Events, Conditional Probability, Independence, Bayes’ Rule
1. Probability Basics Review
2. Conditional Probability
3. Independence of Events
4. 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. \( \sum_{\omega \in \Omega} \Pr[\omega] = 1 \).
▶ Event.
Set of the outcomes.
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. \( \sum_{\omega \in \Omega} \Pr[\omega] = 1 \).

Event. Set of the outcomes.
Probability Basics Review

Setup:

- Random Experiment.

Probability Space:

Sample Space: Set of outcomes, \( \Omega \).

\( \Omega = \{ \text{HH}, \text{HT}, \text{TH}, \text{TT} \} \)

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

Probability:

\[ \Pr[\omega] \text{ for all } \omega \in \Omega. \]

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

1. \( 0 \leq \Pr[\omega] \leq 1 \).

2. \( \sum_{\omega \in \Omega} \Pr[\omega] = 1 \).

Event:

Set of the outcomes.
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, \( \Omega \).

\[ \Omega = \{ HH, HT, TH, TT \} \] (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, $\Omega$.
    $\Omega = \{HH, HT, TH, TT\}$
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 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 \( \omega \in \Omega \).
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 \( \omega \in \Omega \).
    \[ Pr[HH] = \cdots = Pr[TT] = \frac{1}{4} \]
    1. \( 0 \leq Pr[\omega] \leq 1 \).
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 $\omega \in \Omega$.
  $Pr[HH] = \cdots = Pr[TT] = 1/4$
  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. $\sum_{\omega \in \Omega} Pr[\omega] = 1$.

- Event.
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. $\sum_{\omega \in \Omega} Pr[\omega] = 1$.
- **Event.** Set of the outcomes.
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. $\sum_{\omega \in \Omega} Pr[\omega] = 1$.

- Event. *Set of the outcomes.*
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.

Can I instead say that $|A \cup B| = |A| + |B|$?

No! We don't know if the sample space is uniform.
Probability is Additive

Theorem

(a) If events $A$ and $B$ are disjoint,
Probability is Additive

**Theorem**

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

$$\Pr[A \cup B] = \Pr[A] + \Pr[B].$$
Probability is Additive

**Theorem**

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

$$Pr[A \cup B] = Pr[A] + Pr[B].$$
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,
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., $A_k \cap A_m = \emptyset$, $\forall k \neq m$, can I instead say that $|A \cup B| = |A| + |B|$?

No! We don't know if the sample space is uniform.
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

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

$$Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] =$$
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.

$$Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] = \sum_{\omega \in A} Pr[\omega] +$$
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.

$$Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] = \sum_{\omega \in A} Pr[\omega] + \sum_{\omega \in B} Pr[\omega]$$

Can I instead say that $|A \cup B| = |A| + |B|$?

No!

We don't know if the sample space is uniform.
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.

$$Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] = \sum_{\omega \in A} Pr[\omega] + \sum_{\omega \in B} Pr[\omega] = Pr[A] + Pr[B].$$
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.

$$Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] = \sum_{\omega \in A} Pr[\omega] + \sum_{\omega \in B} Pr[\omega] = Pr[A] + Pr[B]$$

Can I instead say that $|A \cup B| = |A| + |B|$?
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.

\[ Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] = \sum_{\omega \in A} Pr[\omega] + \sum_{\omega \in B} Pr[\omega] = Pr[A] + Pr[B] \]

Can I instead say that $|A \cup B| = |A| + |B|$?

No!
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.

$$Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega] = \sum_{\omega \in A} Pr[\omega] + \sum_{\omega \in B} Pr[\omega] = Pr[A] + Pr[B]$$

Can I instead say that $|A \cup B| = |A| + |B|$?

**No!** We don’t know if the sample space is uniform.
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. See next two slides for (a) and (c).
Consequences of Additivity

Theorem

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

Proof: (b) is obvious. See next two slides for (a) and (c).
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) \( Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n] \);
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)
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 \( \Omega \),
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 $\bigcup_{m=1}^{N} A_m = \Omega$, 
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
    \( Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N]. \)
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 \( \cup_{m=1}^{N} A_m = \Omega \), then

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

\text{(law of total probability)}

Proof: 
(b) is obvious. See next two slides for (a) and (c).
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.

See next two slides for (a) and (c).
Inclusion/Exclusion

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

Can I instead say that \( |A \cup B| = |A| + |B| - |A \cap B| \)?

No! We don't know if the sample space is uniform.
Inclusion/Exclusion

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

Can I instead say that
\[ |A \cup B| = |A| + |B| - |A \cap B| \]

No!

We don't know if the sample space is uniform.
Inclusion/Exclusion

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

Can I instead say that \( |A \cup B| = |A| + |B| - |A \cap B| \)?
Inclusion/Exclusion

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

Can I instead say that \(|A \cup B| = |A| + |B| - |A \cap B|\)?

No!
Inclusion/Exclusion

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

Can I instead say that \(|A \cup B| = |A| + |B| - |A \cap B|\)?

No! We don’t know if the sample space is uniform.
Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$. 

\[
\text{Total probability} \\

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

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 $n = 1, \ldots, N$. 

Roll a Red and a Blue Die.

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

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

\[ E_1 = \text{'Red die shows 6'} \]
\[ E_2 = \text{'Blue die shows 6'} \]
\[ E_1 \cup E_2 = \text{'At least one die shows 6'} \]
\[ \Pr[E_1] = \frac{1}{36}, \quad \Pr[E_2] = \frac{1}{36}, \quad \Pr[E_1 \cup E_2] = \frac{11}{36} \]

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

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

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

$\Pr[E_1] = \frac{1}{6}$, $\Pr[E_2] = \frac{1}{6}$, $\Pr[E_1 \cup E_2] = \frac{11}{36}$. 

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

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

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

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

\[ E_1 = 'Red die shows 6'; E_2 = 'Blue die shows 6' \]
\[ E_1 \cup E_2 = 'At least one die shows 6' \]
\[ Pr[E_1] = \frac{6}{36}, \]
\[ Pr[E_2] = \frac{6}{36}, \]
\[ Pr[E_1 \cup E_2] = \frac{11}{36}. \]
Roll a Red and a Blue Die.

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

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

\[ Pr[E_1] = \frac{6}{36}, \ Pr[E_2] = \frac{6}{36}, \]

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

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

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

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

Two coin flips (fair coin).
Conditional probability: example.

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

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

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

Two coin flips (fair coin). First flip is heads. Probability of two heads?

\[ \Omega = \{ HH, HT, TH, TT \}; \text{ Uniform probability space.} \]
Conditional probability: example.

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

Two coin flips (fair coin). First flip is heads. Probability of two heads?

\( \Omega = \{ HH, HT, TH, TT \} \); Uniform probability space.

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

\( \Omega : \text{uniform} \)

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

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

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

New sample space: A;
Conditional probability: example.

Two coin flips (fair coin). First flip is heads. Probability of two heads?

\( \Omega = \{ HH, HT, TH, TT \} \); Uniform probability space.

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

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

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

\(A\) : uniform

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

Two coin flips (fair coin). First flip is heads. Probability of two heads?

\( \Omega = \{ HH, HT, TH, TT \} \); Uniform probability space.

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

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

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

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

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

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

Two coin flips (fair coin). First flip is heads. Probability of two heads?

\( \Omega = \{HH, HT, TH, TT\} \); Uniform probability space.

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

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

Event \( B = \) two heads.

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

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

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

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 1/2.
A similar example.

Two coin flips (fair coin).
A similar example.

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

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

Two coin flips (fair coin). At least one of the flips is heads.

→ Probability of two heads?

\( \Omega = \{ HH, HT, TH, TT \} \);
A similar example.

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

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

Two coin flips (fair coin). 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 similar example.

Two coin flips (fair coin). At least one of the flips is heads. 
$$\rightarrow$$ 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 (fair coin). 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. \( A = \{ HH, HT, TH \} \).
A similar example.

Two coin flips (fair coin). 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: \(A\);
A similar example.

Two coin flips (fair coin). 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 (fair coin). 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.
A similar example.

Two coin flips (fair coin). 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.

Event \( B = \) two heads.
A similar example.

Two coin flips (fair coin). 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.

Event \( B \) = two heads.

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

Two coin flips (fair coin). 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.

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 (fair coin). 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.

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

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

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

Physical experiment

Probability model

Ω = {Red, Green, Yellow, Blue}

\[Pr[\omega]\]

- Red 3/10
- Green 4/10
- Yellow 2/10
- Blue 1/10

\[Pr[\text{Red} | \text{Red or Green}] = \frac{3}{7} = Pr[\text{Red} \cap (\text{Red or Green})] / Pr[\text{Red or Green}]\]
Conditional Probability: A non-uniform example

Physical experiment

Probability model

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

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

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

Physical experiment

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

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

Probability model

- Red
  - 3/10
- Green
  - 4/10
- Yellow
  - 2/10
- Blue
  - 1/10
Conditional Probability: A non-uniform example

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

\[ Pr[\text{Red}|\text{Red or Green}] = \frac{3}{7} = \frac{Pr[\text{Red} \cap (\text{Red or Green})]}{Pr[\text{Red or Green}]} \]
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 $A = \{3, 4\}, B = \{1, 2, 3\}$. 

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

[Diagram showing a circle divided into segments labeled 1, 2, 3, 4, and a variable \( \omega \), with probabilities labeled as \( p_1, p_2, p_3, p_4, p_\omega \).]
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 $Pr[n] = p_n$. 
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 \( A = \{2, 3, 4\}, B = \{1, 2, 3\} \).
Yet another non-uniform example

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

\[
Pr[A|B] = Pr[A \cap B] / Pr[B].
\]
Yet another non-uniform example

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

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

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

- **In $A$!**
- **In $B$?**
- **Must be in $A \cap B$.**
Conditional Probability.

**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?

\[ P(B \mid A) = \frac{B \cap A}{|A|} = \frac{1}{3}; \]

versus \[ P(B) = \frac{1}{6}. \]

\( B \) 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?

\[ Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{3}, \]

\( \Omega \): Uniform

\( \Omega = \{1, \ldots, 6\}^2 \)

\( A = \{(1,3), (2,2), (3,1)\} \)

\( B = \{(1,1), \ldots, (1,6)\} \)

\( B \) 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?

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

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

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

Observing \( A \) does not change your mind about the likelihood of \( B \).
Yet more fun with conditional probability.

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

\[
\begin{align*}
\Pr[B|A] &= \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{1}{6};
\end{align*}
\]

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

Observing \(A\) does not change your mind about the likelihood of \(B\).
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?

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

Suppose I toss 3 balls into 3 bins.
Suppose I toss 3 balls into 3 bins. 
\( A \) = “1st bin empty”;

\[ \text{Pr}[A|B] = \frac{\text{Pr}[A \cap B]}{\text{Pr}[B]} = \frac{1/27}{8/27} = \frac{1}{8}; \]

\[ \text{vs.} \quad \text{Pr}[A] = \frac{8}{27}. \]

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

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

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

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

$Pr[B]$
Emptiness..

Suppose I toss 3 balls into 3 bins.

$A =$“1st bin empty”; $B =$“2nd bin empty.” What is $Pr[A|B]$?

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

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

$Pr[B] = Pr[\{(a, b, c) \mid a, b, c \in \{1, 3\}\}] =$
Suppose I toss 3 balls into 3 bins. 

$A = \text{“1st bin empty”}$; $B = \text{“2nd bin empty.”}$ What is $Pr[A|B]$?

![Diagram showing possible outcomes](image)

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

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

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

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

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{\frac{1}{27}}{\frac{8}{27}} = \frac{1}{8}$; 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.
Suppose I toss 3 balls into 3 bins. 

$A =$ “1st bin empty”; $B =$ “2nd bin empty.” What is $Pr[A|B]$?

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

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

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

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

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

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

Suppose I toss 3 balls into 3 bins.

\(A = \text{“1st bin empty”}; \ B = \text{“2nd bin empty.”} \) What is \(Pr[A|B]\)?

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

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

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

\[Pr[A \cap B] = Pr[(3, 3, 3)] =\]
Suppose I toss 3 balls into 3 bins. 

A = “1st bin empty”; B = “2nd bin empty.” What is $Pr[A \mid B]$?

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$
Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{\left(\frac{1}{27}\right)}{\left(\frac{8}{27}\right)} = \frac{1}{8}; \text{ vs. } Pr[A] = \frac{8}{27}.
$$
Suppose I toss 3 balls into 3 bins. 
\(A = \text{“1st bin empty”}; \ B = \text{“2nd bin empty.”}\) What is \(Pr[A|B]\)?

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

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

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

\(A\) is less likely given \(B\):
Emptiness..

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

\[
\begin{align*}
\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; \text{ vs. } Pr[A] = \frac{8}{27}.
\end{align*}
\]

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.
$A = \text{“first 50 flips are heads”}$
Flap a fair coin 51 times.
$A =$ "first 50 flips are heads"
$B =$ "the 51st is heads"
Flip a fair coin 51 times.
$A = \text{“first 50 flips are heads”}$
$B = \text{“the 51st is heads”}$
$Pr[B|A]$ ?
Gambler’s fallacy.

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

$A = \{HH \cdots HT, HH \cdots HH\}$
Gambler’s fallacy.

Flip a fair coin 51 times.

A = “first 50 flips are heads”
B = “the 51st is heads”

\[ \Pr[B|A] \]

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

\[ B \cap A = \{ HH \cdots HH \} \]
Gambler’s fallacy.

Flip a fair coin 51 times.

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

$B = \text{“the 51st is heads”}$

$Pr[B|A]$?

$A = \{HH\cdots HT, HH\cdots HH\}$

$B \cap A = \{HH\cdots HH\}$

Uniform probability space.
Gambler’s fallacy.

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

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

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”
$Pr[B|A]$?

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

Uniform probability space.

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

Same as $Pr[B]$.

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

Recall the definition:
Product Rule

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,

\[ Pr[A \cap B] = Pr[A]Pr[B|A] . \]
Recall the definition:

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

Hence,

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

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,

\[
Pr[A \cap B \cap C] = Pr[(A \cap B) \cap C] \\
= Pr[A \cap B]Pr[C|A \cap B]
\]
Recall the definition:
\[
Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}.
\]
Hence,
\[
Pr[A \cap B] = Pr[A]Pr[B|A].
\]
Consequently,
\[
Pr[A \cap B \cap C] = Pr[(A \cap B)\cap C] = Pr[A \cap B]Pr[C|A \cap B] = Pr[A]Pr[B|A]Pr[C|A \cap B].
\]
**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

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

**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 \( n + 1 \). \( \square \)
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.

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

Conclusion:
▶ Smoking increases the probability of lung cancer by 17%.
▶ Smoking causes lung cancer.
An example.
Random experiment: Pick a person at random.

Pr[A | B] = 1.17 \times \text{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.
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.

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

\[ \Pr[A \mid B] = 1.17 \times \Pr[A]. \]

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.

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

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


Conclusion:

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

$\blacktriangleright$ 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:

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

▶ Lung cancer causes smoking. Really?
Correlation


A second look.

Note that

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

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

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.
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.
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 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).
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.
  - (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.)
Proving Causality

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.)
Total probability

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

$\Omega$

$A_1$

$A_2$

$A_N$

$B$
Total probability

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

Then,

\[ Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B]. \]
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 $n = 1, \ldots, N$. Thus,

Total probability

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

\[
\]
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

$$Pr[A \cap B] = Pr[A]Pr[B].$$
Definition: Two events $A$ and $B$ are independent if

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

Examples:
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

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

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
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, $A =$ 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
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

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

Indeed:

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

so that

$$\Pr[A | B] = \Pr[A] \iff \Pr[A \cap B] = \Pr[A] \Pr[B].$$
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
**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

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

We know $P[B|A] = 1/2$, $P[B|\neg A] = 0.6$, $Pr[A] = 1/2 = Pr[\neg A]$.

Now,

$Pr[B] = Pr[A \cap B] + Pr[\neg A \cap B] = Pr[A]Pr[B|A] + Pr[\neg A]Pr[B|\neg A] = (1/2)(1/2) + (1/2)(0.6) = 0.55$.

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:

$$A = \text{‘coin is fair’}, \quad B = \text{‘outcome is heads’}$$
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:

$A = \text{\textquoteleft coin is fair\textquoteleft}, \ B = \text{\textquoteleft outcome is heads\textquoteleft}$

We want to calculate $P[A|B]$. 
Is you 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 \( P[B|A] = \)
Is you 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 \( P[B|A] = 1/2 \), \( P[B|\bar{A}] = \)
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 $P[B|A] = 1/2$, $P[B|\bar{A}] = 0.6$,
Is you 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 \( P[B|A] = 1/2, P[B|\bar{A}] = 0.6, Pr[A] = \)
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:**

$A = \text{‘coin is fair’}, B = \text{‘outcome is heads’}$

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:**

$A = \text{`coin is fair'}, B = \text{`outcome is heads'}$

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,

$$Pr[B] = Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}]$$
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,

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

$= (1/2)(1/2) + (1/2)0.6 = 0.55.$
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:
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.
Is your coin loaded?

A picture:

Imagine 100 situations, among which 
\[ m := 100(\frac{1}{2})(\frac{1}{2}) \] are such that \( A \) and \( B \) occur and 
\[ n := 100(\frac{1}{2})(0.6) \] are such that \( \bar{A} \) and \( B \) occur.

Thus, among the \( m + n \) situations where \( B \) occurred, there are \( m \) where \( A \) occurred.
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 $\overline{A}$ and $B$ occur.

Thus, among the $m + n$ situations where $B$ occurred, there are $m$ where $A$ occurred.

Hence,

$$Pr[A|B] = \frac{m}{m+n} = \frac{(1/2)(1/2)}{(1/2)(1/2) + (1/2)0.6}.$$
Why do you have a fever?

Pr[Flu|High Fever] = 0.15 × 0.80 = 0.12

Pr[Ebola|High Fever] = 10^{-8} × 1 = 10^{-8}

Pr[Other|High Fever] = 0.85 × 0.15 × 0.80 + 10^{-8} × 1 = 0.10

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

\[
\begin{align*}
\Pr[\text{Flu} \mid \text{High Fever}] &= 0.15 \times 0.80 + 10^{-8} \\
\Pr[\text{Ebola} \mid \text{High Fever}] &= 0.80 \\
\Pr[\text{Other} \mid \text{High Fever}] &= 0.85 \times 0.15 + 10^{-8}
\end{align*}
\]

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

\[
Pr[Flu|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58
\]

\[
Pr[Ebola|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}
\]
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

[Environment]

Priors: $Pr[A_n]$
Observe $B$

Bayes’ Rule

Posteriors: $Pr[A_n|B]$

Conditional: $Pr[B|A_n]$

[Model of system]
Bayes’ Rule is the canonical example of how information changes our opinions.
Thomas Bayes

<table>
<thead>
<tr>
<th>Born</th>
<th>c. 1701</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>London, England</td>
</tr>
<tr>
<td>Died</td>
<td>7 April 1761 (aged 59)</td>
</tr>
<tr>
<td>Residence</td>
<td>Tunbridge Wells, Kent, England</td>
</tr>
<tr>
<td>Nationality</td>
<td>English</td>
</tr>
<tr>
<td>Known for</td>
<td>Bayes' theorem</td>
</tr>
</tbody>
</table>

A Bayesian picture of Thomas Bayes.

Figure 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:

A - prostate cancer.
B - positive PSA test.

\[ Pr[A] = 0.0016, \text{(0.16% of the male population is affected.)} \]
\[ Pr[B|A] = 0.80 \text{ (80% chance of positive test with disease.)} \]
\[ Pr[B|\neg A] = 0.10 \text{ (10% chance of positive test without disease.)} \]


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.
Testing for disease.

Let’s watch TV!!
Random Experiment: Pick a random male.
Outcomes: \((test, \text{disease})\)

\(A\) - prostate cancer.
\(B\) - positive PSA test.

- \(Pr[A] = 0.0016, (.16 \% \text{ of the male population is affected.})\)
- \(Pr[B|A] = 0.80 \text{ (80\% chance of positive test with disease.)}\)
- \(Pr[B|\bar{A}] = 0.10 \text{ (10\% chance of positive test without disease.)}\)
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: \((test, disease)\)

\(A\) - prostate cancer.
\(B\) - positive PSA test.

\[ Pr[A] = 0.0016, \text{ (.16 \% of the male population is affected.)} \]
\[ Pr[B|A] = 0.80 \text{ (80\% chance of positive test with disease.)} \]
\[ Pr[B|\overline{A}] = 0.10 \text{ (10\% chance of positive test without disease.)} \]


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?

\[ Pr[A|B] \]
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.
Bayes Rule.

Using Bayes’ rule, we find

\[
P(A | B) = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} \approx 0.13
\]

A 1.3% chance of prostate cancer with a positive PSA test.
Using Bayes’ rule, we find

\[
P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10}
\]

A 1.3% chance of prostate cancer with a positive PSA test.
Bayes Rule.

Using Bayes’ rule, we find

\[ P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = 0.013. \]
Bayes Rule.

Using Bayes’ rule, we find

\[
P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = .013.
\]

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. !!!!
Using Bayes’ rule, we find

$$P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = .013.$$ 

A 1.3% chance of prostate cancer with a positive PSA test. !!!!
Monty Hall.
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:**
  \[ Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]} \]
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:**
  \[ Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]} \]

\( Pr[A_n|B] = \text{posterior probability}; \ Pr[A_n] = \text{prior probability} \).
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: