Lecture 16: Continuing Probability.

Wrap up: Probability Formalism.
Lecture 16: Continuing Probability.

Wrap up: Probability Formalism.

Events, Conditional Probability, Independence, Bayes’ Rule
Probability Space: Formalism

Simplest physical model of a uniform probability space:
Probability Space: Formalism
Simplest physical model of a uniform probability space:

Ω

$Pr[\omega]$

- Red 1/8
- Green 1/8
- Maroon 1/8

Physical experiment
Probability model
Probability Space: Formalism

Simplest physical model of a uniform probability space:

A bag of identical balls, except for their color (or a label).
A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.
A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.

\[ \Omega = \{ \text{white, red, yellow, grey, purple, blue, maroon, green} \} \]
Probability Space: Formalism
Simplest physical model of a uniform probability space:

A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.

\( \Omega = \{ \text{white, red, yellow, grey, purple, blue, maroon, green} \} \)

\( Pr[\text{blue}] = \frac{1}{8} \)
A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.

\[ \Omega = \{ \text{white, red, yellow, grey, purple, blue, maroon, green} \} \]

\[ Pr[\text{blue}] = \frac{1}{8}. \]
Simplest physical model of a non-uniform probability space:
Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[
\begin{align*}
\text{Pr}[\omega] &= \frac{3}{10}, \\
\text{Pr}[\text{Green}] &= \frac{4}{10}, \\
\text{Pr}[\text{Yellow}] &= \frac{2}{10}, \\
\text{Pr}[\text{Blue}] &= \frac{1}{10}.
\end{align*}
\]

Note: Probabilities are restricted to rational numbers: \( \mathbb{Q} \).
Probability Space: Formalism

Simplest physical model of a **non-uniform** probability space:

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$
Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10} \]

\[ Pr[\text{Green}] = \frac{4}{10} \]

\[ Pr[\text{Yellow}] = \frac{2}{10} \]

\[ Pr[\text{Blue}] = \frac{1}{10} \]
Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \]

\[ Pr[\text{Green}] = \frac{4}{10}, \]

\[ Pr[\text{Yellow}] = \frac{2}{10}, \]

\[ Pr[\text{Blue}] = \frac{1}{10}. \]
Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad Pr[\text{Yellow}] = \frac{2}{10}, \quad Pr[\text{Blue}] = \frac{1}{10} \]
Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad \text{etc.} \]
Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad \text{etc.} \]

Note: Probabilities are restricted to rational numbers: \( \frac{N_k}{N} \).
Probability Space: Formalism

Physical model of a general non-uniform probability space:
Physical model of a general **non-uniform** probability space:

- Fraction $p_1$ of circumference
- Fraction $p_2$ of circumference
- Fraction $p_3$ of circumference

- Physical experiment
- Probability model

The roulette wheel stops in sector $\omega$ with probability $p_\omega$.

$\Omega = \{1, 2, 3, \ldots, N\}$,

$Pr[\omega] = p_\omega$.

- Green = 1
- Purple = 2
- Yellow

### Probability Space: Formalism
Probability Space: Formalism

Physical model of a general non-uniform probability space:

The roulette wheel stops in sector $\omega$ with probability $p_\omega$. 

Diagram: 

\begin{itemize}
  \item Green = 1
  \item Purple = 2
  \item Yellow
  \item $\omega$
\end{itemize}

\[ \Omega = \{1, 2, 3, \ldots, N\}, \quad Pr[\omega] = p_\omega. \]
Probability Space: Formalism

Physical model of a general non-uniform probability space:

Physical experiment

Probability model

The roulette wheel stops in sector \( \omega \) with probability \( p_\omega \).

\[ \Omega = \{1, 2, 3, \ldots, N\}, \]
Probability Space: Formalism

Physical model of a general non-uniform probability space:

\[
\Omega = \{1, 2, 3, \ldots, N\}, \Pr[\omega] = p_\omega.
\]
An important remark

- The random experiment selects one and only one outcome in $\Omega$.
An important remark

- The random experiment selects **one and only one** outcome in $\Omega$.
- For instance, when we flip a fair coin **twice**
An important remark

- The random experiment selects *one and only one* outcome in $\Omega$.
- For instance, when we flip a fair coin *twice*
  
  - $\Omega = \{HH, TH, HT, TT\}$
An important remark

► The random experiment selects **one and only one** outcome in $\Omega$.
► For instance, when we flip a fair coin **twice**
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects **one** of the elements of $\Omega$. 
An important remark

- The random experiment selects **one and only one** outcome in $\Omega$.
- For instance, when we flip a fair coin **twice**
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects **one** of the elements of $\Omega$.
- In this case, it's wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
An important remark

- The random experiment selects **one and only one** outcome in $\Omega$.
- For instance, when we flip a fair coin  twice
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects **one** of the elements of $\Omega$.
- In this case, it's wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
- Why?
An important remark

- The random experiment selects **one and only one** outcome in $\Omega$.
- For instance, when we flip a fair coin **twice**
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects **one** of the elements of $\Omega$.
- In this case, it's wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
- Why? Because this would not describe how the two coin flips are related to each other.
An important remark

- The random experiment selects **one and only one** outcome in \( \Omega \).
- For instance, when we flip a fair coin twice
  - \( \Omega = \{HH, TH, HT, TT\} \)
  - The experiment selects **one** of the elements of \( \Omega \).
- In this case, its wrong to think that \( \Omega = \{H, T\} \) and that the experiment selects two outcomes.
- Why? Because this would not describe how the two coin flips are related to each other.
- For instance, say we glue the coins side-by-side so that they face up the same way.
An important remark

- The random experiment selects **one and only one** outcome in $\Omega$.
- For instance, when we flip a fair coin **twice**
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects **one** of the elements of $\Omega$.
- In this case, it's wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
- Why? Because this would not describe how the two coin flips are related to each other.
- For instance, say we glue the coins side-by-side so that they face up the same way. Then one gets $HH$ or $TT$ with probability 50% each.
An important remark

- The random experiment selects **one and only one** outcome in $\Omega$.
- For instance, when we flip a fair coin **twice**
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects **one** of the elements of $\Omega$.
- In this case, it's wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
- Why? Because this would not describe how the two coin flips are related to each other.
- For instance, say we glue the coins side-by-side so that they face up the same way. Then one gets $HH$ or $TT$ with probability 50% each. This is not captured by ‘picking two outcomes.’
Lecture 15: Summary

Modeling Uncertainty: Probability Space
Lecture 15: Summary

Modeling Uncertainty: Probability Space

1. Random Experiment
Modeling Uncertainty: Probability Space

1. Random Experiment
2. Probability Space: $\Omega; \Pr[\omega] \in [0, 1]; \sum_\omega \Pr[\omega] = 1$. 
Lecture 15: Summary

Modeling Uncertainty: Probability Space

1. Random Experiment
2. Probability Space: $\Omega; Pr[\omega] \in [0, 1]; \sum_\omega Pr[\omega] = 1$.
3. Uniform Probability Space: $Pr[\omega] = 1/|\Omega|$ for all $\omega \in \Omega$. 
1. Random Experiment

2. Probability Space: $\Omega; Pr[\omega] \in [0, 1]; \sum_\omega Pr[\omega] = 1.$

3. Uniform Probability Space: $Pr[\omega] = 1/|\Omega|$ for all $\omega \in \Omega.$
CS70: On to Calculation.

Events, Conditional Probability, Independence, Bayes’ Rule
CS70: On to Calculation.

Events, Conditional Probability, Independence, Bayes’ Rule

1. Probability Basics Review
2. Events
3. Conditional Probability
4. Independence of Events
5. Bayes’ Rule
Probability Basics Review

Setup:

- Random Experiment. Flip a fair coin twice.

- Probability Space.

  - Sample Space: Set of outcomes, $\Omega$.
    
    $\Omega = \{HH, HT, TH, TT\}$

    (Note: Not $\Omega = \{H, T\}$ with two picks!)

  - Probability: $\Pr[\omega]$ for all $\omega \in \Omega$.

    $\Pr[HH] = \cdots = \Pr[TT] = \frac{1}{4}$

1. $0 \leq \Pr[\omega] \leq 1$.

2. $\sum_{\omega \in \Omega} \Pr[\omega] = 1$. 
Probability Basics Review

Setup:
Probability Basics Review

Setup:

▶ Random Experiment.
Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.
Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
  - **Sample Space**: Set of outcomes, $\Omega$.
Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
  - **Sample Space**: Set of outcomes, $\Omega$.
    $\Omega = \{HH, HT, TH, TT\}$
Setup:

- Random Experiment. Flip a fair coin twice.
- Probability Space.
  - **Sample Space**: Set of outcomes, $\Omega$.
    $\Omega = \{HH, HT, TH, TT\}$
    (Note: Not $\Omega = \{H, T\}$ with two picks!)

1. $0 \leq \Pr[\omega] \leq 1$.
2. $\sum_{\omega \in \Omega} \Pr[\omega] = 1$. 
Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
  - **Sample Space:** Set of outcomes, \( \Omega \).
    \[ \Omega = \{HH, HT, TH, TT\} \]
    (Note: Not \( \Omega = \{H, T\} \) with two picks!)

- **Probability:** \( Pr[\omega] \) for all \( \omega \in \Omega \).
Setup:

- Random Experiment.
  Flip a fair coin twice.
- Probability Space.
  - **Sample Space:** Set of outcomes, $\Omega$.
    $\Omega = \{HH, HT, TH, TT\}$
    (Note: Not $\Omega = \{H, T\}$ with two picks!)
  - **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
    $Pr[HH] = \cdots = Pr[TT] = 1/4$
Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
  - **Sample Space:** Set of outcomes, $\Omega$.
    $\Omega = \{HH, HT, TH, TT\}$
    (Note: Not $\Omega = \{H, T\}$ with two picks!)

  - **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
    $Pr[HH] = \cdots = Pr[TT] = 1/4$
    1. $0 \leq Pr[\omega] \leq 1$. 
Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
  
  - **Sample Space:** Set of outcomes, \( \Omega \).
    \[
    \Omega = \{ HH, HT, TH, TT \}
    \]
    (Note: Not \( \Omega = \{ H, T \} \) with two picks!)

  - **Probability:** \( Pr[\omega] \) for all \( \omega \in \Omega \).
    \[
    Pr[HH] = \cdots = Pr[TT] = \frac{1}{4}
    \]
    
    1. \( 0 \leq Pr[\omega] \leq 1 \).
    2. \( \sum_{\omega \in \Omega} Pr[\omega] = 1 \).
Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.

- Probability Space.
  - **Sample Space**: Set of outcomes, $\Omega$.
    $$\Omega = \{HH, HT, TH, TT\}$$
    (Note: Not $\Omega = \{H, T\}$ with two picks!)

  - **Probability**: $Pr[\omega]$ for all $\omega \in \Omega$.
    $$Pr[HH] = \cdots = Pr[TT] = 1/4$$

  1. $0 \leq Pr[\omega] \leq 1$.
  2. $\sum_{\omega \in \Omega} Pr[\omega] = 1$. 
Set notation review

Figure: Two events

Figure: Complement (not)

Figure: Union (or)

Figure: Intersection (and)

Figure: Difference (A, not B)

Figure: Symmetric difference (only one)
Set notation review

\[ \Omega \]

\[ A \cap B \]

Figure: Two events
Set notation review

Figure: Two events

Figure: Complement (not)
Set notation review

Figure: Two events

Figure: Union (or)

Figure: Complement (not)
Set notation review

- **Two events**
  - \( A \) and \( B \)

- **Union (or)**
  - \( A \cup B \)

- **Complement (not)**
  - \( \bar{A} \)

- **Intersection (and)**
  - \( A \cap B \)
Set notation review

Figure: Two events

Figure: Union (or)

Figure: Difference (A, not B)

Figure: Complement (not)

Figure: Intersection (and)
Set notation review

Figure: Two events

Figure: Union (or)

Figure: Difference (A, not B)

Figure: Complement (not)

Figure: Intersection (and)

Figure: Symmetric difference (only one)
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: \(HT, TH\).

This leads to a definition!

Definition:
- An event, \(E\), is a subset of outcomes: \(E \subset \Omega\).
- The probability of \(E\) is defined as \(\Pr[E] = \sum_{\omega \in E} \Pr[\omega]\).
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: \( HT, TH \).
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: $HT$, $TH$.

This leads to a definition!
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: $HT, TH$.

This leads to a definition!

**Definition:**
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: $HT, TH$.

This leads to a definition!

**Definition:**

- An event, $E$, is a subset of outcomes: $E \subseteq \Omega$. 
Probability of exactly one ‘heads’ in two coin flips?
Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: $HT$, $TH$.

This leads to a definition!

Definition:

- An **event**, $E$, is a subset of outcomes: $E \subset \Omega$.
- The **probability of** $E$ is defined as $Pr[E] = \sum_{\omega \in E} Pr[\omega]$. 
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: $HT, TH$.

This leads to a definition!

**Definition:**

- An **event**, $E$, is a subset of outcomes: $E \subset \Omega$.
- The **probability of** $E$ is defined as $Pr[E] = \sum_{\omega \in E} Pr[\omega]$. 
Probability of exactly one ‘heads’ in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one ‘heads’: $HT$, $TH$.

This leads to a definition!

**Definition:**

- An event, $E$, is a subset of outcomes: $E \subset \Omega$.
- The probability of $E$ is defined as $Pr[E] = \sum_{\omega \in E} Pr[\omega]$. 

![Sample Space Diagram]

[Diagram showing sample space $\Omega$, event $E$, samples (outcomes), and uniform probability space]
Event: Example

Ω = {Red, Green, Yellow, Blue}

Pr[Red] = \frac{3}{10}, Pr[Green] = \frac{4}{10}, etc.

E = \{Red, Green\} \Rightarrow Pr[E] = \frac{3}{10} + \frac{4}{10} = Pr[Red] + Pr[Green].
Event: Example

Physical experiment

Probability model

$\Omega = \{\text{Red, Green, Yellow, Blue}\}$

$Pr[\omega] = \frac{3}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10}$

$E = \{\text{Red, Green}\} \Rightarrow Pr[E] = \frac{3}{10} + \frac{4}{10} = \frac{7}{10}$

$Pr[\text{Red}] + Pr[\text{Green}]$
Event: Example

\( \Omega = \{ \text{Red, Green, Yellow, Blue} \} \)

Physical experiment

\[ \begin{array}{c}
\Omega \\
Pr[\omega] \\
\text{Red} & 3/10 \\
\text{Green} & 4/10 \\
\text{Yellow} & 2/10 \\
\text{Blue} & 1/10 \\
\end{array} \]

Probability model
Event: Example

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ \Pr[\text{Red}] = \]
Event: Example

Physical experiment

Ω = \{Red, Green, Yellow, Blue\}

Pr[Red] = \frac{3}{10},

Probability model

Ω

\begin{array}{c}
\bullet \text{Red} \\
\bullet \text{Green} \\
\bullet \text{Yellow} \\
\bullet \text{Blue}
\end{array}

Pr[\omega]

\begin{array}{c}
3/10 \\
4/10 \\
2/10 \\
1/10
\end{array}
Event: Example

Physical experiment

Probability model

$\Omega = \{\text{Red, Green, Yellow, Blue}\}$

$Pr[\text{Red}] = \frac{3}{10}$, $Pr[\text{Green}] =$
Event: Example

Physical experiment

Probability model

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad \text{etc.} \]
Event: Example

\[ \Omega = \{\text{Red, Green, Yellow, Blue}\} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad \text{etc.} \]

\[ E = \{\text{Red, Green}\} \]
Event: Example

$\Omega = \{\text{Red, Green, Yellow, Blue}\}$

$Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}$, etc.

$E = \{\text{Red, Green}\}$ $\Rightarrow$ $Pr[E] =$
Event: Example

Physical experiment

\[ \Omega = \{\text{Red, Green, Yellow, Blue}\} \]

\[
Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad \text{etc.}
\]

\[ E = \{\text{Red, Green}\} \Rightarrow Pr[E] = \frac{3 + 4}{10} = \]

Probability model
Event: Example

Physical experiment

Probability model

\[ \Omega = \{\text{Red, Green, Yellow, Blue}\} \]

\[ Pr[\text{Red}] = \frac{3}{10}, \quad Pr[\text{Green}] = \frac{4}{10}, \quad \text{etc.} \]

\[ E = \{\text{Red, Green}\} \Rightarrow Pr[E] = \frac{3 + 4}{10} = \frac{3}{10} + \frac{4}{10} = \]
Event: Example

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}, \text{ etc.} \]

\[ E = \{ \text{Red, Green} \} \Rightarrow Pr[E] = \frac{3 + 4}{10} = \frac{3}{10} + \frac{4}{10} = Pr[\text{Red}] + Pr[\text{Green}]. \]
Probability of exactly one heads in two coin flips?

Sample Space, $\Omega = \{HH, HT, TH, TT\}$.

Uniform probability space: \[ \text{Pr}[HH] = \text{Pr}[HT] = \text{Pr}[TH] = \text{Pr}[TT] = \frac{1}{4}. \]

Event, $E$, "exactly one heads": $\{TH, HT\}$.

\[ \text{Pr}[E] = \sum_{\omega \in E} \text{Pr}[\omega] = \frac{|E|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}. \]
Probability of exactly one heads in two coin flips?

Sample Space, $\Omega = \{HH, HT, TH, TT\}$. 
Probability of exactly one heads in two coin flips?

Sample Space, $\Omega = \{HH, HT, TH, TT\}$.

Uniform probability space: $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$. 
Probability of exactly one heads in two coin flips?

Sample Space, \( \Omega = \{ HH, HT, TH, TT \} \).

Uniform probability space: \( Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4} \).

Event, \( E \), “exactly one heads”: \( \{ TH, HT \} \).
Probability of exactly one heads in two coin flips?

Sample Space, \( \Omega = \{HH, HT, TH, TT\} \).

Uniform probability space: \( Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4} \).

Event, \( E \), “exactly one heads”: \( \{TH, HT\} \).

\[
Pr[E] = \sum_{\omega \in E} Pr[\omega]
\]
Probability of exactly one heads in two coin flips?

Sample Space, $\Omega = \{HH, HT, TH, TT\}$.

Uniform probability space: $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$.

Event, $E$, “exactly one heads”: $\{TH, HT\}$.

$$Pr[E] = \sum_{\omega \in E} Pr[\omega] = \frac{|E|}{|\Omega|}$$
Probability of exactly one heads in two coin flips?

Sample Space, \( \Omega = \{ HH, HT, TH, TT \} \).

Uniform probability space: \( Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4} \).

Event, \( E \), “exactly one heads”: \( \{ TH, HT \} \).

\[
Pr[E] = \sum_{\omega \in E} Pr[\omega] = \frac{|E|}{|\Omega|} = \frac{2}{4}
\]
Probability of exactly one heads in two coin flips?

Sample Space, $\Omega = \{HH, HT, TH, TT\}$.

Uniform probability space: $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$.

Event, $E$, “exactly one heads”: $\{TH, HT\}$.

$$Pr[E] = \sum_{\omega \in E} Pr[\omega] = \frac{|E|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}.$$
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega =$ set of 20 fair coin tosses.
Example: 20 coin tosses.

20 coin tosses

Sample space: $\Omega = \text{set of 20 fair coin tosses.}$

$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20};$
Example: 20 coin tosses.

20 coin tosses

Sample space: \( \Omega = \text{set of 20 fair coin tosses.} \)
\( \Omega = \{ T, H \}^{20} \equiv \{ 0, 1 \}^{20}; \ |\Omega| = 2^{20}. \)
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses.}$
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$

▶ What is more likely?
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses.}$
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \ |\Omega| = 2^{20}.$

▶ What is more likely?

▶ $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \text{ or}$
Example: 20 coin tosses.

20 coin tosses

Sample space: $\Omega =$ set of 20 fair coin tosses.
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \ |\Omega| = 2^{20}.$

▶ What is more likely?

▶ $\omega_1 := (1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1), \text{ or}$
▶ $\omega_2 := (1,0,1,1,0,0,0,1,0,1,0,1,0,1,1,0,0,0)?$
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses.} \quad \Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}$.

▶ What is more likely?

▶ $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$, or
▶ $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0)$?

Answer:
Example: 20 coin tosses.

20 coin tosses

Sample space: $\Omega =$ set of 20 fair coin tosses.
$\Omega = \{T,H\}^{20} \equiv \{0,1\}^{20}$; $|\Omega| = 2^{20}$.

What is more likely?

$\omega_1 := (1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)$, or
$\omega_2 := (1,0,1,1,0,0,0,1,0,1,0,1,1,0,1,1,0,0,0)$?

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$. 
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses.}$
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$

What is more likely?

- $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \text{ or}$
- $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0)?$

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$.

What is more likely?
Example: 20 coin tosses.

20 coin tosses
Sample space: \( \Omega = \text{set of 20 fair coin tosses.} \)
\[
\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.
\]

What is more likely?

\[ \omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \] or
\[ \omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0)? \]

Answer: Both are equally likely: \( Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}. \)

What is more likely?

\[ (E_1) \text{ Twenty Hs out of twenty, or} \]
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega =$ set of 20 fair coin tosses.
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$

What is more likely?

$\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \quad \text{or}$
$\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0)?$

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{\Omega}.$

What is more likely?

$(E_1)$ Twenty Hs out of twenty, or
$(E_2)$ Ten Hs out of twenty?
Example: 20 coin tosses.

20 coin tosses

Sample space: $\Omega = \text{set of 20 fair coin tosses}$. 
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}$.

What is more likely?

- $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$, or
- $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0)$?

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$.

What is more likely?

- $(E_1)$ Twenty Hs out of twenty, or
- $(E_2)$ Ten Hs out of twenty?

Answer: Ten Hs out of twenty.
Example: 20 coin tosses.

20 coin tosses
Sample space: Ω = set of 20 fair coin tosses.
Ω = \{ T, H \}^{20} \equiv \{0, 1\}^{20}; \ |\Omega| = 2^{20}.

▶ What is more likely?
▶ \( \omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1) \), or
▶ \( \omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0) \)?

Answer: Both are equally likely: \( Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}. \)

▶ What is more likely?

\( (E_1) \) Twenty Hs out of twenty, or
\( (E_2) \) Ten Hs out of twenty?

Answer: Ten Hs out of twenty.

Why?
Example: 20 coin tosses.

20 coin tosses

Sample space: $\Omega =$ set of 20 fair coin tosses.
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \ |\Omega| = 2^{20}.$

What is more likely?

$\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \text{ or }$
$\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0)?$

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$.

What is more likely?

$(E_1)$ Twenty Hs out of twenty, or
$(E_2)$ Ten Hs out of twenty?

Answer: Ten Hs out of twenty.

Why? There are many sequences of 20 tosses with ten Hs;
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses.}$
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$

- What is more likely?
  - $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \text{ or}$
  - $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0)?$

  Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}.$

- What is more likely?
  - $(E_1)$ Twenty Hs out of twenty, or
  - $(E_2)$ Ten Hs out of twenty?

  Answer: Ten Hs out of twenty.

  Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs.
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses.}$
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$

- What is more likely?
  - $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \text{ or}$
  - $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0)?$

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}.$

- What is more likely?
  - $(E_1)$ Twenty Hs out of twenty, or
  - $(E_2)$ Ten Hs out of twenty?

Answer: Ten Hs out of twenty.

Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs. $\Rightarrow Pr[E_1] = \frac{1}{|\Omega|} \ll Pr[E_2] = \frac{|E_2|}{|\Omega|}.$
Example: 20 coin tosses.

**20 coin tosses**

Sample space: \( \Omega = \text{set of 20 fair coin tosses.} \)

\[ \Omega = \{ T, H \}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}. \]

- What is more likely?
  - \( \omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \) or
  - \( \omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0) \)?

Answer: Both are equally likely: \( Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}. \)

- What is more likely?
  
  (\( E_1 \)) Twenty Hs out of twenty, or
  (\( E_2 \)) Ten Hs out of twenty?

Answer: Ten Hs out of twenty.

Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs. \( \Rightarrow Pr[E_1] = \frac{1}{|\Omega|} \ll Pr[E_2] = \frac{|E_2|}{|\Omega|}. \)

\[ |E_2| = \]
Example: 20 coin tosses.

20 coin tosses
Sample space: $\Omega = \text{set of 20 fair coin tosses}$. 
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \ |\Omega| = 2^{20}$.

▶ What is more likely?
- $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$, or
- $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0)$?

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$.

▶ What is more likely?
(E₁) Twenty Hs out of twenty, or
(E₂) Ten Hs out of twenty?

Answer: Ten Hs out of twenty.

Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs. $\Rightarrow Pr[E_1] = \frac{1}{|\Omega|} \ll Pr[E_2] = \frac{|E_2|}{|\Omega|}$.

$|E_2| = \binom{20}{10} =$
Example: 20 coin tosses.

Sample space: $\Omega =$ set of 20 fair coin tosses.
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \ |\Omega| = 2^{20}.$

What is more likely?

- $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \text{ or }$
- $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0)?$

Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$.

What is more likely?

$(E_1)$ Twenty Hs out of twenty, or
$(E_2)$ Ten Hs out of twenty?

Answer: Ten Hs out of twenty.

Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs. $\Rightarrow Pr[E_1] = \frac{1}{|\Omega|} \ll Pr[E_2] = \frac{|E_2|}{|\Omega|}$.

$|E_2| = \binom{20}{10} = 184,756.$
Probability of $n$ heads in 100 coin tosses.
Probability of $n$ heads in 100 coin tosses.

$$\Omega = \{H, T\}^{100};$$
Probability of $n$ heads in 100 coin tosses.

$$\Omega = \{H, T\}^{100}; \ |\Omega| = 2^{100}.$$
Probability of $n$ heads in 100 coin tosses.

$\Omega = \{H, T\}^{100}; \ |\Omega| = 2^{100}$. 

Observe:
- Concentration around mean: Law of Large Numbers;
- Bell-shape: Central Limit Theorem.
Probability of $n$ heads in 100 coin tosses.

$\Omega = \{H, T\}^{100}; \ |\Omega| = 2^{100}.$

Event $E_n = \text{‘}n\text{ heads’};$
Probability of \( n \) heads in 100 coin tosses.

\[
\Omega = \{H, T\}^{100}; \quad |\Omega| = 2^{100}.
\]

Event \( E_n = \text{‘} n \text{ heads’} \); \( |E_n| = \)
Probability of $n$ heads in 100 coin tosses.

$\Omega = \{H, T\}^{100}; \ |\Omega| = 2^{100}.$

Event $E_n = \text{‘}n\text{ heads’}; \ |E_n| = \binom{100}{n}$
Probability of $n$ heads in 100 coin tosses.

\[ \Omega = \{H, T\}^{100}; \ |\Omega| = 2^{100}. \]

Event \( E_n = \text{‘}n\text{ heads’}; \ |E_n| = \binom{100}{n} \)

\[ p_n := \Pr[E_n] = \]
Probability of $n$ heads in 100 coin tosses.

$\Omega = \{H, T\}^{100}; |\Omega| = 2^{100}.$

Event $E_n = \text{‘}n\text{ heads’}; |E_n| = \binom{100}{n}$

$p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} =$
Probability of \( n \) heads in 100 coin tosses.

\[ \Omega = \{ H, T \}^{100}; \quad |\Omega| = 2^{100}. \]

Event \( E_n = 'n \) heads'; \( |E_n| = \binom{100}{n} \)

\[ p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}} \]
Probability of $n$ heads in 100 coin tosses.

$\Omega = \{H, T\}^{100}; \quad |\Omega| = 2^{100}.$

Event $E_n = \text{‘n heads’}; \quad |E_n| = \binom{100}{n}$

$p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}}$

Observe:
Probability of \( n \) heads in 100 coin tosses.

\[ \Omega = \{H, T\}^{100}; \; |\Omega| = 2^{100}. \]

Event \( E_n = \text{‘} n \text{ heads’}; \; |E_n| = \binom{100}{n} \]

\[ p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}} \]

Observe:

- Concentration around mean:
Probability of \( n \) heads in 100 coin tosses.

\[ \Omega = \{H, T\}^{100}; \quad |\Omega| = 2^{100}. \]

Event \( E_n = \text{‘}n \text{ heads’}; \quad |E_n| = {100 \choose n} \]

\[ p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{{100 \choose n}}{2^{100}} \]

Observe:

- Concentration around mean: Law of Large Numbers;
Probability of \( n \) heads in 100 coin tosses.

\[ \Omega = \{H, T\}^{100}; \quad |\Omega| = 2^{100}. \]

Event \( E_n = 'n \text{ heads}' \); \( |E_n| = \binom{100}{n} \)

\[ p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}} \]

Observe:

- Concentration around mean: Law of Large Numbers;
- Bell-shape:
Probability of \( n \) heads in 100 coin tosses.

\[ \Omega = \{H, T\}^{100}; \quad |\Omega| = 2^{100}. \]

Event \( E_n = \text{‘}n\text{ heads’}; \quad |E_n| = \binom{100}{n} \]

\[ p_n := \Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}} \]

Observe:

- Concentration around mean: Law of Large Numbers;
- Bell-shape: Central Limit Theorem.
Roll a red and a blue die.
Roll a red and a blue die.

\[
\begin{align*}
Pr[\text{Sum to 7}] &= \frac{6}{36} \\
Pr[\text{Sum to 10}] &= \frac{3}{36}
\end{align*}
\]
Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega = \text{set of 100 coin tosses}$
Exactly 50 heads in 100 coin tosses.

Sample space: \( \Omega = \text{set of 100 coin tosses} = \{H, T\}^{100} \).
Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega = \text{set of 100 coin tosses} = \{H, T\}^{100}$. 
$|\Omega| = 2 \times 2 \times \cdots \times 2$
Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega =$ set of 100 coin tosses $= \{H, T\}^{100}$. 
$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}$. 
Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega = \text{set of 100 coin tosses} = \{H, T\}^{100}$. 
$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}$.

Uniform probability space: $Pr[\omega] = \frac{1}{2^{100}}$.
Exactly 50 heads in 100 coin tosses.

Sample space: \( \Omega = \text{set of 100 coin tosses} = \{H, T\}^{100} \).
\(|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}.\)

Uniform probability space: \( Pr[\omega] = \frac{1}{2^{100}}.\)

Event \( E = \text{“100 coin tosses with exactly 50 heads”} \)
Exactly 50 heads in 100 coin tosses.

Sample space: \( \Omega = \text{set of 100 coin tosses} = \{H, T\}^{100} \).
\(|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100} \).

Uniform probability space: \( Pr[\omega] = \frac{1}{2^{100}} \).

Event \( E = \text{“100 coin tosses with exactly 50 heads”} \)

Choose 50 positions out of 100 to be heads.
Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega = \text{set of 100 coin tosses} = \{H, T\}^{100}$.  
$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}$.

Uniform probability space: $Pr[\omega] = \frac{1}{2^{100}}$.

Event $E = \text{“100 coin tosses with exactly 50 heads”}$  
$|E|$?  
Choose 50 positions out of 100 to be heads.  
$|E| = \binom{100}{50}$. 
Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega = \text{set of 100 coin tosses} = \{H, T\}^{100}$.  
$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}$.

Uniform probability space: $Pr[\omega] = \frac{1}{2^{100}}$.

Event $E = \text{“100 coin tosses with exactly 50 heads”}$

$|E|$?

Choose 50 positions out of 100 to be heads.

$|E| = \binom{100}{50}$.

$$Pr[E] = \frac{\binom{100}{50}}{2^{100}}.$$
Calculation.
Stirling formula (for large $n$):

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$
Calculation.

Stirling formula (for large $n$):

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$ 

$$\binom{2n}{n} \approx \frac{\sqrt{4\pi n(2n/e)^{2n}}}{[\sqrt{2\pi n(e/n)^n}]^2}.$$
Calculation.

Stirling formula (for large $n$):

\[ n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \]

\[ \binom{2n}{n} \approx \frac{\sqrt{4\pi n}(2n/e)^{2n}}{[\sqrt{2\pi n}(n/e)^n]^2} \approx \frac{4^n}{\sqrt{\pi n}}. \]
Calculation.
Stirling formula (for large $n$):

$$n! \approx \sqrt{2\pi n} \left( \frac{n}{e} \right)^n.$$

$$\binom{2n}{n} \approx \frac{\sqrt{4\pi n(2n/e)^{2n}}}{[\sqrt{2\pi n(n/e)^n}]^2} \approx \frac{4^n}{\sqrt{\pi n}}.$$

$$Pr[E] = \frac{|E|}{|\Omega|} =$$
Calculation.

Stirling formula (for large $n$):

$$n! \approx \sqrt{2\pi n} \left( \frac{n}{e} \right)^n.$$ 

$$\binom{2n}{n} \approx \frac{\sqrt{4\pi n (2n/e)^{2n}}}{[\sqrt{2\pi n (n/e)^n}]^2} \approx \frac{4^n}{\sqrt{\pi n}}.$$ 

$$Pr[E] = \frac{|E|}{|\Omega|} = \frac{|E|}{2^{2n}} =$$
Calculation.
Stirling formula (for large $n$):

\[ n! \approx \sqrt{2\pi n} \left( \frac{n}{e} \right)^n. \]

\[ \binom{2n}{n} \approx \frac{\sqrt{4\pi n}(2n/e)^{2n}}{[\sqrt{2\pi n}(n/e)^n]^2} \approx \frac{4^n}{\sqrt{\pi n}}. \]

\[ Pr[E] = \frac{|E|}{|\Omega|} = \frac{|E|}{2^{2n}} = \frac{1}{\sqrt{\pi n}} = \]
Calculation.
Stirling formula (for large \( n \)):

\[
n! \approx \sqrt{2\pi n} \left( \frac{n}{e} \right)^n.
\]

\[
\binom{2n}{n} \approx \frac{\sqrt{4\pi n}(2n/e)^{2n}}{[\sqrt{2\pi n}(n/e)^n]^2} \approx \frac{4^n}{\sqrt{\pi n}}.
\]

\[
Pr[E] = \frac{|E|}{|\Omega|} = \frac{|E|}{2^{2n}} = \frac{1}{\sqrt{\pi n}} = \frac{1}{\sqrt{50\pi}} \approx .08.
\]
Exactly 50 heads in 100 coin tosses.

\[ Pr[n \text{ Heads out of } 2n] = \frac{\binom{2n}{n}}{2^{2n}} \]
Probability is Additive

Theorem

(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then $\Pr[A \cup B] = \Pr[A] + \Pr[B]$.

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint, i.e., $A_k \cap A_m = \emptyset$, $\forall k \neq m$, then $\Pr[A_1 \cup \cdots \cup A_n] = \Pr[A_1] + \cdots + \Pr[A_n]$.

Proof: Obvious.
Theorem
(a) If events $A$ and $B$ are disjoint,
Probability is Additive

Theorem
(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, 

Proof: Obvious.
Theorem

(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$
Theorem

(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint,
Theorem

(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint, i.e., $A_k \cap A_m = \emptyset$, $\forall k \neq m$, then

$$Pr[A_1 \cup \cdots \cup A_n] = Pr[A_1] + \cdots + Pr[A_n].$$
Theorem

(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint, i.e., $A_k \cap A_m = \emptyset$, $\forall k \neq m$, then

$$Pr[A_1 \cup \cdots \cup A_n] = Pr[A_1] + \cdots + Pr[A_n].$$
Theorem

(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint, i.e., $A_k \cap A_m = \emptyset$, $\forall k \neq m$, then

$$Pr[A_1 \cup \cdots \cup A_n] = Pr[A_1] + \cdots + Pr[A_n].$$

Proof:
Probability is Additive

Theorem
(a) If events $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint, i.e., $A_k \cap A_m = \emptyset, \forall k \neq m$, then

$$Pr[A_1 \cup \cdots \cup A_n] = Pr[A_1] + \cdots + Pr[A_n].$$

Proof:
Obvious.
Consequences of Additivity

Theorem

(a) \[ \Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]; \]
(inclusion-exclusion property)

(b) \[ \Pr[A_1 \cup \cdots \cup A_n] \leq \Pr[A_1] + \cdots + \Pr[A_n]; \]
(union bound)

(c) If \( A_1, \ldots, A_N \) are a partition of \( \Omega \), i.e., pairwise disjoint and \( \bigcup_{m=1}^{N} A_m = \Omega \), then
\[ \Pr[B] = \Pr[B \cap A_1] + \cdots + \Pr[B \cap A_N]. \]
(law of total probability)

Proof: (b) is obvious. Proofs for (a) and (c)? Next...
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;

(inclusion-exclusion property)
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;

   (inclusion-exclusion property)

(b) $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;

Proof: (b) is obvious. Proofs for (a) and (c)? Next...
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;  
    \hspace{1cm} (inclusion-exclusion property)

(b) $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;  
    \hspace{1cm} (union bound)
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;
    
    (inclusion-exclusion property)

(b) $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;
    
    (union bound)

(c) If $A_1, \ldots, A_N$ are a partition of $\Omega$, 

Proof: (b) is obvious. Proofs for (a) and (c)? Next...
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;

\hspace{1cm} \text{(inclusion-exclusion property)}

(b) $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;

\hspace{1cm} \text{(union bound)}

(c) If $A_1, \ldots, A_N$ are a partition of $\Omega$, i.e.,

pairwise disjoint and $\bigcup_{m=1}^N A_m = \Omega$,
Consequences of Additivity

Theorem

(a) \( \Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] \);  
    (inclusion-exclusion property)

(b) \( \Pr[A_1 \cup \cdots \cup A_n] \leq \Pr[A_1] + \cdots + \Pr[A_n] \);  
    (union bound)

(c) If \( A_1, \ldots, A_N \) are a partition of \( \Omega \), i.e.,  
    pairwise disjoint and \( \cup_{m=1}^{N} A_m = \Omega \), then  
    \[ \Pr[B] = \Pr[B \cap A_1] + \cdots + \Pr[B \cap A_N]. \]
Consequences of Additivity

Theorem

(a) \( Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]; \)

\hspace{2cm} (inclusion-exclusion property)

(b) \( Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]; \)

\hspace{2cm} (union bound)

(c) If \( A_1, \ldots A_N \) are a \textit{partition} of \( \Omega \), i.e.,

pairwise disjoint and \( \cup_{m=1}^{N} A_m = \Omega \), then

\[ Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N]. \]

\hspace{2cm} (law of total probability)
Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;
   (inclusion-exclusion property)

(b) $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;
   (union bound)

(c) If $A_1, \ldots, A_N$ are a partition of $\Omega$, i.e.,
   pairwise disjoint and $\bigcup_{m=1}^{N} A_m = \Omega$, then
   $Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N]$.
   (law of total probability)

Proof:
Consequences of Additivity

Theorem

(a) \( Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]; \)
   (inclusion-exclusion property)

(b) \( Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]; \)
   (union bound)

(c) If \( A_1, \ldots, A_N \) are a partition of \( \Omega \), i.e.,
   pairwise disjoint and \( \bigcup_{m=1}^{N} A_m = \Omega \), then
   \( Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N]. \)
   (law of total probability)

Proof:

(b) is obvious.
Consequences of Additivity

Theorem

(a) \( \Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] \);  
    (inclusion-exclusion property)

(b) \( \Pr[A_1 \cup \cdots \cup A_n] \leq \Pr[A_1] + \cdots + \Pr[A_n] \);  
    (union bound)

(c) If \( A_1, \ldots A_N \) are a partition of \( \Omega \), i.e., pairwise disjoint and \( \bigcup_{m=1}^{N} A_m = \Omega \), then

\[ \Pr[B] = \Pr[B \cap A_1] + \cdots + \Pr[B \cap A_N]. \]
    (law of total probability)

Proof:

(b) is obvious.

Proofs for (a) and (c)?
Consequences of Additivity

Theorem

(a) \( Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B] \);
   \text{(inclusion-exclusion property)}

(b) \( Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n] \);
   \text{(union bound)}

(c) If \( A_1, \ldots, A_N \) are a partition of \( \Omega \), i.e.,
   pairwise disjoint and \( \bigcup_{m=1}^{N} A_m = \Omega \), then
   \[ Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N]. \]
   \text{(law of total probability)}

Proof:

(b) is obvious.

Proofs for (a) and (c)? Next...
Inclusion/Exclusion

\[ Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B] \]
Inclusion/Exclusion

\[ Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B] \]

Another view. Any \( \omega \in A \cup B \) is in \( A \cap B \), \( A \cup B \), or \( A \cap B \). So, add it up.

\[
\begin{align*}
Pr[A] &= x + y \\
Pr[B] &= y + z \\
Pr[A \cap B] &= y \\
Pr[A \cup B] &= x + y + z
\end{align*}
\]
Inclusion/Exclusion

\[ Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B] \]

Another view.

\[ Pr[A] = x + y \\
Pr[B] = y + z \\
Pr[A \cap B] = y \\
Pr[A \cup B] = x + y + z \]
Inclusion/Exclusion

\[ Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B] \]

Another view. Any \( \omega \in A \cup B \) is in \( A \cap \overline{B} \), \( A \cup B \), or \( \overline{A} \cap B \). So, add it up.
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$. 

In "math":

$\omega \in B$ is in exactly one of $A_i \cap B$.

Adding up probability of them, get $\Pr[\omega]$ in sum.

..Did I say...

Add it up.
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$. 
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$.

In “math”: $\omega \in B$ is in exactly one of $A_i \cap B$. 
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$.

In “math”: $\omega \in B$ is in exactly one of $A_i \cap B$.

Adding up probability of them, get $Pr[\omega]$ in sum.
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$.

In “math”: $\omega \in B$ is in exactly one of $A_i \cap B$.

Adding up probability of them, get $Pr[\omega]$ in sum.

..Did I say...
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$.

In “math”: $\omega \in B$ is in exactly one of $A_i \cap B$.

Adding up probability of them, get $Pr[\omega]$ in sum.

..Did I say...

Add it up.
Roll a Red and a Blue Die.

\[ E_1 = \text{Red die shows 6}; \quad E_2 = \text{Blue die shows 6} \]

\[ E_1 \cup E_2 = \text{At least one die shows 6} \]

\[ \Pr[E_1] = \frac{1}{6}, \quad \Pr[E_2] = \frac{1}{6}, \quad \Pr[E_1 \cup E_2] = \frac{11}{36}. \]
Roll a Red and a Blue Die.

$E_1 = \text{\'Red die shows 6\'}$; $E_2 = \text{\'Blue die shows 6\'}$.

$E_1 \cup E_2 = \text{\'At least one die shows 6\'}$.

$\Pr[E_1] = \frac{6}{36}, \quad \Pr[E_2] = \frac{6}{36}, \quad \Pr[E_1 \cup E_2] = \frac{11}{36}$.

$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$
Roll a Red and a Blue Die.

\[ E_1 = \text{`Red die shows 6'}; \]
Roll a Red and a Blue Die.

$E_1 = \text{`Red die shows 6'}$; $E_2 = \text{`Blue die shows 6'}$

$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$
Roll a Red and a Blue Die.

$E_1 = 'Red die shows 6'; E_2 = 'Blue die shows 6'$

$E_1 \cup E_2 = 'At least one die shows 6'$
Roll a Red and a Blue Die.

\[ E_1 = \text{`Red die shows 6'} \; \text{; } E_2 = \text{`Blue die shows 6'} \]
\[ E_1 \cup E_2 = \text{`At least one die shows 6'} \]
\[ Pr[E_1] = \frac{6}{36} \]
Roll a Red and a Blue Die.

$E_1 = \text{‘Red die shows 6’}; E_2 = \text{‘Blue die shows 6’}

E_1 \cup E_2 = \text{‘At least one die shows 6’}

Pr[E_1] = \frac{6}{36}, Pr[E_2] = \frac{6}{36}$
Roll a Red and a Blue Die.

$E_1 = \text{Red die shows 6}$; $E_2 = \text{Blue die shows 6}$

$E_1 \cup E_2 = \text{At least one die shows 6}$

$Pr[E_1] = \frac{6}{36}, Pr[E_2] = \frac{6}{36}, Pr[E_1 \cup E_2] = \frac{11}{36}$. 
Conditional probability: example.

Two coin flips.

Ω = {HH, HT, TH, TT}; uniform probability space.

Event \( A \) = first flip is heads: \( A = \{HH, HT\} \).

New sample space: \( A \); uniform still.

Event \( B \) = two heads. The probability of two heads if the first flip is heads.

The probability of \( B \) given \( A \) is \( \frac{1}{2} \).
Conditional probability: example.

Two coin flips. First flip is heads.
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? 
\[\Omega = \{HH, HT, TH, TT\};\]
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? 
\( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space.
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
Ω = \{HH, HT, TH, TT\}; Uniform probability space.
Event A = first flip is heads:
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
\[ \Omega = \{HH, HT, TH, TT\}; \] Uniform probability space.
Event \( A = \) first flip is heads: \( A = \{HH, HT\}. \)
Two coin flips. First flip is heads. Probability of two heads?
\( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space.
Event \( A = \) first flip is heads: \( A = \{HH, HT\} \).

\[ \Omega : \text{uniform} \]

\( \bullet TH \)
\( \bullet TT \)
\( \bullet HH \)
\( \bullet HT \)
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? \( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space. Event \( A = \) first flip is heads: \( A = \{HH, HT\} \).

\[ \Omega : \text{uniform} \]

\[ \bullet TH \quad \bullet HH \]
\[ \bullet TT \quad \bullet HT \]

New sample space: \( A \);
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? 
\( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space. 
Event \( A = \) first flip is heads: \( A = \{HH, HT\} \).

\[ \Omega : \text{uniform} \]

\[ \bullet TH \quad \bullet HH \]
\[ \bullet TT \quad \bullet HT \]

New sample space: \( A \); uniform still.
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? 
\[ \Omega = \{HH, HT, TH, TT\}; \] Uniform probability space. 
Event \( A = \) first flip is heads: \( A = \{HH, HT\} \).

\[ \Omega : \text{uniform} \]

\[ \bullet TH \quad \bullet HH \]
\[ \bullet TT \quad \bullet HT \]

New sample space: \( A \); uniform still.

\[ \bullet HH \quad A : \text{uniform} \]
\[ \bullet HT \]
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? \( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space. Event \( A = \text{first flip is heads} \): \( A = \{HH, HT\} \).

\[
\Omega: \text{uniform}
\]

\[
\bullet TH \quad \bullet HH \\
\bullet TT \quad \bullet HT
\]

New sample space: \( A \); uniform still.

\[
\bullet HH \quad \bullet HT
\]

Event \( B = \text{two heads} \).
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? 
\(\Omega = \{HH, HT, TH, TT\}\); Uniform probability space.
Event \(A = \) first flip is heads: \(A = \{HH, HT\}\).

\(\Omega : \) uniform

\(\bullet TH \quad \bullet HH \quad \bullet TT \quad \bullet HT\)

New sample space: \(A\); uniform still.

\(\bullet HH \quad \bullet HT\) \(A : \) uniform

Event \(B = \) two heads.

The probability of two heads if the first flip is heads.
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
\( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space.
Event \( A = \) first flip is heads: \( A = \{HH, HT\} \).

\[ \Omega : \text{uniform} \]

\[ \bullet TH \quad \bullet HH \]
\[ \bullet TT \quad \bullet HT \]

\( A \)

New sample space: \( A \); uniform still.

\[ \bullet HH \quad \bullet HT \]

\( A \): uniform

Event \( B = \) two heads.

The probability of two heads if the first flip is heads.

**The probability of \( B \) given \( A \)**
Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads? \( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space. Event \( A = \) first flip is heads: \( A = \{HH, HT\} \).

\[ \Omega : \text{uniform} \]

\[ \begin{array}{ccc}
  & TH & \bullet HH \\
\bullet TT & \bullet HT & A
\end{array} \]

New sample space: \( A \); uniform still.

\[ \begin{array}{ccc}
  & HH \\
\bullet HT & A : \text{uniform}
\end{array} \]

Event \( B = \) two heads.

The probability of two heads if the first flip is heads. **The probability of \( B \) given \( A \) is 1/2.**
A similar example.

Two coin flips.
A similar example.

Two coin flips. At least one of the flips is heads.
A similar example.

Two coin flips. At least one of the flips is heads. → Probability of two heads?
A similar example.

Two coin flips. At least one of the flips is heads.
→ Probability of two heads?

Ω = \{HH, HT, TH, TT\};
A similar example.

Two coin flips. At least one of the flips is heads.

→ Probability of two heads?

\[ \Omega = \{HH, HT, TH, TT\}; \text{ uniform.} \]
A similar example.

Two coin flips. At least one of the flips is heads.
$\implies$ Probability of two heads?

$\Omega = \{HH, HT, TH, TT\}$; uniform.
Event $A =$ at least one flip is heads.
A similar example.

Two coin flips. At least one of the flips is heads.
→ Probability of two heads?

\( \Omega = \{HH, HT, TH, TT\} \); uniform.
Event \( A = \) at least one flip is heads. \( A = \{HH, HT, TH\} \).
A similar example.

Two coin flips. At least one of the flips is heads.
→ Probability of two heads?

$\Omega = \{HH, HT, TH, TT\}$; uniform.
Event $A =$ at least one flip is heads. $A = \{HH, HT, TH\}$. 

![Diagram of coin flips](https://example.com/diagram.png)

$\Omega : \text{uniform}$
A similar example.

Two coin flips. At least one of the flips is heads.  
\[ \rightarrow \text{Probability of two heads?} \]

\[ \Omega = \{HH, HT, TH, TT\}; \text{uniform.} \]
Event \( A = \) at least one flip is heads. \( A = \{HH, HT, TH\} \).

New sample space: \( A \);
Two coin flips. At least one of the flips is heads. 
→ Probability of two heads?

\[ \Omega = \{HH, HT, TH, TT\}; \text{ uniform.} \]
Event \( A = \) at least one flip is heads. \( A = \{HH, HT, TH\} \).

New sample space: \( A \); uniform still.
A similar example.

Two coin flips. At least one of the flips is heads. → Probability of two heads?

Ω = \{HH, HT, TH, TT\}; uniform.

Event \(A\) = at least one flip is heads. \(A = \{HH, HT, TH\}\).

New sample space: \(A\); uniform still.
A similar example.

Two coin flips. At least one of the flips is heads. → Probability of two heads?

\[ \Omega = \{HH, HT, TH, TT\}; \] uniform.

Event \( A \) = at least one flip is heads. \( A = \{HH, HT, TH\} \).

New sample space: \( A \); uniform still.

Event \( B \) = two heads.
Two coin flips. At least one of the flips is heads. → Probability of two heads?

$\Omega = \{HH, HT, TH, TT\}$; uniform.

Event $A =$ at least one flip is heads. $A = \{HH, HT, TH\}$.

New sample space: $A$; uniform still.

Event $B =$ two heads.

The probability of two heads if at least one flip is heads.
A similar example.

Two coin flips. At least one of the flips is heads. → Probability of two heads?

\[ \Omega = \{ HH, HT, TH, TT \} \]; uniform.

Event \( A = \) at least one flip is heads. \( A = \{ HH, HT, TH \} \).

New sample space: \( A \); uniform still.

Event \( B = \) two heads.

The probability of two heads if at least one flip is heads. **The probability of \( B \) given \( A \)**
A similar example.

Two coin flips. At least one of the flips is heads. 
→ Probability of two heads?

\[ \Omega = \{HH, HT, TH, TT\}; \] uniform.
Event \( A = \) at least one flip is heads. \( A = \{HH, HT, TH\} \).

New sample space: \( A \); uniform still.

Event \( B = \) two heads.

The probability of two heads if at least one flip is heads. **The probability of \( B \) given \( A \) is** \( 1/3 \).
Conditional Probability: A non-uniform example
Conditional Probability: A non-uniform example

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ \Pr[\omega] = \begin{array}{c}
\text{Red} \\
\text{Green} \\
\text{Yellow} \\
\text{Blue}
\end{array} \begin{array}{c}
3/10 \\
4/10 \\
2/10 \\
1/10
\end{array} \]
Conditional Probability: A non-uniform example

Physical experiment

Probability model

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]
Conditional Probability: A non-uniform example

Physical experiment

Ω = {Red, Green, Yellow, Blue}

\[ Pr[\text{Red}|\text{Red or Green}] = \]
Conditional Probability: A non-uniform example

Physical experiment

Probability model

\[ \Omega = \{ \text{Red, Green, Yellow, Blue} \} \]

\[ Pr[\text{Red}|\text{Red or Green}] = \frac{3}{7} = \]

\[ \Omega \]

\[
\begin{array}{c}
\text{Pr}[\omega] \\
\text{Red} & 3/10 \\
\text{Green} & 4/10 \\
\text{Yellow} & 2/10 \\
\text{Blue} & 1/10 \\
\end{array}
\]
Conditional Probability: A non-uniform example

\[ \Omega = \{\text{Red, Green, Yellow, Blue}\} \]

\[ \Pr[\text{Red} \mid \text{Red or Green}] = \frac{3}{7} = \frac{\Pr[\text{Red} \cap (\text{Red or Green})]}{\Pr[\text{Red or Green}]} \]
Another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. 

Another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.
Let $A = \{3, 4\}, B = \{1, 2, 3\}$.
Another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. Let $A = \{3, 4\}, B = \{1, 2, 3\}$. 

\[
Pr[A|B] = p_3 p_1 + p_2 p_3 = Pr[A \cap B] Pr[B].
\]
Another non-uniform example

Consider \( \Omega = \{1, 2, \ldots, N\} \) with \( Pr[n] = p_n \).
Let \( A = \{3, 4\}, B = \{1, 2, 3\} \).

\[
Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}
\]
Another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. Let $A = \{3, 4\}, B = \{1, 2, 3\}.$

$$Pr[A|B] = \frac{p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$
Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. 

Let $A = \{2, 3, 4\}$, $B = \{1, 2, 3\}$.

$$Pr[A | B] = p_2 + p_3 p_1 + p_2 p_3 = Pr[A \cap B] Pr[B].$$
Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. Let $A = \{2, 3, 4\}, B = \{1, 2, 3\}$. 

$$Pr[A | B] = p_2 + p_3 p_1 + p_2 + p_3 = Pr[A \cap B] Pr[B]$$
Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.
Let $A = \{2, 3, 4\}, B = \{1, 2, 3\}$.
Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. Let $A = \{2, 3, 4\}, B = \{1, 2, 3\}$.

\[
Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}
\]
Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$. Let $A = \{2, 3, 4\}, B = \{1, 2, 3\}$.

$$Pr[A|B] = \frac{p_2 + p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$
Conditional Probability.

**Definition:** The conditional probability of $B$ given $A$ is

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$$
Conditional Probability.

**Definition:** The conditional probability of $B$ given $A$ is

$$Pr[B | A] = \frac{Pr[A \cap B]}{Pr[A]}$$
**Definition:** The conditional probability of $B$ given $A$ is

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$$

In $A!$

In $B?$
Definition: The conditional probability of B given A is

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$$

A must be in A ∩ B.

In A!

In B?

Must be in A ∩ B.
**Definition:** The **conditional probability** of $B$ given $A$ is

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$$

In $A$!
In $B$?
Must be in $A \cap B$.

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}.$$
More fun with conditional probability.

Toss a red and a blue die, sum is 4,
More fun with conditional probability.

Toss a red and a blue die, sum is 4,
What is probability that red is 1?
More fun with conditional probability.

Toss a red and a blue die, sum is 4,
What is probability that red is 1?

\[
\Pr[B \mid A] = \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{1}{3};
\]

versus 
\[
\Pr[B] = \frac{1}{6}.
\]

B is more likely given A.
More fun with conditional probability.

Toss a red and a blue die, sum is 4, What is probability that red is 1?

\[
Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{3};
\]
More fun with conditional probability.

Toss a red and a blue die, sum is 4, What is probability that red is 1?

\[ Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{3}; \text{ versus } Pr[B] = \frac{1}{6}. \]
More fun with conditional probability.

Toss a red and a blue die, sum is 4, What is probability that red is 1?

\[ \Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{3}; \text{ versus } \Pr[B] = \frac{1}{6}. \]

\( B \) is more likely given \( A \).
Yet more fun with conditional probability.

Toss a red and a blue die, sum is 7, what is probability that red is 1?
Yet more fun with conditional probability.

Toss a red and a blue die, sum is 7, what is probability that red is 1?

$$\Pr[B \mid A] = \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{1}{6};$$

versus

$$\Pr[B] = \frac{1}{6}.$$
Yet more fun with conditional probability.

Toss a red and a blue die, sum is 7, what is probability that red is 1?

\[
Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{6};
\]
Yet more fun with conditional probability.

Toss a red and a blue die, sum is 7, what is probability that red is 1?

\[
\begin{align*}
\Omega : \text{Uniform} \\
\Omega = \{1, \ldots, 6\}^2 \\
A = \{(1, 6), \ldots, (6,1)\} \\
B = \{(1, 1), \ldots, (1, 6)\} \\
A = \text{`sum is 7'} \\
\end{align*}
\]

\[
Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{6}; \text{ versus } Pr[B] = \frac{1}{6}.
\]
Yet more fun with conditional probability.

Toss a red and a blue die, sum is 7, what is probability that red is 1?

\[
Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{6}; \text{ versus } Pr[B] = \frac{1}{6}.
\]

Observing \( A \) does not change your mind about the likelihood of \( B \).
Suppose I toss 3 balls into 3 bins.
Suppose I toss 3 balls into 3 bins.
A = “1st bin empty”;

\[
\begin{align*}
\Pr[A | B] &= \Pr[A \cap B] / \Pr[B] \\
\Pr[B] &= \Pr[\{(a, b, c) | a, b, c \in \{1, 3\}\}] = \frac{8}{27} \\
\Pr[A \cap B] &= \Pr[(3, 3, 3)] = \frac{1}{27} \\
\Pr[A | B] &= \left(\frac{1}{27}\right) \left(\frac{8}{27}\right) = \frac{1}{8};
\end{align*}
\]

A is less likely given B: If second bin is empty the first is more likely to have balls in it.
Suppose I toss 3 balls into 3 bins. 
$A = \text{“1st bin empty”}; \ B = \text{“2nd bin empty.”}$
Suppose I toss 3 balls into 3 bins. 
$A =$“1st bin empty”; $B =$“2nd bin empty.” What is $Pr[A|B]$?
Emptiness..

Suppose I toss 3 balls into 3 bins. 
A = “1st bin empty”; B = “2nd bin empty.” What is $Pr[A|B]$?

$\Omega = \{1, 2, 3\}^3$

$\omega = (\text{bin of red ball, bin of blue ball, bin of green ball})$
Emptiness.

Suppose I toss 3 balls into 3 bins. $A$ = “1st bin empty”; $B$ = “2nd bin empty.” What is $Pr[A|B]$?

$Pr[B] = \frac{8}{27}$.

$Pr[A \cap B] = Pr[(3, 3, 3)] = \frac{1}{27}$.

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{\frac{1}{27}}{\frac{8}{27}} = \frac{1}{8}$. 

$A$ is less likely given $B$: If second bin is empty the first is more likely to have balls in it.
Suppose I toss 3 balls into 3 bins. 
A = “1st bin empty”; B = “2nd bin empty.” What is \( Pr[A|B] \)?

\[
\Omega = \{1, 2, 3\}^3
\]

\[
\omega = (\text{bin of red ball}, \text{bin of blue ball}, \text{bin of green ball})
\]

\[
Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\} =
\]
Suppose I toss 3 balls into 3 bins. 
\( A = \text{“1st bin empty”}; \ B = \text{“2nd bin empty.”} \) What is \( Pr[A|B] \)?

\[
\begin{align*}
\Omega &= \{1, 2, 3\}^3 \\
\omega &= (\text{bin of red ball, bin of blue ball, bin of green ball})
\end{align*}
\]

\[
Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] = Pr[\{1, 3\}^3] =
\]

\[
Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{Pr[(3, 2, 3)]}{Pr[(3, 2, 3)]} = \frac{1}{27} \]
Suppose I toss 3 balls into 3 bins. 
\( A = \text{“1st bin empty”}; \ B = \text{“2nd bin empty.”} \) What is \( Pr[A|B] \)?

\[
Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] = Pr[\{1, 3\}^3] = \frac{8}{27}
\]
Suppose I toss 3 balls into 3 bins.  
\( A = \text{“1st bin empty”; } B = \text{“2nd bin empty.”} \) What is \( Pr[A|B] \)?

\[
\begin{align*}
\Omega &= \{1, 2, 3\}^3 \\
\omega &= (\text{bin of red ball, bin of blue ball, bin of green ball})
\end{align*}
\]

\[
\begin{align*}
Pr[B] &= Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] = Pr[\{1, 3\}^3] = \frac{8}{27} \\
Pr[A \cap B] &= \frac{1}{8}
\end{align*}
\]
Suppose I toss 3 balls into 3 bins. 
$A =$ “1st bin empty”; $B =$ “2nd bin empty.” What is $Pr[A | B]$?

$\Omega = \{1, 2, 3\}^3$

$\omega = (\text{bin of red ball}, \text{bin of blue ball}, \text{bin of green ball})$

$Pr[B] = Pr[(a, b, c) | a, b, c \in \{1, 3\}] = Pr[\{1, 3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3, 3, 3)] =$
Suppose I toss 3 balls into 3 bins. 
A = “1st bin empty”; B = “2nd bin empty.” What is $Pr[A|B]$?

$Pr[B] = Pr[\{(a,b,c) | a,b,c \in \{1,3\}\}] = Pr[\{1,3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3,3,3)] = \frac{1}{27}$
Suppose I toss 3 balls into 3 bins. 
$A =$“1st bin empty”; $B =$“2nd bin empty.” What is $Pr[A|B]$?

$Pr[B] = Pr[(a, b, c) \mid a, b, c \in \{1, 3\}] = Pr[\{1, 3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3, 3, 3)] = \frac{1}{27}$

$Pr[A|B]$
Suppose I toss 3 balls into 3 bins. 
A = “1st bin empty”; B = “2nd bin empty.” What is $Pr[A|B]$?

$Pr[B] = Pr[(a, b, c) | a, b, c \in \{1, 3\}] = Pr[\{1, 3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3, 3, 3)] = \frac{1}{27}$

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

$Pr[A|B] = \frac{\frac{1}{27}}{\frac{8}{27}} = \frac{1}{8}$
Emptiness..

Suppose I toss 3 balls into 3 bins. 
\( A = \text{“1st bin empty”}; \ B = \text{“2nd bin empty.”} \) What is \( Pr[A|B] \)?

\( \Omega = \{1, 2, 3\}^3 \)

\[ \omega = (\text{bin of red ball, bin of blue ball, bin of green ball}) \]

\( Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] = Pr[\{1,3\}^3] = \frac{8}{27} \)

\( Pr[A \cap B] = Pr[(3, 3, 3)] = \frac{1}{27} \)

\( Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{(1/27)}{(8/27)} = 1/8; \)
Emptiness..

Suppose I toss 3 balls into 3 bins. 
A = “1st bin empty”; B = “2nd bin empty.” What is \( Pr[A|B] \)?

\[
\Omega = \{1, 2, 3\}^3
\]

\[
\omega = (\text{bin of red ball}, \text{bin of blue ball}, \text{bin of green ball})
\]

\[
Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] = Pr[\{1, 3\}^3] = \frac{8}{27}
\]

\[
Pr[A \cap B] = Pr[(3, 3, 3)] = \frac{1}{27}
\]

\[
Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{(1/27)}{(8/27)} = 1/8; \text{ vs. } Pr[A] = \frac{8}{27}.
\]
Emptiness..

Suppose I toss 3 balls into 3 bins. 
$A =$“1st bin empty”; $B =$“2nd bin empty.” What is $Pr[A|B]$?

$Pr[B] = Pr\{(a,b,c) \mid a,b,c \in \{1,3\}\} = Pr[\{1,3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3,3,3)] = \frac{1}{27}$

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{(1/27)}{(8/27)} = 1/8$; vs. $Pr[A] = \frac{8}{27}$.

$A$ is less likely given $B$: 
Emptiness..

Suppose I toss 3 balls into 3 bins. 
A = “1st bin empty”; B = “2nd bin empty.” What is $Pr[A|B]$?

$\Omega = \{1, 2, 3\}^3$

$\omega = (\text{bin of red ball, bin of blue ball, bin of green ball})$

$Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] = Pr[\{1, 3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3, 3, 3)] = \frac{1}{27}$

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{(1/27)}{(8/27)} = 1/8$; vs. $Pr[A] = \frac{8}{27}$.

A is less likely given B: If second bin is empty the first is more likely to have balls in it.
Gambler’s fallacy.

Flip a fair coin 51 times.
Gambler’s fallacy.

Flip a fair coin 51 times.
A = “first 50 flips are heads”
Gambler’s fallacy.

Flip a fair coin 51 times.

\[ A = \text{“first 50 flips are heads”} \]

\[ B = \text{“the 51st is heads”} \]
Gambler’s fallacy.

Flip a fair coin 51 times.
A = “first 50 flips are heads”
B = “the 51st is heads”
\( Pr[B|A] \) ?
Gambler’s fallacy.

Flip a fair coin 51 times.
\( A = \text{“first 50 flips are heads”} \)
\( B = \text{“the 51st is heads”} \)
\( Pr[B|A] \)?
\( A = \{HH\cdots HT, HH\cdots HH\} \)
Flip a fair coin 51 times.

A = “first 50 flips are heads”

B = “the 51st is heads”

\[ Pr[B|A] \]?

\[ A = \{ HH \cdots HT, HH \cdots HH \} \]

\[ B \cap A = \{ HH \cdots HH \} \]

The likelihood of 51st heads does not depend on the previous flips.
Gambler’s fallacy.

Flip a fair coin 51 times.
A = “first 50 flips are heads”
B = “the 51st is heads”

\[ \Pr[B|A] \]

\[ A = \{HH\ldots HT, HH\ldots HH\} \]
\[ B \cap A = \{HH\ldots HH\} \]

Uniform probability space.
Gambler’s fallacy.

Flip a fair coin 51 times.
A = “first 50 flips are heads”
B = “the 51st is heads”

\[ Pr[B|A] \]

\[
A = \{HH \cdots HT, HH \cdots HH\}
\]

\[
B \cap A = \{HH \cdots HH\}
\]

Uniform probability space.

\[
Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{2}.
\]
Gambler’s fallacy.

Flip a fair coin 51 times.
A = “first 50 flips are heads”
B = “the 51st is heads”
Pr[B|A] ?

A = \{HH\cdots HT, HH\cdots HH\}
B \cap A = \{HH\cdots HH\}

Uniform probability space.

Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{2}.

Same as Pr[B].
Gambler’s fallacy.

Flip a fair coin 51 times.
$A =$ “first 50 flips are heads”
$B =$ “the 51st is heads”
$Pr[B|A]$ ?

$A = \{HH\ldots HT, HH\ldots HH\}$
$B \cap A = \{HH\ldots HH\}$

Uniform probability space.

$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{2}.$

Same as $Pr[B]$.

The likelihood of 51st heads does not depend on the previous flips.
Product Rule

Recall the definition:
Recall the definition:

\[ Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]} \]
Recall the definition:

\[ Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]} . \]

Hence,

\[ Pr[A \cap B] = Pr[A] Pr[B|A]. \]
Product Rule

Recall the definition:

\[ Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]} \]

Hence,

\[ Pr[A \cap B] = Pr[A] Pr[B|A]. \]

Consequently,

\[ Pr[A \cap B \cap C] = Pr[(A \cap B) \cap C] \]
Recall the definition:

\[ Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]} \]

Hence,

\[ Pr[A \cap B] = Pr[A] Pr[B|A] \]

Consequently,

\[ Pr[A \cap B \cap C] = Pr[(A \cap B) \cap C] = Pr[A \cap B] Pr[C|A \cap B] \]
Product Rule

Recall the definition:

\[ \Pr[B|A] = \frac{\Pr[A \cap B]}{\Pr[A]} \]

Hence,

\[ \Pr[A \cap B] = \Pr[A] \Pr[B|A]. \]

Consequently,

\[
\begin{align*}
\Pr[A \cap B \cap C] &= \Pr[(A \cap B) \cap C] \\
&= \Pr[A \cap B] \Pr[C|A \cap B] \\
&= \Pr[A] \Pr[B|A] \Pr[C|A \cap B].
\end{align*}
\]
**Theorem** Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then
Theorem Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}] .$$
Theorem Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

Proof:
**Theorem** Product Rule
Let \(A_1, A_2, \ldots, A_n\) be events. Then

\[
Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1] Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].
\]

**Proof:** By induction.
**Theorem** Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1] Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}] .$$

**Proof:** By induction.
Assume the result is true for $n$. 
**Theorem** Product Rule

Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1] Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}] .$$

**Proof:** By induction.

Assume the result is true for $n$. (It holds for $n = 2$.)
**Theorem** Product Rule

Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

**Proof:** By induction.

Assume the result is true for $n$. (It holds for $n = 2$.) Then,

$$Pr[A_1 \cap \cdots \cap A_n \cap A_{n+1}]$$

$$= Pr[A_1 \cap \cdots \cap A_n] Pr[A_{n+1}|A_1 \cap \cdots \cap A_n]$$
**Theorem** Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1] Pr[A_2 | A_1] \cdots Pr[A_n | A_1 \cap \cdots \cap A_{n-1}].$$

**Proof:** By induction. Assume the result is true for $n$. (It holds for $n = 2$.) Then,

$$Pr[A_1 \cap \cdots \cap A_n \cap A_{n+1}]$$

$$= Pr[A_1 \cap \cdots \cap A_n] Pr[A_{n+1} | A_1 \cap \cdots \cap A_n]$$

$$= Pr[A_1] Pr[A_2 | A_1] \cdots Pr[A_n | A_1 \cap \cdots \cap A_{n-1}] Pr[A_{n+1} | A_1 \cap \cdots \cap A_n],$$
**Theorem** Product Rule

Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1] Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

**Proof:** By induction.

Assume the result is true for $n$. (It holds for $n = 2$.) Then,

$$Pr[A_1 \cap \cdots \cap A_n \cap A_{n+1}]$$

$$= Pr[A_1 \cap \cdots \cap A_n] Pr[A_{n+1}|A_1 \cap \cdots \cap A_n]$$

$$= Pr[A_1] Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}] Pr[A_{n+1}|A_1 \cap \cdots \cap A_n],$$

so that the result holds for $n+1$. 

\[ \square \]
An example.

Correlation

Random experiment: Pick a person at random.

Event $A$: the person has lung cancer.

Event $B$: the person is a heavy smoker.

Fact: $\Pr[A | B] = 1.17 \times \Pr[A]$.

Conclusion:

▶ Smoking increases the probability of lung cancer by 17%.

▶ Smoking causes lung cancer.
An example.
Random experiment: Pick a person at random.
An example.
Random experiment: Pick a person at random.
Event $A$: the person has lung cancer.
An example.
Random experiment: Pick a person at random.
Event $A$: the person has lung cancer.
Event $B$: the person is a heavy smoker.
An example.
Random experiment: Pick a person at random.
Event $A$: the person has lung cancer.
Event $B$: the person is a heavy smoker.

Fact:

$$Pr[A|B] = 1.17 \times Pr[A].$$

Conclusion:
▶ Smoking increases the probability of lung cancer by 17%.
▶ Smoking causes lung cancer.
Correlation

An example.
Random experiment: Pick a person at random.
Event $A$: the person has lung cancer.
Event $B$: the person is a heavy smoker.

Fact:
\[ Pr[A|B] = 1.17 \times Pr[A]. \]

Conclusion:
- Smoking increases the probability of lung cancer by 17%.
Correlation

An example.
Random experiment: Pick a person at random.
Event $A$: the person has lung cancer.
Event $B$: the person is a heavy smoker.

Fact:

$$Pr[A|B] = 1.17 \times Pr[A].$$

Conclusion:

- Smoking increases the probability of lung cancer by 17%.
- Smoking causes lung cancer.
Correlation


A second look.

Note that $Pr[A|B] = 1.17 \times Pr[A] \iff Pr[A \cap B] = 1.17 \times Pr[A] \iff Pr[B|A] = 1.17 \times Pr[B]$. 

Conclusion:

▶ Lung cancer increases the probability of smoking by 17%.
▶ Lung cancer causes smoking.

Really?


**Correlation**


A second look.
Correlation


A second look.

Note that

$$Pr[A|B] = 1.17 \times Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A]$$

Conclusion:

$\uparrow$ Lung cancer increases the probability of smoking by 17%.

$\uparrow$ Lung cancer causes smoking.

Really?
Correlation


A second look.

Note that

$$Pr[A|B] = 1.17 \times Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A]$$

$$\iff Pr[A \cap B] = 1.17 \times Pr[A] Pr[B]$$

Conclusion:

▶ Lung cancer increases the probability of smoking by 17%.

▶ Lung cancer causes smoking. Really?
Correlation


A second look.

Note that

$$Pr[A|B] = 1.17 \times Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A]$$

$$\iff Pr[A \cap B] = 1.17 \times Pr[A]Pr[B]$$

$$\iff Pr[B|A] = 1.17 \times Pr[B].$$

Conclusion:

▶ Lung cancer increases the probability of smoking by 17\%.

▶ Lung cancer causes smoking.

Really?
Correlation


A second look.

Note that

\[ Pr[A|B] = 1.17 \times Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A] \]
\[ \iff Pr[A \cap B] = 1.17 \times Pr[A] Pr[B] \]
\[ \iff Pr[B|A] = 1.17 \times Pr[B]. \]

Conclusion:

- Lung cancer increases the probability of smoking by 17%.
Correlation


A second look.

Note that

$$Pr[A|B] = 1.17 \times Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A]$$

$$\iff Pr[A \cap B] = 1.17 \times Pr[A]Pr[B]$$

$$\iff Pr[B|A] = 1.17 \times Pr[B].$$

Conclusion:

- Lung cancer increases the probability of smoking by 17%.
- Lung cancer causes smoking.
Correlation


A second look.

Note that

$$Pr[A|B] = 1.17 \times Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A]$$

$$\iff Pr[A \cap B] = 1.17 \times Pr[A]Pr[B]$$

$$\iff Pr[B|A] = 1.17 \times Pr[B].$$

Conclusion:

- Lung cancer increases the probability of smoking by 17%.
- Lung cancer causes smoking. Really?
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

\[
Pr[A \cap B] > Pr[A]Pr[B].
\]

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

▶ Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.

▶ People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.

▶ Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$. 

Other examples:

▶ Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.

▶ People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.

▶ Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

▶ Tesla owners are more likely to be rich.

▶ People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.

▶ Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?
Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career.
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses.
Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

\[ Pr[A \cap B] > Pr[A]Pr[B]. \]

(E.g., smoking and lung cancer.)

$A$ and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?
Proving Causality

Proving causality is generally difficult.
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:

- $A$ and $B$ may be positively correlated because they have a common cause.
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:

- $A$ and $B$ may be positively correlated because they have a common cause. (E.g., being a rabbit.)
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:

- $A$ and $B$ may be positively correlated because they have a common cause. (E.g., being a rabbit.)

- If $B$ precedes $A$, then $B$ is more likely to be the cause. (E.g., smoking.)
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:

- $A$ and $B$ may be positively correlated because they have a common cause. (E.g., being a rabbit.)

- If $B$ precedes $A$, then $B$ is more likely to be the cause. (E.g., smoking.) However, they could have a common cause that induces $B$ before $A$. 
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:

- $A$ and $B$ may be positively correlated because they have a common cause. (E.g., being a rabbit.)

- If $B$ precedes $A$, then $B$ is more likely to be the cause. (E.g., smoking.) However, they could have a common cause that induces $B$ before $A$. (E.g., smart, CS70, Tesla.)
Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:

- $A$ and $B$ may be positively correlated because they have a common cause. (E.g., being a rabbit.)

- If $B$ precedes $A$, then $B$ is more likely to be the cause. (E.g., smoking.) However, they could have a common cause that induces $B$ before $A$. (E.g., smart, CS70, Tesla.)

More about such questions later. For fun, check “N. Taleb: Fooled by randomness.”
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$. 

\[ \Pr[B] = \Pr[A_1 \cap B] + \cdots + \Pr[A_N \cap B]. \] 

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$. 

Thus, 

\[ \Pr[B] = \Pr[A_1] \Pr[B \mid A_1] + \cdots + \Pr[A_N] \Pr[B \mid A_N]. \]
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

![Diagram showing $\Omega$, $A_1, A_2, A_N$ intersecting with $B$]

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$. 
Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$. Thus,

Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

\[ Pr[B] = Pr[A_1]Pr[B|A_1] + \cdots + Pr[A_N]Pr[B|A_N]. \]
Is your coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

---

Analysis:

$A = \text{'coin is fair'}$, $B = \text{'outcome is heads'}$

We want to calculate $Pr[A|B]$.

We know $Pr[B|A] = 1/2$, $Pr[B|\bar{A}] = 0.6$, $Pr[A] = 1/2 = Pr[\bar{A}]$.

Now, $Pr[B] = Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}] = (1/2)(1/2) + (1/2)(0.6) = 0.55$.

Thus, $Pr[A|B] = Pr[A]Pr[B|A]/Pr[B] = (1/2)(1/2)/(0.55) \approx 0.45$. 
Is your coin loaded?

Your coin is fair w.p. 1/2 or such that \( Pr[H] = 0.6 \), otherwise.

You flip your coin and it yields heads.
Is your coin loaded?

- Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.
- You flip your coin and it yields heads.
- What is the probability that it is fair?
Is your coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**
Is your coin loaded?

Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**

$$A = \text{‘coin is fair’},$$
Is your coin loaded?

Your coin is fair w.p. 1/2 or such that \( Pr[H] = 0.6 \), otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

Analysis:

\[ A = \text{‘coin is fair’}, B = \text{‘outcome is heads’} \]
Is your coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**

$$A = \text{‘coin is fair’}, B = \text{‘outcome is heads’}$$

We want to calculate $P[A|B]$. 

Is your coin loaded?

Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**

$A = \text{\textquoteleft} \text{coin is fair}'$, $B = \text{\textquoteleft} \text{outcome is heads}'$

We want to calculate $P[A|B]$.

We know $P[B|A] =$

$=$
Is your coin loaded?
Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.
You flip your coin and it yields heads.
What is the probability that it is fair?

Analysis:

$A = \text{‘coin is fair’}, B = \text{‘outcome is heads’}$

We want to calculate $P[A|B]$.
We know $P[B|A] = 1/2$, $P[B|\bar{A}] = \ldots$
Is your coin loaded?

Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

Analysis:

$A = \text{‘coin is fair’}, B = \text{‘outcome is heads’}$

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2, P[B|\bar{A}] = 0.6,$
Is your coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**

$$A = \text{‘coin is fair’}, \quad B = \text{‘outcome is heads’}$$

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2$, $P[B|\bar{A}] = 0.6$, $Pr[A] =$
Is your coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.
What is the probability that it is fair?

Analysis:

$A = \text{‘coin is fair’}, \; B = \text{‘outcome is heads’}

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2, \; P[B|\bar{A}] = 0.6, \; Pr[A] = 1/2$
Is your coin loaded?

Your coin is fair w.p. 1/2 or such that \( Pr[H] = 0.6 \), otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**

\[ A = \text{‘coin is fair’}, \quad B = \text{‘outcome is heads’} \]

We want to calculate \( P[A|B] \).

We know \( P[B|A] = 1/2 \), \( P[B|\bar{A}] = 0.6 \), \( Pr[A] = 1/2 = Pr[\bar{A}] \)
Is your coin loaded?

Your coin is fair w.p. 1/2 or such that \( Pr[H] = 0.6 \), otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

**Analysis:**

\[ A = \text{‘coin is fair’}, \ B = \text{‘outcome is heads’} \]

We want to calculate \( P[A|B] \).

We know \( P[B|A] = 1/2, \ P[B|\bar{A}] = 0.6, \ Pr[A] = 1/2 = Pr[\bar{A}] \)

Now,

\[
Pr[B] = Pr[A \cap B] + Pr[\bar{A} \cap B] =
\]
Is your coin loaded?
Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.
You flip your coin and it yields heads.
What is the probability that it is fair?

Analysis:

Let $A = \text{\textquoteleft coin is fair\textquoteright}$, $B = \text{\textquoteleft outcome is heads\textquoteright}$

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2$, $P[B|\bar{A}] = 0.6$, $Pr[A] = 1/2 = Pr[\bar{A}]$

Now,

$$Pr[B] = Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}]$$
Is your coin loaded?
Your coin is fair w.p. 1/2 or such that \( Pr[H] = 0.6 \), otherwise.
You flip your coin and it yields heads.
What is the probability that it is fair?

**Analysis:**

\[ A = \text{‘coin is fair’}, B = \text{‘outcome is heads’} \]

We want to calculate \( P[A|B] \).
We know \( P[B|A] = 1/2, P[B|\bar{A}] = 0.6, Pr[A] = 1/2 = Pr[\bar{A}] \)
Now,

\[
Pr[B] = Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}]
= (1/2)(1/2) + (1/2)(1/2)0.6 = 0.55.
\]
Is your coin loaded?
Your coin is fair w.p. 1/2 or such that $Pr[H] = 0.6$, otherwise.
You flip your coin and it yields heads.
What is the probability that it is fair?

Analysis:

$A = \text{‘coin is fair’}, B = \text{‘outcome is heads’}$

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2, P[B|\bar{A}] = 0.6, Pr[A] = 1/2 = Pr[\bar{A}]$

Now,

$$Pr[B] = Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}] = (1/2)(1/2) + (1/2)0.6 = 0.55.$$ 

Thus,

$$Pr[A|B] = \frac{Pr[A]Pr[B|A]}{Pr[B]} = \frac{(1/2)(1/2)}{(1/2)(1/2) + (1/2)0.6} \approx 0.45.$$
Is your coin loaded?

A picture:
Is your coin loaded?

A picture:

Imagine 100 situations, among which $m = 100 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2$ are such that $A$ and $B$ occur and $n = 100 \left(\frac{1}{2}\right)^2 \left(0.6\right)$ are such that $\bar{A}$ and $B$ occur. Thus, among the $m + n$ situations where $B$ occurred, there are $m$ where $A$ occurred. Hence, $\Pr[A | B] = \frac{m}{m + n} = \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \left(0.6\right)$. 
Imagine 100 situations, among which
\( m := 100(1/2)(1/2) \) are such that \( A \) and \( B \) occur and
\( n := 100(1/2)(0.6) \) are such that \( \bar{A} \) and \( B \) occur.
Is your coin loaded?

A picture:

Imagine 100 situations, among which

\[ m := 100 \left( \frac{1}{2} \right) \left( \frac{1}{2} \right) \] are such that \( A \) and \( B \) occur and

\[ n := 100 \left( \frac{1}{2} \right) (0.6) \] are such that \( \bar{A} \) and \( B \) occur.

Thus, among the \( m + n \) situations where \( B \) occurred, there are \( m \) where \( A \) occurred.
Is your coin loaded?

A picture:

Imagine 100 situations, among which $m := 100(1/2)(1/2)$ are such that $A$ and $B$ occur and $n := 100(1/2)(0.6)$ are such that $\bar{A}$ and $B$ occur.

Thus, among the $m + n$ situations where $B$ occurred, there are $m$ where $A$ occurred.

Hence,

$$Pr[A|B] = \frac{m}{m+n} = \frac{(1/2)(1/2)}{(1/2)(1/2) + (1/2)0.6}.$$
Independence

**Definition:** Two events $A$ and $B$ are **independent** if

\[
\Pr[A \cap B] = \Pr[A] \Pr[B].
\]
Independence

**Definition:** Two events $A$ and $B$ are **independent** if

$$Pr[A \cap B] = Pr[A] Pr[B].$$
Independence

**Definition:** Two events $A$ and $B$ are independent if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

Examples:

- When rolling two dice, $A = \text{sum is 7}$ and $B = \text{red die is 1}$ are independent;
- When rolling two dice, $A = \text{sum is 3}$ and $B = \text{red die is 1}$ are not independent;
- When flipping coins, $A = \text{coin 1 yields heads}$ and $B = \text{coin 2 yields tails}$ are independent;
- When throwing 3 balls into 3 bins, $A = \text{bin 1 is empty}$ and $B = \text{bin 2 is empty}$ are not independent;
Independence

**Definition:** Two events $A$ and $B$ are independent if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

Examples:

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are not independent;
- When flipping coins, $A =$ coin 1 yields heads and $B =$ coin 2 yields tails are independent;
- When throwing 3 balls into 3 bins, $A =$ bin 1 is empty and $B =$ bin 2 is empty are not independent;
Independence

**Definition:** Two events $A$ and $B$ are *independent* if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

Examples:

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
Independence

**Definition:** Two events $A$ and $B$ are independent if

$$Pr[A \cap B] = Pr[A] \cdot Pr[B].$$

Examples:

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are not independent;
Independence

**Definition:** Two events $A$ and $B$ are **independent** if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

Examples:

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are **not** independent;
Independence

**Definition:** Two events \( A \) and \( B \) are **independent** if

\[
Pr[A \cap B] = Pr[A]Pr[B].
\]

**Examples:**

- When rolling two dice, \( A = \) sum is 7 and \( B = \) red die is 1 are independent;
- When rolling two dice, \( A = \) sum is 3 and \( B = \) red die is 1 are **not** independent;
- When flipping coins, \( A = \) coin 1 yields heads and \( B = \) coin 2 yields tails are
Independence

**Definition:** Two events $A$ and $B$ are **independent** if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

**Examples:**

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are **not** independent;
- When flipping coins, $A =$ coin 1 yields heads and $B =$ coin 2 yields tails are independent;
Independence

**Definition:** Two events $A$ and $B$ are **independent** if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

**Examples:**

- When rolling two dice, $A = \text{sum is 7}$ and $B = \text{red die is 1}$ are independent;
- When rolling two dice, $A = \text{sum is 3}$ and $B = \text{red die is 1}$ are **not** independent;
- When flipping coins, $A = \text{coin 1 yields heads}$ and $B = \text{coin 2 yields tails}$ are independent;
- When throwing 3 balls into 3 bins, $A = \text{bin 1 is empty}$ and $B = \text{bin 2 is empty}$ are
Independence

**Definition:** Two events $A$ and $B$ are **independent** if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

**Examples:**

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are *not* independent;
- When flipping coins, $A =$ coin 1 yields heads and $B =$ coin 2 yields tails are independent;
- When throwing 3 balls into 3 bins, $A =$ bin 1 is empty and $B =$ bin 2 is empty are *not* independent;
Fact: Two events $A$ and $B$ are independent if and only if

$$
\Pr[A|B] = \Pr[A] \iff \Pr[A \cap B] = \Pr[A] \Pr[B].
$$
Fact: Two events $A$ and $B$ are independent if and only if

$$Pr[A|B] = Pr[A].$$
Independence and conditional probability

Fact: Two events $A$ and $B$ are independent if and only if

$$Pr[A|B] = Pr[A].$$

Indeed:
Fact: Two events $A$ and $B$ are independent if and only if

$$Pr[A|B] = Pr[A].$$

Indeed: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$, so that
Independence and conditional probability

**Fact:** Two events $A$ and $B$ are independent if and only if

$$Pr[A|B] = Pr[A].$$

Indeed: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$, so that

$$Pr[A|B] = Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = Pr[A].$$
Fact: Two events $A$ and $B$ are independent if and only if

$$Pr[A|B] = Pr[A].$$

Indeed: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$, so that

$$Pr[A|B] = Pr[A] \iff \frac{Pr[A \cap B]}{Pr[B]} = Pr[A] \iff Pr[A \cap B] = Pr[A]Pr[B].$$
Bayes Rule

Another picture: We imagine that there are $N$ possible causes $A_1, \ldots, A_N$. 

$$
\text{Pr}[A_n | B] = \frac{p_n q_n}{\sum_m p_m q_m}.
$$
Bayes Rule

Another picture: We imagine that there are $N$ possible causes $A_1, \ldots, A_N$. 

$A_1, \ldots, A_N$ disjoint

$A_1 \cup \cdots \cup A_N = \Omega$
Bayes Rule

Another picture: We imagine that there are $N$ possible causes $A_1, \ldots, A_N$.

Imagine 100 situations, among which $100p_nq_n$ are such that $A_n$ and $B$ occur, for $n = 1, \ldots, N$.

Thus, among the $100\sum_m p_m q_m$ situations where $B$ occurred, there are $100p_nq_n$ where $A_n$ occurred.
Bayes Rule

Another picture: We imagine that there are \( N \) possible causes \( A_1, \ldots, A_N \).

Imagine 100 situations, among which 100\( p_nq_n \) are such that \( A_n \) and \( B \) occur, for \( n = 1, \ldots, N \).

Thus, among the 100\( \sum_m p_mq_m \) situations where \( B \) occurred, there are 100\( p_nq_n \) where \( A_n \) occurred.

Hence,

\[
Pr[A_n|B] = \frac{p_nq_n}{\sum_m p_mq_m}.
\]
Why do you have a fever?

Using Bayes' rule, we find:

\[
\Pr[\text{Flu} | \text{High Fever}] = 0.15 \times 0.80 / \left(0.15 \times 0.80 + 10^{-8} \times 0.85 \times 0.10\right) \approx 0.58
\]

\[
\Pr[\text{Ebola} | \text{High Fever}] = 10^{-8} \times 0.85 / \left(0.15 \times 0.80 + 10^{-8} \times 0.85 \times 0.10\right) \approx 5 \times 10^{-8}
\]

\[
\Pr[\text{Other} | \text{High Fever}] = 0.85 \times 0.15 / \left(0.15 \times 0.80 + 10^{-8} \times 0.85 \times 0.10\right) \approx 0.42
\]

These are the posterior probabilities. One says that 'Flu' is the Most Likely a Posteriori (MAP) cause of the high fever.
Why do you have a fever?

Using Bayes’ rule, we find

\[
\begin{align*}
\Pr[\text{Flu} \mid \text{High Fever}] &= 0.15 \times 0.80 \approx 0.12 \approx 0.58, \\
\Pr[\text{Ebola} \mid \text{High Fever}] &= 10^{-8} \times 1 \approx 0.0000000010, \\
\Pr[\text{Other} \mid \text{High Fever}] &= 0.85 \times 0.1 \approx 0.085 \approx 0.42.
\end{align*}
\]

These are the posterior probabilities. One says that ‘Flu’ is the Most Likely a Posteriori (MAP) cause of the high fever.
Why do you have a fever?

Using Bayes’ rule, we find

\[
Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58
\]
Why do you have a fever?

Using Bayes’ rule, we find

\[
Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58
\]

\[
Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}
\]
Why do you have a fever?

Using Bayes’ rule, we find

\[
Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58
\]

\[
Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}
\]

\[
Pr[\text{Other}|\text{High Fever}] = \frac{0.85 \times 0.1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.42
\]
Why do you have a fever?

Using Bayes’ rule, we find

\[
Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58
\]

\[
Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}
\]

\[
Pr[\text{Other}|\text{High Fever}] = \frac{0.85 \times 0.1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.42
\]

These are the posterior probabilities.
Why do you have a fever?

Using Bayes’ rule, we find

\[
Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58
\]

\[
Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}
\]

\[
Pr[\text{Other}|\text{High Fever}] = \frac{0.85 \times 0.1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.42
\]

These are the posterior probabilities. One says that ‘Flu’ is the Most Likely a Posteriori (MAP) cause of the high fever.
Bayes’ Rule Operations
Bayes’ Rule Operations

Bayes’ Rule is the canonical example of how information changes our opinions.

Priors: $Pr[A_n]$  
Observe $B$  
Conditional: $Pr[B|A_n]$  
[Model of system]  

Posterior: $Pr[A_n|B]$  

[Environment]
Bayes’ Rule is the canonical example of how information changes our opinions.
Thomas Bayes

Portrait used of Bayes in a 1936 book,[1] but it is doubtful whether the portrait is actually of him.[2]
No earlier portrait or claimed portrait survives.

**Born**  
c. 1701  
London, England

**Died**  
7 April 1761 (aged 59)  
Tunbridge Wells, Kent, England

**Residence**  
Tunbridge Wells, Kent, England

**Nationality**  
English

**Known for**  
Bayes' theorem

Thomas Bayes

A Bayesian picture of Thomas Bayes.

Fig. 3. Joshua Bayes (1671–1746).
Testing for disease.

Let’s watch TV!!
Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.

Outcomes: 
- test
- disease

A - prostate cancer.
B - positive PSA test.

▶ Pr[A] = 0.0016, (0.16% of the male population is affected.)
▶ Pr[B|A] = 0.80 (80% chance of positive test with disease.)
▶ Pr[B|A'] = 0.10 (10% chance of positive test without disease.)


Positive PSA test (B). Do I have disease? Pr[A|B]??
Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.
Outcomes: \((test, disease)\)
Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.
Outcomes: \((test, disease)\)
\(A\) - prostate cancer.
\(B\) - positive PSA test.

\[
\begin{align*}
\Pr[A] &= 0.0016, \quad (0.16\% \text{ of the male population is affected.}) \\
\Pr[B|A] &= 0.80 \quad (80\% \text{ chance of positive test with disease.}) \\
\Pr[B|\overline{A}] &= 0.10 \quad (10\% \text{ chance of positive test without disease.})
\end{align*}
\]


Positive PSA test \((B)\). Do I have disease? \(\Pr[A|B]??\)
Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.
Outcomes: \((test, disease)\)
\(A\) - prostate cancer.
\(B\) - positive PSA test.

- \(Pr[A] = 0.0016\), (.16 % of the male population is affected.)
- \(Pr[B|A] = 0.80\) (80% chance of positive test with disease.)
- \(Pr[B|\overline{A}] = 0.10\) (10% chance of positive test without disease.)
Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.
Outcomes: (test, disease)
A - prostate cancer.
B - positive PSA test.

- $Pr[A] = 0.0016$, (.16 % of the male population is affected.)
- $Pr[B|A] = 0.80$ (80% chance of positive test with disease.)
- $Pr[B|\overline{A}] = 0.10$ (10% chance of positive test without disease.)

Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.
Outcomes: \((\text{test, disease})\)
\(A\) - prostate cancer.
\(B\) - positive PSA test.

- \(Pr[A] = 0.0016\), (.16 % of the male population is affected.)
- \(Pr[B|A] = 0.80\) (80% chance of positive test with disease.)
- \(Pr[B|\overline{A}] = 0.10\) (10% chance of positive test without disease.)


Positive PSA test \((B)\). Do I have disease?
Testing for disease.

Let's watch TV!!
Random Experiment: Pick a random male.
Outcomes: (test, disease)
A - prostate cancer.
B - positive PSA test.

- Pr[A] = 0.0016, (.16 % of the male population is affected.)
- Pr[B|A] = 0.80 (80% chance of positive test with disease.)
- Pr[B|\overline{A}] = 0.10 (10% chance of positive test without disease.)


Positive PSA test (B). Do I have disease?

Pr[A|B]???
Bayes Rule.

Using Bayes' rule, we find

\[ P[A|B] = 0.0016 \times 0.80 + 0.9984 \times 0.10 = 0.013 \]

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?

Impotence...

Incontinence..

Death.
Bayes Rule.

Using Bayes’ rule, we find

\[ P[A|B] = 0.0016 \times 0.80 + 0.9984 \times 0.10 = 0.013 \]

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?

Impotence...

Incontinence...

Death.
Using Bayes’ rule, we find

\[
P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10}
\]
Bayes Rule.

Using Bayes’ rule, we find

\[ P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = .013. \]
Using Bayes’ rule, we find

\[
P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = 0.013.
\]

A 1.3% chance of prostate cancer with a positive PSA test.
Bayes Rule.

Using Bayes’ rule, we find

\[
P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = 0.013.
\]

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?
Bayes Rule.

Using Bayes’ rule, we find

\[ P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = 0.013. \]

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?
Impotence...
Using Bayes’ rule, we find

\[
P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = 0.013.
\]

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?

Impotence...

Incontinence..
Bayes Rule.

Using Bayes’ rule, we find

\[ P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = 0.013. \]

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?

Impotence...

Incontinence..

Death.
Summary

Events, Conditional Probability, Independence, Bayes’ Rule
Summary

Events, Conditional Probability, Independence, Bayes’ Rule

Key Ideas:

- Conditional Probability:
  
  \[ Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} \]
Summary

Key Ideas:

- **Conditional Probability:**
  \[ Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} \]

- **Independence:** \( Pr[A \cap B] = Pr[A]Pr[B] \).
Summary

Events, Conditional Probability, Independence, Bayes’ Rule

Key Ideas:

- **Conditional Probability:**
  \[ Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} \]

- **Independence:** \( Pr[A \cap B] = Pr[A]Pr[B] \).

- **Bayes’ Rule:**
  \[ Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]} \].
Summary

**Events, Conditional Probability, Independence, Bayes’ Rule**

**Key Ideas:**

- **Conditional Probability:**
  \[ Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} \]

- **Independence:**
  \[ Pr[A \cap B] = Pr[A]Pr[B]. \]

- **Bayes’ Rule:**
  \[ Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]} \cdot \]
  
  \[ Pr[A_n|B] = \text{posterior probability}; \ Pr[A_n] = \text{prior probability} \]
Summary

Events, Conditional Probability, Independence, Bayes’ Rule

Key Ideas:

▶ **Conditional Probability:**

\[ Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} \]

▶ **Independence:**

\[ Pr[A \cap B] = Pr[A]Pr[B] \]

▶ **Bayes’ Rule:**

\[ Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]} \]

\[ Pr[A_n|B] = \text{posterior probability}; \ Pr[A_n] = \text{prior probability} \]

▶ All these are possible:

\[ Pr[A|B] < Pr[A]; \ Pr[A|B] > Pr[A]; \ Pr[A|B] = Pr[A] \]