PART 1: MATH REVIEW

Sets are simply collections of elements: $A = \{x, y, z\}$ (so $x \in A$)

- A subset is a set with (weakly) fewer elements ($B \subset A$)

Set Theory

Sets are simply collections of elements: $A = \{x, y, z\}$ (so $x \in A$)

- A subset is a set with (weakly) fewer elements ($B \subset A$)
- Sets are combined through set operations:
  - **Union**: $A \cup B$
  - **Intersection**: $A \cap B$
  - **Complement**: A^c, $\overline{A}$ or A'
  - **Set Difference**: $A \setminus B$

Set Theory

Sets are simply collections of elements: $A = \{x, y, z\}$ (so $x \in A$)

- A subset is a set with (weakly) fewer elements ($B \subset A$)
- Sets are combined through set operations:
  - **Union**: $A \cup B$
  - **Intersection**: $A \cap B$
  - **Complement**: A^c, $\overline{A}$ or A'
  - **Set Difference**: $A \setminus B$
- DeMorgan's Laws combines these operations:
  - $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$
  - $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$

Sets are simply collections of elements: $A = \{x, y, z\}$ (so $x \in A$)

- A subset is a set with (weakly) fewer elements ($B \subset A$)
- Sets are combined through set operations:
  - **Union**: $A \cup B$
  - **Intersection**: $A \cap B$
  - **Complement**: A^c, $\overline{A}$ or A'
  - **Set Difference**: $A \setminus B$
- DeMorgan's Laws combines these operations:
  - $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$
  - $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$
- Set with no elements: $\emptyset$
- Don't disrespect the Venn Diagram!

- **Factorial**: $n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 2 \cdot 1$
  - Example: $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

Mathematical Operations Used in Probability

- **Factorial**: $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$
  - Example: $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$
- **Exponents**
  - $e \approx 2.17$ (used in growth and construction of key distributions)
  - **Negative exponents** are interpreted as **reciprocals**: $x^{-2} = \frac{1}{x^2}$
  - $x^0 = 1$
  - **Other exponent rules**: $c^x \cdot c^y, c^x/c^y, (c^x)^y$

Mathematical Operations Used in Probability

- **Factorial**: $n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 2 \cdot 1$
  - Example: $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$
- **Exponents**
  - $e \approx 2.17$ (used in growth and construction of key distributions)
  - **Negative exponents** are interpreted as **reciprocals**: $x^{-2} = \frac{1}{x^2}$
  - $x^0 = 1$
  - **Other exponent rules**: $c^x \cdot c^y, c^x/c^y, (c^x)^y$
- **Logarithms**
  - $\log_x(y) =? \Rightarrow$ how many powers of x give you y?
  - Example: $\log_{10}(100) =?$
  - **Natural logarithm** is the inverse of the exponential, e.x.: $\ln(e^5) = e^{\ln(5)} = 5$
  - **Other log properties**: $\ln(xy), \ln(x/y), \ln(x^2), \ln(x+y)$

Summation Notation

- Suppose you have a data of $n = 500$ observations, and want to find the **mean**

Summation Notation

- Suppose you have a data of $n = 500$ observations, and want to find the **mean**
- Denote each observation by x_i for $i = 1, 2, ..., 500$
- Then $\overline{x} = \frac{1}{500}(x_1 + x_2 + ... + x_n)$, which is simplified by

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Summation Notation

- Suppose you have a data of $n = 500$ observations, and want to find the **mean**
- Denote each observation by x_i for $i = 1, 2, ..., 500$
- Then $\overline{x} = \frac{1}{500}(x_1 + x_2 + ... + x_n)$, which is simplified by

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- **Some properties of summation notation**:
  - $\sum_{i=1}^{n}(ax_i + by_i) = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i$
  - $\sum_{i=1}^{n} x = nx$
  - **Geometric sums** $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$ as long as $|r| < 1$

Product Notation (and More!)

- Any linear operation can be generalized in this way:

$$\Pi_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n, \ \bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n$$

Product Notation (and More!)

- Any linear operation can be generalized in this way:

$$\prod_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n, \quad \bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n$$

- **Some properties of product notation**:
  - $\prod_{i=1}^{n} x = x^n$
  - $\prod_{i=1}^{n} e^{x_i} = e^{\sum_{i=1}^{n} x_i}$
  - **Important!** The log of products becomes summation

$$\begin{aligned} \ln\left(\prod_{i=1}^{n} x_i\right) &= \ln(x_1 \cdot x_2 \cdot \ldots \cdot x_n) \\ &= \ln(x_1) + \ln(x_2) + \ldots + \ln(x_n) \\ &= \sum_{i=1}^{n} \ln(x_i) \end{aligned}$$

Functions

Functions are pathways between two spaces

- **Domain**: the origin space
- **Co-Domain/Range**: the destination space
- A function $f : D \to R$ has one (and only one!) output in R for any element in D

Functions

Functions are **pathways** between two spaces

- **Domain**: the origin space
- **Co-Domain/Range**: the destination space
- A function $f : D \to R$ has one (and only one!) output in R for any element in D
- Examples of **real-valued** functions
  - $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$
  - $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + y^2$
  - $f : \ell^\infty \to \mathbb{R}$ given by $U(c_1, c_2, ...) = \sum_{i=1}^{\infty} u(c_i)$

Functions

Functions are pathways between two spaces

- **Domain**: the origin space
- **Co-Domain/Range**: the destination space
- A function $f : D \to R$ has one (and only one!) output in R for any element in D
- Examples of real-valued functions
  - $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$
  - $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + y^2$
  - $f : \ell^\infty \to \mathbb{R}$ given by $U(c_1, c_2, ...) = \sum_{i=1}^{\infty} u(c_i)$
- Combining functions:
  - **Linear operations**: $(f \pm g) = f(x) \pm g(x)$
  - **Composition**: $f(g(x))$ (functions must match on spaces)

Functions

Functions are pathways between two spaces

- **Domain**: the origin space
- **Co-Domain/Range**: the destination space
- A function $f : D \to R$ has one (and only one!) output in R for any element in D
- Examples of real-valued functions
  - $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$
  - $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + y^2$
  - $f : \ell^\infty \to \mathbb{R}$ given by $U(c_1, c_2, ...) = \sum_{i=1}^\infty u(c_i)$
- Combining functions:
  - **Linear operations**: $(f \pm g) = f(x) \pm g(x)$
  - **Composition**: $f(g(x))$ (functions must match on spaces)

Some things to review for other classes: continuity, injective/surjective/bijective. Not needed here.

Calculus: Derivatives

- Derivatives capture notions of rates of change or slope
  - In economics, critical for optimization
  - To solve $\max_x f(x)$, need a first-order condition ($f'(x) = 0$) and a second-order condition ($f''(x) < 0$)

Calculus: Derivatives

- Derivatives capture notions of **rates of change** or slope
  - In economics, critical for optimization
  - To solve $\max_x f(x)$, need a first-order condition ($f'(x) = 0$) and a second-order condition ($f''(x) < 0$)
- Refresher on **common derivatives**:
  - $f(x) = x^a \rightarrow f'(x) = ax^{a-1}$
  - $f(x) = a \cdot x \rightarrow f'(x) = a$
  - $f(x) = a \rightarrow f'(x) = 0$
  - $f(x) = x^{-a} \rightarrow f'(x) = -ax^{-a-1}$
  - $f(x) = e^x \rightarrow f'(x) = e^x$
  - $f(x) = \ln(x) \rightarrow f'(x) = \frac{1}{x}$

Calculus: Derivatives

- Derivatives capture notions of rates of change or slope
  - In economics, critical for optimization
  - To solve $\max_x f(x)$, need a first-order condition ($f'(x) = 0$) and a second-order condition ($f''(x) < 0$)
- Refresher on common derivatives:
  - $f(x) = x^a \rightarrow f'(x) = ax^{a-1}$
  - $f(x) = a \cdot x \rightarrow f'(x) = a$
  - $f(x) = a \rightarrow f'(x) = 0$
  - $f(x) = x^{-a} \rightarrow f'(x) = -ax^{-a-1}$
  - $f(x) = e^x \rightarrow f'(x) = e^x$
  - $f(x) = \ln(x) \rightarrow f'(x) = \frac{1}{x}$
- Other derivative rules to know:
  - **Product rule**: $\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$
  - **Chain rule**: $\frac{d}{dx}[f(g(x))] = f'(g(x))g'(x)$
  - **Quotient rule**: $\frac{d}{dx}[f(x)/g(x)] = \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}$
  - $f^{(k)}(x)$ denotes the k-th derivative of $f(x)$

Writing Proofs

“The most important step in a proof is the step you **don't know** you're taking"

Writing Proofs

"The most important step in a proof is the step you **don't know** you're taking"

Useful tips:

- Organization helps you *and* me!
  - Start your proof with **Proof.** and end with □. Use sentences.
  - Take steps one line at a time
  - Justify everything!
  - Similar to commenting on code—don't want to have to rework

Writing Proofs

“The most important step in a proof is the step you **don’t know** you’re taking"

Useful tips:

- Organization helps you *and* me!
  - Start your proof with **Proof.** and end with $\square$. Use sentences.
  - Take steps one line at a time
  - Justify everything!
  - Similar to commenting on code—don’t want to have to rework
- When in doubt, write down what you need to show (WTS)
- If still in doubt, think about other ways of attack:
  - Proof by contradiction (never don’t forget this!)
  - Induction
  - Contrapositive

Writing Proofs

“The most important step in a proof is the step you **don’t know** you’re taking"

Useful tips:

- Organization helps you *and* me!
  - Start your proof with **Proof.** and end with □. Use sentences.
  - Take steps one line at a time
  - Justify everything!
  - Similar to commenting on code—don’t want to have to rework
- When in doubt, write down what you need to show (WTS)
- If still in doubt, think about other ways of attack:
  - Proof by contradiction (never don’t forget this!)
  - Induction
  - Contrapositive

“Anything is possible when you fly on the wings of mathematics!"

SECTION 2.1-2.3: BASICS OF PROBABILITY

Building blocks in the language of probability:

1. **Experiment**: The action/process of uncertainty

Setting the Stage: Terminology

Building blocks in the language of probability:

1. **Experiment**: The action/process of uncertainty
2. **Sample Space**, $\mathcal{S}$: The set of all possible outcomes

Setting the Stage: Terminology

Building blocks in the language of probability:

1. **Experiment**: The action/process of uncertainty
2. **Sample Space**, $\mathcal{S}$: The set of all possible outcomes
3. **Event**, E: Any subset of outcomes in $\mathcal{S}$. Events are stored in a collection, $E \in \mathcal{E}$

Setting the Stage: Terminology

Building blocks in the language of probability:

1. **Experiment**: The action/process of uncertainty
2. **Sample Space**, $\mathcal{S}$: The set of all possible outcomes
3. **Event**, E: Any subset of outcomes in $\mathcal{S}$. Events are stored in a collection, $E \in \mathcal{E}$
4. Assigning relative likelihood of events in $\mathcal{S}$ is **probability** (P)

A triple $(\mathcal{S}, \mathcal{E}, P)$ is a **probability space**

Examples of A Sample Space

- Flipping a coin $\Rightarrow$ $\mathcal{S} = \{H, T\}$

Examples of A Sample Space

- Flipping a coin $\Rightarrow\ \mathcal{S} = \{H, T\}$
- Flipping that coin 3 times
  $\Rightarrow\ \mathcal{S} = \{HHH, HHT, HTT, TTT, TTH, THH, HTH, THT\}$

Examples of A Sample Space

- Flipping a coin $\Rightarrow \ \mathcal{S} = \{H, T\}$
- Flipping that coin 3 times
  $\Rightarrow \ \mathcal{S} = \{HHH, HHT, HTT, TTT, TTH, THH, HTH, THT\}$
- What about rolling dice?

Examples of A Sample Space

- Flipping a coin $\Rightarrow\ \mathcal{S} = \{H, T\}$
- Flipping that coin 3 times
  $\Rightarrow\ \mathcal{S} = \{HHH, HHT, HTT, TTT, TTH, THH, HTH, THT\}$
- What about rolling dice?
- Two gas stations with 6 pumps each. We measure how many pumps are in use at each station at 9:17 am on a Saturday.
  What does the sample space look like?

A Complicated Sample Space

	Second Station						
First Station	**0**	**1**	**2**	**3**	**4**	**5**	**6**
0	(0, 0)	(0, 1)	(0, 2)	(0, 3)	(0, 4)	(0, 5)	(0, 6)
1	(1, 0)	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 0)	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 0)	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 0)	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 0)	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 0)	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Events

Events are **any subset** of outcomes in $\mathcal{S}$

- Flipping at least one tails
- Rolling only even numbers
- Exactly 4 pumps in use at the first station

Events

Events are **any subset** of outcomes in $\mathcal{S}$

- Flipping at least one tails
- Rolling only even numbers
- Exactly 4 pumps in use at the first station

An event can be **simple** ($|E| = 1$) or **compound** ($|E| > 1$).

The collection of all events, $\mathcal{E}$, is the **power set** of $\mathcal{S}$ when the state space is discrete[1]

[1] When $\mathcal{S}$ is not discrete, $\mathcal{E}$ needs to satisfy some special closure conditions, so that it is a σ-algebra.

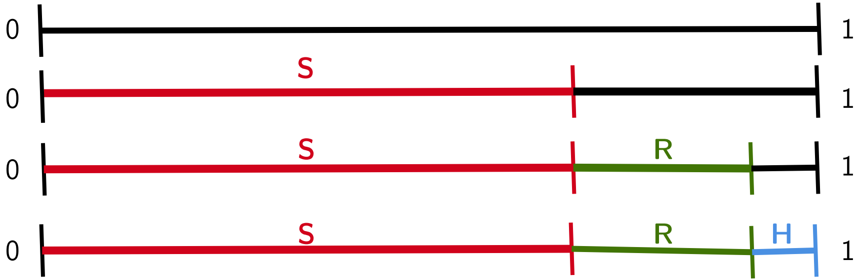
Okay, so what?

The goal of probability is to map **events** to **numbers** in a way that conveys information about likelihoods

The goal of probability is to map **events** to **numbers** in a way that conveys information about likelihoods

The goal of probability is to map events to numbers in a way that conveys information about likelihoods

The goal of probability is to map events to numbers in a way that conveys information about likelihoods



Axioms of Probability

For an event A, denote the probability that A occurs by $P(A)$

A **probability measure** is a function $p : \mathcal{S} \to \mathbb{R}$ that satisfies 3 axioms (building blocks of theory):

A1 For all $A \in \mathcal{S}$, $P(A) \geq 0$

A2 $P(\mathcal{S}) = 1$

A3 For any **infinite** collection of **disjoint** events $\{A_1, A_2, ...\}$, $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ (countable additivity)

Questions about the axioms: What do we think? Why don't we need finiteness in A3? Etc?

Proposition: Null Events have 0 Probability

What we want to show: $P(\emptyset) = 0$
Proof: Use Axiom 3, taking as our infinite collection $\{\emptyset, \emptyset, ...\}$.
Notice that $\emptyset \cap \emptyset = \emptyset$, so the events are **disjoint**.
Then, $P\left(\bigcup_{i=1}^{\infty} \emptyset\right) = P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$.
But we can't have $x = \sum_{i=1}^{\infty} x$ unless $x = 0$. Hence, $P(\emptyset) = 0$.
$\square$

Proposition: Axiom 3 holds for *finite* collections

What we want to show: If $\{A_1, ..., A_k\}$ is a **finite** collection of disjoint events, then $P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i)$.

Proof: Take a finite collection of events $\{A_1, ..., A_k\}$.
Make this an **infinite** collection by appending $\emptyset$'s!
That is, $\{A_1, ..., A_k\} \Rightarrow \{A_1, A_2, ..., A_k, \emptyset, \emptyset, ...\}$.
Then the third axiom says

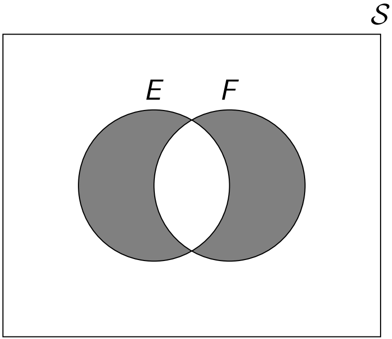
$$P\left(\bigcup_{i=1}^{k} A_i\right) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{k} P(A_i) + 0.$$

$\square$

More Probability Properties

1. For all $E \in \mathcal{E}$, $P(E) = 1 - P(\overline{E})$
2. For all $E \in \mathcal{E}$, $P(E) \leq 1$
3. For any two events E, F,
   $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

Can you prove these?



This is something to **watch out for**!

Some examples

- If you flip a fair coin twice, what's the probability of getting at least one head?
- What's the probability of drawing either a spade or an ace from a deck of cards?

This is something to **watch out for**!

Some examples

- If you flip a fair coin twice, what's the probability of getting at least one head?
- What's the probability of drawing either a spade or an ace from a deck of cards?

When in doubt, **list and count**!

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183
Weekends	48 hrs/wk	2496	2687

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183
Weekends	48 hrs/wk	2496	2687
Vacation	2 wks	336	2361

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183
Weekends	48 hrs/wk	2496	2687
Vacation	2 wks	336	2361
Holidays	5 days/yr	120	2231

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183
Weekends	48 hrs/wk	2496	2687
Vacation	2 wks	336	2361
Holidays	5 days/yr	120	2231
Commute	2	500	1731

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183
Weekends	48 hrs/wk	2496	2687
Vacation	2 wks	336	2361
Holidays	5 days/yr	120	2231
Commute	2	500	1731
Working	8	2000	**-269**

How to use this to your advantage: there are 8760 hours in a year. **But** you need time for...

	Hrs Needed (day)	Hrs Needed (year)	Hours Left
			8760
Sleeping	8	2920	5840
Eating	$.5 * 3 = 1.5$	547.5	5250.5
Bathing	$\frac{1}{3}$	109.5	5183
Weekends	48 hrs/wk	2496	2687
Vacation	2 wks	336	2361
Holidays	5 days/yr	120	2231
Commute	2	500	1731
Working	8	2000	**-269**

So you have to be **about an hour late** to work every day!

SECTION 2.4: CONDITIONAL PROBABILITY

Suppose we're on a research team for the FDA. We're trying to evaluate **Panacea™**, a new SSRI for depression

What questions can probability answer about Panacea? What would we like to know in order to approve it?

Suppose we're on a research team for the FDA. We're trying to evaluate **Panacea^{TM}**, a new SSRI for depression

What questions can probability answer about Panacea? What would we like to know in order to approve it?

- What is the probability that Panacea reduces symptoms?
- How likely is it that taking Panacea causes migraines?

Suppose we're on a research team for the FDA. We're trying to evaluate **Panacea™**, a new SSRI for depression

What questions can probability answer about Panacea? What would we like to know in order to approve it?

- What is the probability that Panacea reduces symptoms?
- How likely is it that taking Panacea causes migraines?

What happens when **new information** comes along?

- Patient started taking Panacea, has canceled 3 appointments
- Migraine patient has a history of headaches

How do we **use this info** to update desired probabilities?

Multiple Events

When we have two events A and B, it is helpful to distinguish:

- **Marginal probabilities**: $P(A)$ and $P(B)$
- **Joint probability**: $P(A \cap B)$
- **Conditional probabilities**: $P(A|B)$, $P(B|A)$

When we have two events A and B, it is helpful to distinguish:

- **Marginal probabilities**: $P(A)$ and $P(B)$
- **Joint probability**: $P(A \cap B)$
- **Conditional probabilities**: $P(A|B)$, $P(B|A)$

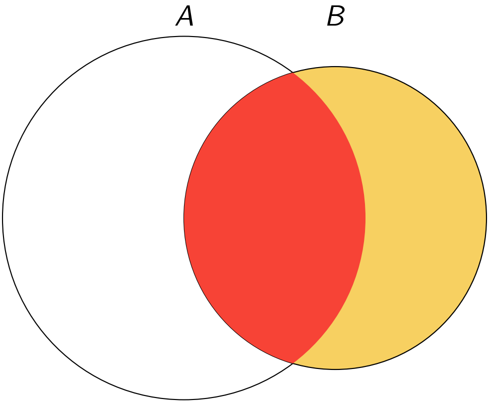
	Side Effects	
Symptom Repression	No	Yes
No	15	10
Yes	70	5

How effective is this drug? (What is $P(SymptomRepression)$?)

What if the patient is reporting side effects? (What is $P(SR|SE)$?)

A Formula for Conditional Probability

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



Example 1

You're on a date and want to choose a restaurant. You're craving a flour-less chocolate cake (ours really is the best!), but your date is insistent on getting french fries.

Example 1

You're on a date and want to choose a restaurant. You're craving a flour-less chocolate cake (ours really is the best!), but your date is insistent on getting french fries.

You know that 70% of the restaurants in Boston sell french fries, and that 40% still make the cake you like. Googling tells you that 30% of restaurants do both.
If you go to a restaurant with french fries, **what is the probability you can still have your cake?**

Example 1

You're on a date and want to choose a restaurant. You're craving a flour-less chocolate cake (ours really is the best!), but your date is insistent on getting french fries.

You know that 70% of the restaurants in Boston sell french fries, and that 40% still make the cake you like. Googling tells you that 30% of restaurants do both.
If you go to a restaurant with french fries, **what is the probability you can still have your cake?**

	C	$\overline{C}$	*Marginal*
FF	—	—	0.7
$\overline{FF}$	—	—	0.3
Marginal	0.4	0.6	1

Example 1

You're on a date and want to choose a restaurant. You're craving a flour-less chocolate cake (ours really is the best!), but your date is insistent on getting french fries.

You know that 70% of the restaurants in Boston sell french fries, and that 40% still make the cake you like. Googling tells you that 30% of restaurants do both.

If you go to a restaurant with french fries, **what is the probability you can still have your cake?**

	C	$\overline{C}$	*Marginal*
FF	0.3	—	0.7
$\overline{FF}$	—	—	0.3
Marginal	0.4	0.6	1

Example 1: Applying conditional probabilities

We know that:

- $P(A) + P(\overline{A}) = 1$
- $P(A \cap B) + P(A \cap \overline{B}) = P(A)$

Example 1: Applying conditional probabilities

We know that:

- $P(A) + P(\overline{A}) = 1$
- $P(A \cap B) + P(A \cap \overline{B}) = P(A)$

Hence,

	C	$\overline{C}$	*Marginal*
FF	0.3	0.4	0.7
$\overline{FF}$	0.1	0.2	0.3
Marginal	0.4	0.6	1

Example 1: Applying conditional probabilities

We know that:

- $P(A) + P(\overline{A}) = 1$
- $P(A \cap B) + P(A \cap \overline{B}) = P(A)$

Hence,

	C	$\overline{C}$	*Marginal*
FF	0.3	0.4	0.7
$\overline{FF}$	0.1	0.2	0.3
Marginal	0.4	0.6	1

Hence, if you choose FF, then $P(C|FF) = 3/7 \approx 0.49$. Those odds are **better** than your marginal odds!

Example 2

Check out this cool site for a visualization of conditional probability

SECTION 2.5: INDEPDENDENCE

Indepdendence

Two events A and B are *independent* if $P(A|B) = P(A)$.
Otherwise, they are *dependent*.

Some consequences of independence:

- $P(B|A) = P(B)$ (symmetry)
- $\overline{A}$ and B are independent
- B and $\overline{A}$ are independent
- $\overline{A}$ and $\overline{B}$ are independent

How do we know when events are independent?

If two events are independent:

$$P(A \cap B) \Leftrightarrow P(A|B)P(B) \Leftrightarrow P(A)P(B)$$

Hence, when the **intersection** of two events has the same probability as the **product** of the two events, they are independent

Example: Playing Cards

Q: What is the prob. of drawing a red card that is **not an ace**?

Q: What is the prob. of drawing a red card that is **not an ace**?

	R	$\overline{R}$	
A	$2/52$	$2/52$	$4/52$
$\overline{A}$	$24/52$	$24/52$	$48/52$
	$26/52$	$26/52$	1

Example: Playing Cards

Q: What is the prob. of drawing a red card that is **not an ace**?

	R	$\overline{R}$	
A	$1/26$	$1/26$	$1/13$
$\overline{A}$	$6/13$	$6/13$	$12/13$
	$1/2$	$1/2$	1

Example: Playing Cards

Q: What is the prob. of drawing a red card that is **not an ace**?

	R	$\overline{R}$	
A	$1/26$	$1/26$	$1/13$
$\overline{A}$	$6/13$	$6/13$	$12/13$
	$1/2$	$1/2$	1

A: Hence, $P(\overline{A} \cap R) = 6/13$.

Example: Playing Cards

Q: Are the events $\overline{A}$, R independent?

Example: Playing Cards

Q: Are the events $\overline{A}, R$ independent?

A: Take product of the marginals:

$$P(\overline{A}) \cdot P(R) = \frac{12}{13}\frac{1}{2} = \frac{6}{13}$$

Example: Playing Cards

Q: Are the events $\overline{A}, R$ independent?

A: Take product of the marginals:

$$P(\overline{A}) \cdot P(R) = \frac{12}{13}\frac{1}{2} = \frac{6}{13}$$

A: Hence, the two events **are independent**.

QUESTIONS?