Key definitions

An experiment in probability:

Sample Space, $S$: The set of all possible outcomes of an experiment
Event, $E$: Some subset of $S$ ($E \subseteq S$).

We have the power to redesign our experiment, provided we can recreate the set of outcomes!
Card Flipping

In a 52 card deck, cards are flipped one at a time. After the first ace (of any suit) appears, consider the next card.

Is $P(\text{next card = Ace Spades}) < P(\text{next card = 2 Clubs})$?
Card Flipping

In a 52 card deck, cards are flipped one at a time. After the first ace (of any suit) appears, consider the next card.

Is \( P(\text{next card} = \text{Ace Spades}) < P(\text{next card} = 2 \text{ Clubs}) \)?

**Sample space** \( S = 52 \) in-order cards (shuffle deck)

**Event**

\( E_{AS}, \text{ next card is Ace Spades} \)

1. Take out Ace of Spades.
2. Shuffle leftover 51 cards.
3. Add Ace Spades after first ace.

\(|E_{AS}| = 51! \cdot 1\)

\( E_{2C}, \text{ next card is 2 Clubs} \)

1. Take out 2 Clubs.
2. Shuffle leftover 51 cards.
3. Add 2 Clubs after first ace.

\(|E_{2C}| = 51! \cdot 1\)

\[ P(E_{AS}) = P(E_{2C}) \]
Today’s plan

Conditional Probability and Chain Rule

Law of Total Probability

Bayes’ Theorem
Roll two 6-sided dice, yielding values $D_1$ and $D_2$.

Let $E$ be event: $D_1 + D_2 = 4$.

What is $P(E)$?

$|S| = 36$

$E = \{(1,3), (2,2), (3,1)\}$

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

Let $F$ be event: $D_1 = 2$.

What is $P(E, \text{given } F \text{ already observed})$?
Conditional Probability

The **conditional probability** of $E$ given $F$ is the probability that $E$ occurs given that $F$ has already occurred. This is known as conditioning on $F$.

Written as: $P(E|F)$

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

Sample space → all possible outcomes consistent with $F$ (i.e. $S \cap F$)

Event space → all outcomes in $E$ consistent with $F$ (i.e. $E \cap F$)
Conditional Probability, equally likely outcomes

The **conditional probability** of $E$ given $F$ is the probability that $E$ occurs given that $F$ has already occurred. This is known as conditioning on $F$.

With **equally likely outcomes**:

$$P(E|F) = \frac{{\# \text{ of outcomes in } E \text{ consistent with } F}}{{\# \text{ of outcomes in } S \text{ consistent with } F}} = \frac{|E \cap F|}{|S \cap F|}$$

$$= \frac{|EF|}{|F|}$$

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

$$P(E|F) = \frac{3}{14} \approx 0.21$$
24 emails are sent, 6 each to 4 users.
- 10 of the 24 emails are spam.
- All possible outcomes are equally likely.

Let $E = \text{user 1 receives 3 spam emails}$. What is $P(E)$?

Let $F = \text{user 2 receives 6 spam emails}$. What is $P(E|F)$?

Let $G = \text{user 3 receives 5 spam emails}$. What is $P(G|F)$?

Slicing up the spam

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

Equally likely outcomes
Quick check

You have a flowering plant.

Let $E =$ Flowers bloom

$F =$ It gets watered

$G =$ It gets sun

In English, how do you interpret $P(E|FG)$?

The probability that...

A. ...flowers bloom given the probability that it gets water and it gets sun

B. ...flowers bloom given it gets watered given it gets sun

C. ...flowers bloom given (it gets watered and it gets sun)

D. All/none/other
Conditional probability in general

General definition of conditional probability:

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

The Chain Rule (aka Product rule):

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

These properties hold even when outcomes are not equally likely.
Let $E$ = a user watches Life is Beautiful.
What is $P(E)$?

\[ P(E) \approx \frac{\text{# people who have watched movie}}{\text{# people on Netflix}} \]

\[ = \frac{10,234,231}{50,923,123} \approx 0.20 \]
Let $E$ be the event that a user watches the given movie.

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

Definition of Cond. Probability

$P(E) = 0.19$  
$P(E) = 0.32$  
$P(E) = 0.20$  
$P(E) = 0.20$  
$P(E) = 0.09$
Netflix and Learn

Let $E$ = a user watches Life is Beautiful.
Let $F$ = a user watches Amelie.

What is the probability that a user watches Life is Beautiful, given they watched Amelie?

$$P(E|F)$$

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{\text{# people who have watched both}}{\text{# people on Netflix}} \times \frac{\text{# people who have watched Amelie}}{\text{# people on Netflix}}$$

$$= \frac{\text{# people who have watched both}}{\text{# people who have watched Amelie}}$$

$$\approx 0.42$$
Netflix and Learn

Let $E$ be the event that a user watches the given movie. Let $F$ be the event that the same user watches Amelie.

Let $E$ be the event that a user watches the given movie. Let $F$ be the event that the same user watches Amelie.

\[
\begin{align*}
P(E) & = 0.19 & P(E) & = 0.32 & P(E) & = 0.20 & P(E) & = 0.20 & P(E) & = 0.09 & P(E) & = 0.20 \\
P(E|F) & = 0.14 & P(E|F) & = 0.35 & P(E|F) & = 0.20 & P(E|F) & = 0.72 & P(E|F) & = 0.42
\end{align*}
\]

Definition of Conditional Probability

\[
P(E|F) = \frac{P(EF)}{P(F)}
\]
Today’s plan

Conditional Probability and Chain Rule

Law of Total Probability

Bayes’ Theorem
Today’s plan in pictures

\[ P(EF) \]

Chain rule (Product rule) \[ \downarrow \]
Definition of conditional probability

\[ P(E|F) \]

Law of Total Probability \[ \downarrow \]
Bayes’ Theorem

\[ P(E) \]

\[ P(F|E) \]
Law of Total Probability

**Thm** Let $F$ be an event where $P(F) > 0$. For any event $E$,

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

**Proof**

1. $F$ and $F^C$ are disjoint s.t. $F \cup F^C = S$  \hspace{1cm} \text{Def. of complement}
2. $E = (EF) \cup (EF^C)$ \hspace{1cm} \text{(see below)}
3. $P(E) = P(EF) + P(EF^C)$  \hspace{1cm} \text{Additivity axiom}
4. $P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$  \hspace{1cm} \text{Chain rule (product rule)}
General Law of Total Probability

Thm  For disjoint events $F_1, F_2, ..., F_n$ s.t. $F_1 \cup F_2 \cup \cdots \cup F_n = S$,

$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$
Finding $P(E)$ from $P(E|F)$

- Flip a fair coin.
- If heads: roll a fair 6-sided die.
- Else: roll a fair 3-sided die.

You win if you roll a 6. What is $P$(winning)?

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

Law of Total Probability
Announcements

**Section sign-ups**
- Results: later today
- Late signups/changes: later today

**Problem set 1 autograder issues**
- Read problem carefully: see pinned Piazza post
- Syntax issue: `np.random.randint(1, 101)`
Today’s plan

Conditional Probability and Chain Rule

Law of Total Probability

Bayes’ Theorem
Today’s tasks

\[ P(EF) \]

Chain rule (Product rule) \[ \uparrow \downarrow \]

Definition of conditional probability

\[ P(E|F) \]

Law of Total Probability \[ \downarrow \]

Bayes’ Theorem

\[ P(E) \]

\[ P(F|E) \]
Detecting spam email

Spam volume as percentage of total email traffic worldwide

We can easily calculate how many spam emails contain “Dear”:

\[ P(E|F) = P(\text{"Dear"} | \text{Spam email}) \]

But what is the probability that an email containing “Dear” is spam?

\[ P(F|E) = P(\text{Spam email} | \text{"Dear"}) \]
Bayes’ Theorem

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

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

**Proof**
Bayes’ Theorem (expanded form)

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

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

**Proof**
Detecting spam email

• 60% of all email in 2019 is spam.
• 20% of spam has the word “Dear”
• 1% of non-spam (aka ham) has the word “Dear”

You get an email with the word “Dear” in it.
What is the probability that the email is spam?

Bayes’ Theorem

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

- 60% of all email in 2019 is spam.
- 20% of spam has the word “Dear”
- 1% of non-spam (aka ham) has the word “Dear”

You get an email with the word “Dear” in it.

What is the probability that the email is spam?  

Want: $P(F|E)$  posterior

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

$P(F)$ prior

$P(E|F)$ likelihood

$P(E|F^C)$ likelihood

normalization constant
Zika, an autoimmune disease

A disease spread through mosquito bites. Usually no symptoms; worst case paralysis. During pregnancy: may cause birth defects

If a test returns positive, what is the likelihood you have the disease?
Taking tests: Confusion matrix

Fact, \( F \) or \( F^C \)

Has disease
No disease

Evidence, \( E \) or \( E^C \)

Test positive
Test negative

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Fact</th>
<th>Evidence, ( E ) or ( E^C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E, ) Test +</td>
<td>( F, ) disease +</td>
<td>True positive ( P(E</td>
</tr>
<tr>
<td>( E^c, ) Test –</td>
<td>( F^C, ) disease –</td>
<td>False positive ( P(E</td>
</tr>
</tbody>
</table>

If a test returns positive, what is the likelihood you have the disease?
Taking tests: Confusion matrix

If a test returns positive, what is the likelihood you have the disease?

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Fact</th>
<th>Evidence, $E$</th>
<th>Fact, $F$ or $F^C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E, \text{Test +}$</td>
<td>$F, \text{disease +}$</td>
<td>True positive $P(E</td>
<td>F)$</td>
</tr>
<tr>
<td>$E^C, \text{Test –}$</td>
<td>$F^C, \text{disease –}$</td>
<td>False negative $P(E^C</td>
<td>F)$</td>
</tr>
</tbody>
</table>
Zika Testing

- A test is 98% effective at detecting Zika ("true positive").
- However, the test has a "false positive" rate of 1%.
- 0.5% of the US population has Zika.

What is the likelihood you have Zika if you test positive?

\[
P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)} \quad \text{Bayes' Theorem}
\]
Bayes’ Theorem intuition

Original question:
What is the likelihood you have Zika if you test positive for the disease?
Bayes’ Theorem intuition

Original question:
What is the likelihood you have Zika if you test positive for the disease?

Interpretation:
Of the people who test positive, how many actually have Zika?

The space of facts, conditioned on a positive test result.
Zika Testing

- A test is 98% effective at detecting Zika (“true positive”).
- However, the test has a “false positive” rate of 1%.
- 0.5% of the US population has Zika.

What is the likelihood you have Zika if you test positive?

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.

\[ \approx 0.333 \]
Update your beliefs with Bayes’ Theorem

\[ E = \text{you test positive for Zika} \]
\[ F = \text{you actually have the disease} \]

I have a 0.5% chance of having Zika.

Take test, results positive

With these test results, I now have a 33% chance of having Zika!!!
Why it’s still good to get tested

- A test is 98% effective at detecting Zika (“true positive”).
- However, the test has a “false positive” rate of 1%.
- 0.5% of the US population has Zika.

Let:

- $E$ = you test positive
- $F$ = you actually have the disease
- $E^C$ = you test negative for Zika with this test.

What is $P(F|E^C)$?

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

<table>
<thead>
<tr>
<th></th>
<th>$F$, disease +</th>
<th>$F^C$, disease -</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$, Test+</td>
<td>True positive</td>
<td>False positive</td>
</tr>
<tr>
<td></td>
<td>$P(E</td>
<td>F) = 0.98$</td>
</tr>
</tbody>
</table>
Why it’s still good to get tested

- A test is 98% effective at detecting Zika (“true positive”).
- However, the test has a “false positive” rate of 1%.
- 0.5% of the US population has Zika.

Let:

- $E$ = you test positive
- $F$ = you actually have the disease
- $E^C$ = you test negative for Zika with this test.

What is $P(F|E^C)$?

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

\[\begin{array}{c|c|c}
| & F, disease + & F^C, disease - \\
| \hline
E, Test + & True positive & False positive \\
& $P(E|F) = 0.98$ & $P(E|F^C) = 0.01$
| \hline
E^C, Test - & False negative & True negative \\
& $P(E^C|F) = 0.02$ & $P(E^C|F^C) = 0.99$
| \hline
\end{array}\]
Why it’s still good to get tested

$E = \text{you test positive for Zika}$

$F = \text{you actually have the disease}$

$E^C = \text{you test negative for Zika}$

---

I have a 0.5% chance of having Zika disease.

$P(F)$

---

With these test results, I now have a 33% chance of having Zika!!!

$P(F|E)$

---

Take test, results positive

---

Take test, results negative

---

With these test results, I now have a 0.01% chance of having Zika disease!!!

$P(F|E^C)$
This class going forward

Last week
Equally likely events

Today and for most of this course
Not equally likely events

\[ P(E = \text{Evidence} \mid F = \text{Fact}) \]
\[ P(F = \text{Fact} \mid E = \text{Evidence}) \]

(collected from data)

(counting, combinatorics)
Another conditional probability example
Monty Hall Problem from Let’s Make a Deal

Behind one door is a car (equally likely to be any door). Behind the other two doors are goats

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

Should you switch?
What happens if you switch

\[
P(\text{win} \mid \text{you picked } \_\_, \text{switched}) = \frac{1}{3}
\]

- **A = prize**
  - Host opens
  - You switch to
  - Result
  \[
P(\text{win} \mid \text{A prize, you picked } \_\_, \text{switch}) =
\]

- **B = prize**
  - Host opens
  - You switch to
  - Result
  \[
P(\text{win} \mid \text{B prize, you picked } \_\_, \text{switch}) =
\]

- **C = prize**
  - Host opens
  - You switch to
  - Result
  \[
P(\text{win} \mid \text{C prize, you picked } \_\_, \text{switch}) =
\]
Monty Hall, 1000 envelope version

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

1. You choose 1 envelope.

\[
\begin{align*}
\frac{1}{1000} &= P(\text{envelope is prize}) \\
\frac{999}{1000} &= P(\text{other 999 envelopes have prize})
\end{align*}
\]

2. I open 998 of remaining 999 (showing they are empty).

\[
\begin{align*}
\frac{999}{1000} &= P(998 \text{ empty envelopes had prize}) \\
+ P(\text{last other envelope has prize}) \\
&= P(\text{last other envelope has prize})
\end{align*}
\]

3. Should you switch?

\[
\begin{align*}
P(\text{you win without switching}) &= \frac{1}{\text{original \# envelopes}} \\
P(\text{you win with switching}) &= \frac{\text{original \# envelopes} - 1}{\text{original \# envelopes}}
\end{align*}
\]