Independence

CS109

Announcements

- **PSet 1** is due Friday @ 2 pm
- **Sections** start today!
  - 4^{th} of July plan: go to today's or Friday's section if you can
  - Otherwise, you can work through problems on your own, then submit answers through the pset app
- **Pset 2** is out
  - Practice using core probability formulas
  - Analyze a dataset

PSet 2: Data Science Problem

PS2

1
2a
2b
3
4a
4b
5
6
7
8
9
10

Bat Ebola

After the Ebola outbreak of 2015, there was an urgent need to learn more about the virus. You have been asked to uncover how a particular group of bat genes impact an important trait: whether the bat can carry Ebola. Nobody knows the underlying mechanism; it is up to you to hypothesize what is going on. For 100,000 independently sampled bats you have collected data of whether or not five genes are expressed, and whether or not the bat can carry Ebola.

BatEbolaData.zip

If a gene is expressed, it can affect both the probability of other genes being expressed and the probability of the trait being expressed. You can find the data in a file called `bats.csv`. Each row in the file corresponds to **one bat** and has 6 Booleans:

1. Gene 1: Whether the 1st gene is expressed in the bat (G_1)
2. Gene 2: Whether the 2nd gene is expressed in the bat (G_2)
3. Gene 3: Whether the 3rd gene is expressed in the bat (G_3)
4. Gene 4: Whether the 4th gene is expressed in the bat (G_4)
5. Gene 5: Whether the 5th gene is expressed in the bat (G_5)
6. Trait: Whether the trait is expressed in the bat; i.e., the bat can carry Ebola (T)

PSet 2: Data Science Problem

PS2

1 2a 2b 3 4a 4b 5 6 7 8 9 10

Bat Ebola

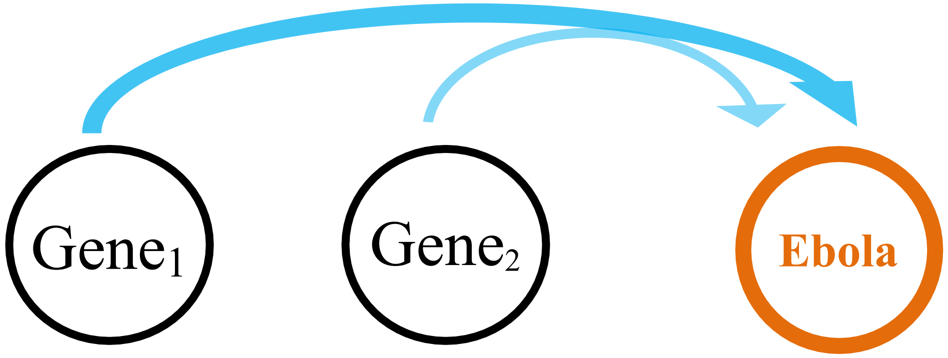
After the Ebola outbreak of 2015, there was an urgent need to learn more about the virus. You have been asked to uncover how a particular group of bat genes impact an important trait: whether the bat can carry Ebola. Nobody knows the underlying mechanism; it is up to you to hypothesize what is going on. For 100,000 independently sampled bats you have collected data of whether or not five genes are expressed, and whether or not the bat can carry Ebola.

[BatEbolaData.zip]

If a gene is expressed, it can affect both the probability of other genes being expressed and the probability of the trait being expressed. You can find the data in a file called `bats.csv`. Each row in the file corresponds to **one bat** and has 6 Booleans:

1. Gene 1: Whether the 1st gene is expressed in the bat (G_1)
2. Gene 2: Whether the 2nd gene is expressed in the bat (G_2)
3. Gene 3: Whether the 3rd gene is expressed in the bat (G_3)
4. Gene 4: Whether the 4th gene is expressed in the bat (G_4)
5. Gene 5: Whether the 5th gene is expressed in the bat (G_5)
6. Trait: Whether the trait is expressed in the bat; i.e., the bat can carry Ebola (T)

$P(E|F) \approx$ **0.42** $\quad P(E|F) \approx 0.17 \quad P(E) \approx 0.18$



Review

CS109: From Counting to Machine Learning

Counting Theory

Core Probability

Random Variables

Probabilistic Models

Uncertainty Theory

Machine Learning

The Core Probability Toolkit

The Law of Total Probability

$$P(E) = P(E \text{ and } F) + P(E \text{ and } F^{C})$$

$$P(E) = P(E|F)\,P(F) + P(E|F^{C})\,P(F^{C})$$

$$P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i) = \sum_{i=1}^{n} P(E|B_i)\,P(B_i)$$

Bayes' Theorem

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|B^{C}) \cdot P(B^{C})}$$

Definition of Conditional Probability

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$$

Axiom 1: $0 \le P(E) \le 1$

Axiom 2: $P(S) = 1$

Axiom 3: If E and F are mutually exclusive, then $P(E \text{ or } F) = P(E) + P(F)$

Otherwise, use Inclusion-Exclusion:

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

$$P(E^{C}) = 1 - P(E)$$

De Morgan's Laws

$(A \text{ or } B)^{C} = A^{C} \text{ and } B^{C}$

$(A \text{ and } B)^{C} = A^{C} \text{ or } B^{C}$

Chain Rule

$$P(E \text{ and } F) = P(E|F) \cdot P(F) = P(F|E) \cdot P(E)$$

Independence

$$P(E|F) = P(E)$$

$$P(E \text{ and } F) = P(E)\,P(F)$$

Definition of Conditional Probability

The **conditional probability** of E given F is the probability that E occurs, *given that F occurs*. This is called "conditioning on F".

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

Definition of Conditional Probability

The **conditional probability** of E given F is the probability that E occurs, *given that F occurs*. This is called "conditioning on F".

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

- This definition works even when outcomes are not equally likely

Definition of Conditional Probability

The **conditional probability** of E given F is the probability that E occurs, *given that F occurs*. This is called "conditioning on F".

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

We can multiply both sides by $P(F)$ to get the **chain rule**:

$$P(E \cap F) = P(E|F)P(F)$$

Definition of Conditional Probability

The **conditional probability** of E given F is the probability that E occurs, *given that F occurs*. This is called "conditioning on F".

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

We can multiply both sides by $P(F)$ to get the **chain rule**:

$$P(E \cap F) = P(E|F)P(F)$$

Remember: $P(E|F) \neq P(E \cap F)$ $\quad$ $P(E|F) \neq P(F|E)$

Definition of Conditional Probability

The **conditional probability** of E given F is the probability that E occurs, *given that F occurs*. This is called "conditioning on F".

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

We can multiply both sides by $P(F)$ to get the **chain rule**:

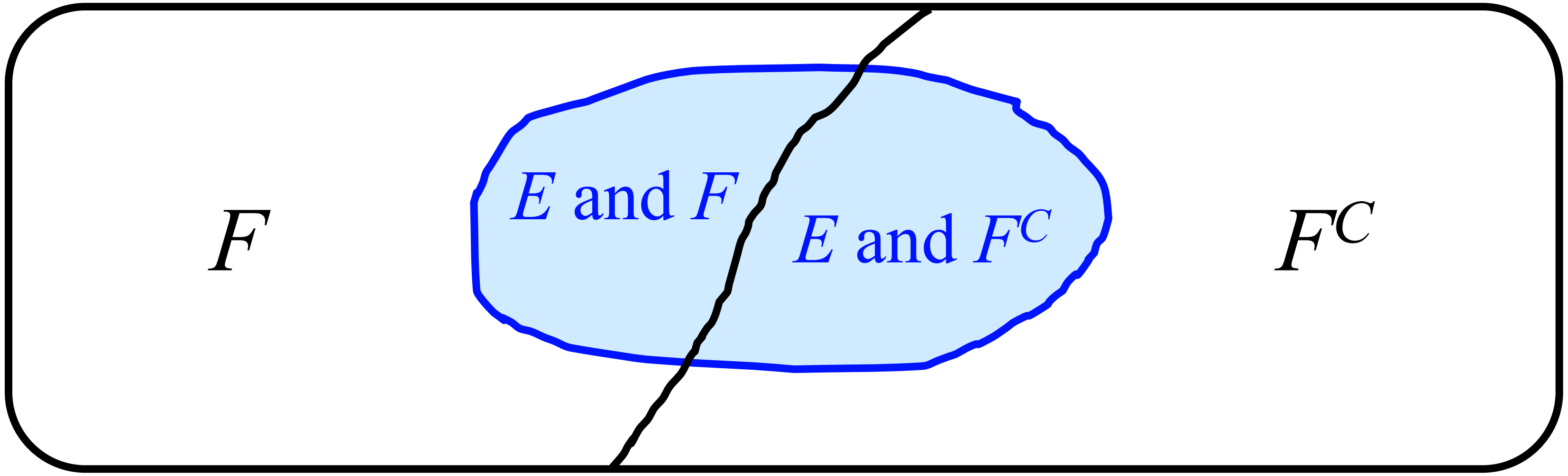
$$P(E \cap F) = P(E|F)P(F)$$

Remember: $P(E|F) \neq P(E \cap F)$ $\quad$ $P(E|F) \neq P(F|E)$

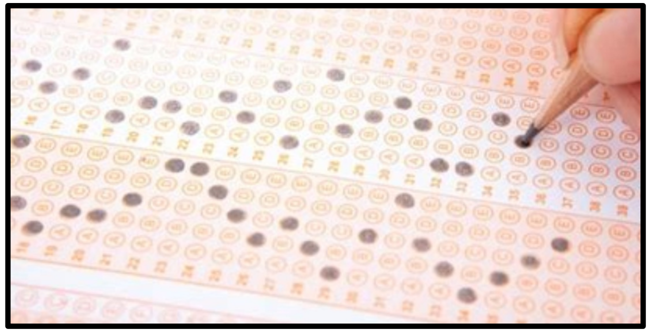
Law of Total Probability

Let E and F be our two events. The **Law of Total Probability** is:

$$\begin{aligned} P(E) &= P(EF) + P(EF^C) \\ &= P(E|F)P(F) + P(E|F^C)P(F^C) \end{aligned}$$



Sample Space



Review: Conditional Probability Formulas

$P(E \text{ and } F)$

Definition of Conditional Probability

Chain Rule

$P(E|F)$

Law of Total Probability

Bayes' Theorem

$P(E)$

$P(F|E)$

Bayes' Theorem

$P(E|F) \Rightarrow P(F|E)$

For any events E and F where $P(E) > 0$ and $P(F) > 0$,

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

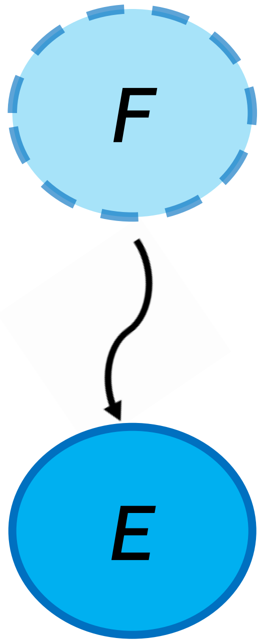
Expanded form:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

Bayes' Theorem

$P(E|F) \Rightarrow P(F|E)$

For any events E and F where $P(E) > 0$ and $P(F) > 0$,



$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

Expanded form:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

Bayes' Theorem

$P(E|F) \Rightarrow P(F|E)$

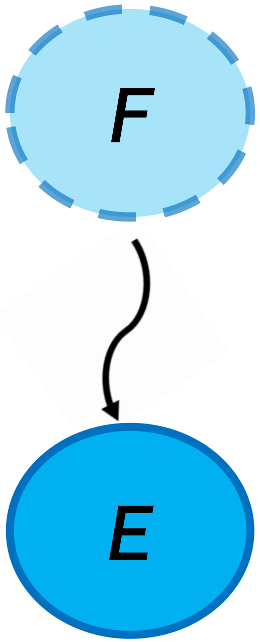
For any events E and F where $P(E) > 0$ and $P(F) > 0$,

likelihood prior

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

posterior

normalization constant



Expanded form:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

Practice: Multiple Choice & Probability

Imagine a multiple choice test, where every question has 4 answer choices.

- Let G be the event that a student guesses an answer to a question. $P(G) = 1/5$.
- Let R be the event that the student gets the answer right.
- Let the probability a student gets the answer right without guessing be $P(R \mid G^C) = 9/10$.

What's the probability a student knows their stuff (doesn't guess), *given* their answer is right?

Practice: Multiple Choice & Probability

Imagine a multiple choice test, where every question has 4 answer choices.

- Let G be the event that a student guesses an answer to a question. $P(G) = 1/5$.
- Let R be the event that the student gets the answer right.
- Let the probability a student gets the answer right without guessing be $P(R \mid G^C) = 9/10$.

What's the probability a student knows their stuff (doesn't guess), *given* their answer is right?

$$P(G^C|R) = \frac{P(R|G^C)P(G^C)}{P(R|G^C)P(G^C) + P(R|G)P(G)}$$

Practice: Multiple Choice & Probability

Imagine a multiple choice test, where every question has 4 answer choices.

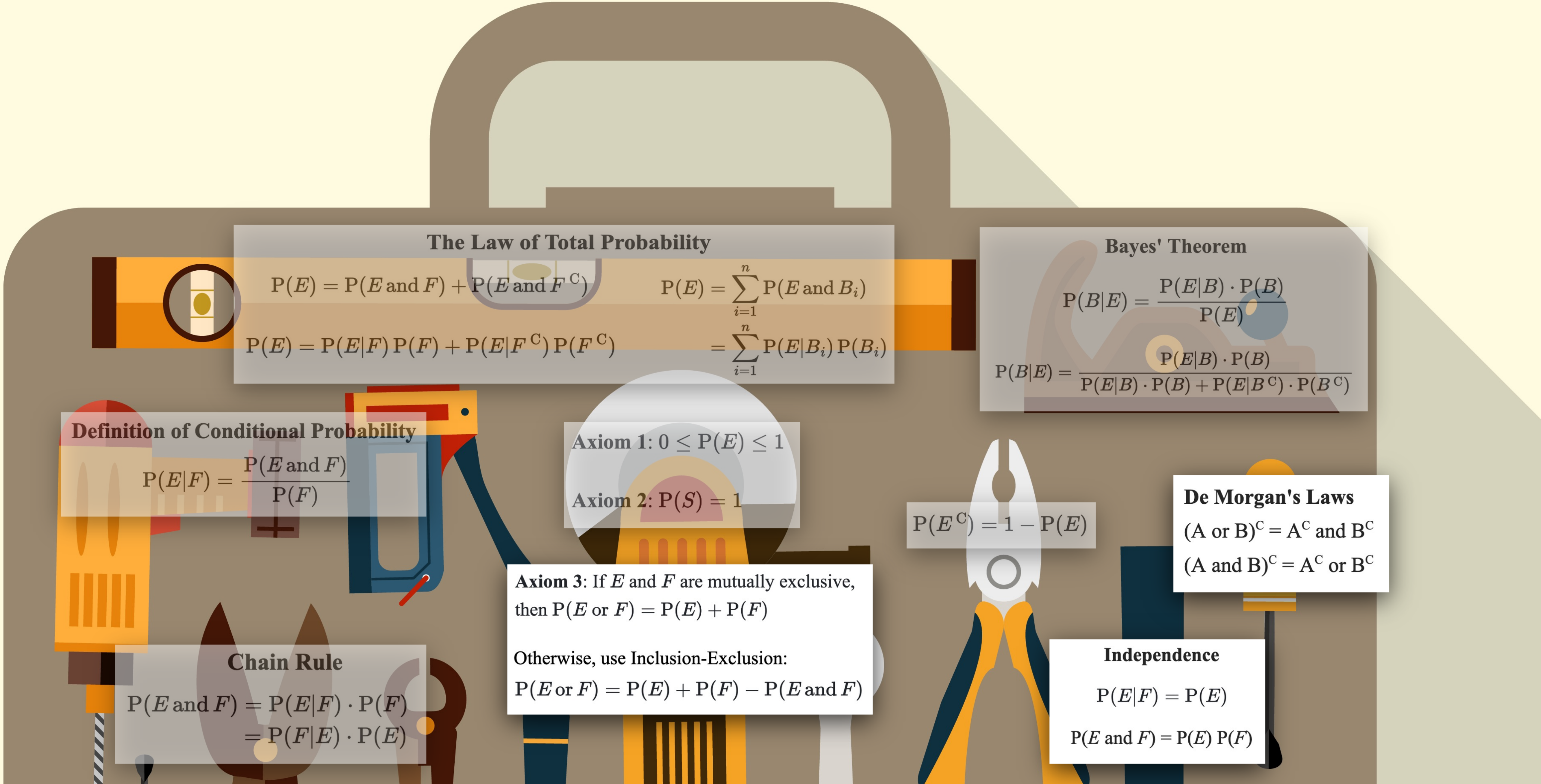
- Let G be the event that a student guesses an answer to a question. $P(G) = 1/5$.
- Let R be the event that the student gets the answer right.
- Let the probability a student gets the answer right without guessing be $P(R \mid G^C) = 9/10$.

What's the probability a student knows their stuff (doesn't guess), *given* their answer is right?

$$P(G^C|R) = \frac{P(R|G^C)P(G^C)}{P(R|G^C)P(G^C) + P(R|G)P(G)} = \frac{\frac{9}{10} \cdot \frac{4}{5}}{\frac{9}{10} \cdot \frac{4}{5} + \frac{1}{4} \cdot \frac{1}{5}}$$

End Review

The Core Probability Toolkit



The Law of Total Probability

$$P(E) = P(E \text{ and } F) + P(E \text{ and } F^{C}) \qquad P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i)$$

$$P(E) = P(E|F)\,P(F) + P(E|F^{C})\,P(F^{C}) \qquad = \sum_{i=1}^{n} P(E|B_i)\,P(B_i)$$

Bayes' Theorem

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|B^{C}) \cdot P(B^{C})}$$

Definition of Conditional Probability

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$$

Axiom 1: $0 \le P(E) \le 1$

Axiom 2: $P(S) = 1$

$$P(E^{C}) = 1 - P(E)$$

De Morgan's Laws

$(A \text{ or } B)^{C} = A^{C} \text{ and } B^{C}$

$(A \text{ and } B)^{C} = A^{C} \text{ or } B^{C}$

Axiom 3: If E and F are mutually exclusive, then $P(E \text{ or } F) = P(E) + P(F)$

Otherwise, use Inclusion-Exclusion:

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

Chain Rule

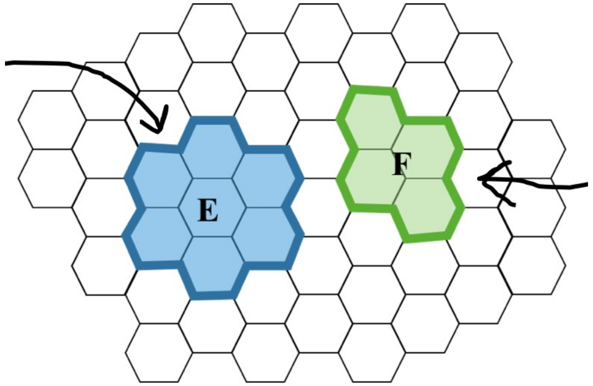
$$P(E \text{ and } F) = P(E|F) \cdot P(F) = P(F|E) \cdot P(E)$$

Independence

$$P(E|F) = P(E)$$

$$P(E \text{ and } F) = P(E)\,P(F)$$

Main Learning Goal For Today



Mutually Exclusive Events

make ***OR*** easy:

$$P(A \text{ or } B) = P(A) + P(B)$$

Independent Events

make ***AND*** easy:

$$P(A \text{ and } B) = P(A) \cdot P(B)$$


We the People
of the
insure domestic Tranquility provide for the common
and our Posterity, do ordain and establish this
Article. I

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

Examples of *independent* events:

- Rolling two dice
- Flipping a coin repeatedly
- Each time you gamble

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

Examples of *independent* events:

- Rolling two dice
- Flipping a coin repeatedly
- Each time you gamble

Sometimes you can argue for independence just by intuition.

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

Examples of *independent* events:

- Rolling two dice
- Flipping a coin repeatedly
- Each time you gamble

Sometimes you can argue for independence just by intuition.

Examples of *dependent* events:

- If it's sunny today, and if it's summer
- Your height vs. your parents' heights
- Going to the moon vs. being named Neil

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

Examples of *independent* events:

- Rolling two dice
- Flipping a coin repeatedly
- Each time you gamble

Sometimes you can argue for independence just by intuition.

Examples of *dependent* events:

- If it's sunny today, and if it's summer
- Your height vs. your parents' heights
- Going to the moon vs. being named Neil

Dependence does not require cause and effect!

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

With independence, we can simplify the chain rule:

$$\begin{aligned} P(A \cap B) &= P(A|B) \cdot P(B) \\ &= P(A) \cdot P(B) \end{aligned}$$

You can also show this ^ to prove independence

Let's Prove: Independence is Reciprocal

Want to show: if A is independent of B, then B is independent of A.

Start: $P(A) = P(A|B)$ Goal: $P(B|A) = P(B)$

Let's Prove: Independence is Reciprocal

Want to show: if A is independent of B, then B is independent of A.

Start: $P(A) = P(A|B)$ Goal: $P(B|A) = P(B)$

Proof:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Start with Bayes' Theorem

Let's Prove: Independence is Reciprocal

Want to show: if A is independent of B, then B is independent of A.

Start: $P(A) = P(A|B)$ Goal: $P(B|A) = P(B)$

Proof:

$$
\begin{aligned}
P(B|A) &= \frac{P(A|B)P(B)}{P(A)} && \text{Start with Bayes' Theorem} \\
&= \frac{P(A)P(B)}{P(A)} && \text{Because A is independent of B} \\
&= P(B)
\end{aligned}
$$

Let's Prove: Independence Of Complements

Want to show: if A is independent of B, then A is independent of B^C.

Start: $P(AB) = P(A) \cdot P(B)$ Goal: $P(AB^C) = P(A) \cdot P(B^C)$

Proof:

$$P(AB) + P(AB^C) = P(A)$$

Start with Law of Total Prob.

Let's Prove: Independence Of Complements

Want to show: if A is independent of B, then A is independent of B^C.

Start: $P(AB) = P(A) \cdot P(B)$ Goal: $P(AB^C) = P(A) \cdot P(B^C)$

Proof:

$$P(AB) + P(AB^C) = P(A)$$ Start with Law of Total Prob.

$$P(AB^C) = P(A) - P(AB)$$ Subtract from both sides

Let's Prove: Independence Of Complements

Want to show: if A is independent of B, then A is independent of B^C.

Start: $P(AB) = P(A) \cdot P(B)$ Goal: $P(AB^C) = P(A) \cdot P(B^C)$

Proof:

$$P(AB) + P(AB^C) = P(A)$$ Start with Law of Total Prob.

$$P(AB^C) = P(A) - P(AB)$$ Subtract from both sides

$$= P(A) - P(A) \cdot P(B)$$ Independence

Let's Prove: Independence Of Complements

Want to show: if A is independent of B, then A is independent of B^C.

Start: $P(AB) = P(A) \cdot P(B)$ Goal: $P(AB^C) = P(A) \cdot P(B^C)$

Proof:

$$P(AB) + P(AB^C) = P(A)$$ Start with Law of Total Prob.

$$P(AB^C) = P(A) - P(AB)$$ Subtract from both sides

$$= P(A) - P(A) \cdot P(B)$$ Independence

$$= P(A)(1 - P(B))$$ Pull out P(A)

$$= P(A) \cdot P(B^C)$$ Identity 3: $P(B^C) = 1 - P(B)$

Independence For 3 Or More Events

Events E_1 , E_2 , ..., E_n are independent if, **for every subset** of size r of the events:

$$P(E_{1'} E_{2'} E_{3'} \ldots E_{r'}) = P(E_{1'})\ P(E_{2'})\ P(E_{3'})\ \ldots\ P(E_{r'})$$

Independence For 3 Or More Events

Events $E_1, E_2, \ldots, E_n$ are independent if, **for every subset** of size r of the events:

$$P(E_{1'}E_{2'}E_{3'}\ldots E_{r'}) = P(E_{1'})\ P(E_{2'})\ P(E_{3'})\ \ldots\ P(E_{r'})$$

For example, for events A, B, C, you would need to show:

$$P(A,B) = P(A)\cdot P(B)$$
$$P(A,C) = P(A)\cdot P(C)$$
$$P(B,C) = P(B)\cdot P(C)$$
$$P(A,B,C) = P(A)\cdot P(B)\cdot P(C)$$

Independence For 3 Or More Events

Events E_1, E_2, ..., E_n are independent if, **for every subset** of size r of the events:

$$P(E_{1'}E_{2'}E_{3'}\ldots E_{r'}) = P(E_{1'})\ P(E_{2'})\ P(E_{3'})\ \ldots\ P(E_{r'})$$

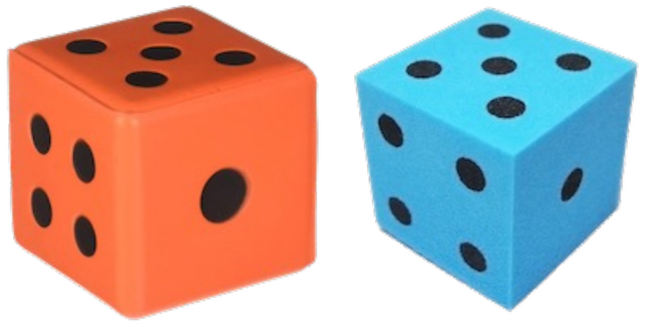
For example, for events A, B, C, you would need to show:

$$P(A,B) = P(A) \cdot P(B)$$
$$P(A,C) = P(A) \cdot P(C)$$
$$P(B,C) = P(B) \cdot P(C)$$
$$P(A,B,C) = P(A) \cdot P(B) \cdot P(C)$$

Good news! if you can argue independence intuitively, the mega-chain rule simplifies:

$$\begin{aligned} P(E_1, E_2, E_3, \ldots, E_n) &= P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1, E_2) \ldots P(E_n|E_1, E_2 \ldots E_{n_1}) \\ &= P(E_1) \cdot P(E_2) \cdot P(E_3) \ldots P(E_n) \end{aligned}$$

Independence & Dice



D_1 D_2

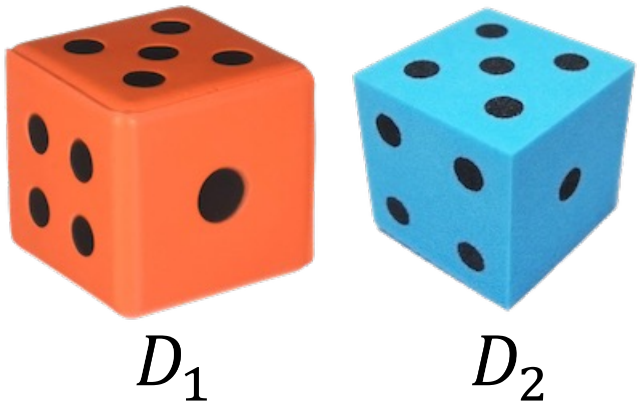
Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$?

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

Independence & Dice



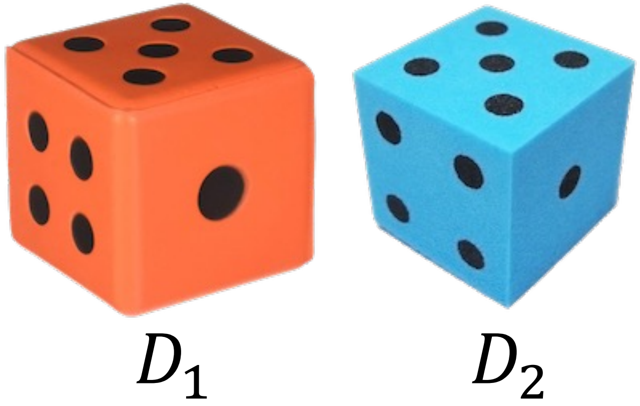
Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

Are E and F independent?

Independence & Dice



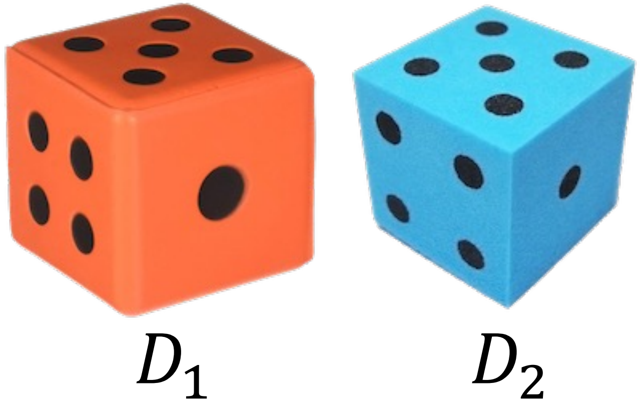
Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

Are E and F independent? **Yes!** Check: $P(EF) = P(E)\,P(F)$

Independence & Dice



Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

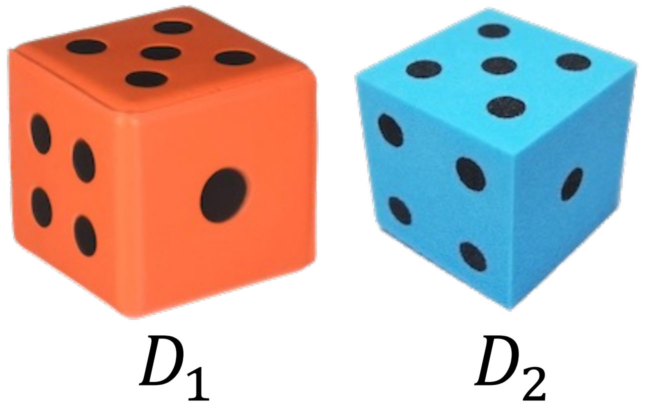
What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

Are E and F independent? **Yes!** Check: $P(EF) = P(E)\,P(F)$

Let G be the event that $D_1 + D_2 = 7$.

What is $P(E)$, $P(G)$, and $P(EG)$?

Independence & Dice



Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

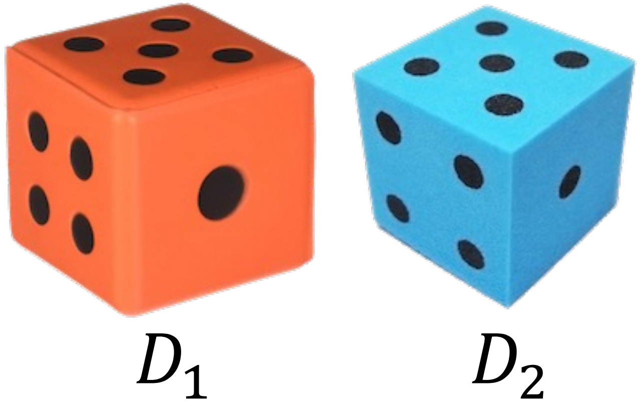
Are E and F independent? Yes! Check: $P(EF) = P(E)\,P(F)$

Let G be the event that $D_1 + D_2 = 7$.

What is $P(E)$, $P(G)$, and $P(EG)$? $P(E) = 1/6$ $P(G) = 1/6$ $P(EG) = 1/36$

Are E and G independent?

Independence & Dice



Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

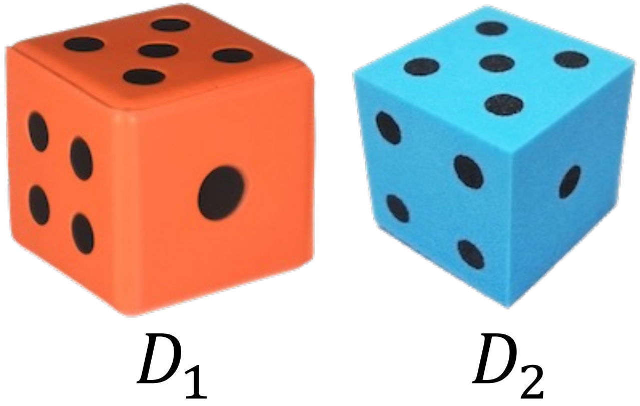
Are E and F independent? Yes! Check: $P(EF) = P(E)\,P(F)$

Let G be the event that $D_1 + D_2 = 7$.

What is $P(E)$, $P(G)$, and $P(EG)$? $P(E) = 1/6$ $P(G) = 1/6$ $P(EG) = 1/36$

Are E and G independent? **Yes?!** Check: $P(EG) = P(E)\,P(G)$

Independence & Dice



Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

Are E and F independent? Yes! Check: $P(EF) = P(E)\,P(F)$

Let G be the event that $D_1 + D_2 = 7$.

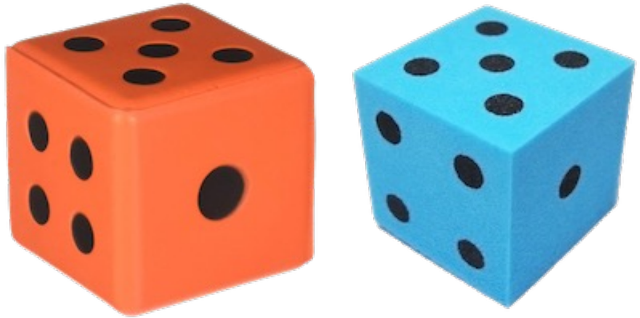
What is $P(E)$, $P(G)$, and $P(EG)$? $P(E) = 1/6$ $P(G) = 1/6$ $P(EG) = 1/36$

Are E and G independent? Yes?! Check: $P(EG) = P(E)\,P(G)$

Let H be the event that $D_1 + D_2 = 4$.

What is $P(E)$, $P(H)$, and $P(EH)$?

Independence & Dice



Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.

What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

Are E and F independent? Yes! Check: $P(EF) = P(E)\,P(F)$

Let G be the event that $D_1 + D_2 = 7$.

What is $P(E)$, $P(G)$, and $P(EG)$? $P(E) = 1/6$ $P(G) = 1/6$ $P(EG) = 1/36$

Are E and G independent? Yes?! Check: $P(EG) = P(E)\,P(G)$

Let H be the event that $D_1 + D_2 = 4$.

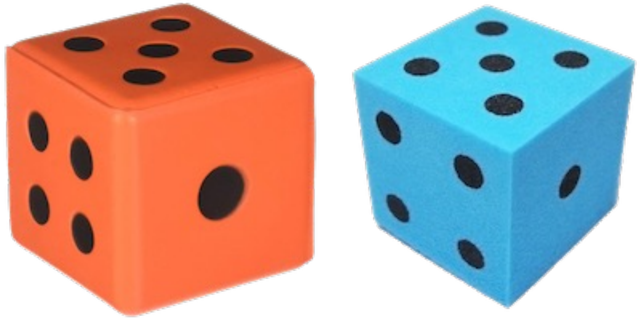
What is $P(E)$, $P(H)$, and $P(EH)$? $P(E) = 1/6$ $P(H) = 3/36$ $P(EH) = 1/36$

Are E and H independent?

Independence & Dice

Roll two 6-sided dice: D_1 and D_2.

Let E be the event $D_1 = 1$, and let F be the event $D_2 = 1$.



What is $P(E)$, $P(F)$, and $P(EF)$? $P(E) = 1/6$ $P(F) = 1/6$ $P(EF) = 1/36$

Are E and F independent? Yes! Check: $P(EF) = P(E)\,P(F)$

Let G be the event that $D_1 + D_2 = 7$.

What is $P(E)$, $P(G)$, and $P(EG)$? $P(E) = 1/6$ $P(G) = 1/6$ $P(EG) = 1/36$

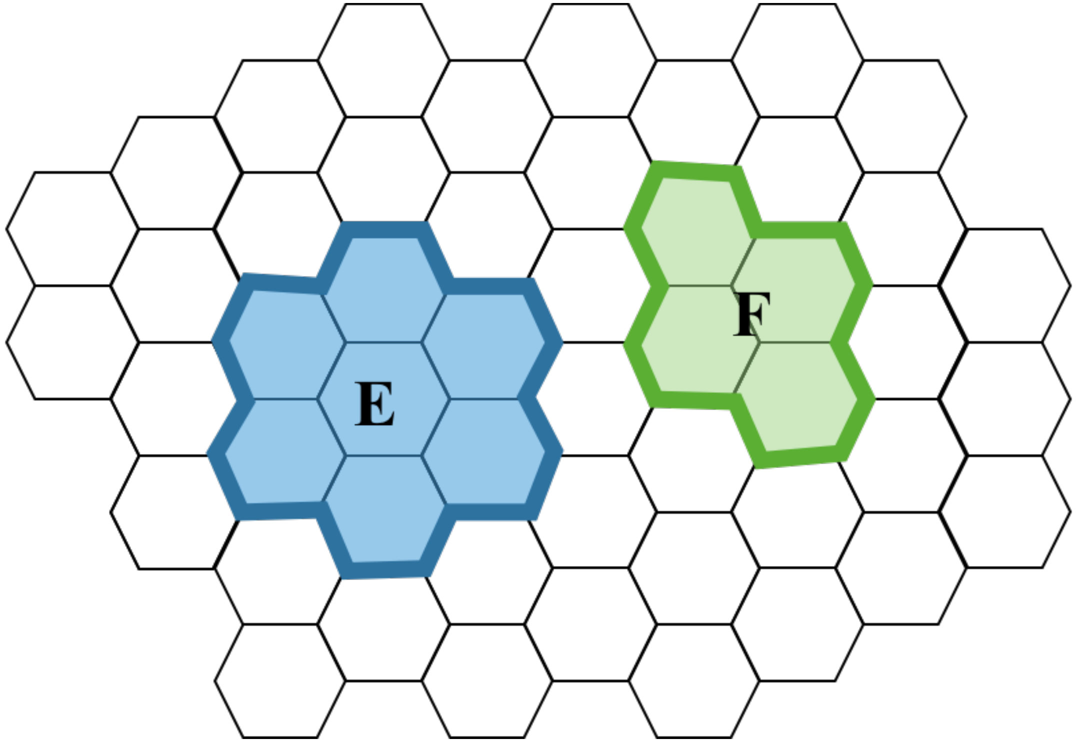
Are E and G independent? Yes?! Check: $P(EG) = P(E)\,P(G)$

Let H be the event that $D_1 + D_2 = 4$.

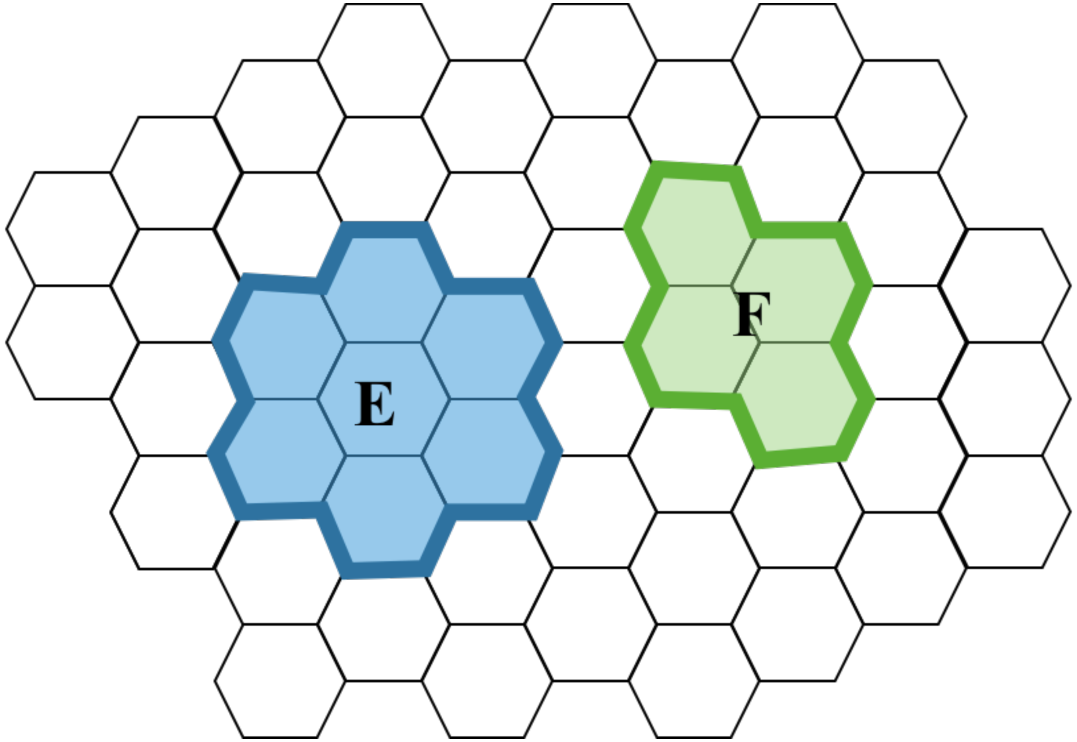
What is $P(E)$, $P(H)$, and $P(EH)$? $P(E) = 1/6$ $P(H) = 3/36$ $P(EH) = 1/36$

Are E and H independent? **No!** Check: $P(EH) \neq P(E)\,P(H)$

Can Mutually Exclusive Events Be Independent?



Can Mutually Exclusive Events Be Independent?

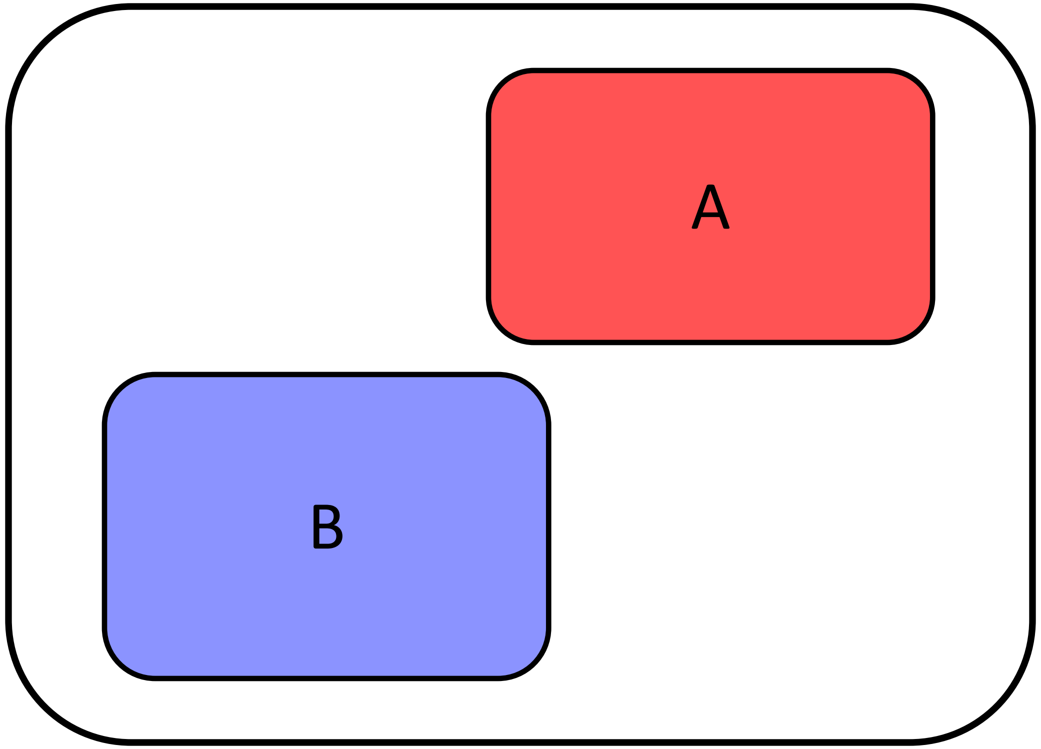


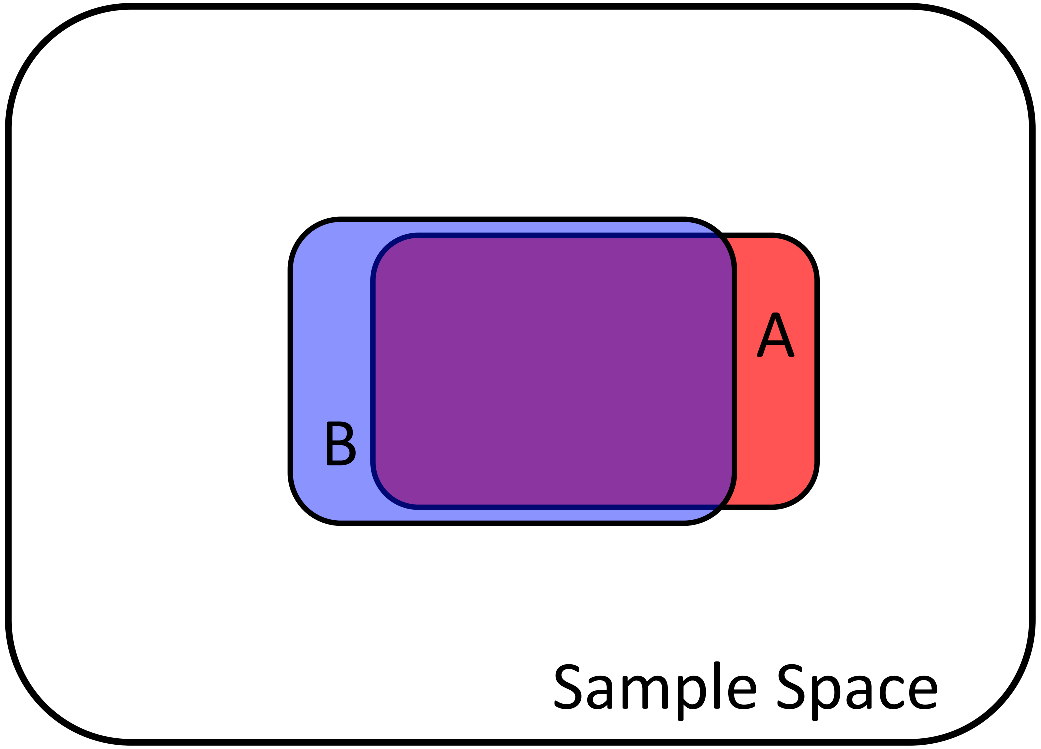
Nope!

Knowing that E happened changes our belief that F happens.

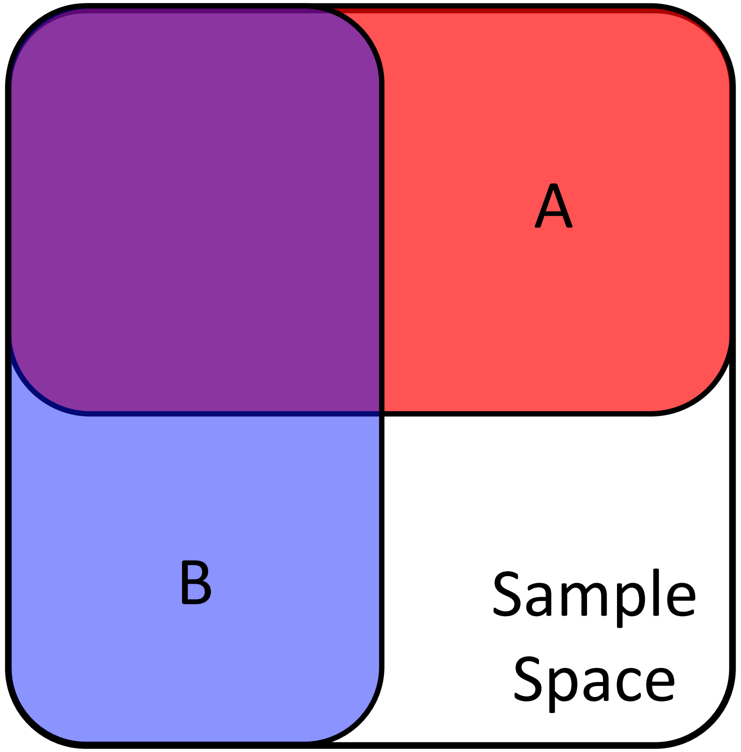
What does independence look like?

What Does *Dependence* Look Like?

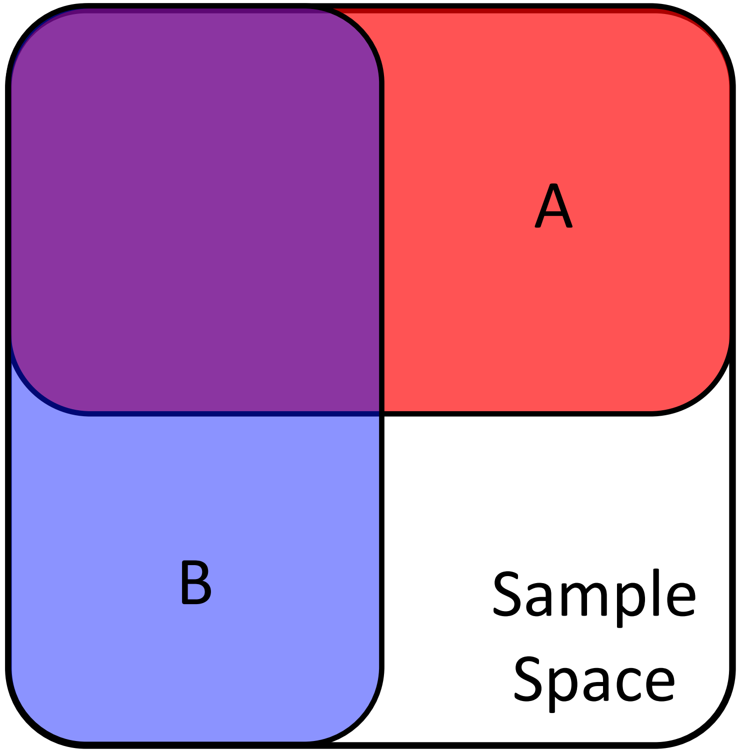


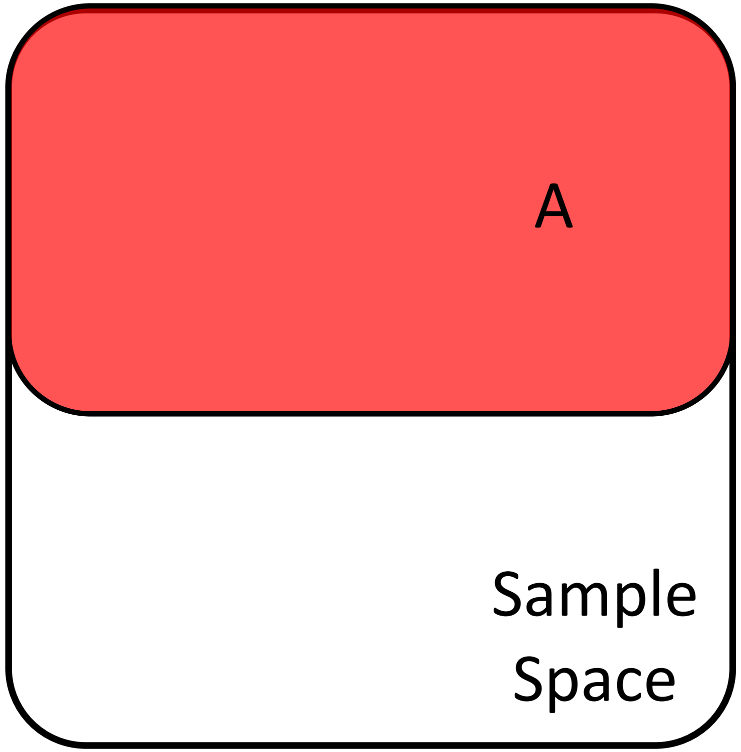


What Does Independence Look Like?



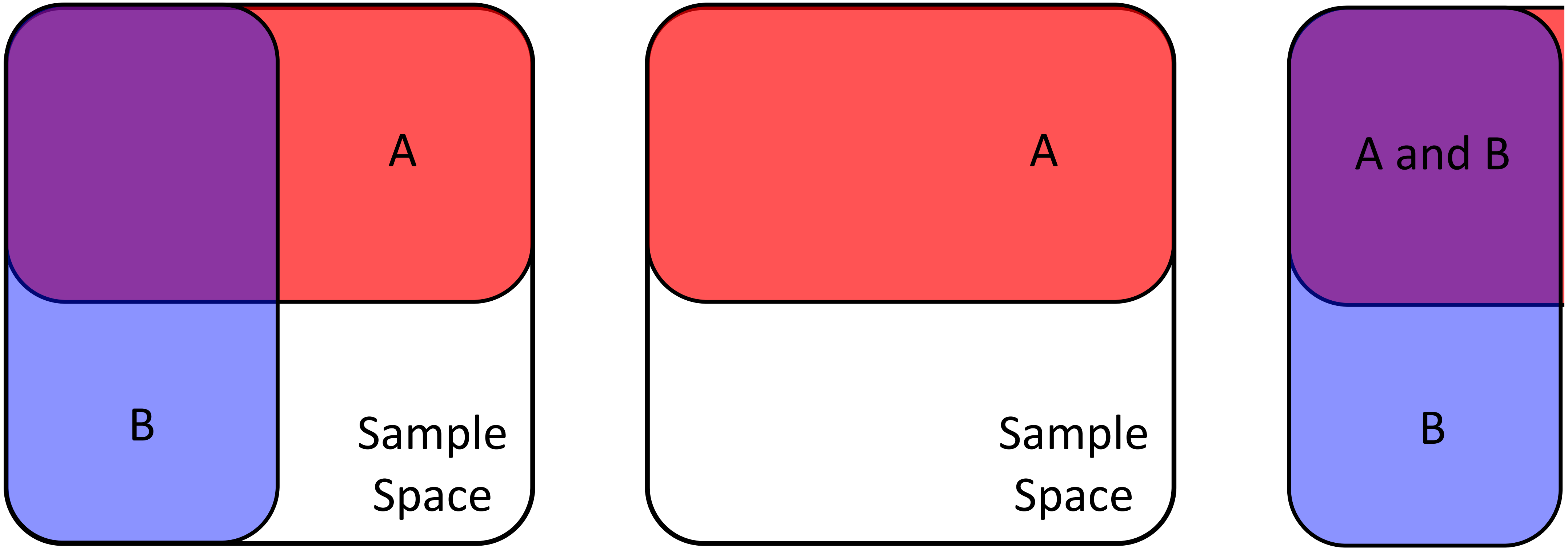
What Does Independence Look Like?





If we look at the whole
sample space, P(A) = 0.5

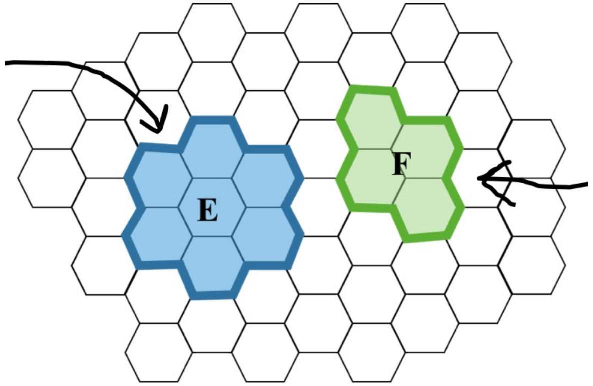
What Does Independence Look Like?



If we look at the whole sample space, P(A) = 0.5

If we shrink the sample space to B, P(A|B) = 0.5

Main Learning Goal For Today



Mutually Exclusive Events

make ***OR*** easy:

$$P(A \text{ or } B) = P(A) + P(B)$$

Independent Events

make ***AND*** easy:

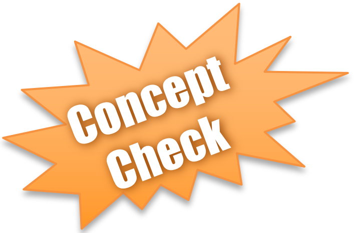
$$P(A \text{ and } B) = P(A) \cdot P(B)$$

Practice: Experiments With Children

Two parents have straight hair, but carry the curly hair gene.

The probability of any single kid having curly hair is 0.25, *independent* of other siblings (curly hair is recessive).

What is the probability that all **three** of their kids have curly hair?



Practice: Experiments With Children

Two parents have straight hair, but carry the curly hair gene.

The probability of any single kid having curly hair is 0.25, *independent* of other siblings (curly hair is recessive).

What is the probability that all **three** of their kids have curly hair?

Let E_1, E_2, E_3 be the events that kid 1, 2, or 3 have curly hair.

Concept Check

Practice: Experiments With Children

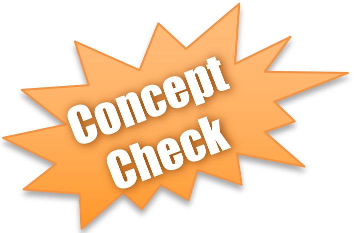
Two parents have straight hair, but carry the curly hair gene.

The probability of any single kid having curly hair is 0.25, *independent* of other siblings (curly hair is recessive).

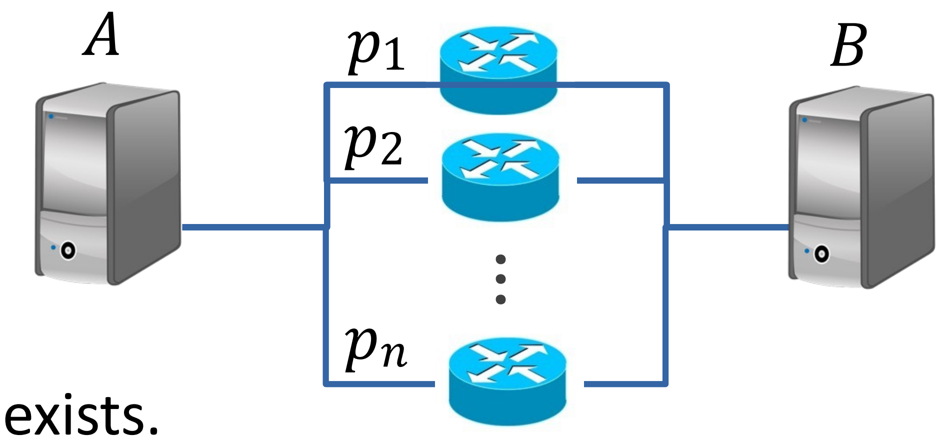
What is the probability that all **three** of their kids have curly hair?

Let E_1, E_2, E_3 be the events that kid 1, 2, or 3 have curly hair.

$$\begin{aligned} P(E_1, E_2, E_3) &= P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1, E_2) \\ &= P(E_1) \cdot P(E_2) \cdot P(E_3) \\ &= (0.25)^3 \approx 0.016 \end{aligned}$$



Practice: Network Reliability



Consider the following parallel network:

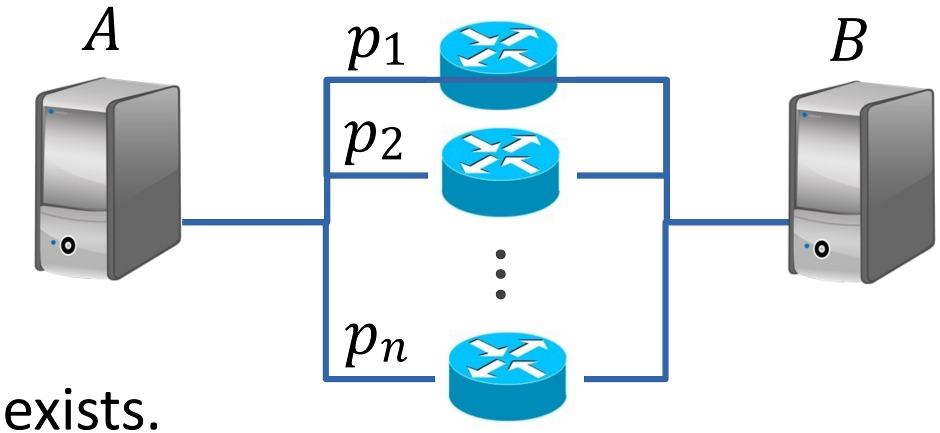
- n independent routers, which each are working with probability p_i $(1 \leq i \leq n)$

Let E be the event that a working path from A to B exists. What is $P(E)$?

Practice: Network Reliability

Consider the following parallel network:

- n independent routers, which each are working with probability p_i ($1 \le i \le n$)

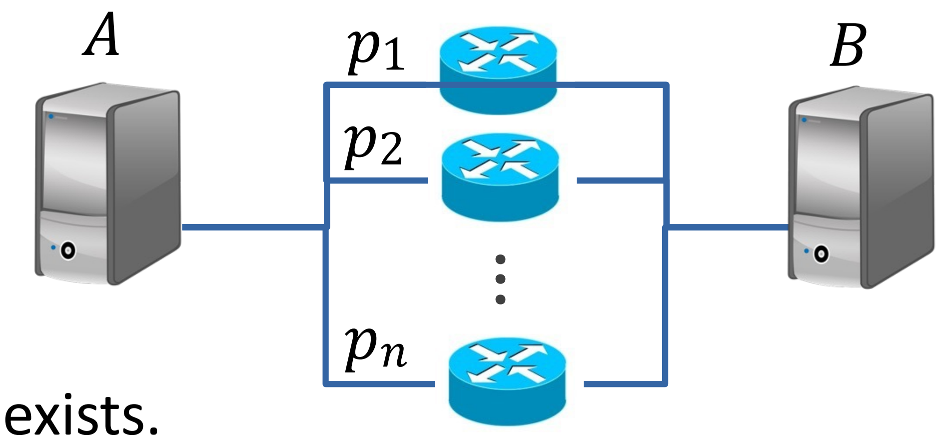


Let E be the event that a working path from A to B exists. What is $P(E)$?

$$P(E) = P(\geq 1 \text{ one router works})$$

Tip: whenever an event is "at least 1", taking the complement will probably help.

Practice: Network Reliability



Consider the following parallel network:

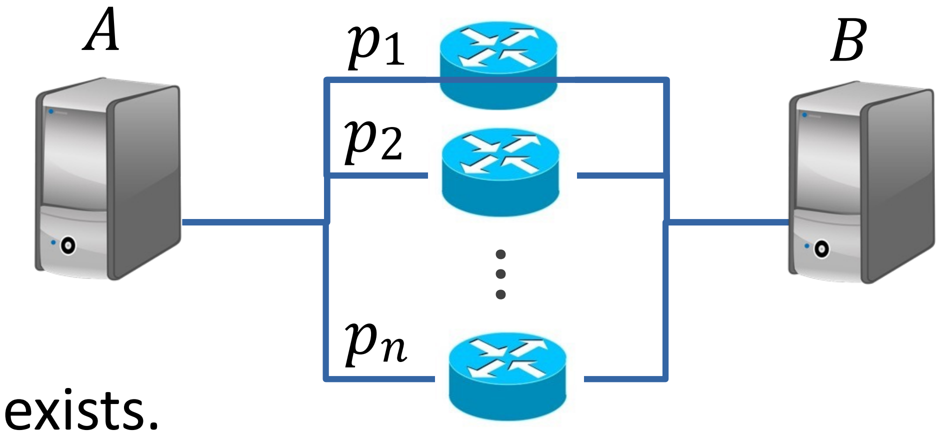
- n independent routers, which each are working with probability p_i $(1 \leq i \leq n)$

Let E be the event that a working path from A to B exists. What is $P(E)$?

$$P(E) = P(\geq 1 \text{ one router works})$$
$$= 1 - P(\text{all routers fail})$$

Tip: whenever an event is "at least 1", taking the complement will probably help.

Practice: Network Reliability



Consider the following parallel network:

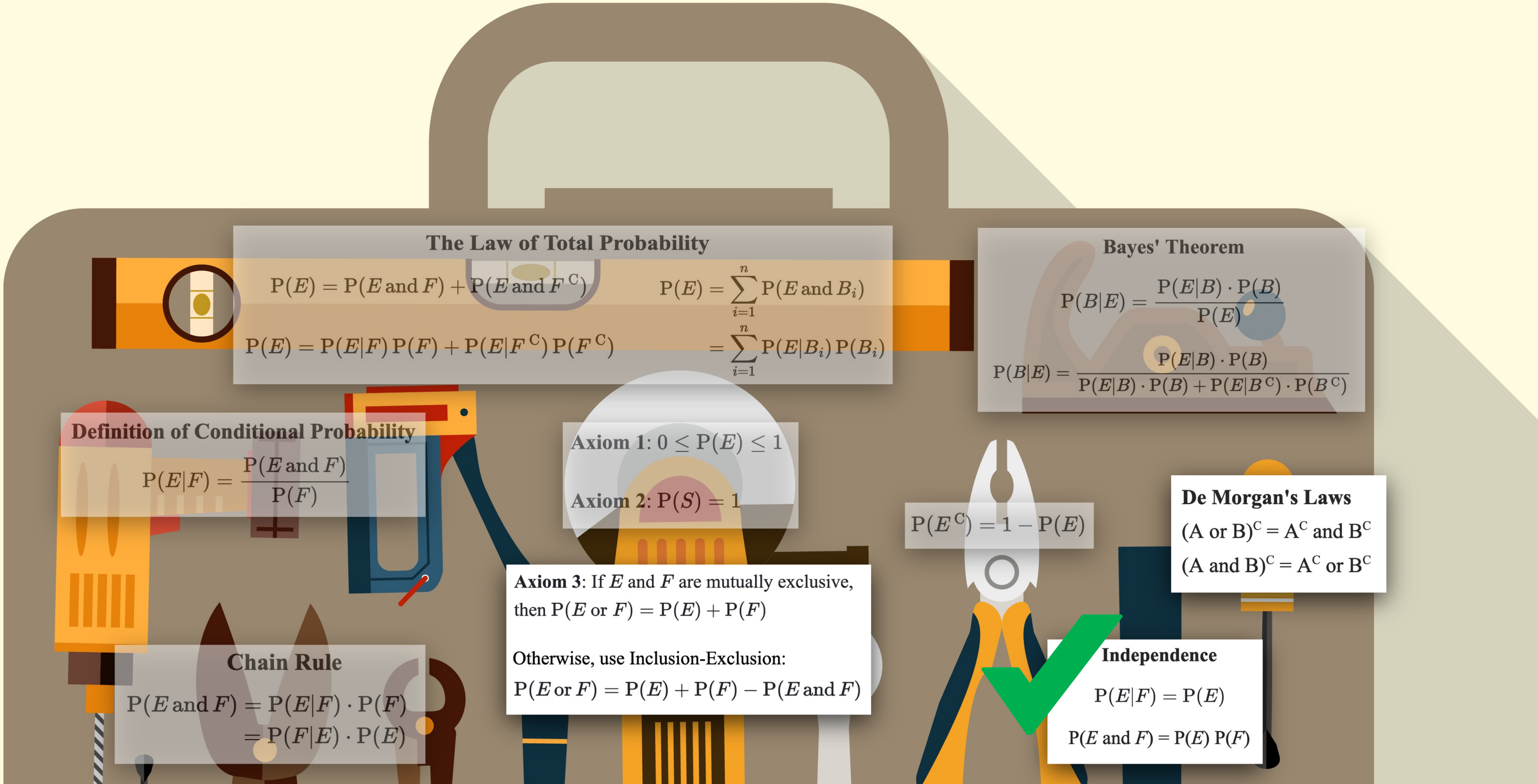
- n independent routers, which each are working with probability p_i $(1 \leq i \leq n)$

Let E be the event that a working path from A to B exists.
What is $P(E)$?

$$
\begin{aligned}
P(E) &= P(\geq 1 \text{ one router works}) \\
&= 1 - P(\text{all routers fail}) \\
&= 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n) \\
&= 1 - \prod_{i=1}^{n} (1 - p_i)
\end{aligned}
$$

Tip: whenever an event is "at least 1", taking the complement will probably help.

The Core Probability Toolkit



The Law of Total Probability

$$P(E) = P(E \text{ and } F) + P(E \text{ and } F^C)$$

$$P(E) = P(E|F)\,P(F) + P(E|F^C)\,P(F^C)$$

$$P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i)$$

$$= \sum_{i=1}^{n} P(E|B_i)\,P(B_i)$$

Bayes' Theorem

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|B^C) \cdot P(B^C)}$$

Definition of Conditional Probability

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$$

Axiom 1: $0 \le P(E) \le 1$

Axiom 2: $P(S) = 1$

$$P(E^C) = 1 - P(E)$$

De Morgan's Laws

$(A \text{ or } B)^C = A^C \text{ and } B^C$

$(A \text{ and } B)^C = A^C \text{ or } B^C$

Axiom 3: If E and F are mutually exclusive, then $P(E \text{ or } F) = P(E) + P(F)$

Otherwise, use Inclusion-Exclusion:

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

Chain Rule

$$P(E \text{ and } F) = P(E|F) \cdot P(F)$$

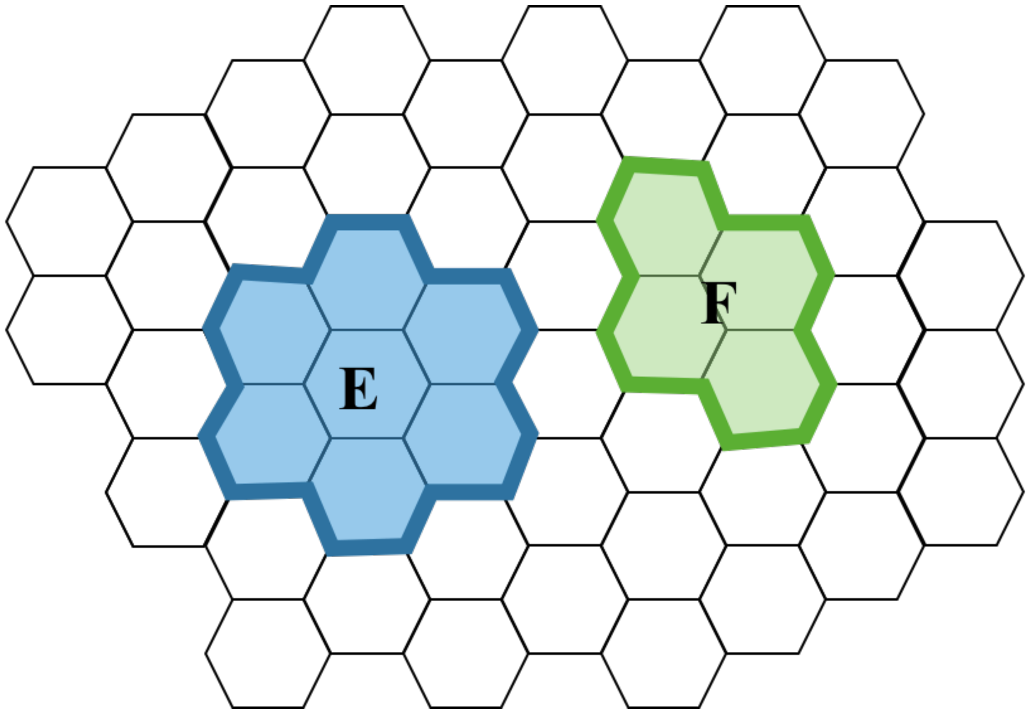
$$= P(F|E) \cdot P(E)$$

Independence

$$P(E|F) = P(E)$$

$$P(E \text{ and } F) = P(E)\,P(F)$$

Probability of Or With Mutually Exclusive Events

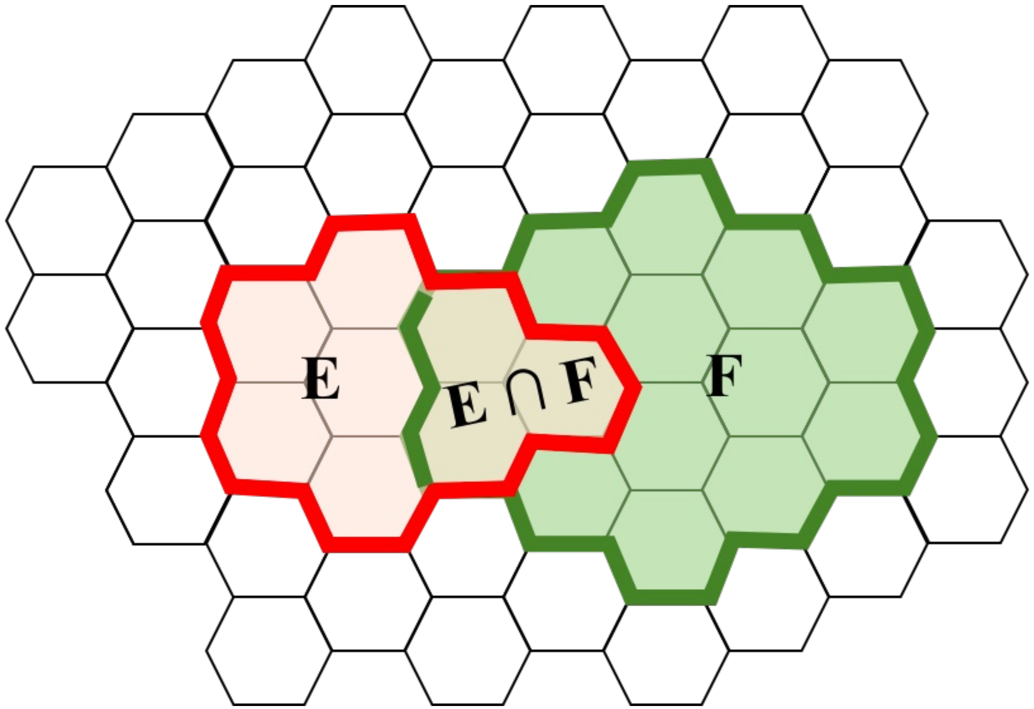


If events have no outcomes in common, probability of OR is simple:

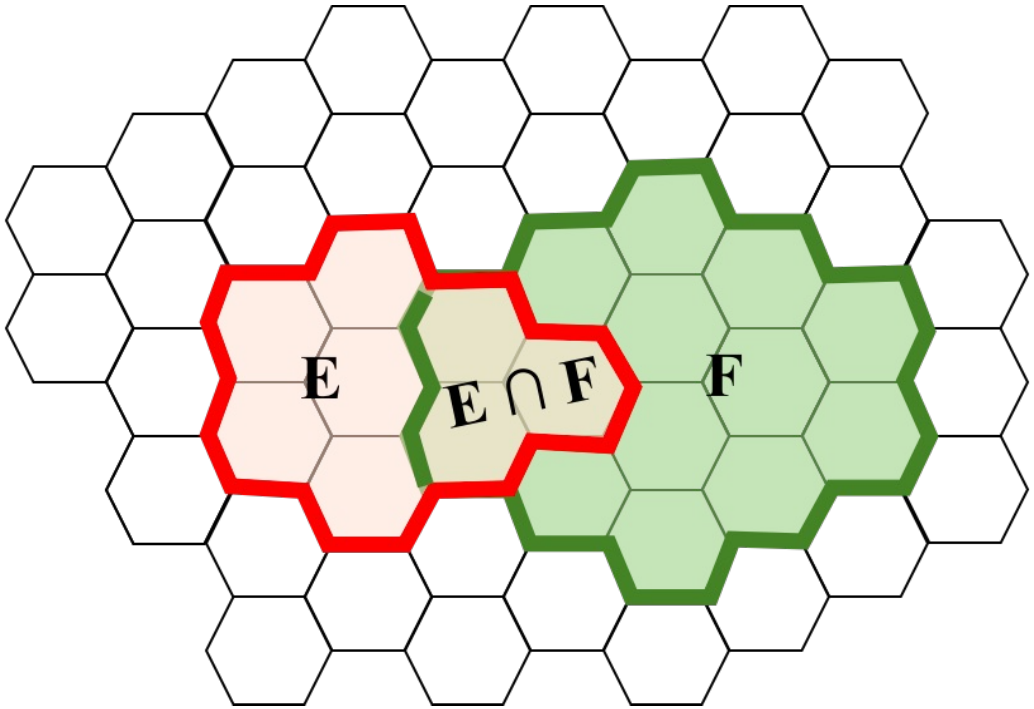
$$P(E \cup F) = P(E) + P(F)$$

What if events are *not* mutually exclusive?

Probability of Or *Without* Mutually Exclusive Events



Probability of Or *Without* Mutually Exclusive Events

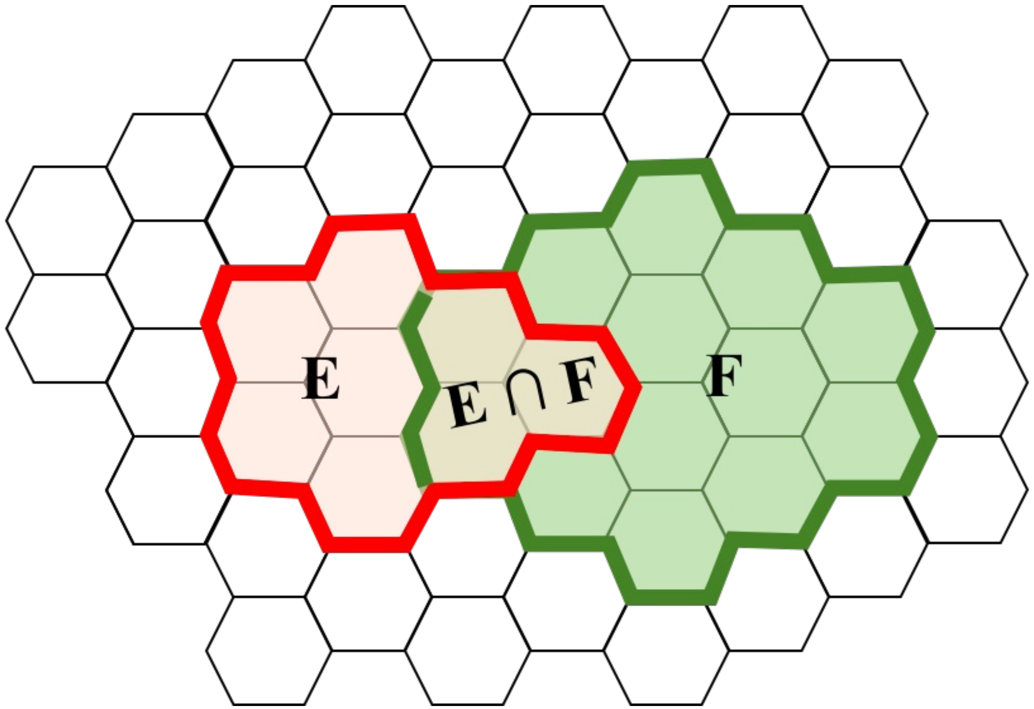


If events have outcomes in common, we correct for double-counting them:

$$P(E \text{ or } F) = P(E) + P(F) - P(EF)$$

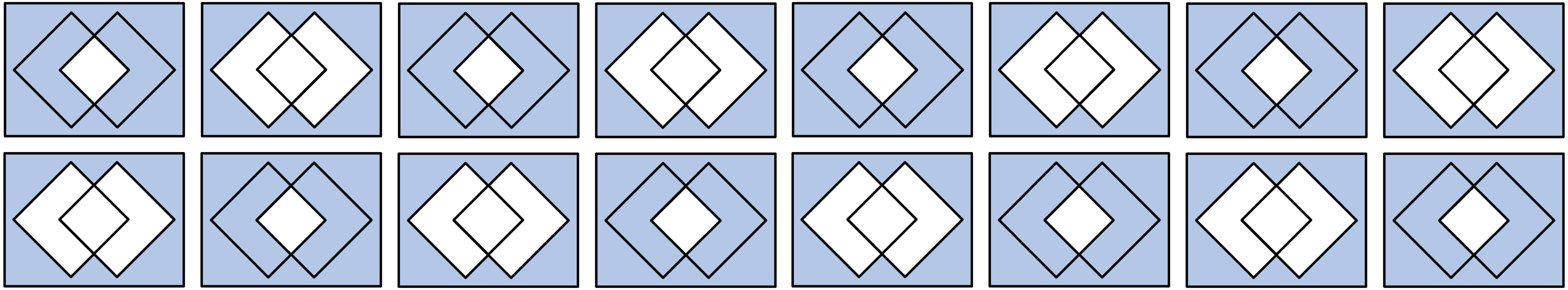
The "Inclusion Exclusion" Rule

Probability of Or *Without* Mutually Exclusive Events

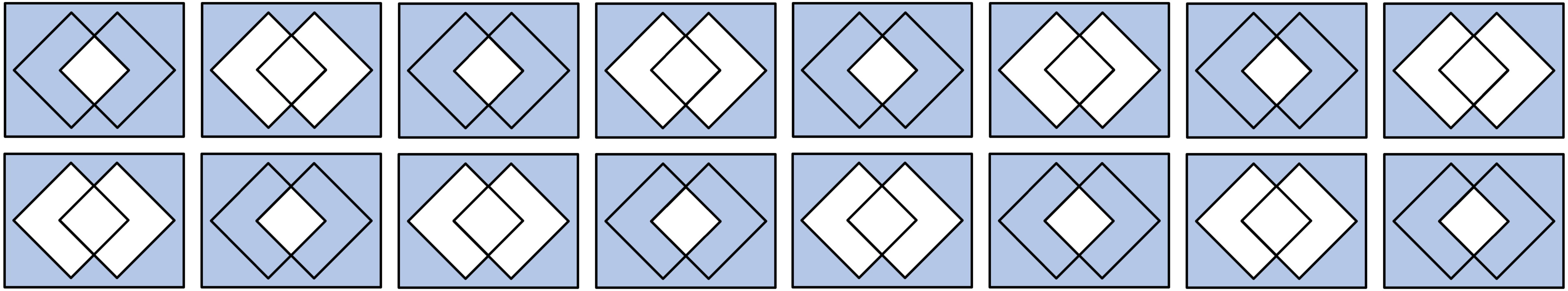


If events have outcomes in common, we correct for double-counting them:

$$P(E \text{ or } F) = \frac{8}{50} + \frac{14}{50} - \frac{3}{50}$$

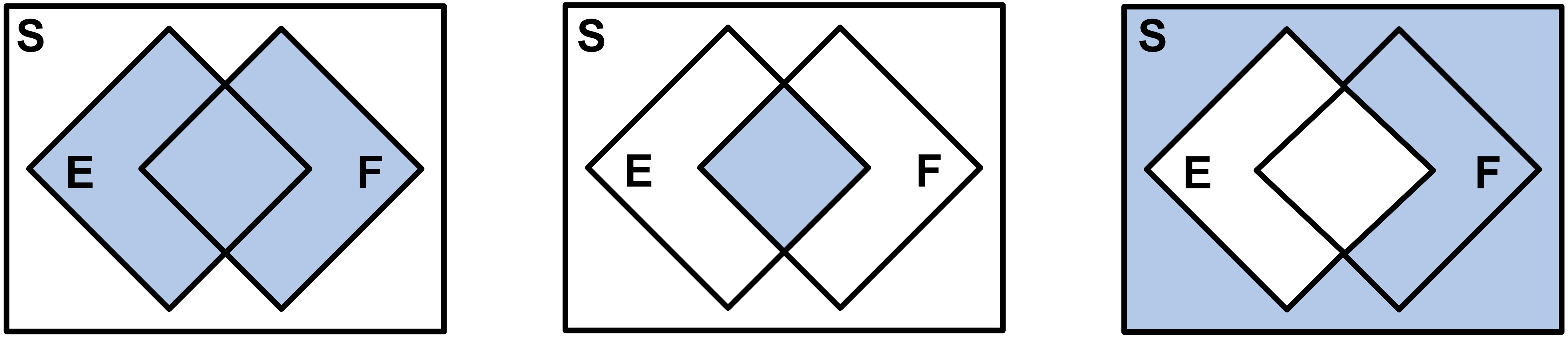


DeMorgan's Laws



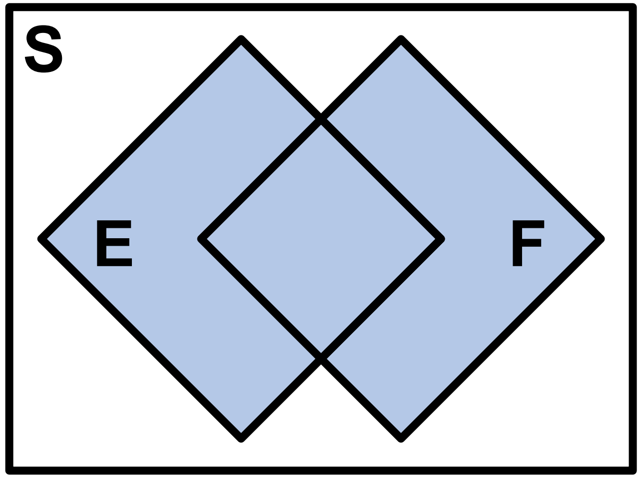
Quick Sets Review

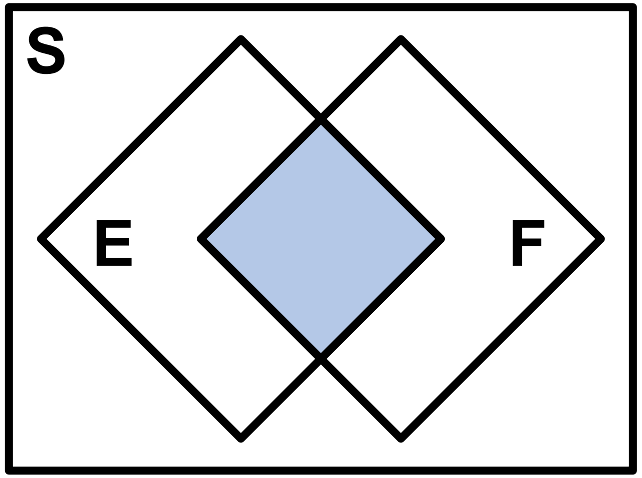
Say *E* and *F* are events in *S*. What set notation describes each picture below?

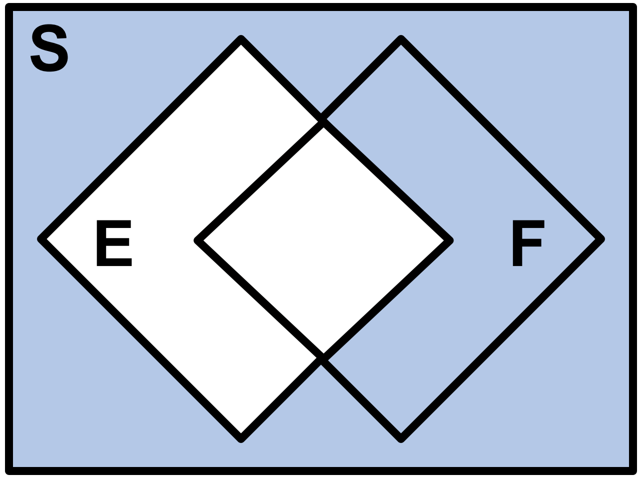


Quick Sets Review

Say *E* and *F* are events in *S*. What set notation describes each picture below?







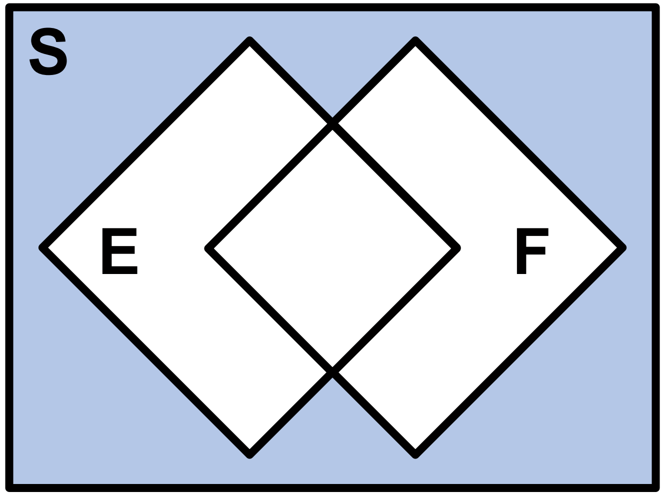
$$E \cup F$$
$$E \text{ or } F$$

$$E \cap F$$
$$E \text{ and } F$$

$$E^C$$
$$\text{not } E$$

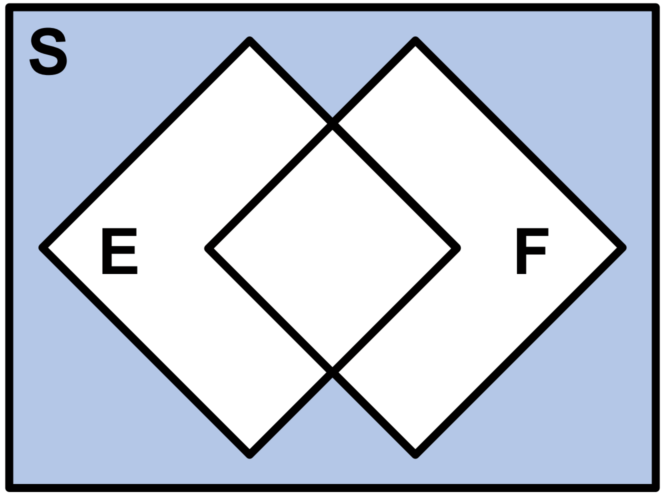
Challenge: Set Notation

Say *E* and *F* are events in *S*. What set notation describes the picture below?



Challenge: Set Notation

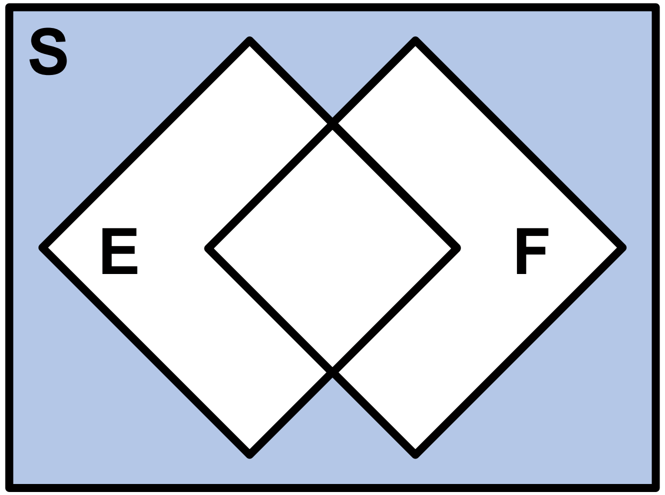
Say *E* and *F* are events in *S*. What set notation describes the picture below?



$(E \text{ or } F)^C$ or $E^C \text{ and } F^C$?

Challenge: Set Notation

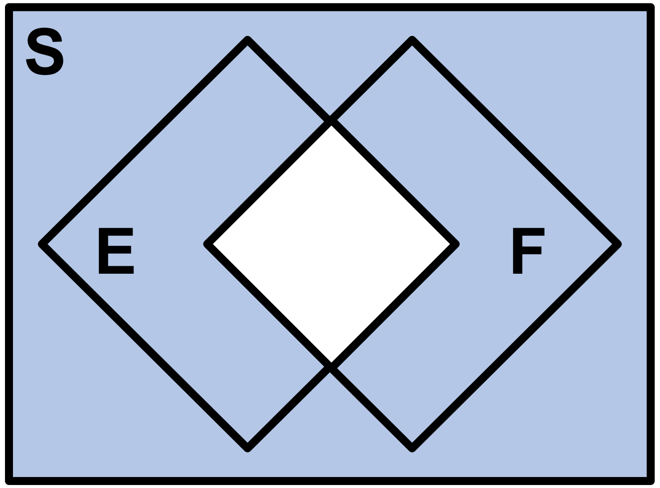
Say *E* and *F* are events in *S*. What set notation describes the picture below?



$$(E \text{ or } F)^C \quad = \quad E^C \text{ and } F^C$$

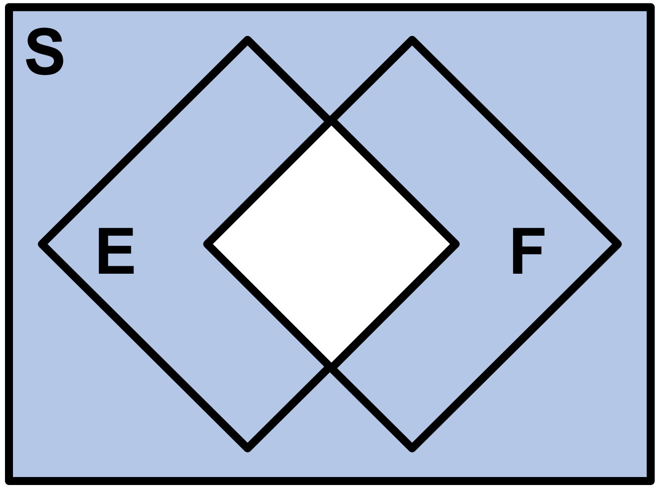
Challenge #2: Set Notation

Say *E* and *F* are events in *S*. What set notation describes the picture below?



Challenge #2: Set Notation

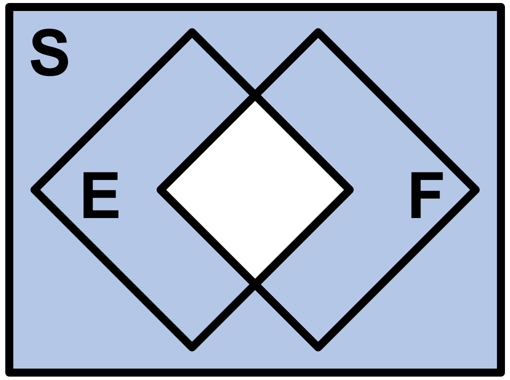
Say *E* and *F* are events in *S*. What set notation describes the picture below?



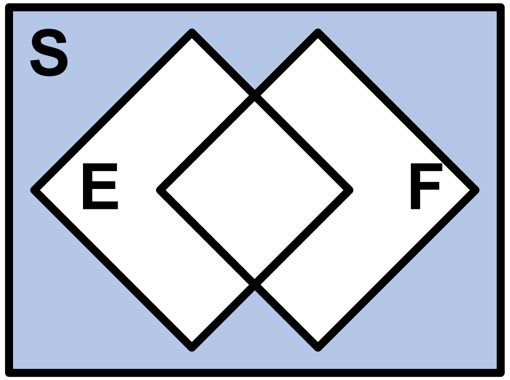
$$(E \text{ and } F)^C \quad = \quad E^C \text{ or } F^C$$

DeMorgan's Laws

These rules let you alternate between AND and OR.



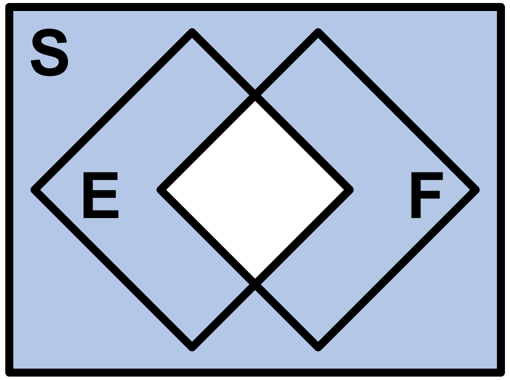
$$(E \text{ and } F)^C = E^C \text{ or } F^C$$



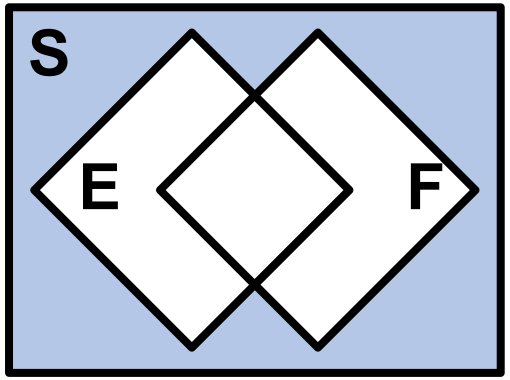
$$(E \text{ or } F)^C = E^C \text{ and } F^C$$

DeMorgan's Laws

These rules let you alternate between AND and OR.



$$(E \text{ and } F)^C = E^C \text{ or } F^C$$



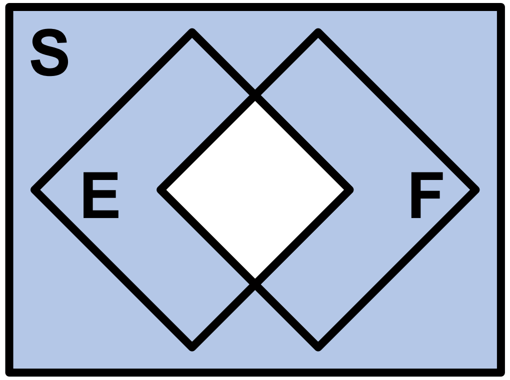
$$(E \text{ or } F)^C = E^C \text{ and } F^C$$

Augustin DeMorgan

Jason Alexander
From Seinfeld

DeMorgan's Laws

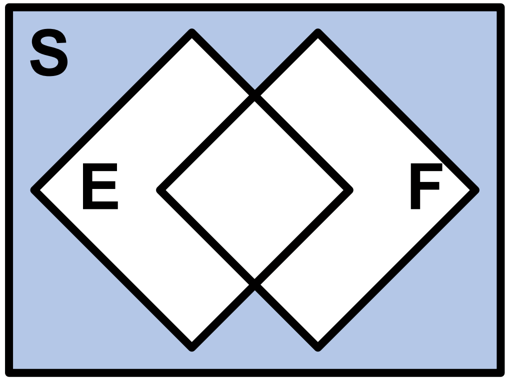
These rules let you alternate between AND and OR.



$$(E \text{ and } F)^C = E^C \text{ or } F^C$$

If events are *dependent*:

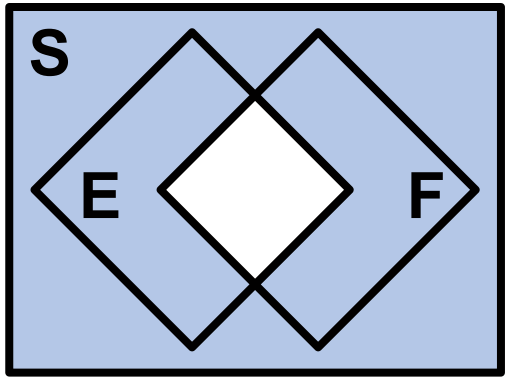
$$P(E_1 E_2 \cdots E_n) = 1 - P((E_1 E_2 \cdots E_n)^C)$$
$$= 1 - P(E_1^C \text{ or } E_2^C \text{ or } \cdots E_n^C)$$



$$(E \text{ or } F)^C = E^C \text{ and } F^C$$

DeMorgan's Laws

These rules let you alternate between AND and OR.

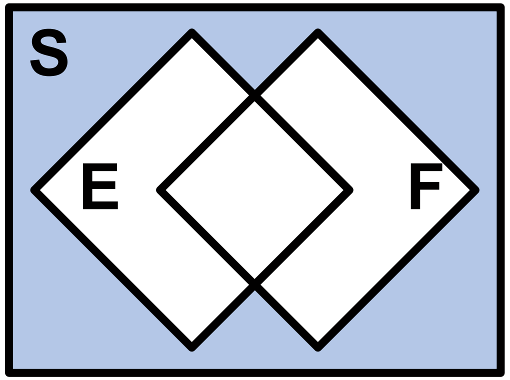


$$(E \text{ and } F)^C = E^C \text{ or } F^C$$

If events are *dependent*:

$$P(E_1 E_2 \cdots E_n) = 1 - P((E_1 E_2 \cdots E_n)^C)$$
$$= 1 - P(E_1^C \text{ or } E_2^C \text{ or } \cdots E_n^C)$$

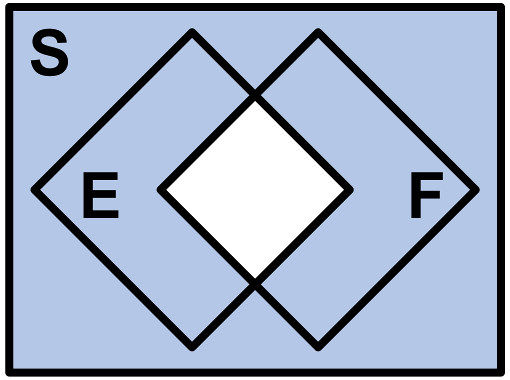
Great if E^C mutually exclusive!



$$(E \text{ or } F)^C = E^C \text{ and } F^C$$

DeMorgan's Laws

These rules let you alternate between AND and OR.

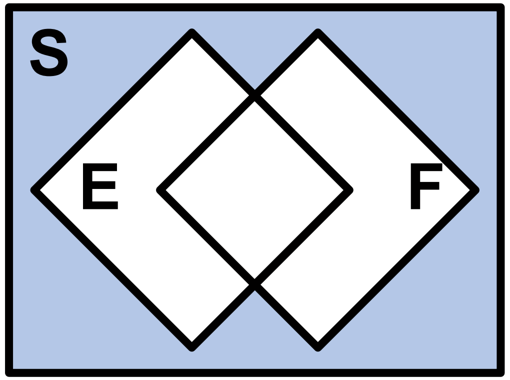


$$(E \text{ and } F)^C = E^C \text{ or } F^C$$

If events are *dependent*:

$$P(E_1 E_2 \cdots E_n) = 1 - P((E_1 E_2 \cdots E_n)^C)$$
$$= 1 - P(E_1^C \text{ or } E_2^C \text{ or } \cdots E_n^C)$$

Great if E^C mutually exclusive!



$$(E \text{ or } F)^C = E^C \text{ and } F^C$$

If events are *not* mutually exclusive:

$$P(E_1 \text{ or } E_2 \text{ or } \cdots E_n) = 1 - P((E_1 \text{ or } E_2 \text{ or } \cdots E_n)^C)$$
$$= 1 - P(E_1^C E_2^C \cdots E_n^C)$$

Great if E_is are independent!

The Core Probability Toolkit

The Law of Total Probability

$$P(E) = P(E \text{ and } F) + P(E \text{ and } F^{C})$$

$$P(E) = P(E|F)\,P(F) + P(E|F^{C})\,P(F^{C})$$

$$P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i) = \sum_{i=1}^{n} P(E|B_i)\,P(B_i)$$

Bayes' Theorem

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|B^{C}) \cdot P(B^{C})}$$

Definition of Conditional Probability

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$$

Axiom 1: $0 \leq P(E) \leq 1$

Axiom 2: $P(S) = 1$

$$P(E^{C}) = 1 - P(E)$$

De Morgan's Laws

$(\text{A or B})^{C} = \text{A}^{C} \text{ and } \text{B}^{C}$

$(\text{A and B})^{C} = \text{A}^{C} \text{ or } \text{B}^{C}$

Axiom 3: If E and F are mutually exclusive, then $P(E \text{ or } F) = P(E) + P(F)$

Otherwise, use Inclusion-Exclusion:

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

Chain Rule

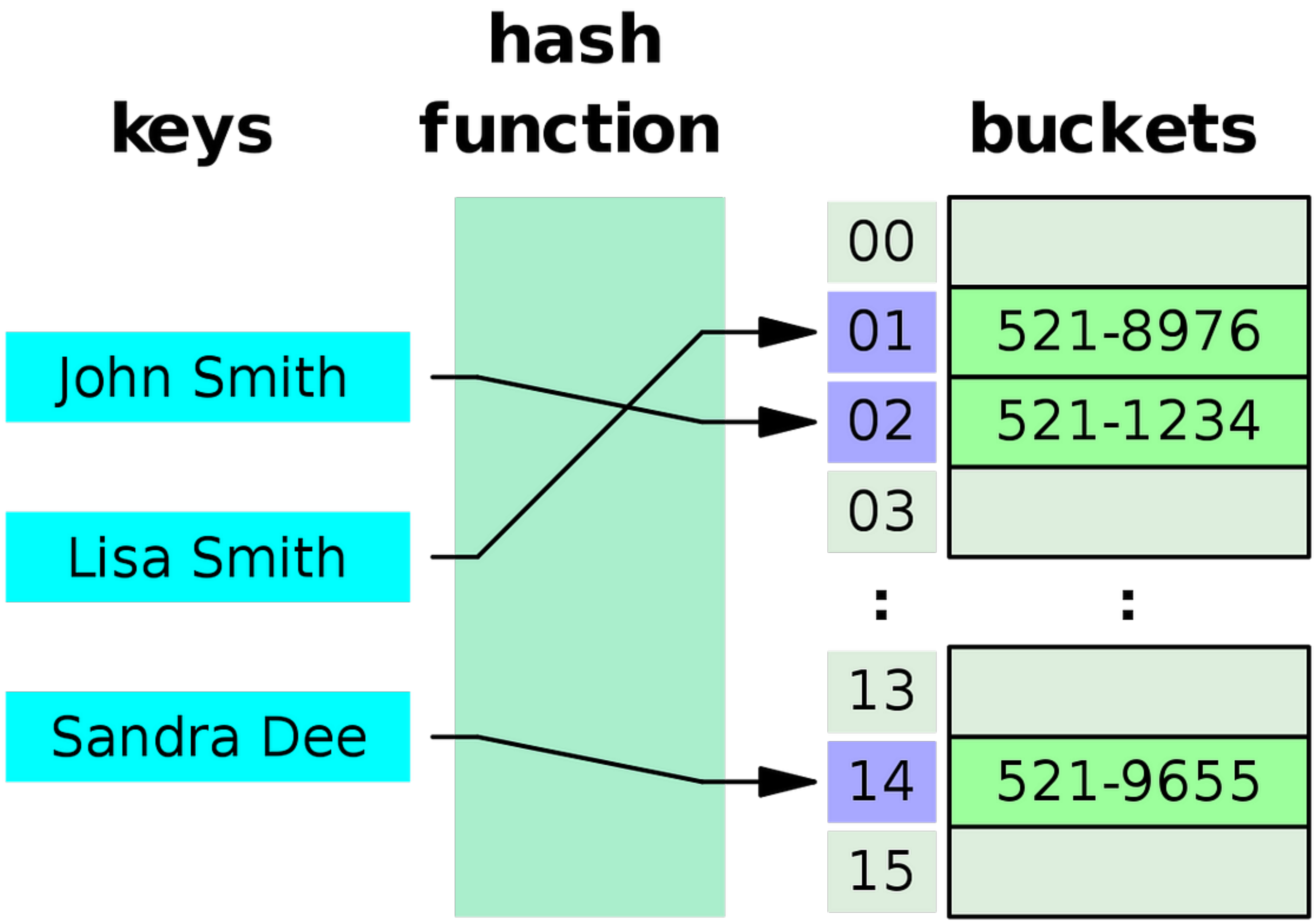
$$P(E \text{ and } F) = P(E|F) \cdot P(F) = P(F|E) \cdot P(E)$$

Independence

$$P(E|F) = P(E)$$

$$P(E \text{ and } F) = P(E)\,P(F)$$

Practice Time: Hash Tables



Hash table fun

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?

Hash table fun

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?

Let S_j be the event that string j hashes into bucket 1.

Hash table fun

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?

Let S_j be the event that string j hashes into bucket 1.

$$
\begin{aligned}
P(E) &= P(S_1 \text{ or } S_2 \text{ or } \dots \text{ or } S_n) \\
&= 1 - P(S_1^C \text{ and } S_2^C \text{ and } \dots \text{ and } S_n^C) \\
&= 1 - P(S_1^C) \cdot P(S_2^C) \cdot \dots \cdot P(S_n^C) \\
&= 1 - (1 - p_1) \cdot (1 - p_1) \cdot \dots \cdot (1 - p_1) \\
&= 1 - (1 - p_1)^n
\end{aligned}
$$

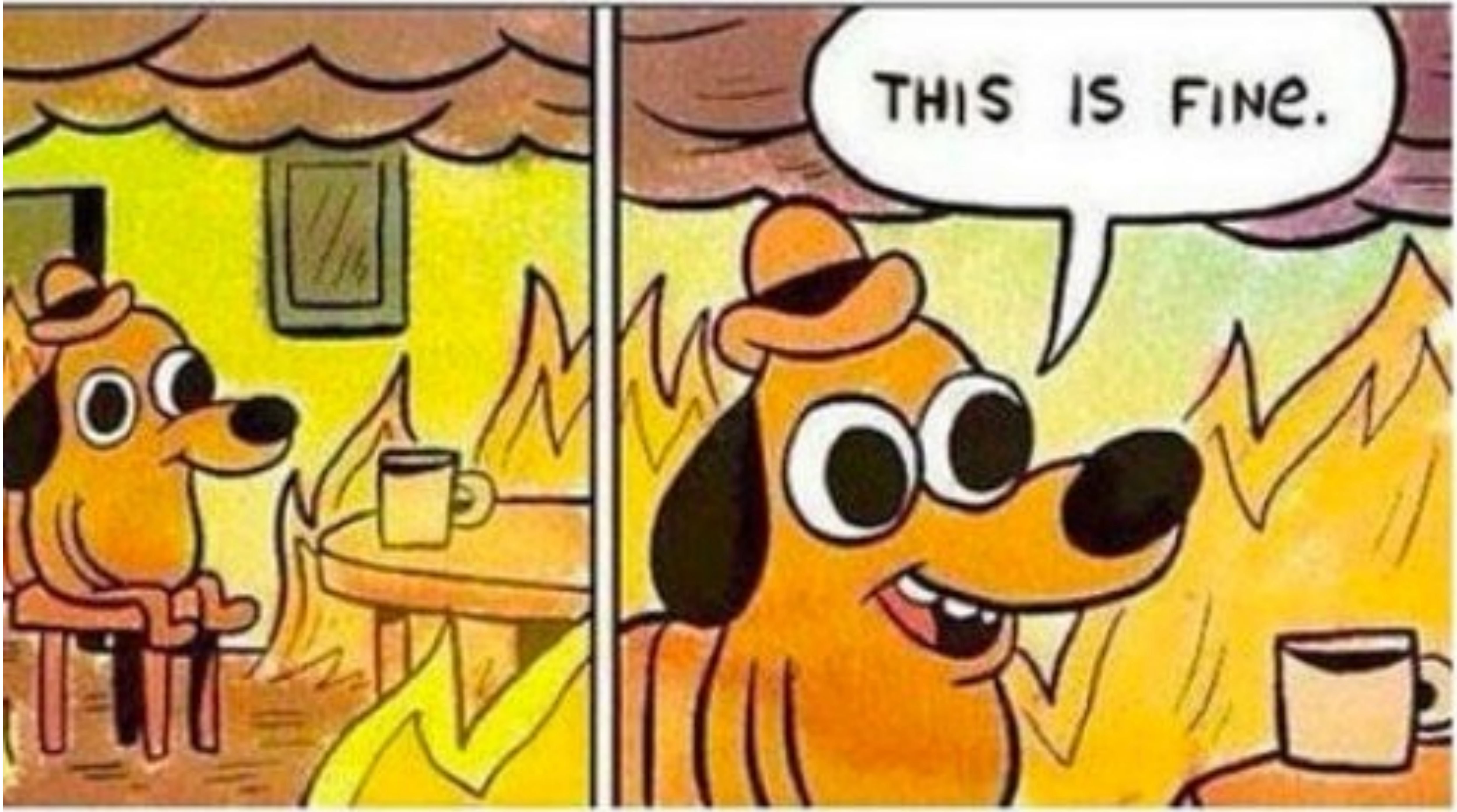
Hash table **fun**

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?
2. …**at least** 1 of buckets 1 to k has ≥ 1 string hashed into it?


THIS IS FINE.

Hash table fun

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?
2. …**at least** 1 of buckets 1 to k has ≥ 1 string hashed into it?

Let F_i be the event that bucket i has 1+ strings.

Hash table fun

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?
2. …**at least** 1 of buckets 1 to k has ≥ 1 string hashed into it?

$$P(E) = P(F_1 \text{ or } F_2 \text{ or } \dots \text{ or } F_k)$$

Let F_i be the event that bucket i has 1+ strings.

⚠ F_i bucket events are *not* mutually exclusive! So we can't just add.

Hash table fun

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?
2. …**at least** 1 of buckets 1 to k has ≥ 1 string hashed into it?

Let F_i be the event that bucket i has 1+ strings.

$$
\begin{aligned}
P(E) &= P(F_1 \text{ or } F_2 \text{ or } \ldots \text{ or } F_k) \\
&= 1 - P(F_1^C \text{ and } F_2^C \text{ and } \ldots \text{ and } F_k^C)
\end{aligned}
$$

Hash table **fun**

We hash n strings into a hash table with r buckets.

- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?
2. …**at least** 1 of buckets 1 to k has ≥ 1 string hashed into it?

Let F_i be the event that bucket i has 1+ strings.

$$
\begin{aligned}
P(E) &= P(F_1 \text{ or } F_2 \text{ or } \ldots \text{ or } F_k) \\
&= 1 - P(F_1^C \text{ and } F_2^C \text{ and } \ldots \text{ and } F_k^C)
\end{aligned}
$$

What is the probability of a string landing in none of the first k buckets?

Hash table fun

We hash n strings into a hash table with r buckets.

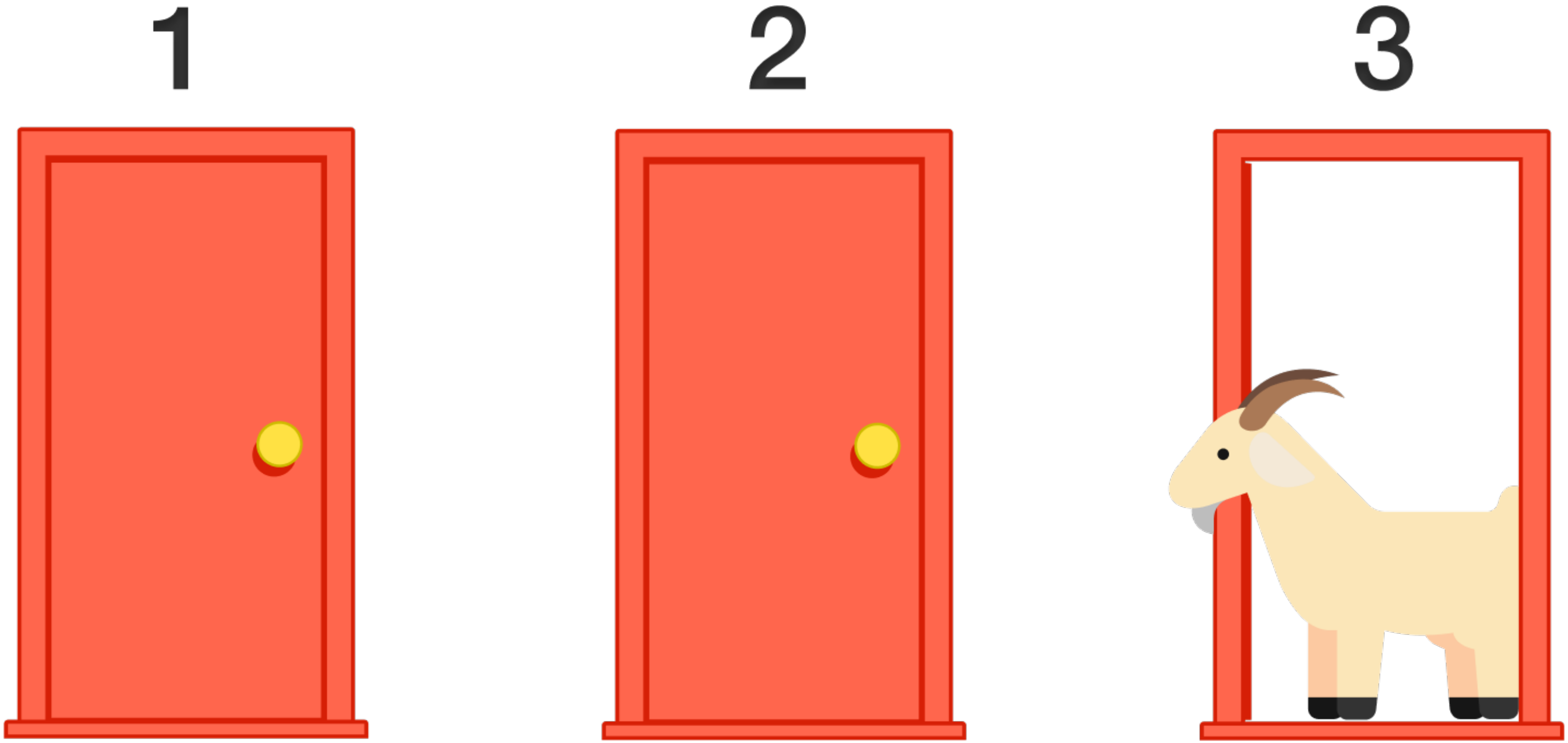
- Each string is hashed independently, with the same p_i of being hashed into bucket i.

What is the probability that…

1. …bucket 1 has ≥ 1 string hashed into it?
2. …**at least** 1 of buckets 1 to k has ≥ 1 string hashed into it?

Let F_i be the event that bucket i has 1+ strings.

$$
\begin{aligned}
P(E) &= P(F_1 \text{ or } F_2 \text{ or } \ldots \text{ or } F_k) \\
&= 1 - P(F_1^C \text{ and } F_2^C \text{ and } \ldots \text{ and } F_k^C) \\
&= 1 - (1 - p_1 - p_2 - \ldots - p_k)^n
\end{aligned}
$$



The Monty Hall Problem

The Monty Hall Problem

Behind one door is a prize (equally likely for each door).

Behind the other two doors are goats.

How to play:

1. We choose a door.
2. Host opens 1 of the other 2 doors, revealing a goat.
3. We are given an option to switch to the other door.

Note: If we don't switch, P(win) = 1/3

We are comparing P(win) vs. P(win|switch)

Should we switch?

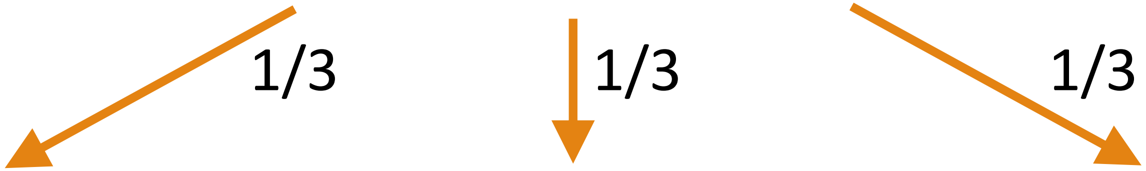
Marilyn Vos Savant: "You Should Always Switch"

Let's Find $P(\text{win} \mid \text{switch})$

Without loss of generality, let's pick door A (out of doors A,B,C).

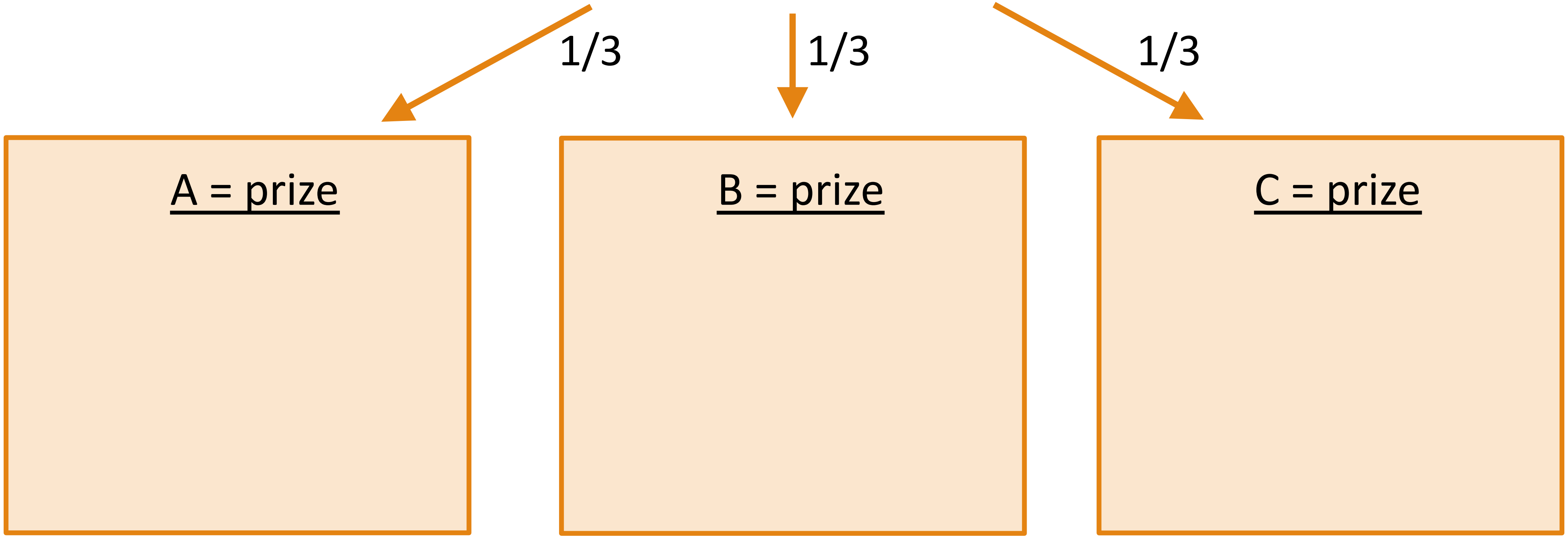
Let's Find $P(\text{win} \mid \text{switch})$

Without loss of generality, let's pick door A (out of doors A,B,C).

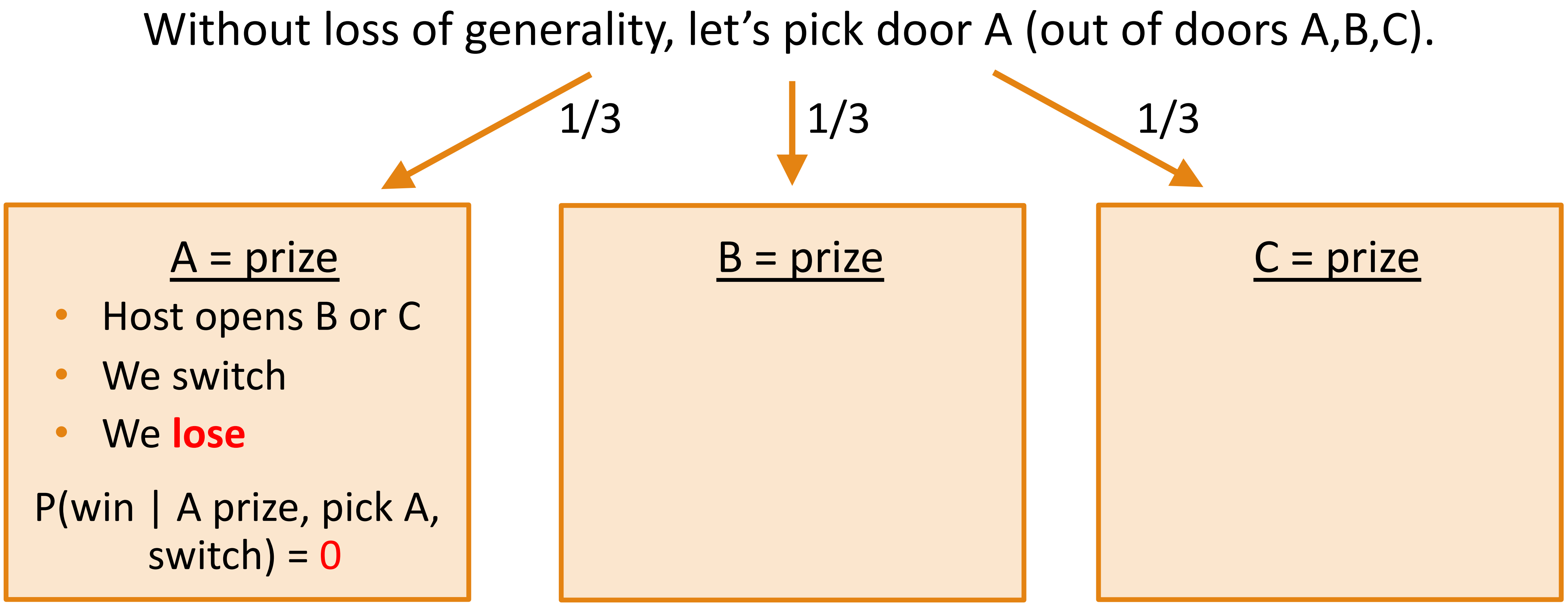


Let's Find $P(\text{win} \mid \text{switch})$

Without loss of generality, let's pick door A (out of doors A,B,C).



Let's Find P(win | switch)



Let's Find P(win | switch)

Without loss of generality, let's pick door A (out of doors A,B,C).

1/3 1/3 1/3

<u>A = prize</u>

- Host opens B or C
- We switch
- We **lose**

P(win | A prize, pick A, switch) = 0

<u>B = prize</u>

- Host must open C
- We switch to B
- We **win**

P(win | B prize, pick A, switch) = 1

<u>C = prize</u>

Let's Find P(win | switch)

Without loss of generality, let's pick door A (out of doors A,B,C).

1/3 → **A = prize**
- Host opens B or C
- We switch
- We lose

P(win | A prize, pick A, switch) = 0

1/3 → **B = prize**
- Host must open C
- We switch to B
- We win

P(win | B prize, pick A, switch) = 1

1/3 → **C = prize**
- Host must open B
- We switch to C
- We win

P(win | C prize, pick A, switch) = 1

Let's Find P(win | switch)

Without loss of generality, let's pick door A (out of doors A,B,C).

1/3 1/3 1/3

A = prize
- Host opens B or C
- We switch
- We **lose**

P(win | A prize, pick A, switch) = 0

B = prize
- Host must open C
- We switch to B
- We **win**

P(win | B prize, pick A, switch) = 1

C = prize
- Host must open B
- We switch to C
- We **win**

P(win | C prize, pick A, switch) = 1

P(win | pick A, switch) = P(win | A prize, pick A, switch) * P(A prize) +
P(win | B prize, pick A, switch) * P(B prize) +
P(win | C prize, pick A, switch) * P(C prize)

= 1/3 * 0 + 1/3 * 1 + 1/3 * 1 = 2/3

You should switch!

Next time:

Next time: X

Have a great 4th of July!