The Core Probability Toolkit
The Core Probability Toolkit

The Law of Total Probability
\[ P(E) = P(E \text{ and } F) + P(E \text{ and } F^C) \]

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

\[ P(E) = P(E|F) P(F) + P(E|F^C) P(F^C) = \sum_{i=1}^{n} P(E|B_i) P(B_i) \]

**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:** \( P(S) = 1 \)

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

Otherwise, use Inclusion-Exclusion:

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

Bayes' Theorem
\[ P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)} \]

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

De Morgan’s Laws
\[ (A \text{ or } B)^C = A^C \text{ and } B^C \]
\[ (A \text{ and } B)^C = A^C \text{ or } B^C \]

Independence
\[ P(E|F) = P(E) \]
\[ P(E \text{ and } F) = P(E) P(F) \]
The Core Probability Toolkit

**The Law of Total Probability**

\[
P(E) = P(E \text{ and } F) + P(E \text{ and } F^C) \quad P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i)
\]

\[
P(E) = P(E|F) P(F) + P(E|F^C) P(F^C) \quad = \sum_{i=1}^{n} P(E|B_i) P(B_i)
\]

**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:** \(P(S) = 1\)

**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**

\[
P(E \text{ and } F) = P(E|F) P(F) = P(F|E) P(E)
\]

**Bayes' Theorem**

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

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

**De Morgan’s Laws**

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

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

**Independence**

\[
P(E|F) = P(E)
\]

\[
P(E \text{ and } F) = P(E) P(F)
\]
Learning Goal for Today: Conditional Probability

\[ P(E \text{ and } F) \]

- Definition of Conditional Probability
- Chain Rule

\[ P(E|F) \]

- Law of Total Probability
- Bayes' Theorem

\[ P(E) \]
\[ P(F|E) \]

Piech & Cain, CS109, Stanford University
Learning Goal for Today: Conditional Probability

\[ P(E \text{ and } F) \]

Definition of Conditional Probability \[ P(E|F) \]

Law of Total Probability

Chain Rule

Bayes’ Theorem

\[ P(E) \]

\[ P(F|E) \]

Piech & Cain, CS109, Stanford University
Review
The Axioms of Probability

• Axiom 1: \(0 \leq P(E) \leq 1\)  
  All probabilities are between 0 and 1

• Axiom 2: \(P(S) = 1\)
  The probability of the sample space is 1

• Axiom 3: If events \(E\) and \(F\) are mutually exclusive,
  \[P(E \cup F) = P(E) + P(F)\]
  Probability of event \(E\) *or* event \(F\)
Axiom 3: Probability of Or, With Mutually Exclusive Events

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

\[ P(E \cup F) = P(E) + P(F) \]
Axiom 3: Probability of Or, With Mutually Exclusive Events

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

\[ P(E \cup F) = \frac{7}{50} + \frac{4}{50} = \frac{11}{50} \]
Axiom 3: Probability of Or, With Mutually Exclusive Events

What is $P(E \cap F)$ in this example?
Axiom 3: Probability of Or, With Mutually Exclusive Events

What is $P(E \cap F)$ in this example?

0! Since E and F are mutually exclusive.
Probability of Or With 3 Or More Events

\[ X_1 \quad X_2 \]

\[ X_3 \]
Probability of Or With 3 Or More Events

\[
P(X_1 \cup X_2 \cup \cdots \cup X_n) = \sum_{i=1}^{n} P(X_i)
\]

Wahoo! All my events are mutually exclusive
The Axioms of Probability

• Axiom 1: $0 \leq P(E) \leq 1$  
  All probabilities are between 0 and 1

• Axiom 2: $P(S) = 1$  
  The probability of the sample space is 1

• Axiom 3: If events $E$ and $F$ are mutually exclusive,

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

  Probability of event $E$ *or* event $F$

• Identity 3*: $P(E^c) = 1 - P(E)$  
  Events either happen...or don't

  "not E"

Piech & Cain, CS109, Stanford University
Where does $P(E^c) = 1 - P(E)$ come from?
Where does $P(E^c) = 1 - P(E)$ come from?

\[ P(E \cup E^c) = P(E) + P(E^c) \]

Axiom 3: $E$ and $E^c$ are mutually exclusive
Where does $P(E^c) = 1 - P(E)$ come from?

\[ P(E \cup E^c) = P(E) + P(E^c) \]

Axiom 3: $E$ and $E^c$ are mutually exclusive

\[ P(S) = P(E) + P(E^c) \]

Since all outcomes are either in $E$ or $E^c$
Where does $P(E^c) = 1 - P(E)$ come from?

\[
P(E \cup E^c) = P(E) + P(E^c)
\]

Axiom 3: $E$ and $E^c$ are mutually exclusive

\[
P(S) = P(E) + P(E^c)
\]

Since all outcomes are either in $E$ or $E^c$

\[
1 = P(E) + P(E^c)
\]

Axiom 2: $P(S) = 1$

\[
P(E^c) = 1 - P(E)
\]

Subtract $P(E)$ from both sides
Chip Defect Detection

Your company has manufactured $n$ chips, 1 of which is defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that the defective chip is in the $k$ selected chips?

$$|S| =$$

$$|E| =$$
Chip Defect Detection

Your company has manufactured $n$ chips, 1 of which is defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that the defective chip is in the $k$ selected chips?

$$|S| = \binom{n}{k}$$

$$|E| = \text{Probability}$$
Chip Defect Detection

Your company has manufactured $n$ chips, 1 of which is defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that the defective chip is in the $k$ selected chips?

$$|S| = \binom{n}{k}$$

$$|E| = \binom{1}{1} \binom{n-1}{k-1} \quad \text{Choose the defective chip, then choose } k - 1 \text{ other chips}$$
Chip Defect Detection

Your company has manufactured $n$ chips, 1 of which is defective. $k$ chips are randomly selected from $n$ for testing.

What is the probability that the defective chip is in the $k$ selected chips?

$$|S| = \binom{n}{k}$$

$$|E| = \left(\frac{1}{1}\right) \left(\binom{n-1}{k-1}\right)$$

Choose the defective chip, then choose $k-1$ other chips

$$P(\text{defective chip is in } k \text{ selected chips}) = \frac{\left(\frac{1}{1}\right) \left(\binom{n-1}{k-1}\right)}{\binom{n}{k}} = \frac{(n-1)!}{(k-1)!(n-k)!} = \frac{k}{n}$$
Chip Defect Detection

Your company has manufactured $n$ chips, five of which are defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that any defective chips are in the $k$ selected chips?
Chip Defect Detection

Your company has manufactured $n$ chips, five of which are defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that any defective chips are in the $k$ selected chips?

Tip: sometimes, it is easier to calculate $P(E^c)$ than $P(E)$. 
Your company has manufactured $n$ chips, five of which are defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that any defective chips are in the $k$ selected chips?

Let’s find $P$(no defective chips chosen) instead:

$$|S| =$$

$$|E| =$$
Chip Defect Detection

Your company has manufactured $n$ chips, five of which are defective. $k$ chips are randomly selected from $n$ for testing. What is the probability that any defective chips are in the $k$ selected chips?

Let’s find $P$(no defective chips chosen) instead:

$$|S| = \binom{n}{k}$$

$$|E| = \binom{n-5}{k}$$

Then,

$$P(\text{any defective}) = 1 - P(\text{no defective}) = 1 - \frac{\binom{n-5}{k}}{\binom{n}{k}}$$
End Review
Conditional Probability
## Conditional Probability: Why?

So far, we’ve only learned how to calculate probability in a global, static way.

<table>
<thead>
<tr>
<th>Day</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rainy</td>
</tr>
<tr>
<td>2</td>
<td>Sunny</td>
</tr>
<tr>
<td>3</td>
<td>Rainy</td>
</tr>
<tr>
<td>4</td>
<td>Cloudy</td>
</tr>
<tr>
<td>5</td>
<td>Rainy</td>
</tr>
<tr>
<td>6</td>
<td>Sunny</td>
</tr>
<tr>
<td>7</td>
<td>Sunny</td>
</tr>
<tr>
<td>8</td>
<td>Sunny</td>
</tr>
<tr>
<td>10000</td>
<td>Cloudy</td>
</tr>
</tbody>
</table>

Let $E$ be the event that it is **Sunny** tomorrow. What is $P(E)$?
Conditional Probability: Why?

So far, we’ve only learned how to calculate probability in a global, static way.

Let $E$ be the event that it is Sunny tomorrow. What is $P(E)$?

- Weather depends on seasons...shouldn’t we only use data from summers to get a probability for tomorrow?
Conditional Probability: Why?

So far, we’ve only learned how to calculate probability in a global, static way.

<table>
<thead>
<tr>
<th>Day</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rainy</td>
</tr>
<tr>
<td>2</td>
<td>Sunny</td>
</tr>
<tr>
<td>3</td>
<td>Rainy</td>
</tr>
<tr>
<td>4</td>
<td>Cloudy</td>
</tr>
<tr>
<td>5</td>
<td>Rainy</td>
</tr>
<tr>
<td>6</td>
<td>Sunny</td>
</tr>
<tr>
<td>7</td>
<td>Sunny</td>
</tr>
<tr>
<td>8</td>
<td>Sunny</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>10000</td>
<td>Cloudy</td>
</tr>
</tbody>
</table>

Let $E$ be the event that it is **Sunny** tomorrow. What is $P(E)$?

- Weather depends on seasons...shouldn’t we only use data from summers to get a probability for tomorrow?
- Weather tomorrow depends on weather today...how can we incorporate today’s weather into our math?
Conditional Probability: Why?

So far, we’ve only learned how to calculate probability in a global, static way. Let $E$ be the event that it is Sunny tomorrow. What is $P(E)$?

- Weather depends on seasons...shouldn’t we only use data from summers to get a probability for tomorrow?
- Weather tomorrow depends on weather today...how can we incorporate today’s weather into our math?
- Generally: how do we calculate probabilities of events under specific conditions, or given specific information?
The conditional probability of $E$ given $F$ is the probability that $E$ occurs, given that $F$ occurs. This is called “conditioning on $F$”.

Written as: $P(E|F)$

Means: “$P(E$, given $F$ is observed)”
The *conditional probability* of $E$ given $F$ is the probability that $E$ occurs, *given that $F$ occurs*. This is called “conditioning on $F$”. 

Written as: $P(E|F)$

Means: “$P(E$, given $F$ is observed)”

Examples:

- $P($sunny $|\ $summer$) = “Probability it’s sunny tomorrow, given that it’s summer”$
- $P($sunny $|\ $rain\ today) = “Probability it’s sunny tomorrow, given that it rained today”
Returning to Rolling Two Dice

You roll two 6-sided dice.

Let $E$ be the event that the dice sum to 7. What is $P(E)$?

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

\[
P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.166666667
\]
You roll two 6-sided dice.

Let $E$ be the event that the dice sum to 4. What is $P(E)$?

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

$$P(E) = \frac{|E|}{|S|} = \frac{3}{36} = \frac{1}{12}$$
You roll two 6-sided dice.

Let $E$ be the event that the dice sum to 4. What is $P(E)$?

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

$$P(E) = \frac{|E|}{|S|} = \frac{3}{36} = \frac{1}{12}$$

Note: if we roll a 4, 5, or 6 on the first dice, getting a sum of 4 becomes impossible.
Stacking The Odds In Our Favor

What outcome for $D_1$ is most likely to lead to a sum of 4?

A. 1 and 3 tie for best
B. 1, 2, and 3 tie for best
C. 2 is the best
D. None of the above
What outcome for $D_1$ is most likely to lead to a sum of 4?

A. 1 and 3 tie for best
B. 1, 2, and 3 tie for best
C. 2 is the best
D. None of the above
Returning to Rolling Two Dice

You roll two 6-sided dice.

Let $F$ be the event that $D_1 = 2$.

Let $E$ be the event that the dice sum to 4. What is $P(E, \text{ given } F)$?

$S = \{ [1,1], [1,2], [1,3], [1,4], [1,5], [1,6], [2,1], [2,2], [2,3], [2,4], [2,5], [2,6], [3,1], [3,2], [3,3], [3,4], [3,5], [3,6], [4,1], [4,2], [4,3], [4,4], [4,5], [4,6], [5,1], [5,2], [5,3], [5,4], [5,5], [5,6], [6,1], [6,2], [6,3], [6,4], [6,5], [6,6] \}$
Returning to Rolling Two Dice

You roll two 6-sided dice.

Let $F$ be the event that $D_1 = 2$.

Let $E$ be the event that the dice sum to 4. What is $P(E, \text{ given } F)$?

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

These are the only possible outcomes now...
Returning to Rolling Two Dice

You roll two 6-sided dice.

Let $F$ be the event that $D_1 = 2$.

Let $E$ be the event that the dice sum to 4. What is $P(E, \text{ given } F)$?


These are the only possible outcomes now....

...so our sample space shrinks!
Returning to Rolling Two Dice

You roll two 6-sided dice.

Let $F$ be the event that $D_1 = 2$.

Let $E$ be the event that the dice sum to 4. What is $P(E, \text{ given } F)$?

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

These are the only possible outcomes now....

$\ldots$so our sample space shrinks!

$$P(E|F) = \frac{|E| \text{ (given } F)}{|S| \text{ (given } F)} = \frac{|E \text{ and } F|}{|F|} = \frac{1}{6}$$
Conditional Probability: Visual Intuition

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

$$P(E) = \frac{8}{50} \approx 0.16$$

$$P(E|F) = ?$$
Conditional Probability: Visual Intuition

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

\[
P(E) = \frac{8}{50} \approx 0.16
\]

\[
P(E|F) = ?
\]
Conditional Probability: Visual Intuition

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

$$P(E) = \frac{8}{50} \approx 0.16$$

$$P(E|F) = \frac{3}{14} \approx 0.21$$

Piech & Cain, CS109, Stanford University
Conditional Probability: Visual Intuition

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

$$\Pr(E|F) = \frac{\# \text{ of outcomes in } E \text{ consistent with } F}{\# \text{ of outcomes in } S \text{ consistent with } F} = \frac{|EF|}{|SF|} = \frac{|EF|}{|F|}$$

Shorthand notation for set intersection (aka set “and”)

Piech & Cain, CS109, Stanford University
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 $F$”.

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

- This definition works even when outcomes are not equally likely
- What if $P(F) = 0$?
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
- What if $P(F) = 0$?? Undefined: *you observed the impossible!*
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
- What if $P(F) = 0$?? Undefined: *you observed the impossible!*

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

\[
P(E \cap F) = P(E|F)P(F)
\]
goodbye, land of equally likely outcomes
Practice: Multiple Choice & Probability

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

- Students can get questions correct by knowing the answer, or just by guessing.
- How can you know if a student really knows their stuff, or if they’re just guessing?
Imagine a multiple choice test, where every question has 4 answer choices.

- Students can get questions correct by knowing the answer, or just by guessing.
- How can you know if a student really knows their stuff, or if they’re just guessing?

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.

1. What is the probability that a student gets a question right, given that they guess?
Imagine a multiple choice test, where every question has 4 answer choices.

- Students can get questions correct by knowing the answer, or just by guessing.
- How can you know if a student really knows their stuff, or if they’re just guessing?

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.

$$P(R|G) = \frac{1}{4}$$

1. What is the probability that a student gets a question right, given that they guess?
Imagine a multiple choice test, where every question has 4 answer choices.

- Students can get questions correct by knowing the answer, or just by guessing.
- How can you know if a student really knows their stuff, or if they’re just guessing?

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.

$$P(R|G) = \frac{1}{4}$$

1. What is the probability that a student gets a question right, given that they guess?
2. What is the probability that a student guesses on a question and gets it right?
Imagine a multiple choice test, where every question has 4 answer choices.

- Students can get questions correct by knowing the answer, or just by guessing.
- How can you know if a student really knows their stuff, or if they're just guessing?

Let $G$ be the event that a student guesses an answer to a question. $P(G) = \frac{1}{5}$.

Let $R$ be the event that the student gets the answer right.

$$P(R|G) = \frac{1}{4}$$

1. What is the probability that a student gets a question right, given that they guess?

2. What is the probability that a student guesses on a question and gets it right?

$$P(G \cap R) = P(R|G)P(G) = \frac{1}{4} \cdot \frac{1}{5} = \frac{1}{20}$$
Common Confusion With Conditional Probability

\[ P(E|F) \neq P(E \cap F) \]

What is the difference between “E, given F”, vs. “E and F”? 
Common Confusion With Conditional Probability

\[ P(E|F) \neq P(E \cap F) \]

What is the difference between “E, given F”, vs. “E and F”?

- Imagine: E is the event that it is hot outside.
Common Confusion With Conditional Probability

\[ P(E|F) \neq P(E \cap F) \]

What is the difference between “E, given F”, vs. “E and F”?

- Imagine: \( E \) is the event that it is hot outside.
- \( F \) is the event that the sun explodes into a massive supernova that engulfs the Earth in flames!!!
- If the Earth is engulfed in flames, it’s definitely hot, so \( P(E|F) \) is near 1.
- But because \( F \) is so unlikely, \( P(F) \) is almost zero, and \( P(E \text{ and } F) \) is near zero too.
Common Confusion With Conditional Probability

\[ P(E|F) \neq P(F|E) \]

What is the difference between “E, given F”, vs. “F, given E”? 
Common Confusion With Conditional Probability

\[ P(E|F) \neq P(F|E) \]

What is the difference between “E, given F”, vs. “F, given E”? 

- If the Earth is engulfed in flames, it’s definitely hot, so \( P(E|F) \) is near 1.
- If it’s hot outside, do we usually believe that a supernova must have happened? No! There are lots of other reasons it could be hot outside. So \( P(F|E) \) is near zero too.

Tip: imagining extremes can help you build intuition!
NETFLIX and Learn
Netflix: Watch Probabilities

What is the probability that a user will watch The Great British Bake-off?
Netflix: Watch Probabilities

What is the probability that a user will watch The Great British Bake-off?

\[
P(E) = \lim_{n \to \infty} \frac{n(E)}{n} \approx \frac{\#\text{people who watched movie}}{\#\text{people on Netflix}}
\]

\[
P(E) \approx \frac{50,234,231}{269,613,547} \approx 0.18
\]
Netflix & Conditional Probability

Let $E$ be the event that a user watches The Great British Bake-off.

Let $F$ be the event that a user watches Chopped.

What is the probability that a user watches The Great British Bake-off, given they watched Chopped?

$P(E|F) = \frac{P(E \cap F)}{P(F)}$
Netflix & Conditional Probability

Let $E$ be the event that a user watches The Great British Bake-off.

Let $F$ be the event that a user watches Chopped.

What is the probability that a user watches The Great British Bake-off, given they watched Chopped?

\[
P(E|F) = \frac{P(E \cap F)}{P(F)} \approx \frac{\text{# people who have watched both}}{\text{# people on Netflix}} = \frac{\text{# people who have watched Chopped}}{\text{# people on Netflix}}
\]
Netflix & Conditional Probability

Let $E$ be the event that a user watches The Great British Bake-off.
Let $F$ be the event that a user watches Chopped.

What is the probability that a user watches The Great British Bake-off, given they watched Chopped?

\[
P(E|F) = \frac{P(E \cap F)}{P(F)} \approx \frac{\text{# people who have watched both}}{\text{# people on Netflix}} \div \frac{\text{# people who have watched Chopped}}{\text{# people on Netflix}}
\]

\[
\approx \frac{\# \text{ people who have watched both}}{\# \text{ people who have watched Chopped}} \approx 0.42
\]
Netflix & Conditional Probability

Let $E$ be the event that a user watches The Great British Bake-off.

Let $F$ be the event that a user watches each of the shows/movies below.

\[
P(E|F) = \frac{P(E \cap F)}{P(F)} \quad \text{Definition of Cond. Probability}
\]
Netflix & Conditional Probability

Let $E$ be the event that a user watches The Great British Bake-off.

Let $F$ be the event that a user watches each of the shows/movies below.

Let $\mathbb{P}(E|F)$ be the probability of event $E$ given that event $F$ has occurred.

Let $\mathbb{P}(F)$ be the probability of event $F$.

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

**Definition of Cond. Probability**

**What does it mean if $\mathbb{P}(E|F) > \mathbb{P}(E)$?** How could Netflix use this for recommendations?
# A Guide To Notation

<table>
<thead>
<tr>
<th>And</th>
<th>Or</th>
<th>Given</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(E \text{ and } F') )</td>
<td>( P(E \text{ or } F') )</td>
<td>( P(E</td>
</tr>
<tr>
<td>( P(E \cap F) )</td>
<td>( P(E \cup F) )</td>
<td>( P(E</td>
</tr>
<tr>
<td>( P(E,F) )</td>
<td></td>
<td>Probability of ( E ) given ( F ) and ( G )</td>
</tr>
<tr>
<td>( P(EF) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Generalized Chain Rule

For just two events, the **chain rule** looks like this:

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

For more than two events:

$$\Pr(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } \ldots E_n)$$

$$= \Pr(E_1) \cdot \Pr(E_2|E_1) \cdot \Pr(E_3|E_1, E_2) \cdot \ldots \cdot \Pr(E_n|E_1, E_2 \ldots E_{n-1})$$

But often, we can simplify this equation a lot...we’ll see how next lecture!
Law of Total Probability
Learning Goal for Today: Conditional Probability

\[ P(E \text{ and } F) \]

Definition of Conditional Probability \[ P(E|F) \] Chain Rule

Law of Total Probability

Bayes' Theorem

\[ P(E) \]

\[ P(F|E) \]

Piech & Cain, CS109, Stanford University
Learning Goal for Today: Conditional Probability

\[ P(E \text{ and } F) \]

- Definition of Conditional Probability
- Chain Rule

\[ P(E | F) \]

- Law of Total Probability
- Bayes' Theorem

\[ P(E) \quad P(F | E) \]

Piech & Cain, CS109, Stanford University
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 is the probability that a student gets an answer right?
Law of Total Probability

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

\[
P(E) = P(EF) + P(EF^C)
\]

\[
= P(E|F)P(F) + P(E|F^C)P(F^C)
\]
Law of Total Probability

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

$$P(E) = P(EF) + P(EF^C)$$

$$= P(E|F)P(F) + P(E|F^C)P(F^C)$$
Law of Total Probability

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

\[ P(E) = P(EF) + P(EF^C) = P(E|F)P(F) + P(E|F^C)P(F^C) \]
Generalized Law of Total Probability

Let $F_1, F_2, \ldots, F_n$ be mutually exclusive events that cover the sample space.

$$P(E) = \sum_{i=1}^{n} P(E \cap F_i)$$

$$= \sum_{i=1}^{n} P(E|F_i)P(F_i)$$
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 is the probability that a student gets an answer right?
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 is the probability that a student gets an answer right?

$$P(R) = P(R \mid G)P(G) + P(R \mid G^C)P(G^C)$$
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 is the probability that a student gets an answer right?

\[
P(R) = P(R\mid G)P(G) + P(R\mid G^c)P(G^c)
= \frac{1}{4} \cdot \frac{1}{5} + \frac{9}{10} \left(1 - \frac{1}{5}\right) = 0.77
\]
You have bacteria in your gut which is causing a disease. 10% of bacteria have a mutation making them resistant to antibiotics. You take half a course of antibiotics...

Probability a bacteria survives, given it has the mutation: 20%
Probability a bacteria survives, given it doesn't have the mutation: 1%

What is the probability that a randomly chosen bacteria survives?

Tip: start problems like this by defining events.
You have bacteria in your gut which is causing a disease. 10% of bacteria have a mutation making them resistant to antibiotics. You take half a course of antibiotics...

Probability a bacteria survives, given it has the mutation: 20%
Probability a bacteria survives, given it doesn't have the mutation: 1%

What is the probability that a randomly chosen bacteria survives?

Let $E$ be the event that a bacteria survives.
Let $M$ be the event that a bacteria has the mutation.
You have bacteria in your gut which is causing a disease.
10% of bacteria have a mutation making them resistant to antibiotics.
You take half a course of antibiotics...

Probability a bacteria survives, given it has the mutation: 20%
Probability a bacteria survives, given it doesn't have the mutation: 1%
What is the probability that a randomly chosen bacteria survives?

Let $E$ be the event that a bacteria survives.
Let $M$ be the event that a bacteria has the mutation.

$$
\Pr(E) = \Pr(E|M)\Pr(M) + \Pr(E|\neg M)\Pr(\neg M) \\
= 0.20 \cdot 0.10 + 0.01 \cdot 0.90 \\
= 0.029
$$
Real question. What is the probability that a surviving bacteria has the mutation?

\[ P(\text{Mutation} \mid \text{Survives}) \]

\[ P(M \mid E) \]
Real Question: $P(M | E)$?

You have bacteria in your gut which is causing a disease. 10% of bacteria have a mutation making them resistant to antibiotics. You take half a course of antibiotics...

$P(E | M) = 0.20$
$P(E | M^c) = 0.01$

What is the probability that a surviving bacteria has the mutation?
Learning Goal for Today: Conditional Probability

\[ P(E \text{ and } F) \]

- Definition of Conditional Probability
- Chain Rule

\[ P(E|F) \]

- Law of Total Probability
- Bayes' Theorem

\[ P(E) \]

\[ P(F|E) \]

Piech & Cain, CS109, Stanford University
Learning Goal for Today: Conditional Probability

\[ P(E \text{ and } F) \]

Definition of Conditional Probability

Law of Total Probability

\[ P(E|F) \]

Chain Rule

\[ P(E) \]

Bayes' Theorem

\[ P(F|E) \]

Piech & Cain, CS109, Stanford University
It’s Time For Bayes’ Theorem
Thomas Bayes

Rev. Thomas Bayes (~1701-1761):
British mathematician and Presbyterian minister
Thomas Bayes

Rev. Thomas Bayes (~1701-1761):
British mathematician and Presbyterian minister

He looked remarkably similar to Sean Astin
I want to calculate $P(\text{State of the world } F \mid \text{Observation } E)$. It seems so tricky!…

The other way around is easy $P(\text{Observation } E \mid \text{State of the world } F)$. What options do I have, chief?

Thomas Bayes
Definition of Conditional Probability

\[ P(F|E) = \frac{P(EF)}{P(E)} \]

Thomas Bayes

\[ P(E|F) \rightarrow P(F|E) \]
Thomas Bayes

\[ P(F|E) = \frac{P(EF)}{P(E)} \]

**Definition of Conditional Probability**

\[ P(E|F) \quad \rightarrow \quad P(F|E) \]

\[ = \frac{P(E|F)P(F)}{P(E)} \]

**The Chain Rule is symmetric**

A little while later…
\[ P(F|E) = \frac{P(EF)}{P(E)} \]

**Definition of Conditional Probability**

A little while later...

\[ = \frac{P(E|F)P(F)}{P(E)} \]

**The Chain Rule is symmetric**

A little while later...

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

**Law of Total Probability**

Piech & Cain, CS109, Stanford University
Bayes’ Theorem

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

For any events $E$ and $F$ where $P(E) > 0$ and $P(F) > 0$, the posterior probability $P(F|E)$ is given by:

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

**Expanded form:**

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

Piech & Cain, CS109, Stanford University
You have bacteria in your gut which is causing a disease. 10% of bacteria have a mutation making them resistant to antibiotics. You take half a course of antibiotics...

\[
P(E \mid M) = 0.20 \\
P(E \mid M^c) = 0.01
\]

What is the probability that a surviving bacteria \textit{has the mutation}?
Real Question: $P(M \mid E)$?

You have bacteria in your gut which is causing a disease. 10% of bacteria have a mutation making them resistant to antibiotics. You take half a course of antibiotics...

$P(E \mid M) = 0.20$
$P(E \mid M^c) = 0.01$

What is the probability that a surviving bacteria has the mutation?

$$P(M \mid E) = \frac{P(E \mid M) \cdot P(M)}{P(E \mid M) \cdot P(M) + P(E \mid M^c) \cdot P(M^c)}$$
Real Question: $P(M | E)$?

You have bacteria in your gut which is causing a disease. 10% of bacteria have a mutation making them resistant to antibiotics. You take half a course of antibiotics...

$P(E | M) = 0.20$
$P(E | M^c) = 0.01$

What is the probability that a surviving bacteria has the mutation?

$$P(M | E) = \frac{P(E | M) P(M)}{P(E | M) P(M) + P(E | M^c) P(M^c)}$$

$$P(M | E) = \frac{0.2 * 0.1}{0.2 * 0.1 + 0.01 * 0.9} \approx 0.69$$
Zika Virus Testing

A test is 98% effective at detecting Zika.

- However, the test has a “false positive” rate of 1%.
- 0.5% of US population has Zika.
- Let $E = \text{you test positive for Zika with this test.}$
- Let $F = \text{you actually have Zika.}$
- What is $P(F \mid E)$?
Zika Virus Testing

A test is 98% effective at detecting Zika.

- However, the test has a “false positive” rate of 1%.
- 0.5% of US population has Zika.
- Let $E =$ you test positive for Zika with this test.
- Let $F =$ you actually have Zika.
- What is $P(F \mid E)$?

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)}
\]
Zika Virus Testing

A test is 98% effective at detecting Zika.

- However, the test has a “false positive” rate of 1%.
- 0.5% of US population has Zika.
- Let E = you test positive for Zika with this test.
- Let F = you actually have Zika.
- What is P(F | E)?

\[
P(F | E) = \frac{P(E | F) P(F)}{P(E | F) P(F) + P(E | F^c) P(F^c)}
\]

\[
P(F | E) = \frac{(0.98)(0.005)}{(0.98)(0.005) + (0.01)(1 - 0.005)} \approx 0.33
\]
Intuition Time
Bayes Theorem Intuition
Bayes Theorem Intuition

People who test positive

All People

Piech & Cain, CS109, Stanford University
Bayes Thorem Intuition

All People

People with Zika

People who test positive

All People

Piech & Cain, CS109, Stanford University
Bayes Theorem Intuition

Conditioning on a positive result changes the sample space to this:

People who test positive

People who test positive and have Zika

\[ \approx 0.330 \]
Bayes Thoren Intuition

Say we have 1000 people:

5 have Zika and test positive. 985 do not have Zika and test negative.
10 do not have Zika and test positive.
Bayes Theorem Intuition

Conditioned on just those that test positive:

Notice that all the people with Zika are here, but the group is still mainly folks without Zika.

5 have Zika and test positive. 985 do not have Zika and test negative. 10 do not have Zika and test positive.
Why it is still good to get tested

- Let $E^c$ = you test negative for Zika with this test.
- Let $F$ = you actually have Zika.
- What is $P(F \mid E^c)$?

<table>
<thead>
<tr>
<th></th>
<th>Zika +</th>
<th>Zika –</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>$0.98 = P(E \mid F)$</td>
<td>$0.01 = P(E \mid F^c)$</td>
</tr>
<tr>
<td>Test –</td>
<td>$0.02 = P(E^c \mid F)$</td>
<td>$0.99 = P(E^c \mid F^c)$</td>
</tr>
</tbody>
</table>
Why it is still good to get tested

- Let \( E^c \) = you test negative for Zika with this test.
- Let \( F \) = you actually have Zika.
- What is \( P(F \mid E^c) \)?

<table>
<thead>
<tr>
<th></th>
<th>Zika +</th>
<th>Zika –</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test +</strong></td>
<td>( 0.98 = P(E \mid F) )</td>
<td>( 0.01 = P(E \mid F^c) )</td>
</tr>
<tr>
<td><strong>Test –</strong></td>
<td>( 0.02 = P(E^c \mid F) )</td>
<td>( 0.99 = P(E^c \mid F^c) )</td>
</tr>
</tbody>
</table>

\[
P(F \mid E^c) = \frac{P(E^c \mid F) \cdot P(F)}{P(E^c \mid F) \cdot P(F) + P(E^c \mid F^c) \cdot P(F^c)}
\]

\[
P(F \mid E^c) = \frac{(0.02)(0.005)}{(0.02)(0.005) + (0.99)(1 - 0.005)} \approx 0.0001
\]
The Monty Hall Problem
Monty Hall Problem aka Let’s Make a Deal

Behind one door is a prize (equally likely to be any door).

Behind the other two doors is nothing

1. We choose a door
2. Host opens 1 of other 2 doors, revealing nothing
3. We are given an option to change to the other door.

Should we switch?

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

We are comparing P(win) and P(win|switch).
Marilyn Vos Savant: “You Should Always Switch”
If we switch

Without loss of generality, say we pick A (out of Doors A, B, C).

A = prize
- Host opens B or C
- We switch
- We lose
P(win | A prize, picked A, switched) = 0

B = prize
- Host must open C
- We switch to B
- We win
P(win | B prize, picked A, switched) = 1

C = prize
- Host must open B
- We switch to C
- We win
P(win | C prize, picked A, switched) = 1

P(win | picked A, switched) = 1/3 * 0 + 1/3 * 1 + 1/3 * 1 = 2/3

You should switch.

Piech & Cain, CS109, Stanford University
Monty Hall, 1000 envelope version

Start with 1000 envelopes (of which 1 is the prize).

1. You choose 1 envelope. \[ \frac{1}{1000} = P(\text{envelope is prize}) \]
   \[ \frac{999}{1000} = P(\text{other 999 envelopes have prize}) \]

2. I open 998 of remaining 999 (showing they are empty). \[ \frac{999}{1000} = P(998 \text{ empty envelopes had prize}) \]
   \[ + P(\text{last other envelope has prize}) \]
   \[ = P(\text{last other envelope has prize}) \]

3. Should you switch? No: \[ P(\text{win without switching}) = \frac{1}{\text{original # envelopes}} \]
   Yes: \[ P(\text{win with new knowledge}) = \frac{1}{\text{original # envelopes} - 1} \]

Piech & Cain, CS109, Stanford University
Next time: gambling!