# Soc400/500: Applied Social Statistics
# Week 1: Introduction and Probability

Brandon Stewart[1]

Princeton

August 31-September 4, 2020

[1] These slides are heavily influenced by Matt Blackwell and Adam Glynn with contributions from Justin Grimmer and Matt Salganik. Illustrations by Shay O'Brien.

# Where We've Been and Where We're Going...

- Last Week
  - living that class-free, quarantine life
- This Week
  - course structure
  - core ideas
  - introduction to probability
  - three big ideas in probability
- Next Week
  - random variables
  - joint distributions
- Long Run
  - probability $\rightarrow$ inference $\rightarrow$ regression $\rightarrow$ causal inference

# Welcome and Introductions

- The tale of two classes: Soc400/Soc500 Applied Social Statistics
- I
  - . . . am an Assistant Professor in Sociology.
  - . . . am trained in political science and statistics
  - . . . do research in methods and statistical text analysis
  - . . . love doing collaborative research
  - . . . talk very quickly
- Your Preceptors
  - Emily Cantrell
  - Alejandro Schugurensky

# Overview

- Goal: train you in statistical thinking.
- Fundamentally a graduate course for sociologists, but also useful for research in other fields, policy evaluation, industry etc.
- Difficult course but with many resources to support you.
- When we are done you will be able to teach yourself many things
- Syllabus is a useful resource including philosophy of the class.

# Specific Goals

- critically read and reason about quantitative social science using linear regression techniques.
- conduct, interpret, and communicate results from analysis using multiple regression.
- explain the limitations of observational data for making causal claims and distinguish between identification and estimation.
- understand the logic and assumptions of several modern designs for making causal claims.
- write clean, reusable, and reliable `R` code in `tidyverse` style.
- feel empowered working with data

# Why R?

- It will give you super powers (but not at first)
- It is free and open source
- It is the *de facto* standard in many applied statistical fields


Artwork by @allison_horst

# Why RMarkdown?


Artwork by @allison_horst

# Mathematical Prerequisites

- No formal pre-requisites.
- Balance of rigor and intuition.
  - no rigor for rigor's sake.
  - we will tell you *why* you need the math, but also feel free to ask.
  - course focus on how to reason about statistics, not just memorize guidelines.
- We will teach you any math you need as we go along
- Crucially though—this class is not about innate statistical aptitude, it is about effort.
- We all come from very different backgrounds. Please have patience with yourself and with others.

# Ways to Learn

- Pre-Recorded Lectures
  learn broad topics (4–8 videos a week, $\approx$2.5 hours)
- Pre-Recorded Precept
  learn data analysis skills, get targeted help on assignments
- Perusall
  an annotation platform for videos
- Course Meetings
  come together and discuss material
- Ed
  ask questions of us and your classmates
- Office Hours
  ask us even more questions, but (sort of) in-person
- Problem Sets
  reinforce understanding of material, practice

# Problem Sets

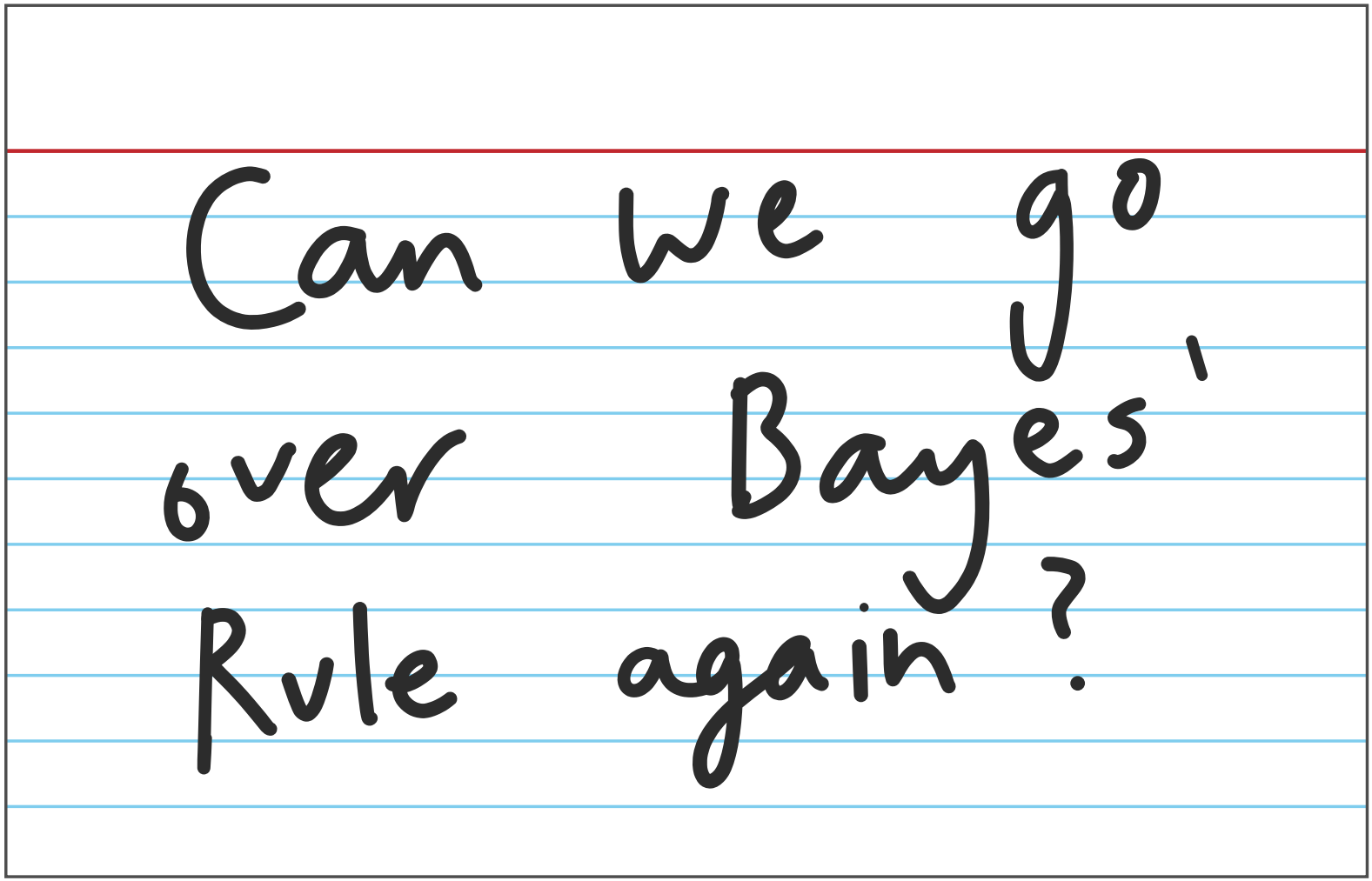
- Schedule (due Friday at 5PM eastern)
- Grading and solutions
- Collaboration policy
- You may find these difficult. Start early and seek help!
- Most important part of the class

# Ways to Learn

- Pre-Recorded Lectures
- Pre-Recorded Precept
- Perusall
- Course Meetings
- Ed
- Office Hours
- Problem Sets
- Instructor Office Hours
- Final Exam Prep
- External Consulting
- Individual and Group Tutoring

Your Job: work hard and get help when you need it!

# Staying in Touch



# A Note on Reading

- Think of the lecture slides as primary reading.
- If you want material to read, come talk to me about recommendations.
- Suggested Books (more in the syllabus!):
  - Angrist and Pischke. 2008. *Mostly Harmless Econometrics*
  - Aronow and Miller. 2019. *Foundations of Agnostic Statistics*
  - Blitzstein and Hwang. 2019. *Introduction to Probability*
- A somewhat obvious tip: don't skip the math!

# Advice from Prior Generations

- Ask questions if you don't know what's going on!
- Investing a considerable amount of time in getting familiar with R and its various tools will pay off in the long run!
- Go over the lecture slides each week. This can be hard when you feel like you're treading water and just staying afloat, but I wish I had done this regularly.
- It's challenging but very doable and rewarding if you put the time in. There are plenty of resources to take advantage of for help.
- I found it helpful to read through the lecture slides again after I had opened the problem set. It made it easier to create connections between what we went through and how to do it.
- Go over your psets and the pset solutions the moment they are graded as a habit and figure out what you don't know.

# Outline of Topics

Outline in reverse order:

- Causal Inference:
  inferring counterfactual effect given association.
- Regression:
  estimate association.
- Inference:
  estimating things we don't know from data.
- Probability:
  learning what data we would expect if we did know the truth.

$$\text{Probability} \rightarrow \text{Inference} \rightarrow \text{Regression} \rightarrow \text{Causal Inference}$$

# Attribution and Thanks

- My philosophy on teaching: don't reinvent the wheel customize, refine, improve.
- Huge thanks to those who have provided slides particularly: Matt Blackwell, Adam Glynn, Justin Grimmer, Jens Hainmueller, Erin Hartman, Kevin Quinn
- Also thanks to those who have discussed with me at length including Dalton Conley, Chad Hazlett, Gary King, Kosuke Imai, Matt Salganik and Teppei Yamamoto.
- Previous generations of preceptors have also been incredible important: Clark Bernier, Elisha Cohen, Ian Lundberg, Simone Zhang, Alex Kindel, Ziyao Tian, Shay O'Brien.
- Shay O'Brien for many hand-drawn illustrations.

# Welcome To Class!

Be sure to read the syllabus for more details.

# Where We've Been and Where We're Going...

- Last Week
  - living that class-free, quarantine life
- This Week
  - course structure
  - core ideas
  - introduction to probability
  - three big ideas in probability
- Next Week
  - random variables
  - joint distributions
- Long Run
  - probability $\rightarrow$ inference $\rightarrow$ regression $\rightarrow$ causal inference

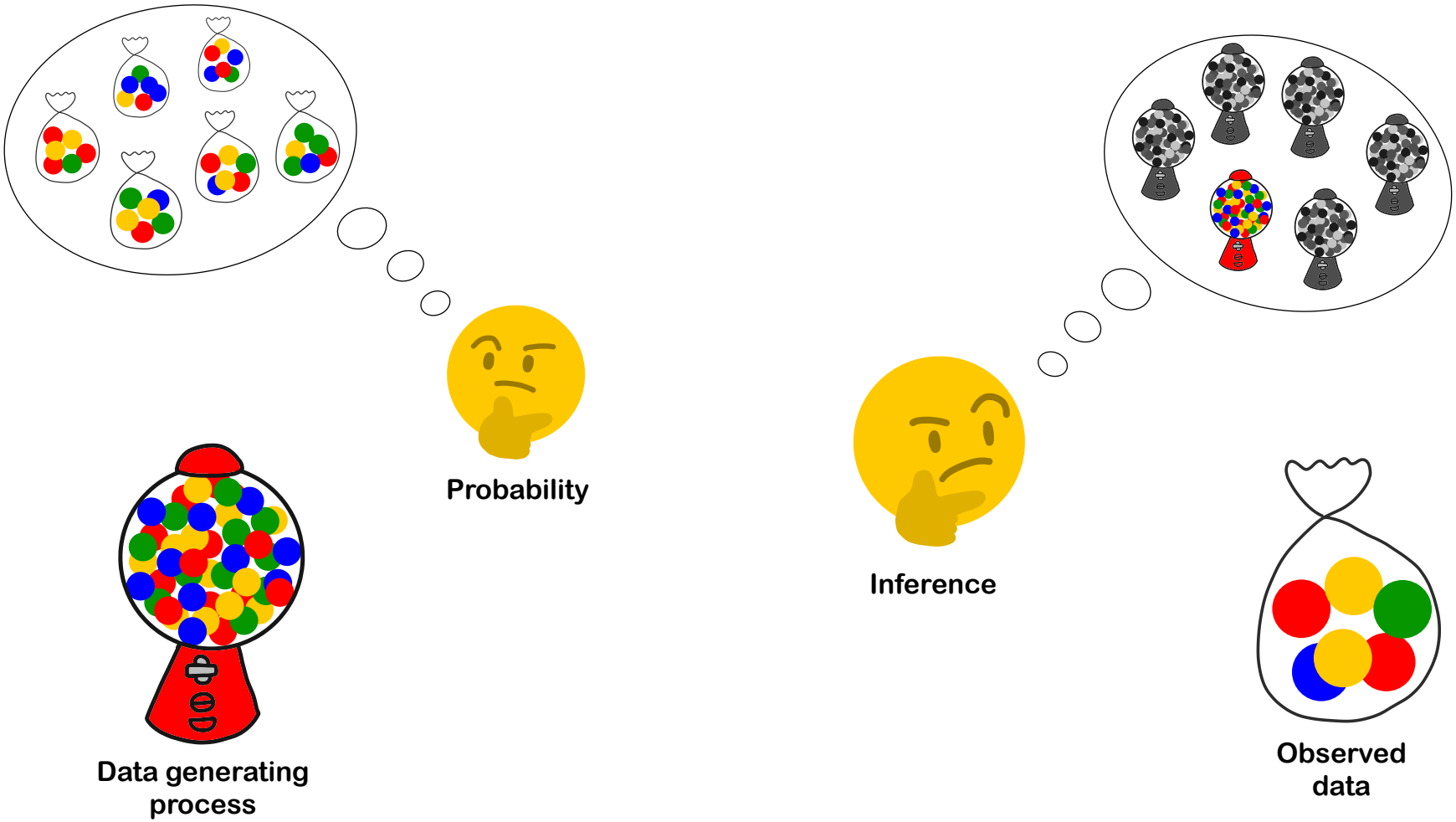
# What is Statistics?

- Branch of mathematics studying collection and analysis of data
- The name statistic comes from the word state
- The arc of developments in statistics
  1) an applied scholar has a problem
  2) they solve the problem by inventing a specific method
  3) statisticians generalize and export the best of these methods
- Relatively recent field (started at end of 19th century)
- Goal: principled guesses based on stated assumptions.
- In practice, an essential part of research, policy making, political campaigns, selling people things...

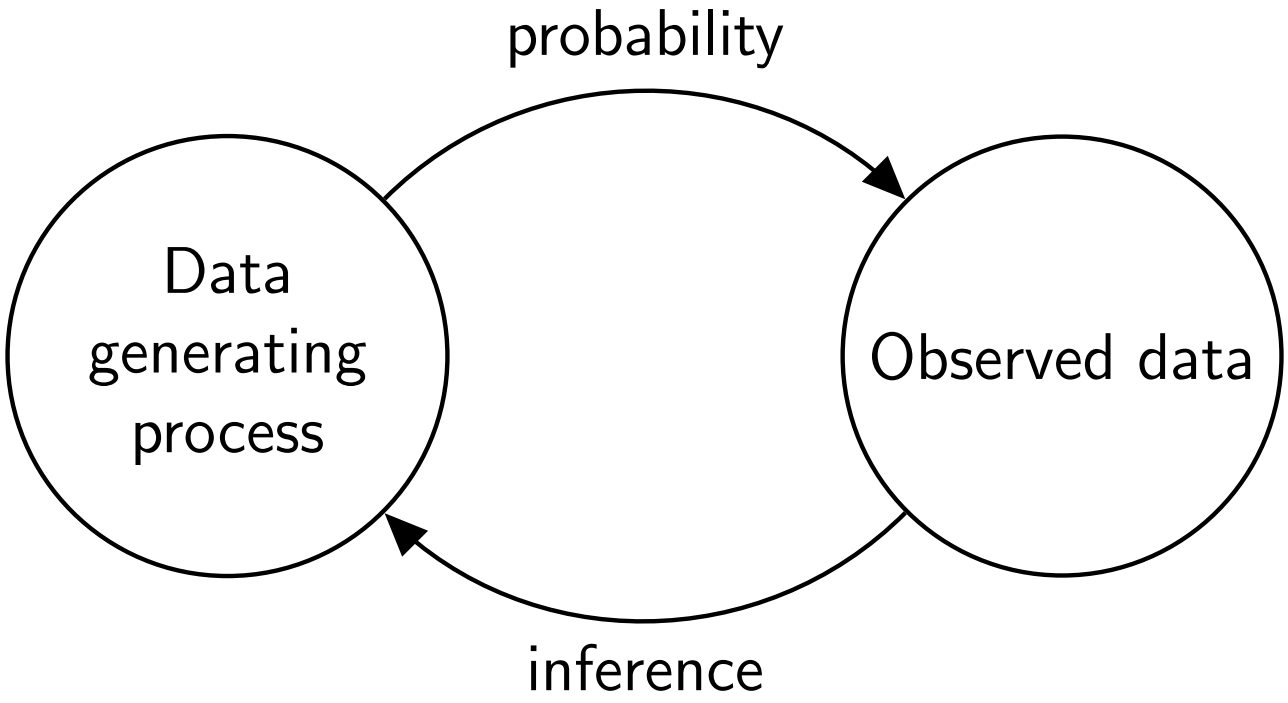
Why study probability?

It enables inference.

# In Picture Form



# In Picture Form



# Statistical Thought Experiments

- We start with probability.
- Allows us to contemplate world under hypothetical scenarios.
  - hypotheticals let us ask- is the observed relationship happening by chance or is it systematic?
  - it tells us what the world would look like under a certain assumption.
- Most of the probability material is in the first two weeks but we will return to these ideas periodically through the semester.

# Example: The Lady Tasting Tea

- The Story Setup
  (lady discerning about tea)
- The Experiment
  (perform a taste test)
- The Hypothetical
  (count possibilities)
- The Result
  (boom she was right)

**Tea-Tasting Distribution**

| Success count | Permutations of selection | Number of permutations |
|---|---|---|
| 0 | oooo | 1 × 1 = 1 |
| 1 | ooox, ooxo, oxoo, xooo | 4 × 4 = 16 |
| 2 | ooxx, oxox, oxxo, xoxo, xxoo, xoox | 6 × 6 = 36 |
| 3 | oxxx, xoxx, xxox, xxxo | 4 × 4 = 16 |
| 4 | xxxx | 1 × 1 = 1 |
| **Total** | | 70 |

This became the Fisher Exact Test.

# A Note on Fisher and the History of Statistics

- The statistician in that story was Sir Ronald Fisher, arguably the most influential statistician of the 20th century.
- Besides founding key areas of statistics, Fisher was also one of the founders of population genetics.
- He was also a eugenicist and a racist.
- Statistics has been used intermittently as a force for progress and a force against progress.

# Preview: Connecting Theory and Evidence

"[Variables] empirically perform as theoretically predicted,
by displaying statistically significant
effects net of other variables in the right direction"

Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics

The **target tautology**:

Research goals are defined by hypotheses about model coefficients

↓

The goal is only defined within the statistical model

↓

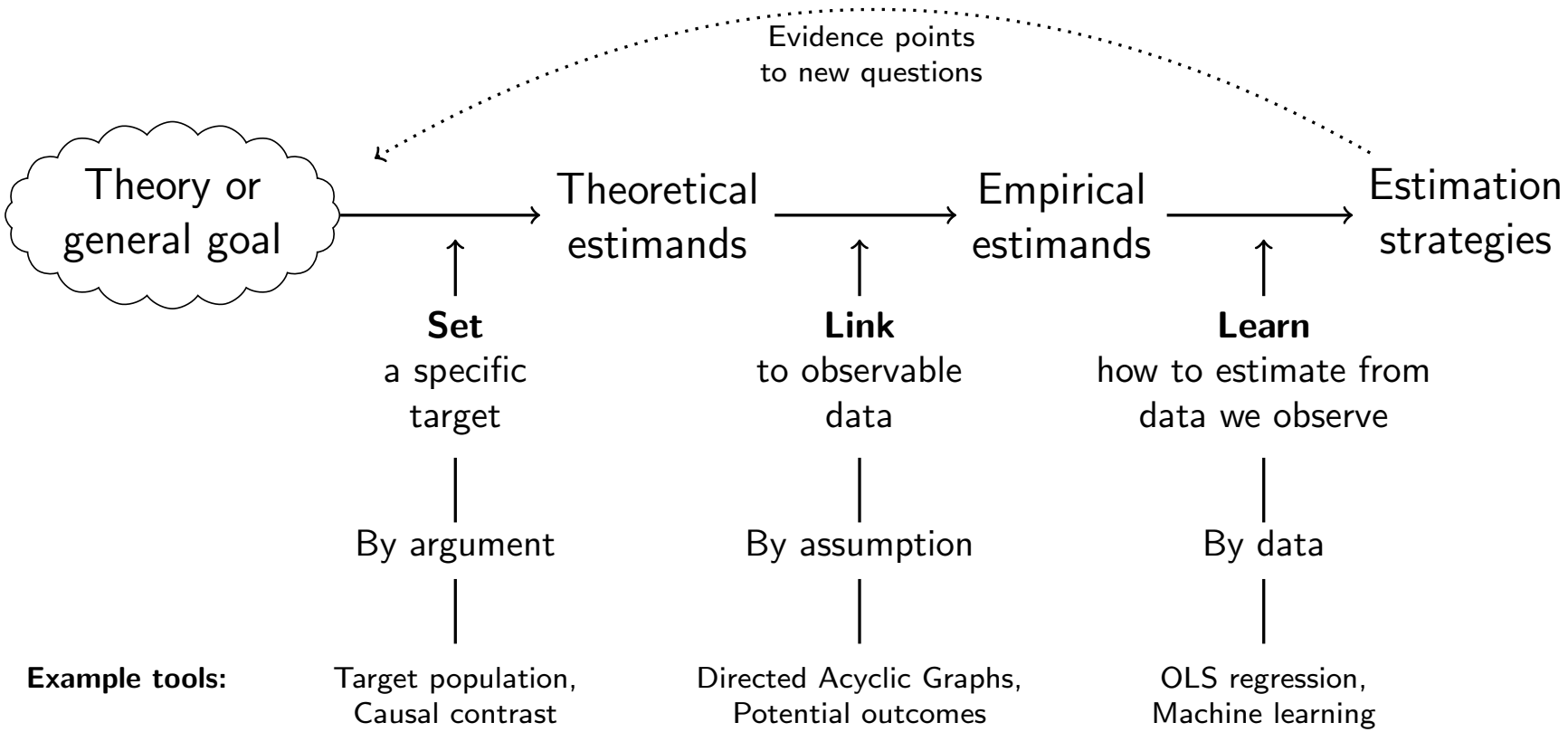
It becomes impossible to reason about other estimation strategies

**Solution**:

Our **diagnosis** for the source of many methodological problems

State the research goal separately from the estimation strategy

# Connecting Theory and Evidence



# We Covered...

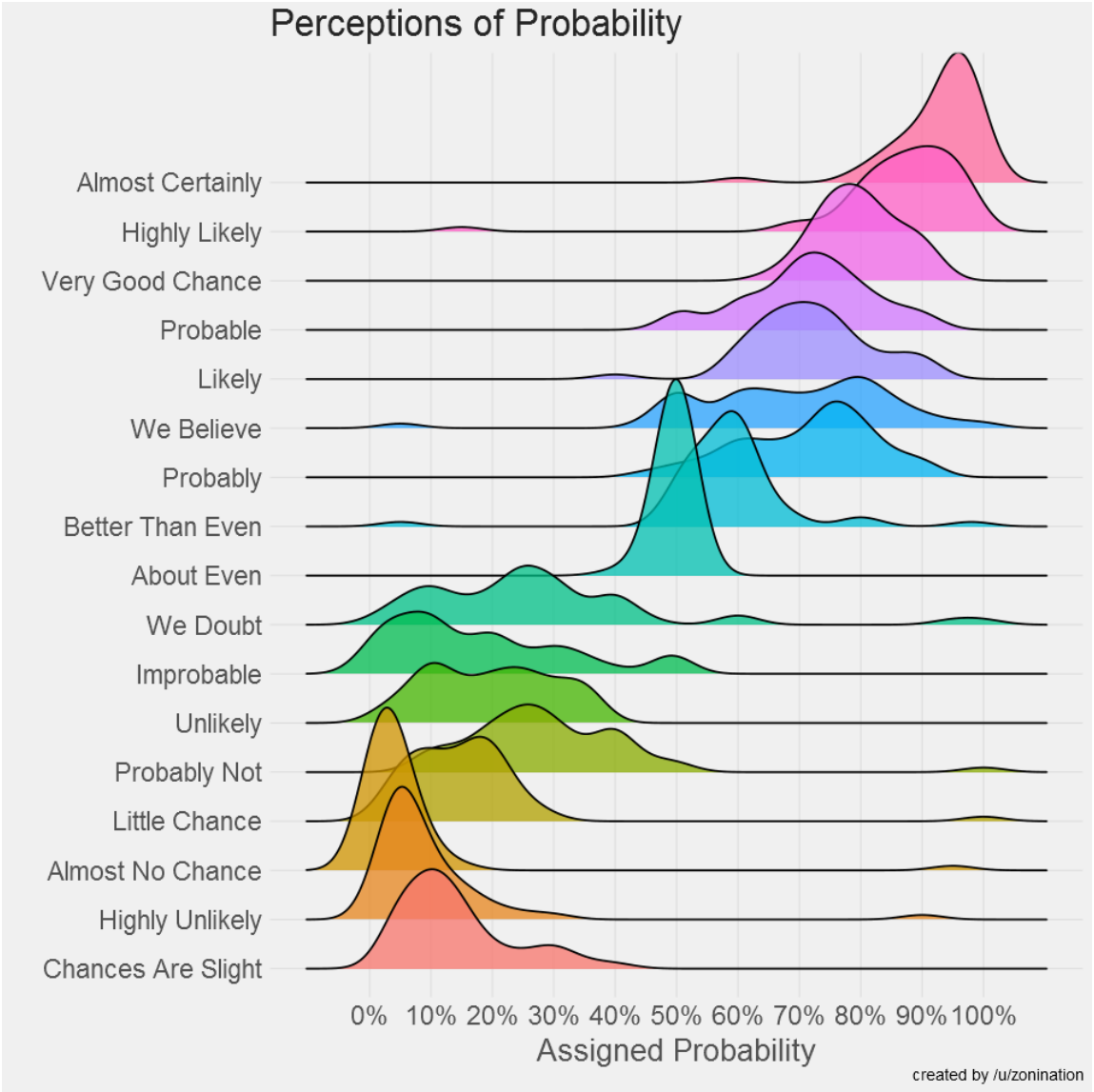
- Statistics as a field (the good, the bad and the ugly)
- The probability and inference loop
- Connecting theory and evidence through estimands

See you next time!

# Where We've Been and Where We're Going...

- Last Week
  - living that class-free, quarantine life
- This Week
  - course structure
  - core ideas
  - introduction to probability
  - three big ideas in probability
- Next Week
  - random variables
  - joint distributions
- Long Run
  - probability $\rightarrow$ inference $\rightarrow$ regression $\rightarrow$ causal inference

# From 'Probably' to Probability



# Why Probability?

- Helps us envision hypotheticals
- Describes uncertainty in how the data is generated
- Estimates probability that something will happen
- Thus: we need to know how probability gives rise to data

# Intuitive Definition of Probability

While there are several interpretations of what probability is, most modern (post 1935 or so) researchers agree on an axiomatic definition of probability.

3 Axioms (Intuitive Version):

1. The probability of any particular event must be non-negative.
2. The probability of anything occurring among all possible events must be 1.
3. The probability of one of many mutually exclusive events happening is the sum of the individual probabilities.

All the rules of probability can be derived from these axioms.
To state them formally, we first need some definitions.

# Sample Spaces

To define probability we need to define the set of possible outcomes.

The sample space is the set of all possible outcomes, and is often written as **S**.

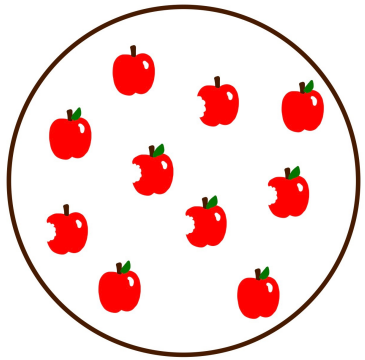
For example, if we flip a coin twice, there are four possible outcomes,

$$\mathbf{S} = \{\{heads, heads\}, \{heads, tails\}, \{tails, heads\}, \{tails, tails\}\}$$

Thus the table in Lady Tasting Tea was defining the sample space. (Note we defined illogical guesses to be prob$= 0$)

# A Running Visual Metaphor

Imagine that we sample one apple from a bag.
Looking in the bag we see:



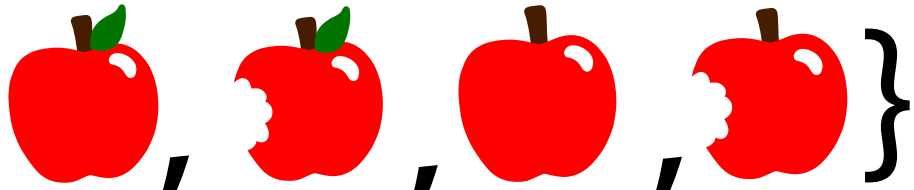
The sample space is:

$$S = \{🍎, 🍎, 🍎, 🍎\}$$

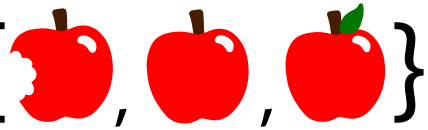
# Events

Events are subsets of the sample space.

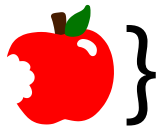
For example, if

$$S = \{$$  $$\}$$

then

$\{$  $\}$

and

$\{$  $\}$

are both **events.**

# Events Are a Kind of Set

Sets are collections of things, in this case collections of outcomes

One way to define an event is to describe the common property that all of the outcomes share. We write this as

$$\{\omega | \omega \text{ satisfies Property they share}\},$$

Example:

If $A = \{\omega | \omega$ has a leaf$\}$:

# Complement

A complement of event $A$, denoted $A^c$, is also a set.

$A^c$, is everything else not in $A$.

and

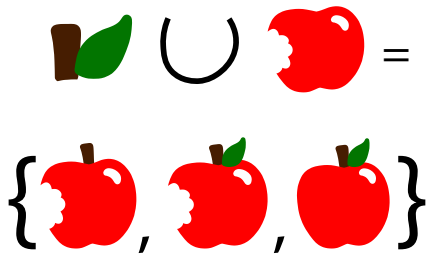
are **complements.**

$A^c = \{\omega \in \mathbf{S} | \omega \notin A\}$.

Important complement: $\mathbf{S}^c = \emptyset$, where $\emptyset$ is the empty set.

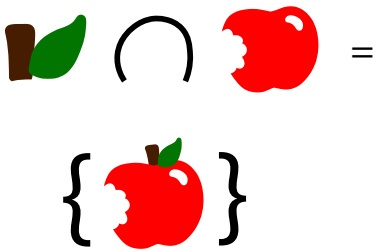
# Unions and Intersections

The union of two events, $A$ and $B$ is the event that $A$ or $B$ occurs:



$$A \cup B = \{\omega | \omega \in A \text{ or } \omega \in B\}.$$

The intersection of two events, $A$ and $B$ is the event that both $A$ and $B$ occur:



$$A \cap B = \{\omega | \omega \in A \text{ and } \omega \in B\}.$$

# Operations on Events

We say that two events $A$ and $B$ are disjoint or mutually exclusive if they don't share any elements or that $A \cap B = \emptyset$.

An event and its complement $A$ and $A^c$ are by definition disjoint.

Sample spaces can have infinite events where we will often write the different events using subscripts of the same letter: $A_1, A_2, \ldots A_\infty$ (e.g. imagine an event that was the count of some object)

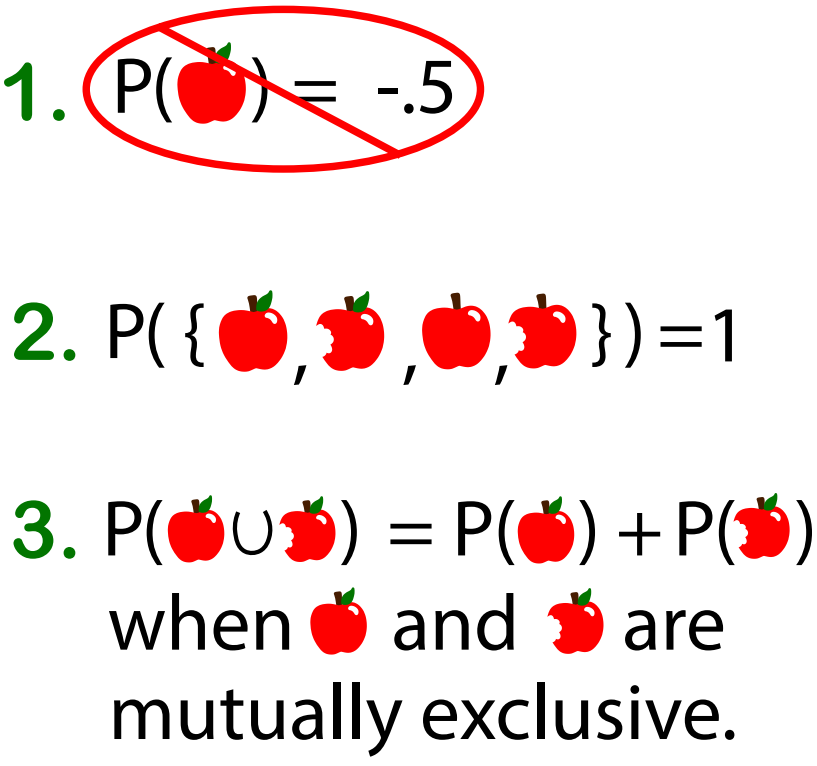
# Probability Function

A probability function $P(\cdot)$ is a function defined over all subsets of a sample space **S** that satisfies the following three axioms:

1) $P(A) \geq 0$ for all $A$ in the set of all events. nonnegativity

2) $P(\mathbf{S}) = 1$ normalization

3) if events $A_1, A_2, \ldots$ are mutually exclusive then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$. additivity




All the rules of probability can be derived from these axioms.
(See Blitzstein & Hwang, Def 1.6.1.)

# A Brief Word on Interpretation

Massive debate on interpretation:

- Subjective Interpretation
  - Example: The probability of drawing 5 red cards out of 10 drawn from a deck of cards is whatever you want it to be. But...
  - If you don't follow the axioms, a bookie can beat you
  - There is a correct way to update your beliefs given your assumptions about the data generating process.
- Frequency Interpretation
  - Probability is the relative frequency with which an event would occur if the process were repeated a large number of times under similar conditions.
  - Example: The probability of drawing 5 red cards out of 10 drawn from a deck of cards is the frequency with which this event occurs in repeated samples of 10 cards.

# We Covered...

- Events and Sample Spaces
- Probability Functions and Three Axioms
- Next: Three Big Ideas derived from the axioms that provide the rules of working with probability.

See you next time!

# Where We've Been and Where We're Going...

- Last Week
  - living that class-free, quarantine life
- This Week
  - course structure
  - core ideas
  - introduction to probability
  - three big ideas in probability
- Next Week
  - random variables
  - joint distributions
- Long Run
  - probability $\rightarrow$ inference $\rightarrow$ regression $\rightarrow$ causal inference

# Three Big Ideas

Marginal, joint, and conditional probabilities

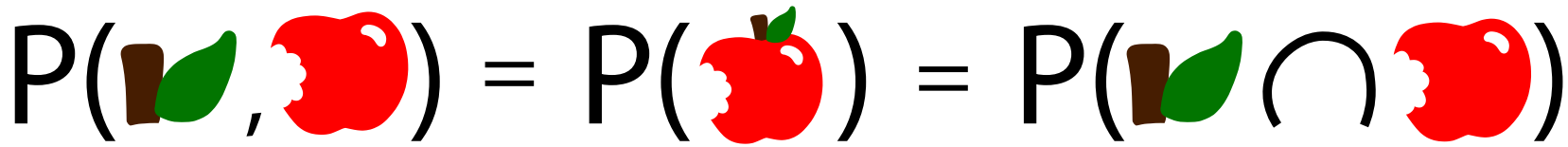
Bayes' rule

Independence

# Marginal and Joint Probability

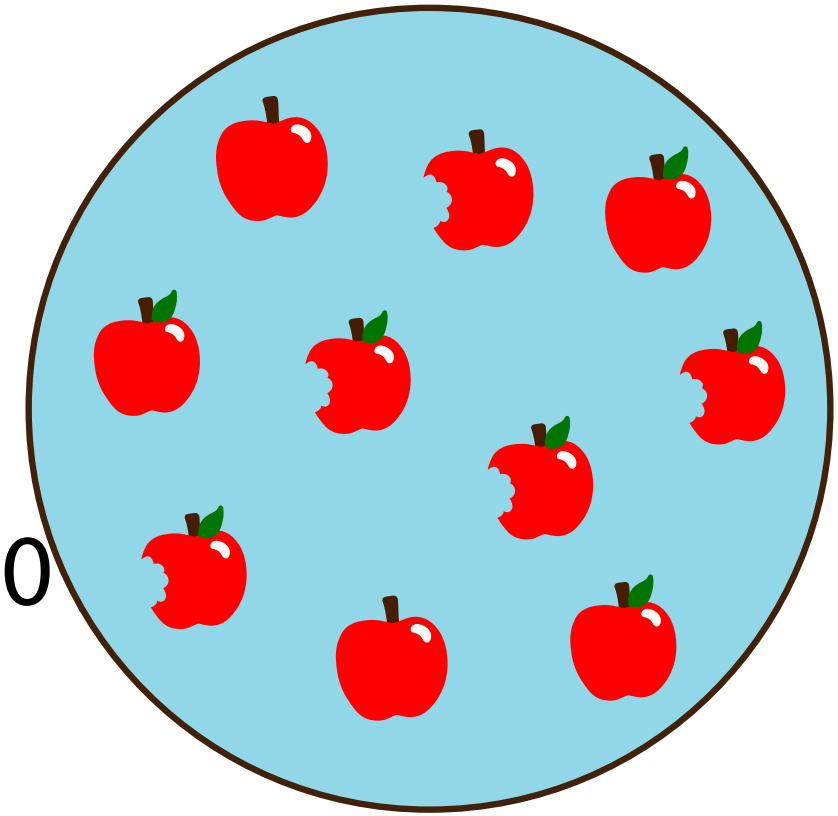
So far we have only considered situations where we are interested in the probability of a single event $A$ occurring. We've denoted this $P(A)$. $P(A)$ is sometimes called a marginal probability.

Suppose we are now in a situation where we would like to express the probability that an event $A$ and an event $B$ occur. This quantity is written as $P(A \cap B)$, $P(B \cap A)$, $P(A, B)$, or $P(B, A)$ and is the joint probability of $A$ and $B$.



$P(\text{leaf}) = 7/10$

$P(\text{leaf}, \text{apple}) = 4/10$



# Conditional Probability

The "soul of statistics"
If $P(A) > 0$ then the probability of $B$ conditional on $A$ is

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

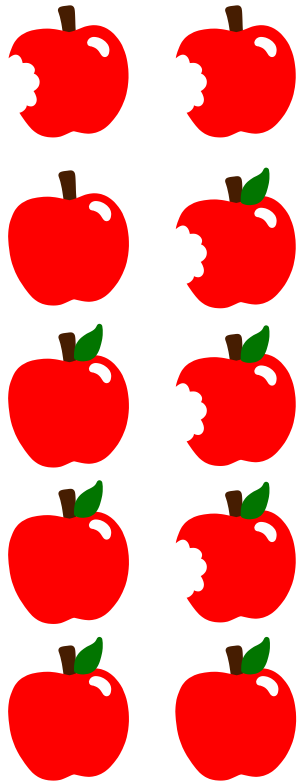
This implies that

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

Hopefully this second formulation is intuitive!

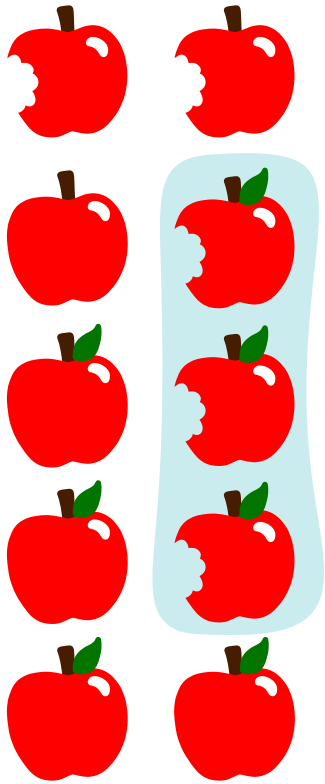
# Conditional Probability: A Visual Example

$$P(\text{bitten} \mid \text{bitten}) = \frac{P(\text{leaf}, \text{bitten})}{P(\text{bitten})}$$



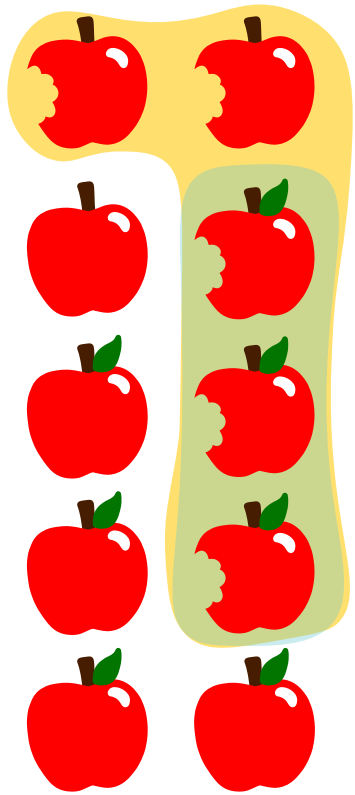
# Conditional Probability: A Visual Example

$$P(\text{leaf} \mid \text{apple}) = \frac{P(\text{leaf}, \text{apple})}{P(\text{apple})}$$



# Conditional Probability: A Visual Example

$$P(\text{leaf} \mid \text{apple}) = \frac{P(\text{leaf}, \text{apple})}{P(\text{apple})}$$



# Law of Total Probability (LTP)

With 2 Events:

$$\begin{aligned} P(B) &= P(B, A) + P(B, A^c) \\ &= P(B|A)P(A) + P(B|A^c)P(A^c) \end{aligned}$$

P( ) = P( ) + P( )

= P( | ) x P( ) + P( | ) x P( )

In general, if $\{A_i : i = 1, 2, 3, \dots\}$ forms a partition of the sample space, then

$$\begin{aligned} P(B) &= \sum_i P(B, A_i) \\ &= \sum_i P(B|A_i)P(A_i) \end{aligned}$$

# Example: Voter Mobilization

Suppose that we have put together a voter mobilization campaign and we want to know what the probability of voting is after the campaign: $P(\text{vote})$. We know the following:

- $P(\text{vote}|\text{mobilized}) = 0.75$
- $P(\text{vote}|\text{not mobilized}) = 0.15$
- $P(\text{mobilized}) = 0.6$ and so $P(\text{not mobilized}) = 0.4$

Note that mobilization partitions the data. Everyone is either mobilized or not. Thus, we can apply the LTP:

$$\begin{aligned} P(\text{vote}) =& P(\text{vote}|\text{mobilized})P(\text{mobilized}) + \\ & P(\text{vote}|\text{not mobilized})P(\text{not mobilized}) \\ =& 0.75 \times 0.6 + 0.15 \times 0.4 \\ =& .51 \end{aligned}$$

# Three Big Ideas

Marginal, joint, and conditional probabilities

Bayes' rule

Independence

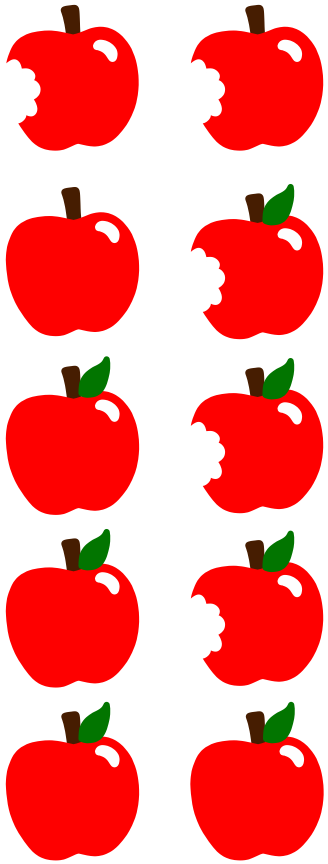
# Bayes' Rule

- Often we have information about $P(B|A)$, but want $P(A|B)$.
- When this happens, always think: Bayes' rule
- Bayes' rule: if $P(B) > 0$, then:

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

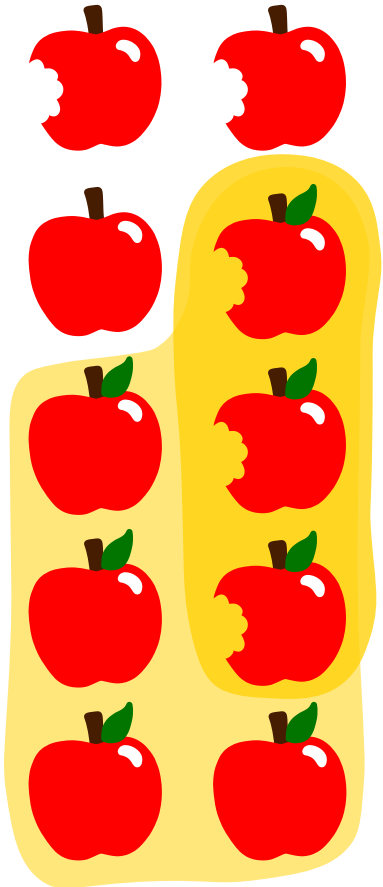
# Bayes' Rule Mechanics

$$P(\text{leaf}|\text{apple}) = \frac{P(\text{apple}|\text{leaf})\,P(\text{leaf})}{P(\text{apple})}$$



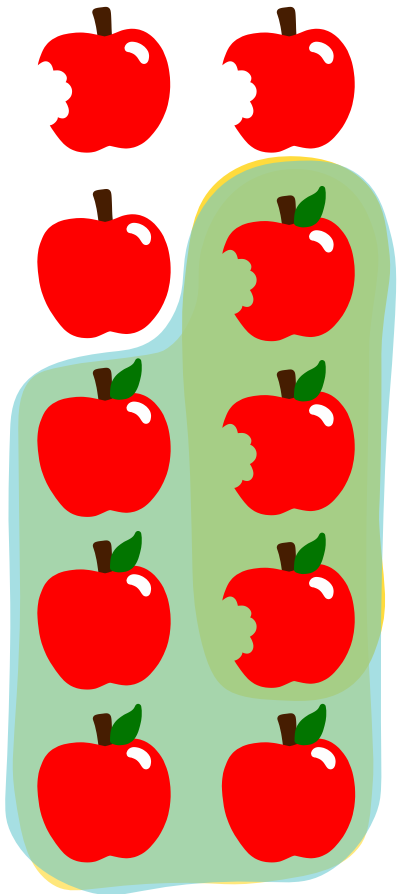
# Bayes' Rule Mechanics

$$P(\text{leaf}|\text{apple}) = \frac{P(\text{apple}|\text{leaf})P(\text{leaf})}{P(\text{apple})}$$



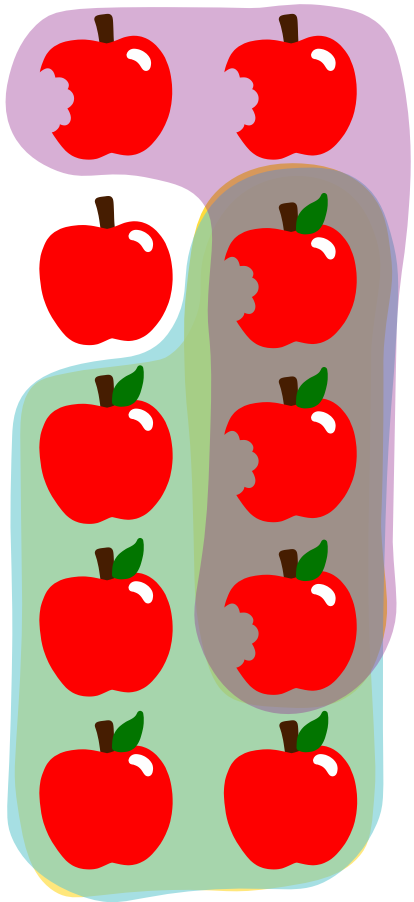
# Bayes' Rule Mechanics

$$P(\text{leaf} \mid \text{apple}) = \frac{P(\text{apple} \mid \text{leaf})\, P(\text{leaf})}{P(\text{apple})}$$



# Bayes' Rule Mechanics

$$P(\text{leaf} \mid \text{apple}) = \frac{P(\text{apple} \mid \text{leaf})\, P(\text{leaf})}{P(\text{apple})}$$



# Example: Race and Names

## What the Demolition of Public Housing Teaches Us about the Impact of Racial Threat on Political Behavior

**Ryan D. Enos** Harvard University

*How does the context in which a person lives affect his or her political behavior? I exploit an event in which demographic context was exogenously changed, leading to a significant change in voters' behavior and demonstrating that voters react strongly to changes in an outgroup population. Between 2000 and 2004, the reconstruction of public housing in Chicago caused the displacement of over 25,000 African Americans, many of whom had previously lived in close proximity to white voters. After the removal of their African American neighbors, the white voters' turnout dropped by over 10 percentage points. Consistent with psychological theories of racial threat, their change in behavior was a function of the size and proximity of the outgroup population. Proximity was also related to increased voting for conservative candidates. These findings strongly suggest that racial threat occurs because of attitude change rather than selection.*

# Example: Race and Names

- Note that the Census collects information on the distribution of names by race.
- For example, Washington is the most common last name among African-Americans in America:
  - $P(\text{AfAm}) = 0.132$
  - $P(\text{not AfAm}) = 1 - P(\text{AfAm}) = .868$
  - $P(\text{Washington}|\text{AfAm}) = 0.00378$
  - $P(\text{Washington}|\text{not AfAm}) = 0.000061$
- We can now use Bayes' Rule

$$P(\text{AfAm}|\text{Wash}) = \frac{P(\text{Wash}|\text{AfAm})P(\text{AfAm})}{P(\text{Wash})}$$

# Example: Race and Names

Note we don't have the probability of the name Washington.

Remember that we can calculate it from the LTP since the sets African-American and not African-American partition the sample space:

$$\begin{aligned} P(\text{AfAm}|\text{Wash}) &= \frac{P(\text{Wash}|\text{AfAm})P(\text{AfAm})}{P(\text{Wash})} \\ &= \frac{P(\text{Wash}|\text{AfAm})P(\text{AfAm})}{P(\text{Wash}|\text{AfAm})P(\text{AfAm}) + P(\text{Wash}|\text{not AfAm})P(\text{not AfAm})} \\ &= \frac{0.132 \times 0.00378}{0.132 \times 0.00378 + .868 \times 0.000061} \\ &\approx 0.9 \end{aligned}$$

# Three Big Ideas

Marginal, joint, and conditional probabilities

Bayes' rule

Independence

# Independence

## Intuitive Definition

Events A and B are independent if knowing whether A occurred provides no information about whether B occurred.

## Formal Definition

$$P(A, B) = P(A)P(B) \implies A \perp\!\!\!\perp B$$

With all the usual $> 0$ restrictions, this implies

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

## Conditional Independence

$$P(A, B|C) = P(A|C)P(B|C) \implies A \perp\!\!\!\perp B|C$$

Independence is a massively important concept in statistics.

# Independence, the Heroic Assumption

Deploy with Caution!

# Advanced Example: Building a Spam Filter

Suppose we have an email $i$, $(i = 1, \ldots, N)$ which we represent with a series of $J$ indicators for whether or not it contains a set of words

$$\boldsymbol{x}_i = (x_{1i}, x_{2i}, \ldots, x_{Ji})$$

We want to classify these into one of two categories: spam or not

$$\{C_{\text{spam}}, C_{\text{not}}\}$$

We have a set of labeled documents $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_N)$ where $Y_i \in \{C_{\text{spam}}, C_{\text{not}}\}$.

Goal: Use what we've learned to build a model which can classify emails into spam and not spam.

# Example: Building a Spam Filter

For each document, we will get to see $x_i$ (the words in the document), and we would like to infer the category.

In other words what we want is $P(C_{\text{spam}}|x_i)$.

Let's use Bayes' Rule!

$$\begin{aligned} P(C_{\text{spam}}|x_i) &= \frac{P(x_i|C_{\text{spam}})P(C_{\text{spam}})}{P(x_i)} \\ &= \frac{P(x_i|C_{\text{spam}})P(C_{\text{spam}})}{P(x_i|C_{\text{spam}})P(C_{\text{spam}}) + P(x_i|C_{\text{not spam}})P(C_{\text{not spam}})} \end{aligned}$$

We used the law of total probability to work out the bottom.

Now there are only 4 pieces we need (2 for spam, 2 for not)

# Estimating the Baseline Prevalence

Let's plug in some estimates based on our labeled emails.

Intuitively, $P(C_{\text{spam}})$ is the probability that a randomly chosen email will be spam.

$$P(C_{\text{spam}}) = \frac{\text{No. Spam Emails}}{\text{No. Emails}}$$

Because 'not spam' is the complement of spam we know that:

$$P(C_{\text{not spam}}) = 1 - P(C_{\text{spam}})$$

Note: this estimate is only good if our labeled emails are a random sample of all emails! More on this in future weeks.

# Estimating the Language Model

Now we need $P(\boldsymbol{x}_i|C_{\text{spam}}$ which we call the language model because it represents the probability of seeing any combination of the $J$ words that we are counting from the emails.

Can we use the same strategy as before (just counting up emails)? No! Remember $\boldsymbol{x}_i$ is a vector of $J$ words, that is $2^J$ possibilities!

We will make the heroic assumption of conditional independence:

$$P(\boldsymbol{x}_i|C_{\text{spam}}) = \prod_{j=1}^{J} P(x_{ij}|C_{\text{spam}})$$

Intuition: count the proportion of spam emails containing each word.

Called Naïve Bayes classifier because the conditional independence assumption is *naïve*.

# Estimating the Naïve Bayes Classifier

$$P(C_{\text{spam}}|\boldsymbol{x}_i) = \frac{P(\boldsymbol{x}_i|C_{\text{spam}})P(C_{\text{spam}})}{P(\boldsymbol{x}_i|C_{\text{spam}})P(C_{\text{spam}}) + P(\boldsymbol{x}_i|C_{\text{not spam}})P(C_{\text{not spam}})}$$

The Naïve Bayes Procedure:

- Learn what spam emails look like to create a function that lets us plug in an email and get out a probability, $P(\boldsymbol{x}_i|C_{\text{spam}})$
- Guess how much spam there is overall, $P(C_{\text{spam}})$
- Plug in values of $\boldsymbol{x}_i$ for new emails to score them by whether they are spam or not.
- . . . Profit?

# Example: Building a Spam Filter

- This was a really advanced example (it is okay if you didn't follow all of it!).
- Draws on all the probabilistic concepts we have introduced:
  - Bayes' Rule
  - Law of Total Probability
  - Conditional Independence
- Shares the basic structure of many models particularly in use of conditional independence.

# This Week in Review

- Course logistics
- Core ideas in statistics
- Foundations of probability
- Three big probability concepts

Going Deeper:

> Blitzstein, Joseph K., and Hwang, Jessica. (2019). *Introduction to Probability*. CRC Press. `http://stat110.net/`

Next week: random variables!

# References


- Enos, Ryan D. "What the demolition of public housing teaches us about the impact of racial threat on political behavior." *American Journal of Political Science* (2015).
- Lundberg, Ian, Rebecca Johnson, and Brandon M. Stewart. "Setting the target: Precise estimands and the gap between theory and empirics." (2020).
- Salsburg, David. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century* (2002).