $k$ successes in $n$ independent Bernoulli trials is

$$C(n,k) \cdot s^k(1-s)^{n-k}$$

where $s$ is the probability of success in each trial.

Let $f = 1 - s$. Apply the Binomial Theorem to $(s+f)^n$:

$$(s+f)^n = 1s^n f^0 + \cdots + C(n,k) \cdot s^k f^{n-k} + \cdots + 1s^0 f^n$$

# Random Variables

# Random Variables

We want to do math on experiments, but some outcomes are not numbers.

So we introduce an indirection: the random variable.

### Definition (Random Variable)

Let $S$ be the sample space for an experiment.

A **random variable** is a function $S \to \mathbb{R}$.

That is, a random variable assigns a real number to each outcome.

**Note:** A random variable is a function, not a variable, and it is not random, but maps "random" results from experiments onto real numbers in a well-defined manner.

# Examples: Random Variables

- **Experiment:** Two dice are rolled. An outcome is a pair of numbers 1–6.

  Suppose we're interested in the sum of the faces of the two dice.

  **Random Variable:** $X(a, b) = a + b$.

- **Experiment:** A card is drawn from a standard deck. An outcome is a pair of face and suit, like $9\heartsuit$ or $\mathrm{K}\spadesuit$. Suppose we're counting cards (Hi-Lo).

  **Random Variable:** $X(f, s) = \begin{cases} +1 & \text{if } f \in \{2, 3, 4, 5, 6\} \\ 0 & \text{if } f \in \{6, 7, 8, 9\} \\ -1 & \text{if } f \in \{10, \mathrm{J}, \mathrm{Q}, \mathrm{K}, \mathrm{A}\} \end{cases}$

- **Experiment:** A single die is rolled. An outcome is a number 1–6.

  If it shows 1–5, then player A wins \$1 from player B.
  If it shows 6, then player B wins \$5 from player A.

  **Random Variable:** $X(d) = \begin{cases} 1 & \text{if } d \in \{1, 2, 3, 4, 5\} \\ -5 & \text{if } d = 6 \end{cases}$

  $X$ represents player A's gain; $-X$ represents player B's gain.

# Expected Value

Once we have defined random variables for the experiment, we can analyze their numeric properties. For example:

**Definition (Expected Value)**

Let $S$ be a sample space and let $X$ be a random variable.

The **expected value** of $X$, written $\mathrm{E}[X]$, is defined as follows:

$$\mathrm{E}[X] = \sum_{s \in S} X(s) \cdot p(s)$$

That is, $\mathrm{E}[X]$ is the average value of $X$ weighted by the probability of each outcome.

The expectation represents what we would expect to get as the mean value of $X$ for a large number of repetitions of the experiment.

# Example: Expected Value (1)

$\mathrm{E}[X] = \sum_{s \in S} X(s) \cdot p(s)$

A six-sided die is rolled.
If it shows 1–5, then player A wins \$1.
If it shows 6, then player B wins \$5.

$$X(d) = \begin{cases} 1 & \text{if } d \in \{1,2,3,4,5\} \\ -5 & \text{if } d = 6 \end{cases}$$

What is player A's expected gain per round? That is, what is $\mathrm{E}[X]$?

# Example: Expected Value (1)

$$\mathrm{E}[X] = \sum_{s \in S} X(s) \cdot p(s)$$

A six-sided die is rolled.
If it shows 1–5, then player A wins $1.
If it shows 6, then player B wins $5.

$$X(d) = \begin{cases} 1 & \text{if } d \in \{1,2,3,4,5\} \\ -5 & \text{if } d = 6 \end{cases}$$

What is player A's expected gain per round? That is, what is $\mathrm{E}[X]$?

**Solution:** There are 6 outcomes, each equally likely (probability $\frac{1}{6}$).

$$\begin{aligned} \mathrm{E}[X] &= \sum_{d \in \{1,2,3,4,5,6\}} X(d) \cdot p(d) \\ &= \frac{1}{6}\big(X(1) + X(2) + X(3) + X(4) + X(5) + X(6)\big) \\ &= \frac{1}{6}\big(1 + 1 + 1 + 1 + 1 + -5\big) \;=\; 0 \end{aligned}$$

# Example: Expected Value (2)

$$\mathrm{E}[X] = \sum_{s \in S} X(s) \cdot p(s)$$

Two six-sided dice are rolled. Let the random variable $X$ be the sum of the two faces. There are $36$ outcomes ($S$ is pairs of numbers from 1 to 6). Each outcome is equally likely (probability $\frac{1}{36}$). What is $\mathrm{E}[X]$?

# Example: Expected Value (2)

$$\mathrm{E}[X] = \sum_{s \in S} X(s) \cdot p(s)$$

Two six-sided dice are rolled. Let the random variable $X$ be the sum of the two faces. There are $36$ outcomes ($S$ is pairs of numbers from 1 to 6). Each outcome is equally likely (probability $\frac{1}{36}$). What is $\mathrm{E}[X]$?

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{(a,b) \in S} p(a,b) \cdot X(a,b) = \frac{1}{36} \sum_{(a,b) \in S} (a+b) = \frac{1}{36} \sum_{a=1}^{6} \sum_{b=1}^{6} (a+b) \\
&= \frac{1}{36} \left( \sum_{a=1}^{6} \sum_{b=1}^{6} a + \sum_{a=1}^{6} \sum_{b=1}^{6} a \right) = \frac{1}{36} \left( 6 \sum_{a=1}^{6} a + 6 \sum_{b=1}^{6} b \right) \\
&= \frac{1}{36} (6 \cdot 21 + 6 \cdot 21) = \frac{252}{36} = 7
\end{aligned}$$

This means that if we roll the dice many times, sum all the numbers that appear and divide the sum by the number of trials, we expect to find a value of 7.

# Linearity of Expected Value

**Theorem**

Suppose that $X$ and $Y$ are random variables on a sample space $S$.

Then $X + Y$ is itself a random variable, and $\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$.

Furthermore, if $a, b \in \mathbb{R}$, then $\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b$.

The theorem also generalizes to more than two random variables:

> *If $X_1, \ldots X_n$ are random variables, then so is $X_1 + \cdots + X_n$, and $\mathrm{E}[X_1 + \cdots + X_n] = \mathrm{E}[X_1] + \ldots \mathrm{E}[X_n]$.*

# Example: Expected Value (2′)

With this theorem, we can solve the previous example more easily.

Let $X_1$ and $X_2$ be the results of the first and the second die, respectively. For each die, there is an equal probability for each of the six numbers to appear. Therefore,

$$\mathrm{E}[X_1] = \mathrm{E}[X_2] = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2}$$

Now we have

$$\mathrm{E}[X_1 + X_2] = \mathrm{E}[X_1] + \mathrm{E}[X_2] = \frac{7}{2} + \frac{7}{2} = 7$$

# Summary

# Topics

- experiment, sample space, outcome, event
- probability of event
- complementary event
- unions of events (disjoint, non-disjoint)
- discrete probability distributions, uniform distribution
- conditional probability, marginal probability
- Bayes' Theorem, prior vs posterior probability
- independent events, Bernoulli trials
- random variables, expected value