1. Conditional Probability

1.1 Learning Objectives

By the end of this lecture, you will be able to:

- Define and compute conditional probabilities
- Apply the multiplication rule
- Understand how conditioning reduces the sample space
- Use conditional probability in sequential experiments

1.2 Motivating Example

You roll a fair die. What's the probability of getting a 6?

1.2 Motivating Example

You roll a fair die. What's the probability of getting a 6?

$$P(\{6\}) = \frac{1}{6}$$

1.2 Motivating Example

You roll a fair die. What's the probability of getting a 6?

$$P(\{6\}) = \frac{1}{6}$$

Now suppose someone tells you: "A die has been rolled and the result was even."

What's the probability of a 6 **given this information**?

$$S = \{\cancel{1}\ \textcircled{2}\ \cancel{3}\ \textcircled{4}\ \cancel{5}\ \textcircled{6}\}$$

$$P(6 \mid \text{even}) = \frac{1}{3}$$

1.2 Motivating Example

You roll a fair die. What's the probability of getting a 6?

$$P(\{6\}) = \frac{1}{6}$$

Now suppose someone tells you: "A die has been rolled and the result was even."

What's the probability of a 6 **given this information**?

$$P(\{6\} \mid \text{even}) = \frac{1}{3}$$

1.3 The Idea of Conditioning

When we learn that event B occurred:

1.3 The Idea of Conditioning

When we learn that event B occurred:

- We **restrict** our attention to outcomes in B

1.3 The Idea of Conditioning

When we learn that event B occurred:

- We **restrict** our attention to outcomes in B
- The sample space effectively becomes B

1.3 The Idea of Conditioning

When we learn that event B occurred:

- We **restrict** our attention to outcomes in B
- The sample space effectively becomes B
- We rescale probabilities so they sum to 1 over B

$S = \{1, 2, 3, 4, 5, 6\}$

$\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$

$S' = B = \{2, 4, 6\}$

$\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$

1.3 The Idea of Conditioning

When we learn that event B occurred:

- We **restrict** our attention to outcomes in B
- The sample space effectively becomes B
- We rescale probabilities so they sum to 1 over B

1.3 The Idea of Conditioning

When we learn that event B occurred:

- We **restrict** our attention to outcomes in B
- The sample space effectively becomes B
- We rescale probabilities so they sum to 1 over B

Definition: Conditional Probability

The conditional probability of A given B is:

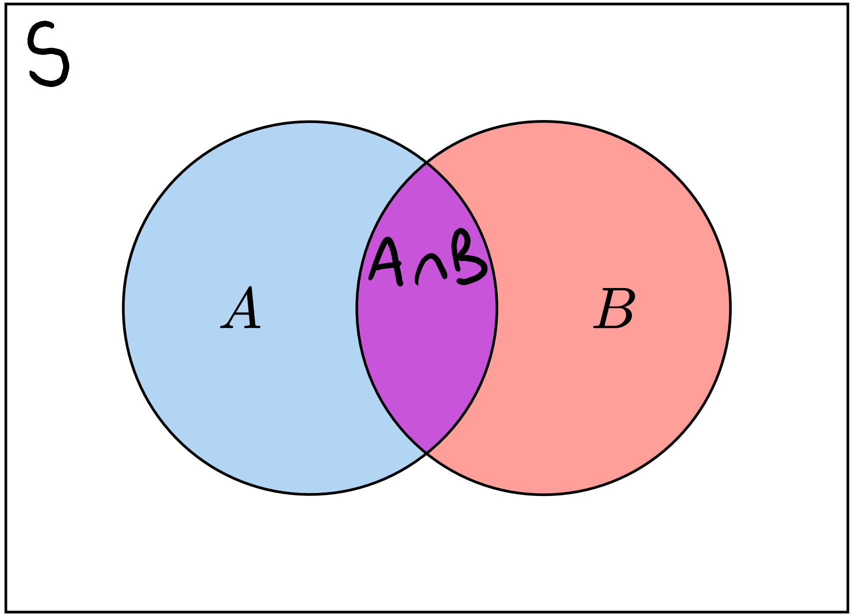
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

provided $P(B) > 0$.

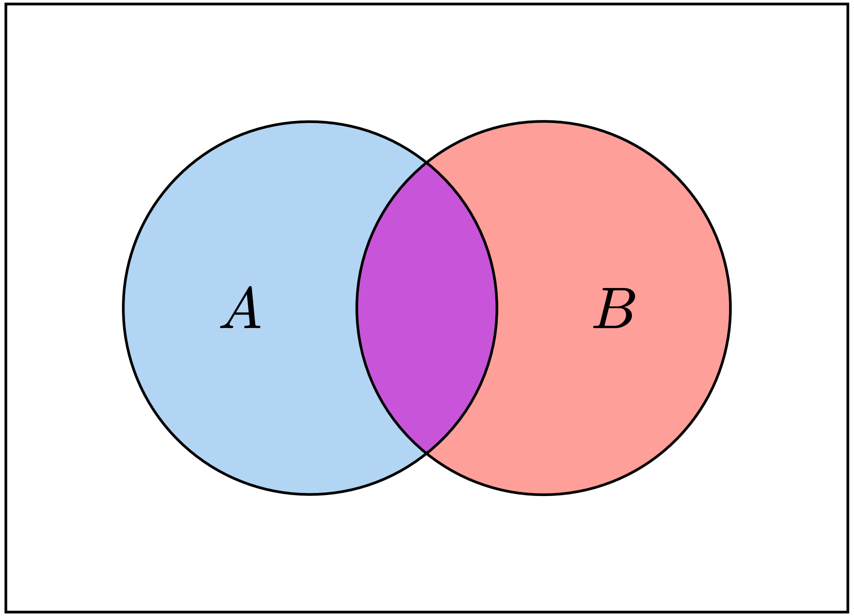
$$P(G \mid \underline{\text{negative}}) \quad P(\text{negative}) = 0$$

$= \text{undef.}$

1.4 Visualizing Conditional Probability

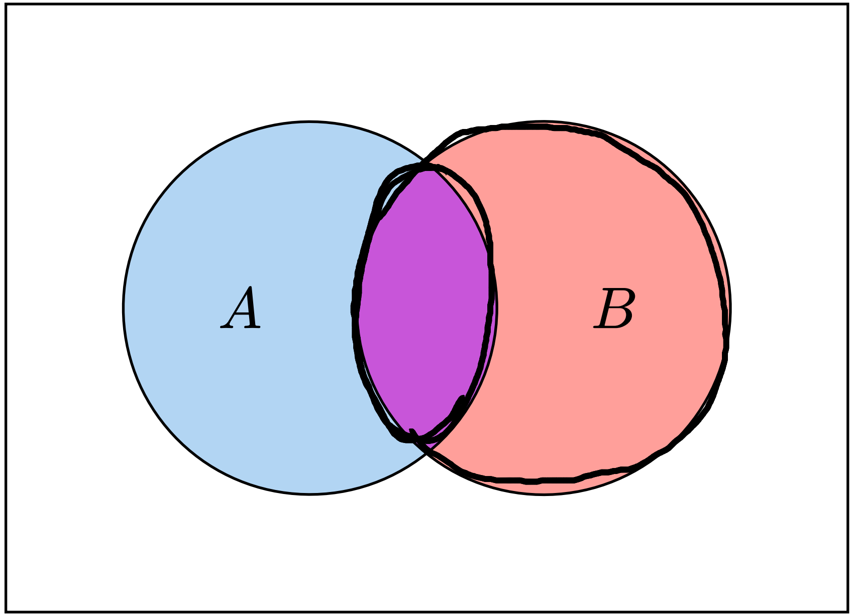


1.4 Visualizing Conditional Probability



- Red region is B (what we know occurred)
- Purple region is $A \cap B$

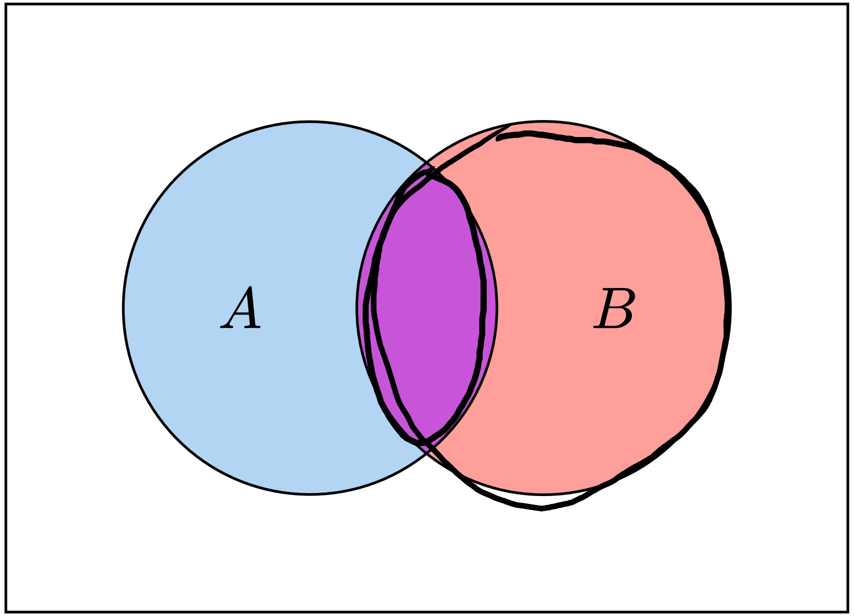
1.4 Visualizing Conditional Probability



- Red region is B (what we know occurred)
- Purple region is $A \cap B$

$P(A \mid B)$ = “What fraction of B is also in A?”

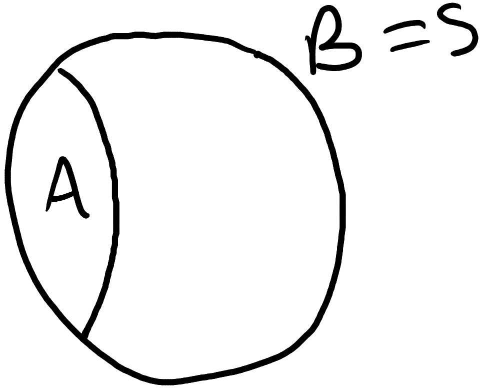
1.4 Visualizing Conditional Probability



- Red region is B (what we know occurred)
- Purple region is $A \cap B$

$P(A \mid B)$ = "What fraction of B is also in A?"

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$



1.5 The Notebook Interpretation

Think of repeating the experiment many, many times, recording each outcome on a line in a notebook.

1.5 The Notebook Interpretation

Think of repeating the experiment many, many times, recording each outcome on a line in a notebook.

For conditional probability $P(A \mid B)$, imagine a column labeled “A | B”:

1.5 The Notebook Interpretation

Think of repeating the experiment many, many times, recording each outcome on a line in a notebook.

For conditional probability $P(A \mid B)$, imagine a column labeled “A | B”:

- If B doesn’t occur on a line: write **“NA”** (not applicable)
- If B does occur on a line: write **“Yes”** if A also occurs, **“No”** otherwise

A	B	A | B	
Yes	Yes	Yes	2
No	Yes	No	
Yes	No	N/A	
No	No	N/A	

A ∩ B is true

$$P(A \mid B) = \frac{1}{2}$$

B true

1.5 The Notebook Interpretation

Think of repeating the experiment many, many times, recording each outcome on a line in a notebook.

For conditional probability $P(A \mid B)$, imagine a column labeled “A | B”:

- If B doesn’t occur on a line: write **“NA”** (not applicable)
- If B does occur on a line: write **“Yes”** if A also occurs, **“No”** otherwise

$$P(A \mid B) = \text{fraction of non-NA lines that say Yes}$$

1.5 The Notebook Interpretation

Think of repeating the experiment many, many times, recording each outcome on a line in a notebook.

For conditional probability $P(A \mid B)$, imagine a column labeled “A | B”:

- If B doesn’t occur on a line: write **“NA”** (not applicable)
- If B does occur on a line: write **“Yes”** if A also occurs, **“No”** otherwise

$$P(A \mid B) = \text{fraction of non-NA lines that say Yes}$$

Note: We only look at lines where B occurred, then ask: among those, what fraction also has A?

Matloff

1.6 Notebook Example: Two Dice

Roll two dice. Let X = first die, S = sum of both dice.

1.6 Notebook Example: Two Dice

Roll two dice. Let X = first die, S = sum of both dice.

$P(X=1 \mid S=6)$

Line	X	S	X = 1 and S = 6	X = 1 | S = 6
1	3	8	No	NA
2	1	6	Yes	Yes
3	2	7	No	NA
4	5	6	No	No
5	1	4	No	NA
6	4	6	No	No

1.6 Notebook Example: Two Dice

Roll two dice. Let X = first die, S = sum of both dice.

Line	X	S	X = 1 and S = 6	X = 1 | S = 6
1	3	8	No	NA
2	1	6	Yes	Yes
3	2	7	No	NA
4	5	6	No	No
5	1	4	No	NA
6	4	6	No	No

- $P(X = 1 \text{ and } S = 6) = \frac{1}{36}$ (1 Yes out of 6 lines total)
- $P(X = 1 \mid S = 6) = \boxed{\frac{1}{3}}$ (1 Yes out of 3 non-NA lines)

1.7 Example: Drawing Cards

Draw one card from a standard deck. Let:

- A = “card is an Ace”
- B = “card is a heart”

1.7 Example: Drawing Cards

Draw one card from a standard deck. Let:

- A = "card is an Ace"
- B = "card is a heart"

13 heart cards

What is $P(A \mid B)$?

$$\frac{1}{13}$$

1.7 Example: Drawing Cards

Draw one card from a standard deck. Let:

- A = “card is an Ace”
- B = “card is a heart”

What is $P(A \mid B)$?

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/\cancel{52}}{13/\cancel{52}} = \boxed{\frac{1}{13}}$$

1.7 Example: Drawing Cards

Draw one card from a standard deck. Let:

- A = “card is an Ace”
- B = “card is a heart”

What is $P(A \mid B)$?

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/52}{13/52} = \frac{1}{13}$$

There are 13 hearts, and exactly 1 is an Ace.

1.8 The Multiplication Rule

Rearranging the definition:

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

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

1.8 The Multiplication Rule

Rearranging the definition:

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

This is the **multiplication rule**: the probability that both A and B occur.

1.8 The Multiplication Rule

Rearranging the definition:

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

This is the **multiplication rule**: the probability that both A and B occur.

Equivalently (by symmetry):

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

1.8 The Multiplication Rule

Rearranging the definition:

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

This is the **multiplication rule**: the probability that both A and B occur.

Equivalently (by symmetry):

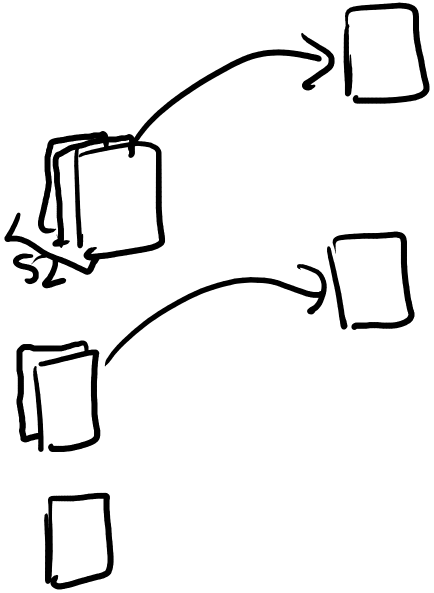
$$P(A \cap B) = P(A) \cdot P(B \mid A)$$

Note: Very useful for sequential experiments!

1.9 Example: Drawing Without Replacement

Draw 2 cards from a deck **without replacement.**

What's $P(\text{both are Aces})$?



1.9 Example: Drawing Without Replacement

Draw 2 cards from a deck **without replacement.**

What's $P(\text{both are Aces})$?

Let A_1 = "first card is Ace", A_2 = "second card is Ace"

1.9 Example: Drawing Without Replacement

Draw 2 cards from a deck **without replacement.**

What's $P(\text{both are Aces})$?

Let A_1 = "first card is Ace", A_2 = "second card is Ace"

$$P(A_1 \cap A_2) = \underline{P(A_1)} \cdot P(A_2 \mid A_1)$$

$$= \frac{\boxed{4}}{\boxed{52}} \cdot \frac{\boxed{3}}{\boxed{51}}$$

$$= \frac{12}{2652} = \frac{1}{221}$$

$$\frac{4}{52} \cdot \frac{4}{52}$$

1.9 Example: Drawing Without Replacement

Draw 2 cards from a deck **without replacement.**

What's $P(\text{both are Aces})$?

Let A_1 = "first card is Ace", A_2 = "second card is Ace"

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1) \cdot P(A_2 \mid A_1) \\ &= \frac{4}{52} \cdot \frac{3}{51} \\ &= \frac{12}{2652} = \frac{1}{221} \end{aligned}$$

After drawing an Ace, only 3 Aces remain among 51 cards.

2. Tree-Based Modeling

2.1 Example: Radar Detection

A radar system monitors airspace for aircraft.

2.1 Example: Radar Detection

A radar system monitors airspace for aircraft.

Given information:

- An aircraft is present with probability 0.05
- If aircraft present: radar detects it with probability 0.99
- If no aircraft: radar falsely registers presence with probability 0.10

classifier

C = { aircraft detected,
aircraft not detected }

2.1 Example: Radar Detection

A radar system monitors airspace for aircraft.

Given information:

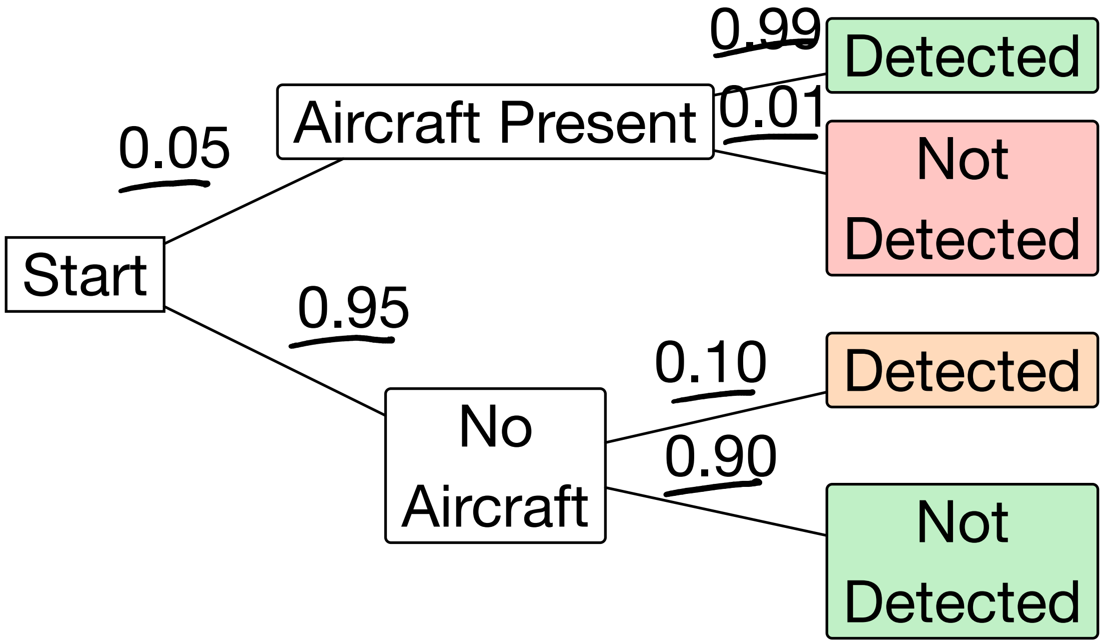
- An aircraft is present with probability 0.05
- If aircraft present: radar detects it with probability 0.99
- If no aircraft: radar falsely registers presence with probability 0.10

Questions:

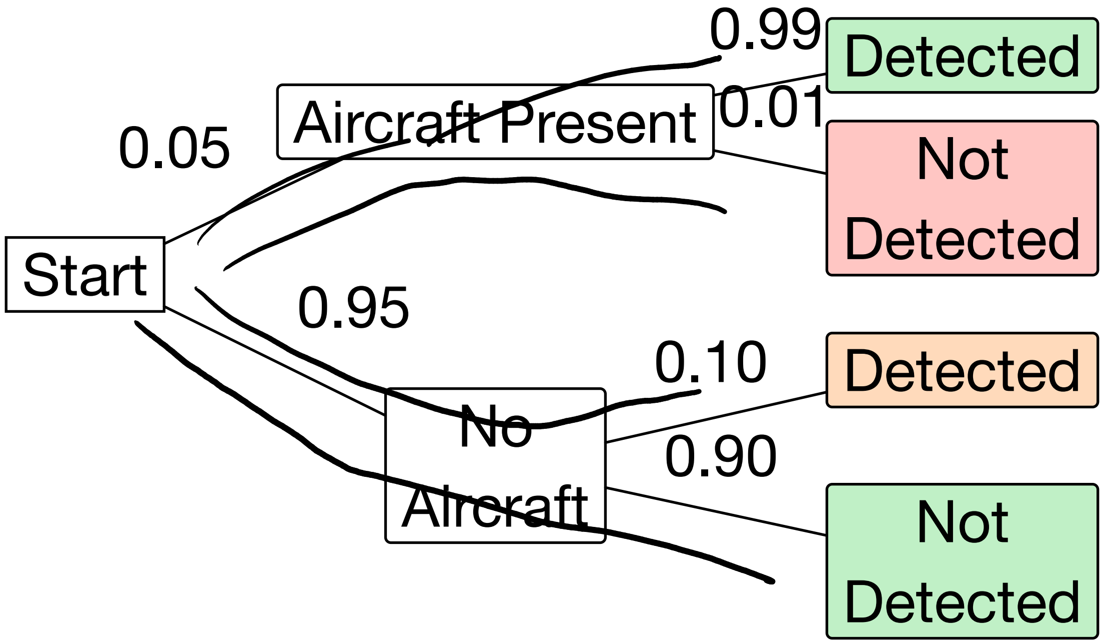
1. What is the probability of a **false alarm**?
2. What is the probability of **missed detection**?

Bertsekas and Tsitsiklis

2.2 Visualizing with a Tree Diagram

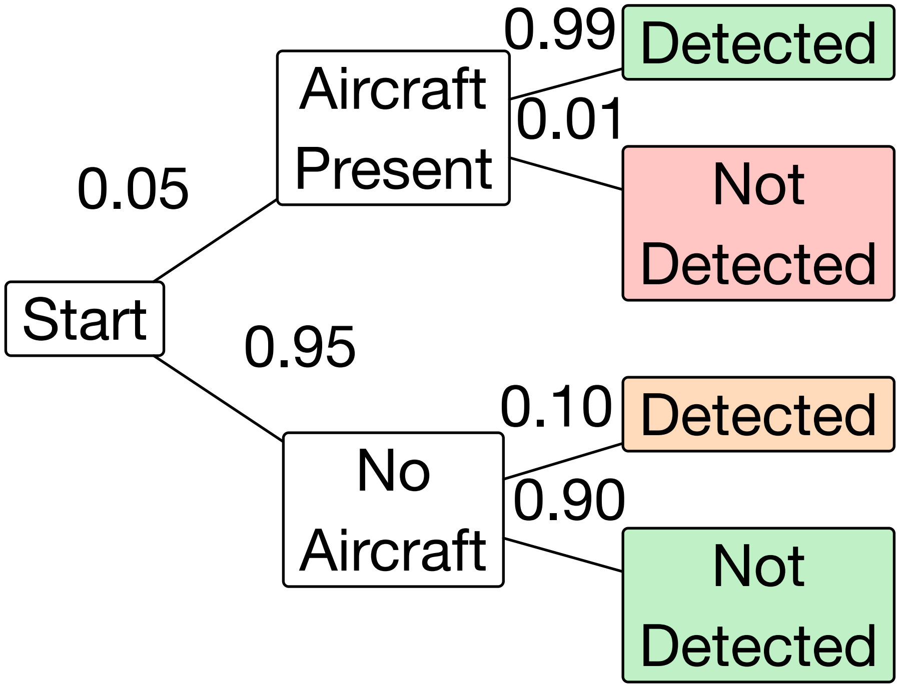


2.2 Visualizing with a Tree Diagram



Each **path** from start to end represents a possible outcome.

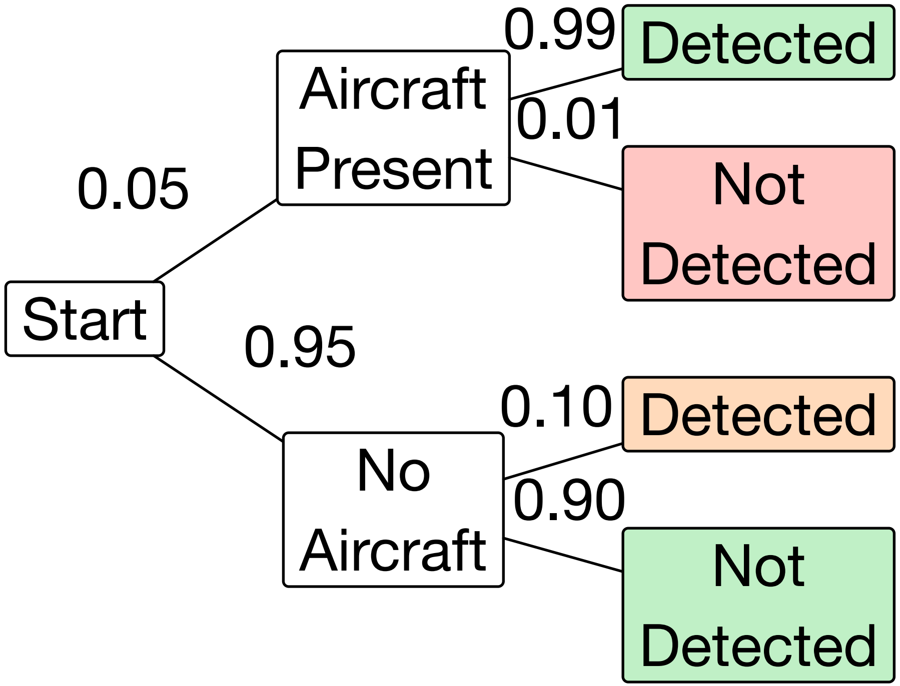
2.3 Defining Events



Let:

- A = “aircraft is present”
- B = “radar registers aircraft presence”

2.3 Defining Events



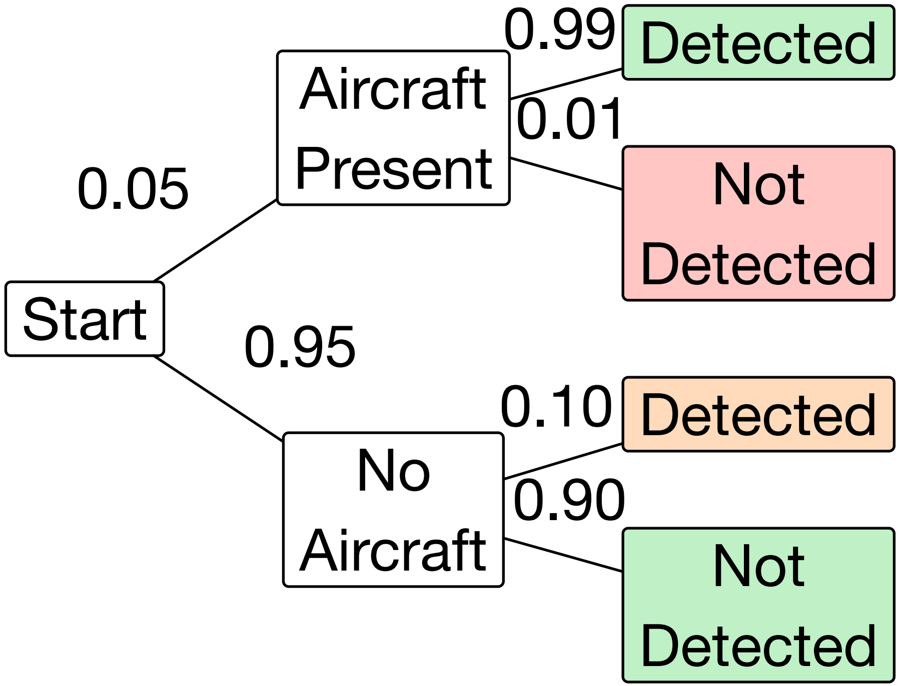
Let:

- A = “aircraft is present”
- B = “radar registers aircraft presence”

Also consider the complements:

- A^c = “aircraft is not present”
- B^c = “radar does not register aircraft presence”

2.3 Defining Events



Let:

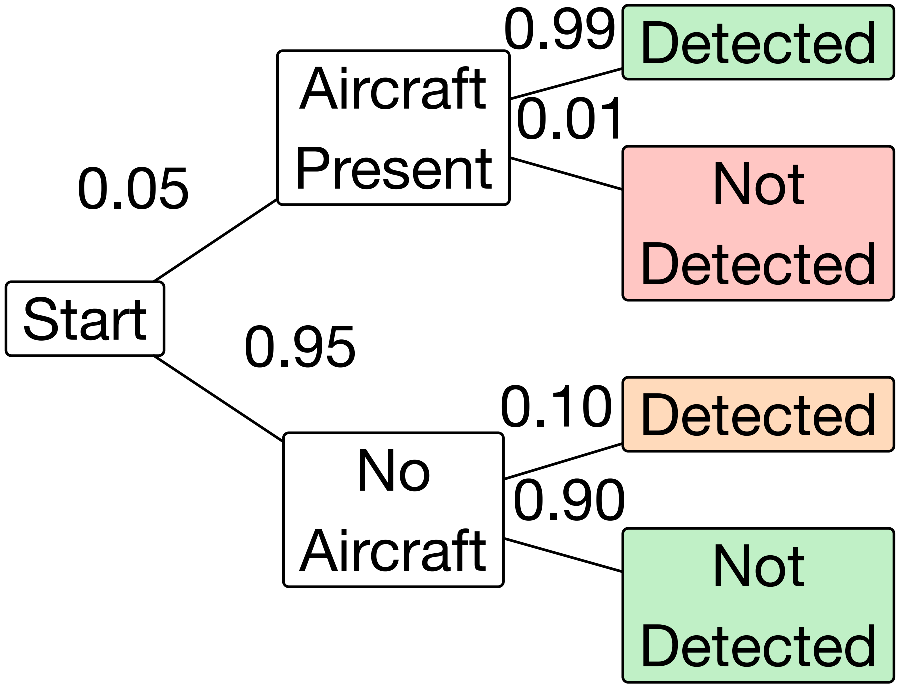
- A = “aircraft is present”
- B = “radar registers aircraft presence”

Also consider the complements:

- A^c = “aircraft is not present”
- B^c = “radar does not register aircraft presence”

False alarm: $A^c \cap B$

2.3 Defining Events



Let:

- A = “aircraft is present”
- B = “radar registers aircraft presence”

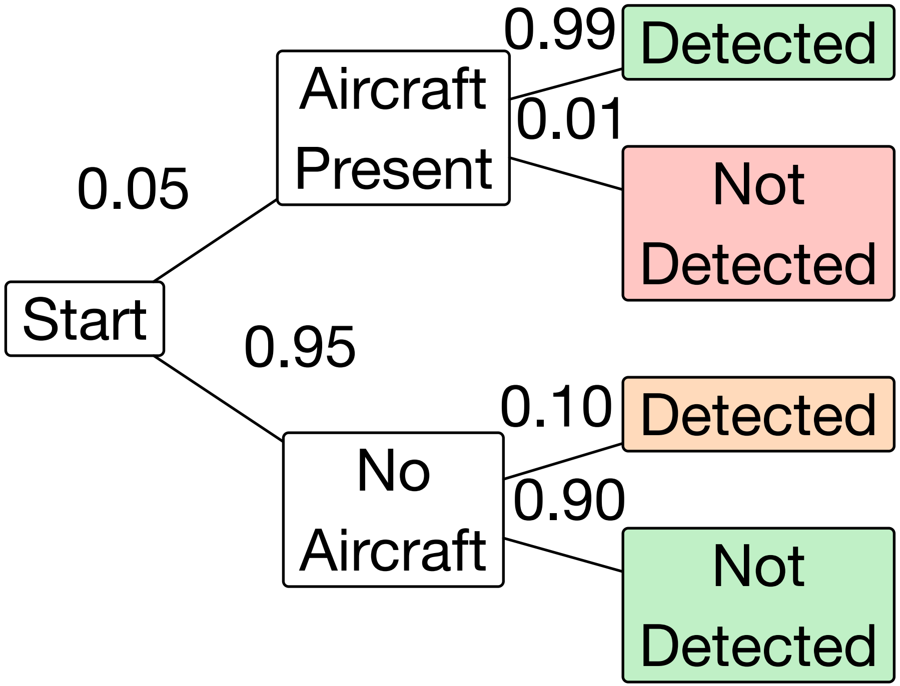
Also consider the complements:

- A^c = “aircraft is not present”
- B^c = “radar does not register aircraft presence”

False alarm: $A^c \cap B$

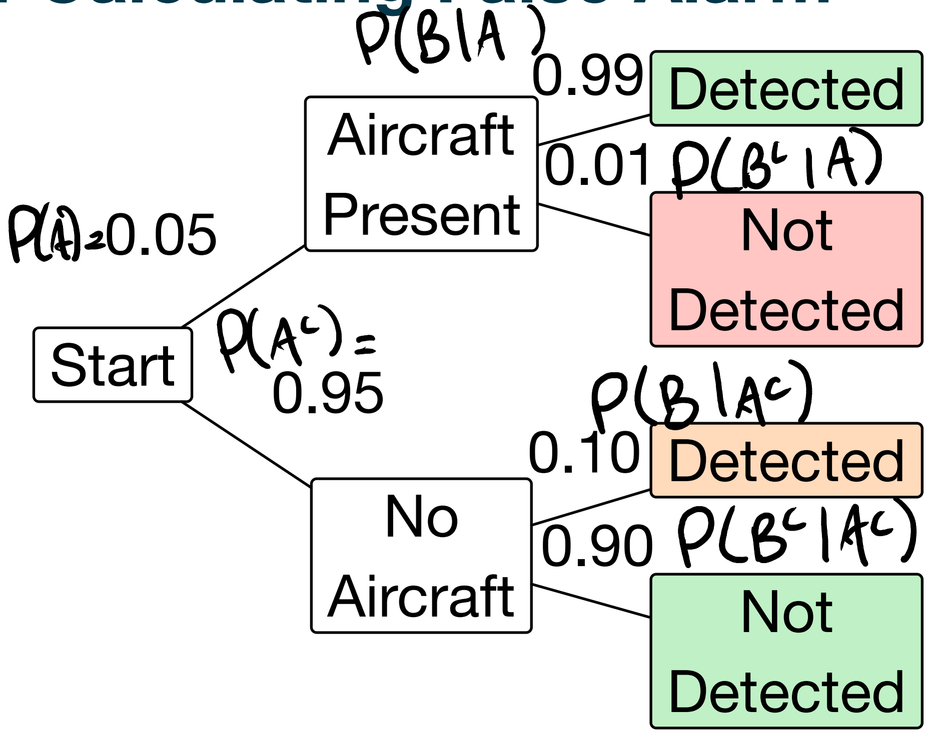
Missed detection: $A \cap B^c$

2.4 Calculating False Alarm



Use the multiplication rule:

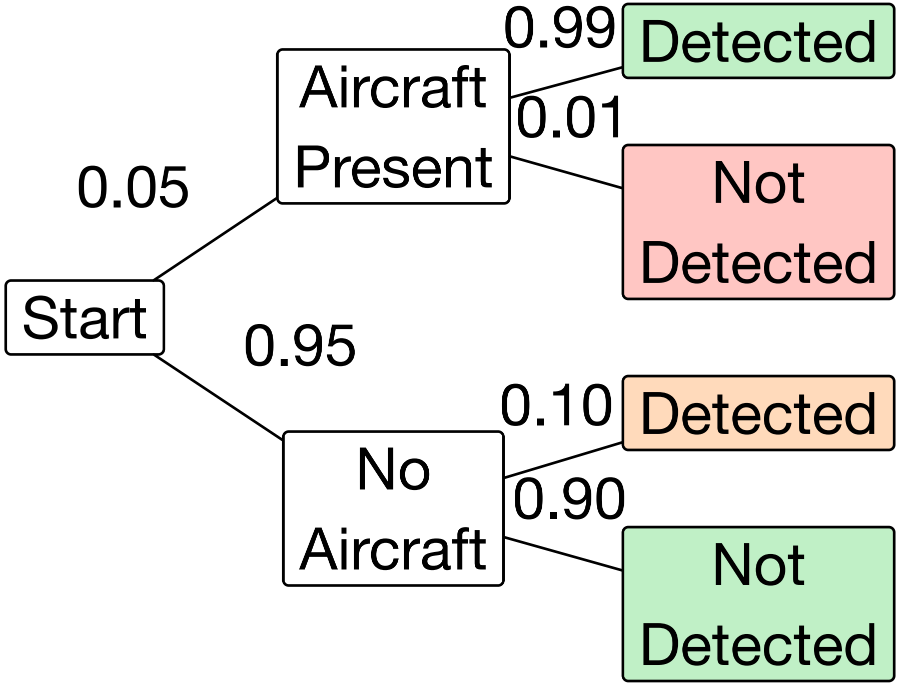
2.4 Calculating False Alarm



Use the multiplication rule:

$$
\begin{aligned}
P(\text{false alarm}) &= P(A^c \cap B) \\
&= P(A^c) \cdot P(B \mid A^c)
\end{aligned}
$$

2.4 Calculating False Alarm

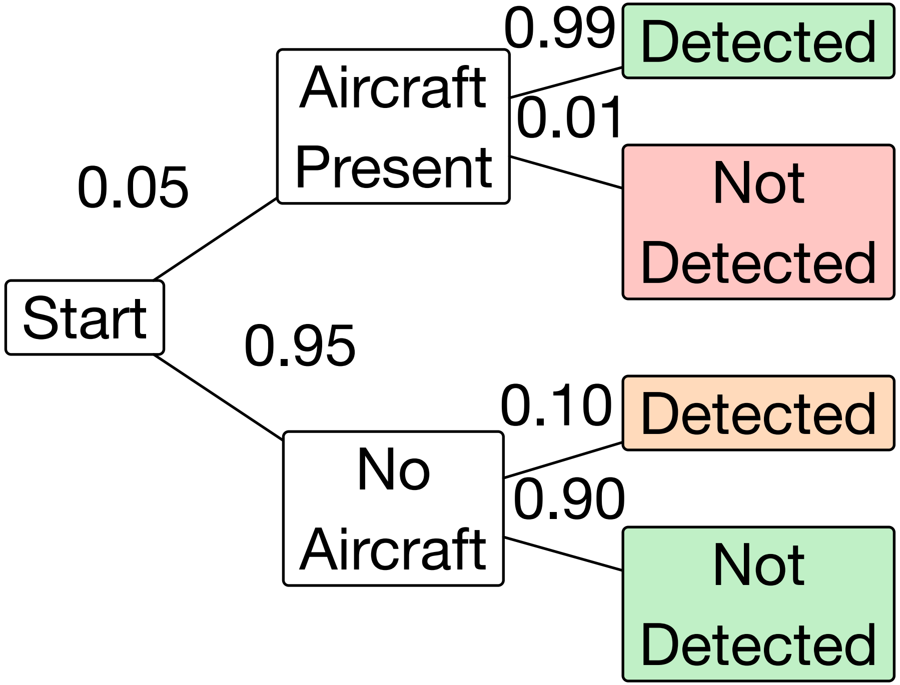


Use the multiplication rule:

$$
\begin{aligned}
P(\text{false alarm}) &= P(A^c \cap B) \\
&= P(A^c) \cdot P(B \mid A^c)
\end{aligned}
$$

$$= 0.95 \cdot 0.10$$

2.4 Calculating False Alarm



Use the multiplication rule:

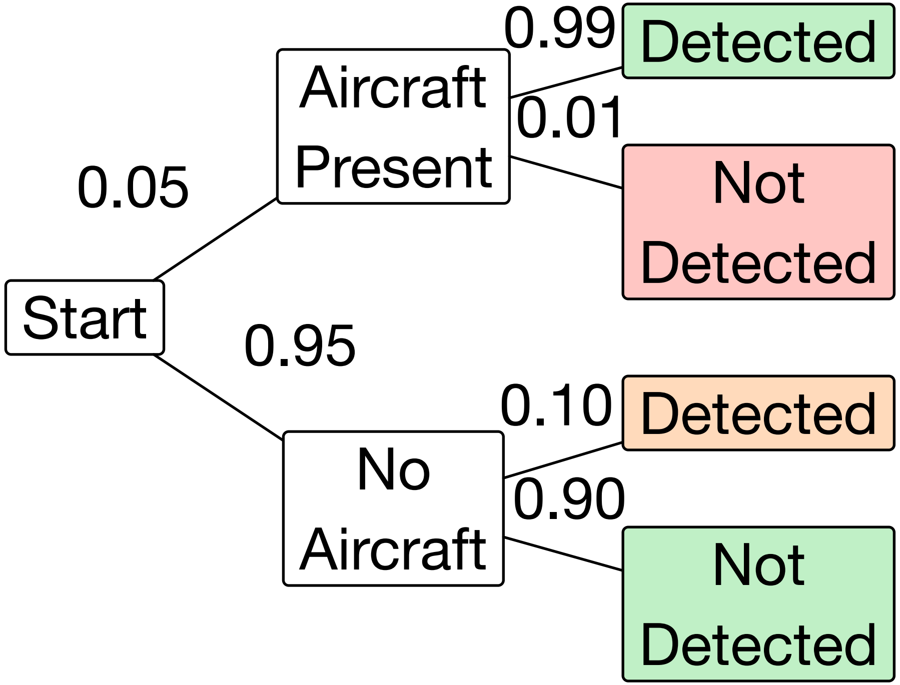
$$P(\text{false alarm}) = P(A^c \cap B)$$

$$= P(A^c) \cdot P(B \mid A^c)$$

$$= 0.95 \cdot 0.10$$

$$= 0.095$$

2.4 Calculating False Alarm



Use the multiplication rule:

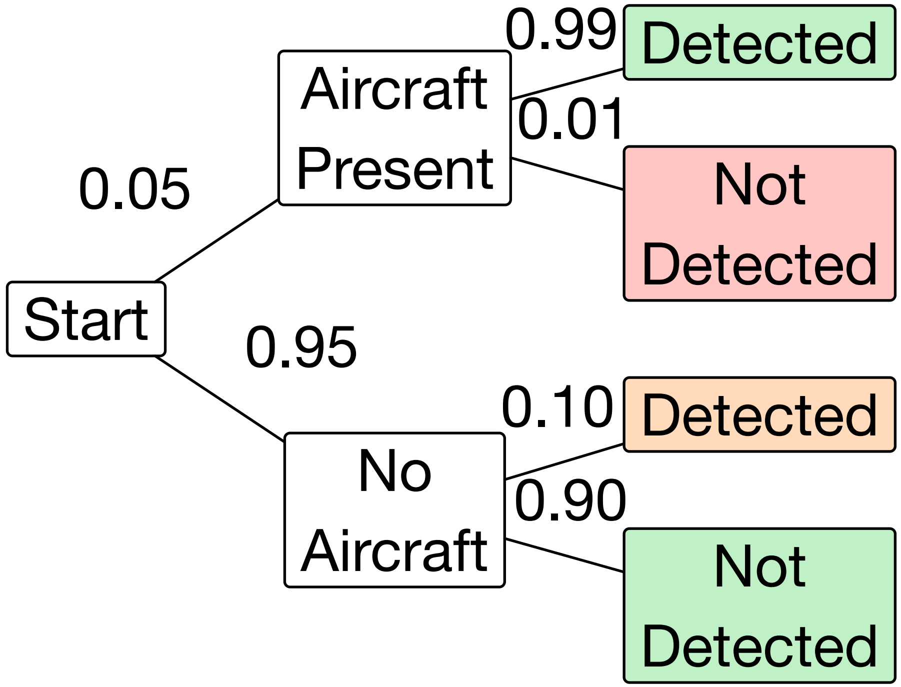
$$\begin{aligned} P(\text{false alarm}) &= P(A^c \cap B) \\ &= P(A^c) \cdot P(B \mid A^c) \end{aligned}$$

$$= 0.95 \cdot 0.10$$

$$= 0.095$$

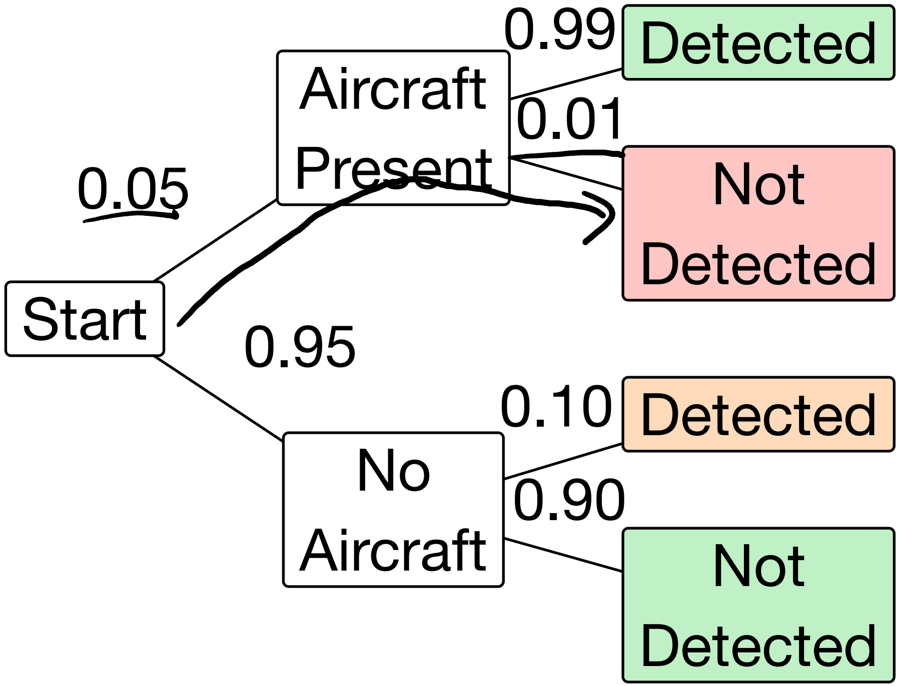
About **9.5% chance** of a false alarm.

2.5 Calculating Missed Detection



Similarly:

2.5 Calculating Missed Detection



Start

0.05

Aircraft Present

0.99 Detected

0.01 Not Detected

0.95

No Aircraft

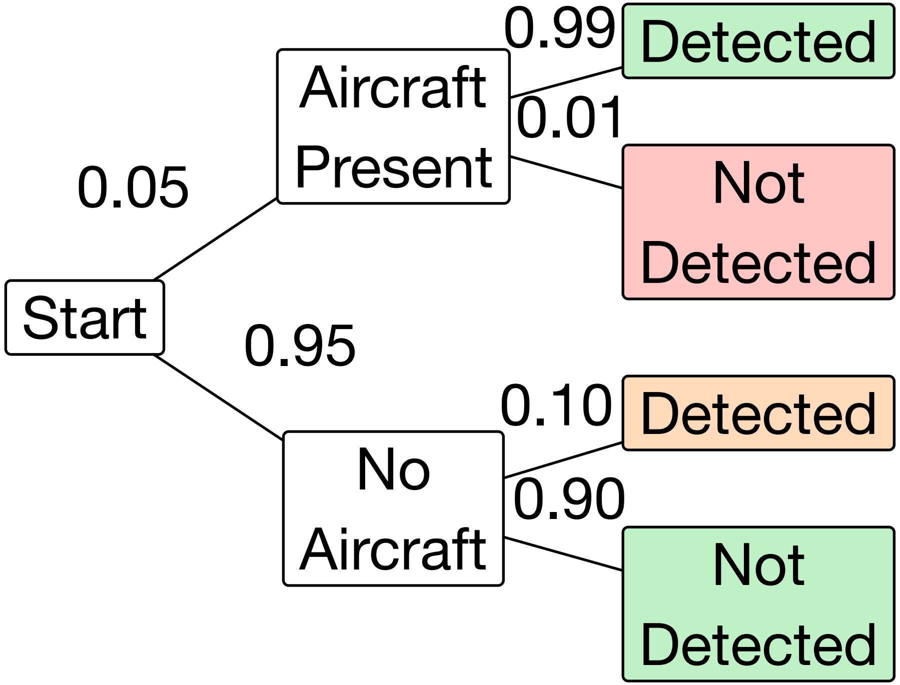
0.10 Detected

0.90 Not Detected

Similarly:

$$
\begin{aligned}
P(\text{missed detection}) &= P(\underline{A} \cap \underline{B^c}) \\
&= P(A) \cdot P(B^c \mid A)
\end{aligned}
$$

2.5 Calculating Missed Detection

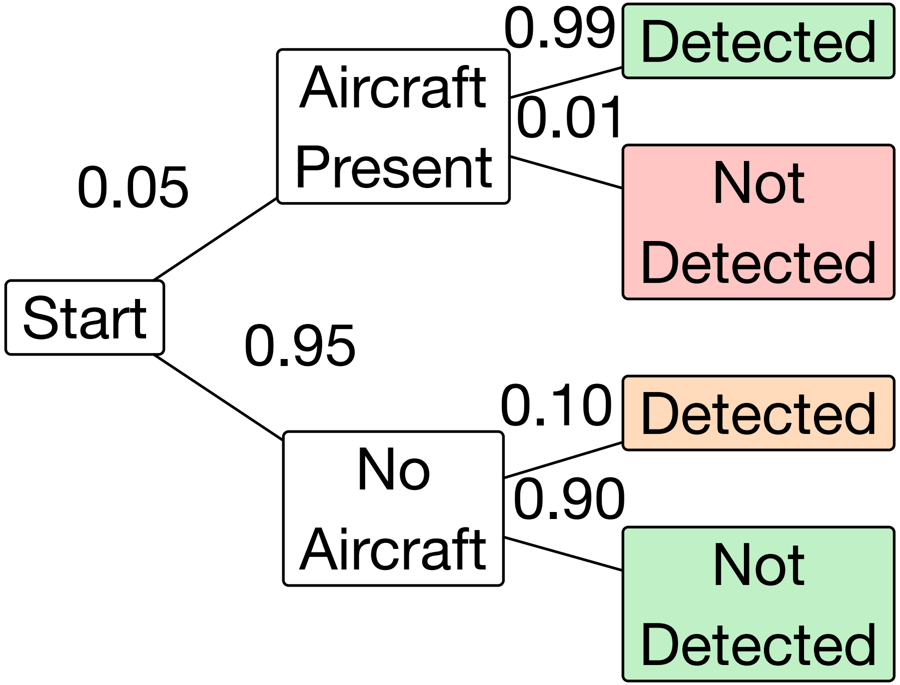


Similarly:

$$\begin{aligned} P(\text{missed detection}) &= P(A \cap B^c) \\ &= P(A) \cdot P(B^c \mid A) \end{aligned}$$

$$= 0.05 \cdot 0.01$$

2.5 Calculating Missed Detection



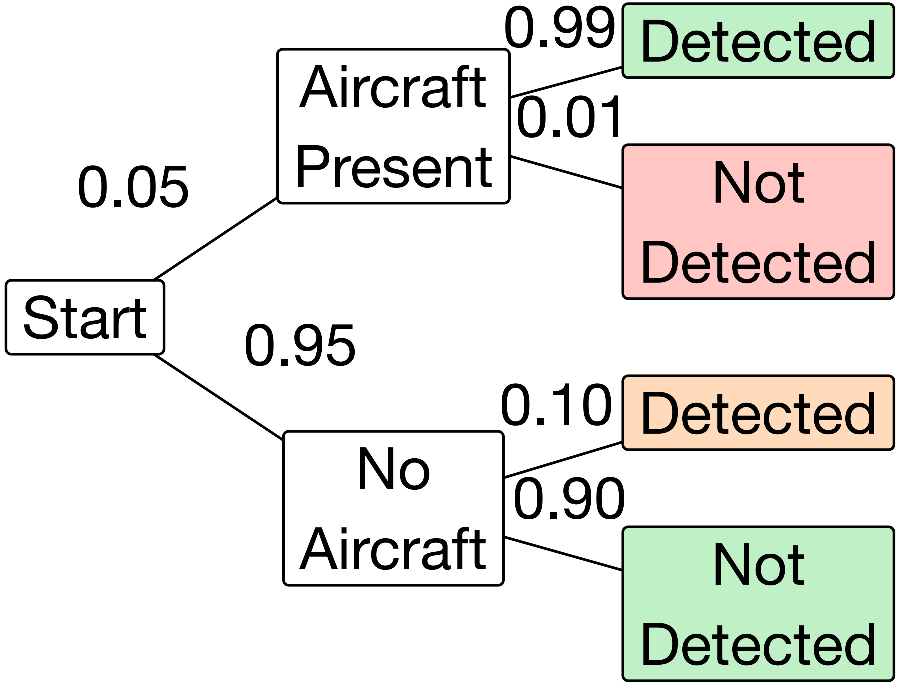
Similarly:

$$
\begin{aligned}
P(\text{missed detection}) &= P(A \cap B^c) \\
&= P(A) \cdot P(B^c \mid A)
\end{aligned}
$$

$$= 0.05 \cdot 0.01$$

$$= 0.0005$$

2.5 Calculating Missed Detection



Similarly:

$$
\begin{aligned}
P(\text{missed detection}) &= P(A \cap B^c) \\
&= P(A) \cdot P(B^c \mid A)
\end{aligned}
$$

$$= 0.05 \cdot 0.01$$

$$= 0.0005$$

Only **0.05% chance** of missing a present aircraft.

2.6 Tree-Based Modeling Recipe

When modeling sequential experiments:

2.6 Tree-Based Modeling Recipe

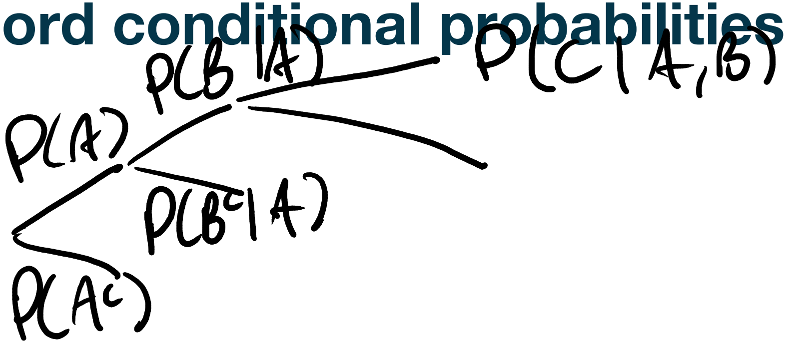
When modeling sequential experiments:

1. **Set up the tree** so events of interest correspond to paths (leaves)

2.6 Tree-Based Modeling Recipe

When modeling sequential experiments:

1. **Set up the tree** so events of interest correspond to paths (leaves)
2. **Record conditional probabilities** on the branches



2.6 Tree-Based Modeling Recipe

When modeling sequential experiments:

1. **Set up the tree** so events of interest correspond to paths (leaves)
2. **Record conditional probabilities** on the branches
3. **Multiply along paths** to get the probability of each outcome

2.6 Tree-Based Modeling Recipe

When modeling sequential experiments:

1. **Set up the tree** so events of interest correspond to paths (leaves)
2. **Record conditional probabilities** on the branches
3. **Multiply along paths** to get the probability of each outcome

This is exactly the multiplication rule, visualized!

3. Problem-Solving Strategies

3.1 Strategy: “How Can It Happen?”

When solving probability problems, ask: **“How can this event happen?”**

3.1 Strategy: “How Can It Happen?”

When solving probability problems, ask: **“How can this event happen?”**

Break complex events into simpler scenarios.

3.1 Strategy: “How Can It Happen?”

When solving probability problems, ask: **“How can this event happen?”**

Break complex events into simpler scenarios.

Example: Drawing Without Replacement

What’s $P(\text{both cards are Aces})$?

3.1 Strategy: “How Can It Happen?”

When solving probability problems, ask: **“How can this event happen?”**

Break complex events into simpler scenarios.

Example: Drawing Without Replacement

What’s $P(\underline{\text{both cards are Aces}})$?

How can this happen?

- First card is Ace **AND** second card is Ace

3.1 Strategy: “How Can It Happen?”

When solving probability problems, ask: **“How can this event happen?”**

Break complex events into simpler scenarios.

Example: Drawing Without Replacement

What’s $P(\text{both cards are Aces})$?

How can this happen?

- First card is Ace **AND** second card is Ace

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 \mid A_1) = \frac{4}{52} \cdot \frac{3}{51}$$

Matloff

3.2 Strategy: Find Equivalent Events

When you write $P(A) = P(B)$, make sure A and B are **equivalent**:

3.2 Strategy: Find Equivalent Events

When you write $P(A) = P(B)$, make sure A and B are **equivalent**:

- If A occurs, then B must occur
- If B occurs, then A must occur

3.2 Strategy: Find Equivalent Events

When you write $P(A) = P(B)$, make sure A and B are **equivalent**:

- If A occurs, then B must occur
- If B occurs, then A must occur

Example: Board Game

Event: “End at square 4 with bonus roll, starting from 0”

3.2 Strategy: Find Equivalent Events

When you write $P(A) = P(B)$, make sure A and B are **equivalent**:

- If A occurs, then B must occur
- If B occurs, then A must occur

Example: Board Game

Event: “End at square 4 with bonus roll, starting from 0”

First guess: $R + B = 4$ and $B > 0$

3.2 Strategy: Find Equivalent Events

When you write $P(A) = P(B)$, make sure A and B are **equivalent**:

- If A occurs, then B must occur
- If B occurs, then A must occur

Example: Board Game

Event: "End at square 4 with bonus roll, starting from 0"

First guess: $R + B = 4$ and $B > 0$

Equivalent simpler event: $R = 3$ and $B = 1$

3.2 Strategy: Find Equivalent Events

When you write $P(A) = P(B)$, make sure A and B are **equivalent**:

- If A occurs, then B must occur
- If B occurs, then A must occur

Example: Board Game

Event: “End at square 4 with bonus roll, starting from 0”

First guess: $R + B = 4$ and $B > 0$

Equivalent simpler event: $R = 3$ and $B = 1$

Check both directions:

- If $\underline{R = 3}, \underline{B = 1}$, then $\underline{R + B = 4}$ and $\underline{B > 0}$ ✓
- If $\underline{R + B = 4}, \underline{B > 0}$, then $\underline{R = 3, B = 1}$ ✓

Matloff

3.3 Strategy: Multiple Approaches

There's rarely one "right" way to solve a problem.

3.3 Strategy: Multiple Approaches

There's rarely one "right" way to solve a problem.

Try different decompositions:

- Condition on different events
- Use complement vs. direct calculation
- Break into cases in different ways

3.3 Strategy: Multiple Approaches

There's rarely one "right" way to solve a problem.

Try different decompositions:

- Condition on different events
- Use complement vs. direct calculation
- Break into cases in different ways

If one approach gets messy, try another!

3.3 Strategy: Multiple Approaches

There's rarely one "right" way to solve a problem.

Try different decompositions:

- Condition on different events
- Use complement vs. direct calculation
- Break into cases in different ways

If one approach gets messy, try another!

Note: Like programming: creative process, not formula-based

Matloff

3.4 Chain Rule (Extended Multiplication)

For multiple events:

$$P(A_1 \cap A_2 \cap A_3) = \underline{P(A_1)} \cdot P(A_2 \mid \overset{\downarrow}{A_1}) \cdot P(A_3 \mid \overset{\downarrow}{A_1} \cap A_2)$$

$$= P(A_1) \cdot P(A_3 \mid \underset{\uparrow}{A_1}) \cdot P(A_2 \mid \underset{\uparrow}{A_1} \cap A_3)$$

3.4 Chain Rule (Extended Multiplication)

For multiple events:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2)$$

In general:

$$P\left(\bigcap_{i=1}^{\boxed{n}} A_i\right) = P(A_1) \cdot P(\underline{A_2} \mid \underline{A_1}) \cdot P(\underline{A_3} \mid \underline{A_1 A_2}) \cdot \ldots \cdot P(\underline{A_n} \mid \underline{A_1 \ldots A_{n-1}})$$

3.4 Chain Rule (Extended Multiplication)

For multiple events:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2)$$

In general:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 A_2) \cdot \ldots \cdot P(A_n \mid A_1 \ldots A_{n-1})$$

This “chains” conditional probabilities together.

3.5 Example: Three Cards

Draw 3 cards without replacement. $P(\text{all three are hearts})$?

3.5 Example: Three Cards

Draw 3 cards without replacement. $P(\text{all three are hearts})$?

$$
\begin{aligned}
P(\underline{H_1} \cap \underline{H_2} \cap \underline{H_3}) &= P(\underline{H_1}) \cdot P(H_2 \mid \underline{H_1}) \cdot P(\underline{H_3} \mid \underline{H_1 H_2}) \\
&= \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \\
&= \frac{1716}{132600} \approx 0.0129
\end{aligned}
$$

3.6 Conditional Probability is a Probability

$P(\cdot \mid B)$ satisfies all three axioms:

3.6 Conditional Probability is a Probability

$P(\cdot \mid B)$ satisfies all three axioms:

1. **Non-negativity**: $P(A \mid B) \geq 0$ ✓

3.6 Conditional Probability is a Probability

$P(\cdot \mid B)$ satisfies all three axioms:

1. **Non-negativity**: $P(A \mid B) \geq 0$ ✓
2. **Normalization**: $P(B \mid B) = \frac{P(B \cap B)}{P(B)} = 1$ ✓

3.6 Conditional Probability is a Probability

$P(\cdot \mid B)$ satisfies all three axioms:

1. **Non-negativity**: $P(A \mid B) \geq 0$ ✓
2. **Normalization**: $P(B \mid B) = \frac{P(B \cap B)}{P(B)} = 1$ ✓
3. **Additivity**: If A_1, A_2 disjoint:

   $$P(A_1 \cup A_2 \mid B) = P(A_1 \mid B) + P(A_2 \mid B)$$

   ✓

3.6 Conditional Probability is a Probability

$P(\cdot \mid B)$ satisfies all three axioms:

1. **Non-negativity**: $P(A \mid B) \geq 0$ ✓
2. **Normalization**: $P(B \mid B) = \frac{P(B \cap B)}{P(B)} = 1$ ✓
3. **Additivity**: If A_1, A_2 disjoint:

   $$P(A_1 \cup A_2 \mid B) = P(A_1 \mid B) + P(A_2 \mid B)$$

   ✓

So all our probability rules apply to conditional probabilities too!

3.7 Common Mistakes with Conditional Probability

Two hugely common errors:

1. $P(A \mid B) \neq P(B \mid A)$ (confusion of the inverse)
2. $P(A \underset{\substack{\uparrow \\ \text{joint}}}{\text{and}} \underset{?.}{\overline{B)}} \neq P(A \underset{\substack{\uparrow \\ \text{conditional}}}{\mid} B)$ ✓ (confusing "joint" and conditional)

3.8 Mistake 1: Confusion of the Inverse

$P(A \mid B) \neq P(B \mid A)$ in general!

3.8 Mistake 1: Confusion of the Inverse

$P(A \mid B) \neq P(B \mid A)$ in general!

Example:

- $P(\text{has fever} \mid \text{has flu}) \approx \underline{0.9}$ (most flu patients have fever)
- $P(\text{has flu} \mid \text{has fever}) \approx \underline{0.1}$ (most fevers aren’t from flu)

3.8 Mistake 1: Confusion of the Inverse

$P(A \mid B) \neq P(B \mid A)$ in general!

Example:

- $P(\text{has fever} \mid \text{has flu}) \approx 0.9$ (most flu patients have fever)
- $P(\text{has flu} \mid \text{has fever}) \approx 0.1$ (most fevers aren’t from flu)

These condition on **different events**, so they’re completely different probabilities.

3.9 Mistake 2: Joint vs. Conditional

$P(A \text{ and } B) \neq P(A \mid B)$ in general!

3.9 Mistake 2: Joint vs. Conditional

$P(A \text{ and } B) \neq P(A \mid B)$ in general!

Roll two dice. Let X = first die, S = sum. Consider $X = 1$ and $S = 6$:

3.9 Mistake 2: Joint vs. Conditional

$P(A \text{ and } B) \neq P(A \mid B)$ in general!

Roll two dice. Let X = first die, S = sum. Consider $X = 1$ and $S = 6$:

- $P(X = 1 \text{ and } S = 6) = \frac{1}{36}$
  - Fraction of **all** rolls where both happen

3.9 Mistake 2: Joint vs. Conditional

$P(A \text{ and } B) \neq P(A \mid B)$ in general!

Roll two dice. Let X = first die, S = sum. Consider $X = 1$ and $S = 6$:

- $P(X = 1 \text{ and } S = 6) = \boxed{\frac{1}{36}}$
  - Fraction of **all** rolls where both happen
- $P(X = 1 \mid \underline{S = 6}) = \boxed{\frac{1}{5}}$
  - Fraction of **$S = 6$ rolls** where $X = 1$ also happens

3.9 Mistake 2: Joint vs. Conditional

$P(A \text{ and } B) \neq P(A \mid B)$ in general!

Roll two dice. Let X = first die, S = sum. Consider $X = 1$ and $S = 6$:

- $P(X = 1 \text{ and } S = 6) = \frac{1}{36}$
  - Fraction of **all** rolls where both happen
- $P(X = 1 \mid S = 6) = \frac{1}{5}$
  - Fraction of $S = 6$ **rolls** where $X = 1$ also happens

Note: Use the notebook view: joint probability counts all lines, conditional counts only non-NA lines

3.10 Connecting to Code

Python:

```python
import random

def simulate_conditional_prob(n_trials=100000
    count_first_ace = 0
    count_both_aces = 0

    for _ in range(n_trials):
        deck = list(range(52))  # 0-3 are Aces
        random.shuffle(deck)

        if deck[0] < 4:  # First card is an Ace
            count_first_ace += 1
            if deck[1] < 4:  # Second card is also an Ace
                count_both_aces += 1

    # P(second ace | first ace) = rate among trials where first is ace
    conditional_prob = count_both_aces / count_first_ace
    return conditional_prob

print(f"P(2nd Ace | 1st Ace) simulated: {simulate_conditional_prob():.4f}")
print(f"P(2nd Ace | 1st Ace) theoretical: {3/51:.4f}")
```

$$\frac{P(A \cap B)}{P(B)}$$

3.10 Connecting to Code

R:

```r
simulate_conditional_prob <- function(n_trials = 100000) {
  count_first_ace <- 0
  count_both_aces <- 0

  for (i in 1:n_trials) {
    deck <- sample(1:52)  # 1-4 are Aces

    if (deck[1] <= 4) {  # First card is an Ace
      count_first_ace <- count_first_ace + 1
      if (deck[2] <= 4) {  # Second card is also an Ace
        count_both_aces <- count_both_aces + 1
      }
    }
  }

  # P(second ace | first ace) = rate among trials where first is ace
  conditional_prob <- count_both_aces / count_first_ace
  return(conditional_prob)
}

cat("P(2nd Ace | 1st Ace) simulated:", simulate_conditional_prob(), "\n")
cat("P(2nd Ace | 1st Ace) theoretical:", 3/51, "\n")
```

3.11 Summary

Concept	**Formula**
Conditional Probability	$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$
Multiplication Rule	$P(A \cap B) = P(A) \cdot P(B \mid A)$
Chain Rule	$P(A_1 \cap \ldots \cap A_n) = \prod P(A_i \mid A_1 \ldots A_{i-1})$

3.13 Recap

Today we covered:

- Conditional probability: probability given new information
- Multiplication rule: $P(A \cap B) = P(A) \cdot P(B \mid A)$
- Chain rule extends to multiple events
- $P(A \mid B) \neq P(B \mid A)$ - order matters!
- Next: Bayes' rule and total probability

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

$P(B) = P(B \mid A)$