2:Conditional Probability

Naijia Liu

Spring 2025

Overview

1. Conditional Probability
2. Bayes's Rule
3. Independence

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?
  - Conditioning our analysis on B having occurred.

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?
  - Conditioning our analysis on B having occurred.
- **Examples:**

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?
  - Conditioning our analysis on B having occurred.
- **Examples:**
  - What is the probability of two states going to war **if** they are both democracies?

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?
    - Conditioning our analysis on B having occurred.
- **Examples:**
    - What is the probability of two states going to war **if** they are both democracies?
    - What is the probability of a judge ruling in a pro-choice direction **conditional** on having daughters?

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?
  - Conditioning our analysis on B having occurred.
- **Examples:**
  - What is the probability of two states going to war **if** they are both democracies?
  - What is the probability of a judge ruling in a pro-choice direction **conditional** on having daughters?
  - What is the probability that there will be a coup in a country **conditional** on having a presidential system?

Conditional probability

- **Conditional probability:** if we know that B has occurred, what is the probability of A?
  - Conditioning our analysis on B having occurred.
- **Examples:**
  - What is the probability of two states going to war **if** they are both democracies?
  - What is the probability of a judge ruling in a pro-choice direction **conditional** on having daughters?
  - What is the probability that there will be a coup in a country **conditional** on having a presidential system?
- Conditional probability is the cornerstone of quantitative social science.

Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

- How often A and B occur divided by how often B occurs.

Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$
- How often A and B occur divided by how often B occurs.
- **WARNING!** $\mathbb{P}(A \mid B)$ does **not**, in general, equal $\mathbb{P}(B \mid A)$.

Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

- How often A and B occur divided by how often B occurs.
- **WARNING!** $\mathbb{P}(A \mid B)$ does **not**, in general, equal $\mathbb{P}(B \mid A)$.
    - $\mathbb{P}(\text{smart} \mid \text{in gov 2001})$ is high.

Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$
- How often A and B occur divided by how often B occurs.
- **WARNING!** $\mathbb{P}(A \mid B)$ does **not**, in general, equal $\mathbb{P}(B \mid A)$.
  - $\mathbb{P}(\text{smart} \mid \text{in gov 2001})$ is high.
  - $\mathbb{P}(\text{in gov 2001} \mid \text{smart})$ is low.

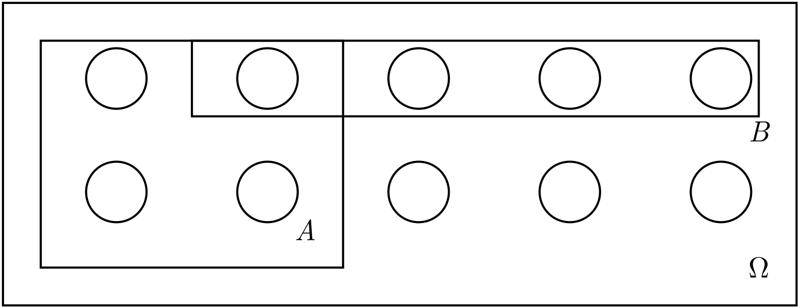
Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$
- How often A and B occur divided by how often B occurs.
- **WARNING!** $\mathbb{P}(A \mid B)$ does **not**, in general, equal $\mathbb{P}(B \mid A)$.
  - $\mathbb{P}(\text{smart} \mid \text{in gov 2001})$ is high.
  - $\mathbb{P}(\text{in gov 2001} \mid \text{smart})$ is low.
  - There are many smart people who are not in this class!

Conditional Probability Definition

- Definition: If $\mathbb{P}(B) > 0$ then the **conditional probability** of A given B is
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$
- How often A and B occur divided by how often B occurs.
- **WARNING!** $\mathbb{P}(A \mid B)$ does **not**, in general, equal $\mathbb{P}(B \mid A)$.
  - $\mathbb{P}(\text{smart} \mid \text{in gov 2001})$ is high.
  - $\mathbb{P}(\text{in gov 2001} \mid \text{smart})$ is low.
  - There are many smart people who are not in this class!
  - Also known as the **prosecutor's fallacy**.

Intuition



Examples

$A = \{\text{you get an A grade}\} \quad B = \{\text{everyone gets an A grade}\}$

Examples

$A = \{\text{you get an A grade}\} \quad B = \{\text{everyone gets an A grade}\}$

- If B occurs then A must also occur, so $\mathbb{P}(A \mid B) = 1$.

Examples

$A = \{\text{you get an A grade}\} \quad B = \{\text{everyone gets an A grade}\}$

- If B occurs then A must also occur, so $\mathbb{P}(A \mid B) = 1$.
  - Does this mean that $\mathbb{P}(B \mid A) = 1$ as well?

Examples

$A = \{\text{you get an A grade}\} \quad B = \{\text{everyone gets an A grade}\}$

- If B occurs then A must also occur, so $\mathbb{P}(A \mid B) = 1$.
  - Does this mean that $\mathbb{P}(B \mid A) = 1$ as well?
- Now let $A = \{\text{you get a B grade}\}$.

Examples

$A = \{\text{you get an A grade}\}$ $\quad B = \{\text{everyone gets an A grade}\}$

- If B occurs then A must also occur, so $\mathbb{P}(A \mid B) = 1$.
  - Does this mean that $\mathbb{P}(B \mid A) = 1$ as well?
- Now let $A = \{\text{you get a B grade}\}$.
  - The intersection $A \cap B = \emptyset$, so that $\mathbb{P}(A \mid B) = 0$.

Examples

$A = \{\text{you get an A grade}\} \quad B = \{\text{everyone gets an A grade}\}$

- If B occurs then A must also occur, so $\mathbb{P}(A \mid B) = 1$.
  - Does this mean that $\mathbb{P}(B \mid A) = 1$ as well?
- Now let $A = \{\text{you get a B grade}\}$.
  - The intersection $A \cap B = \emptyset$, so that $\mathbb{P}(A \mid B) = 0$.
  - Intuitively, it's because B occurring precludes A from occurring.

U.S. Senate Example

	Democrats	**Republicans**	**Independents**	**Total**
Men	33	40	2	75
Women	15	9	1	25
Total	48	49	3	100

U.S. Senate Example

	Democrats	**Republicans**	**Independents**	**Total**
Men	33	40	2	75
Women	15	9	1	25
Total	48	49	3	100

- Choose one senator at random from this population.

U.S. Senate Example

	Democrats	**Republicans**	**Independents**	**Total**
Men	33	40	2	75
Women	15	9	1	25
Total	48	49	3	100

- Choose one senator at random from this population.
- What is the probability that a randomly selected Republican is a woman:

$$\mathbb{P}(\text{Woman} \mid \text{Republican}) = \frac{\mathbb{P}(\text{Woman} \cap \text{Republican})}{\mathbb{P}(\text{Republican})} = \frac{9/100}{49/100} = \frac{9}{49} \approx 0.184.$$

U.S. Senate Example

	Democrats	Republicans	Independents	Total
Men	33	40	2	75
Women	15	9	1	25
Total	48	49	3	100

- Choose one senator at random from this population.
- What is the probability that a randomly selected Republican is a woman:

$$\mathbb{P}(\text{Woman} \mid \text{Republican}) = \frac{\mathbb{P}(\text{Woman} \cap \text{Republican})}{\mathbb{P}(\text{Republican})} = \frac{9/100}{49/100} = \frac{9}{49} \approx 0.184.$$

- Choose two senators at random:

U.S. Senate Example

	Democrats	Republicans	Independents	Total
Men	33	40	2	75
Women	15	9	1	25
Total	48	49	3	100

- Choose one senator at random from this population.
- What is the probability that a randomly selected Republican is a woman:

$$\mathbb{P}(\text{Woman} \mid \text{Republican}) = \frac{\mathbb{P}(\text{Woman} \cap \text{Republican})}{\mathbb{P}(\text{Republican})} = \frac{9/100}{49/100} = \frac{9}{49} \approx 0.184.$$

- Choose two senators at random:
  - $\mathbb{P}(\text{2 women} \mid \text{one draw is a woman})$?

U.S. Senate Example

	Democrats	Republicans	Independents	Total
Men	33	40	2	75
Women	15	9	1	25
Total	48	49	3	100

- Choose one senator at random from this population.
- What is the probability that a randomly selected Republican is a woman:

$$\mathbb{P}(\text{Woman} \mid \text{Republican}) = \frac{\mathbb{P}(\text{Woman} \cap \text{Republican})}{\mathbb{P}(\text{Republican})} = \frac{9/100}{49/100} = \frac{9}{49} \approx 0.184.$$

- Choose two senators at random:
  - $\mathbb{P}(\text{2 women} \mid \text{one draw is a woman})$?
  - $\mathbb{P}(\text{2 women} \mid \text{one draw is Liz Warren})$?

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:
  1. $\mathbb{P}(A \mid B) \geq 0$

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:
    1. $\mathbb{P}(A \mid B) \geq 0$
    2. $\mathbb{P}(\Omega \mid B) = 1$

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:
  1. $\mathbb{P}(A \mid B) \geq 0$
  2. $\mathbb{P}(\Omega \mid B) = 1$
  3. If A_1 and A_2 are disjoint, then
     $\mathbb{P}(A_1 \cup A_2 \mid B) = \mathbb{P}(A_1 \mid B) + \mathbb{P}(A_2 \mid B)$

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:
  1. $\mathbb{P}(A \mid B) \geq 0$
  2. $\mathbb{P}(\Omega \mid B) = 1$
  3. If A_1 and A_2 are disjoint, then $\mathbb{P}(A_1 \cup A_2 \mid B) = \mathbb{P}(A_1 \mid B) + \mathbb{P}(A_2 \mid B)$
- $\sim$ Rules of probability apply to the left-hand side of the conditioning bar (A):

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:
  1. $\mathbb{P}(A \mid B) \geq 0$
  2. $\mathbb{P}(\Omega \mid B) = 1$
  3. If A_1 and A_2 are disjoint, then $\mathbb{P}(A_1 \cup A_2 \mid B) = \mathbb{P}(A_1 \mid B) + \mathbb{P}(A_2 \mid B)$
- $\sim$ Rules of probability apply to the left-hand side of the conditioning bar (A):
  - All probabilities **normalized** to event B, $\mathbb{P}(B \mid B) = 1$.

Conditional Probabilities Are Probabilities

- Conditional probabilities $\mathbb{P}(A \mid B)$ are valid probability functions:
  1. $\mathbb{P}(A \mid B) \geq 0$
  2. $\mathbb{P}(\Omega \mid B) = 1$
  3. If A_1 and A_2 are disjoint, then $\mathbb{P}(A_1 \cup A_2 \mid B) = \mathbb{P}(A_1 \mid B) + \mathbb{P}(A_2 \mid B)$
- $\sim$ Rules of probability apply to the left-hand side of the conditioning bar (A):
  - All probabilities **normalized** to event B, $\mathbb{P}(B \mid B) = 1$.
- Not for the right-hand side, so even if B and C are disjoint,

$$\mathbb{P}(A \mid B \cup C) \neq \mathbb{P}(A \mid B) + \mathbb{P}(A \mid C).$$

Joint Probabilities from Conditionals

- **Joint probabilities:** probability of two events occurring (intersections)

Joint Probabilities from Conditionals

- **Joint probabilities:** probability of two events occurring (intersections)
  - Often replace $\cap$ with commas: $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A, B, C)$

Joint Probabilities from Conditionals

- **Joint probabilities:** probability of two events occurring (intersections)
  - Often replace $\cap$ with commas: $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A, B, C)$
- Definition of conditional probability implies:

$$\mathbb{P}(A \cap B) \equiv \mathbb{P}(A, B) = \mathbb{P}(B)\mathbb{P}(A \mid B) = \mathbb{P}(A)\mathbb{P}(B \mid A)$$

Joint Probabilities from Conditionals

- **Joint probabilities:** probability of two events occurring (intersections)
  - Often replace $\cap$ with commas: $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A, B, C)$
- Definition of conditional probability implies:

$$\mathbb{P}(A \cap B) \equiv \mathbb{P}(A, B) = \mathbb{P}(B)\mathbb{P}(A \mid B) = \mathbb{P}(A)\mathbb{P}(B \mid A)$$

- What about three events?

$$\mathbb{P}(A, B, C) = \mathbb{P}(A)\mathbb{P}(B \mid A)\mathbb{P}(C \mid A, B)$$

Joint Probabilities from Conditionals

- **Joint probabilities:** probability of two events occurring (intersections)
  - Often replace $\cap$ with commas: $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A, B, C)$
- Definition of conditional probability implies:

$$\mathbb{P}(A \cap B) \equiv \mathbb{P}(A, B) = \mathbb{P}(B)\mathbb{P}(A \mid B) = \mathbb{P}(A)\mathbb{P}(B \mid A)$$

- What about three events?

$$\mathbb{P}(A, B, C) = \mathbb{P}(A)\mathbb{P}(B \mid A)\mathbb{P}(C \mid A, B)$$

- Generalize to the intersection of N events:

$$\mathbb{P}(A_1, \ldots, A_N) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1, A_2) \ldots \mathbb{P}(A_N \mid A_1, \ldots, A_{N-1}).$$

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.
- What's the probability that we draw three Aces?

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \mathbb{P}(\text{Ace}_1)\mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1)\mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1)$$

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.
- What's the probability that we draw three Aces?

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \mathbb{P}(\text{Ace}_1)\mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1)\mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1)$$

- What are these probabilities?

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.
- What's the probability that we draw three Aces?

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \mathbb{P}(\text{Ace}_1)\mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1)\mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1)$$

- What are these probabilities?
  - 4 Aces to pick out of 52 cards $\sim \mathbb{P}(\text{Ace}_1) = \frac{4}{52}$

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.
- What's the probability that we draw three Aces?

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \mathbb{P}(\text{Ace}_1)\mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1)\mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1)$$

- What are these probabilities?
  - 4 Aces to pick out of 52 cards $\sim \mathbb{P}(\text{Ace}_1) = \frac{4}{52}$
  - 3 Aces left in the 51 remaining cards $\sim \mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1) = \frac{3}{51}$

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.
- What's the probability that we draw three Aces?

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \mathbb{P}(\text{Ace}_1)\mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1)\mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1)$$

- What are these probabilities?
  - 4 Aces to pick out of 52 cards $\sim \mathbb{P}(\text{Ace}_1) = \frac{4}{52}$
  - 3 Aces left in the 51 remaining cards $\sim \mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1) = \frac{3}{51}$
  - 2 Aces left in the 50 remaining cards $\sim \mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1) = \frac{2}{50}$

Joint Probabilities, Example

- Draw three cards at random from a deck without replacement.
- What's the probability that we draw three Aces?

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \mathbb{P}(\text{Ace}_1)\mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1)\mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1)$$

- What are these probabilities?
  - 4 Aces to pick out of 52 cards $\sim \mathbb{P}(\text{Ace}_1) = \frac{4}{52}$
  - 3 Aces left in the 51 remaining cards $\sim \mathbb{P}(\text{Ace}_2 \mid \text{Ace}_1) = \frac{3}{51}$
  - 2 Aces left in the 50 remaining cards $\sim \mathbb{P}(\text{Ace}_3 \mid \text{Ace}_2 \cap \text{Ace}_1) = \frac{2}{50}$
- Thus,

$$\mathbb{P}(\text{Ace}_1 \cap \text{Ace}_2 \cap \text{Ace}_3) = \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50} \approx 0.00018.$$

Probability of War Resolution

- Suppose we observed country-dyads over 3 years.
- In each year the dyad can be at war (W_t) or at peace (P_t).
- What's the probability that a war starts in year 1 and ends after 2 years?

$$\mathbb{P}(W_1, W_2, P_3) = \mathbb{P}(W_1)\mathbb{P}(W_2 \mid W_1)\mathbb{P}(P_3 \mid W_1, W_2).$$

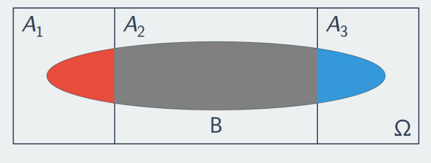
Probability of War Resolution

- Suppose we observed country-dyads over 3 years.
- In each year the dyad can be at war (W_t) or at peace (P_t).
- What's the probability that a war starts in year 1 and ends after 2 years?

$$\mathbb{P}(W_1, W_2, P_3) = \mathbb{P}(W_1)\mathbb{P}(W_2 \mid W_1)\mathbb{P}(P_3 \mid W_1, W_2).$$

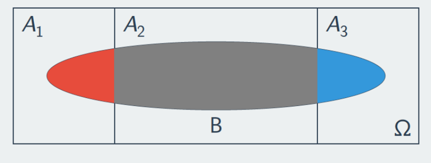
- **Actual Research Question™:** modeling the continuation probability of war, $\mathbb{P}(W_2 \mid W_1)$, and the probability of conflict resolution, $\mathbb{P}(P_3 \mid W_1, W_2)$.

Law of Total Probability



- Often we only have disaggregated probabilities.
  - $B =$ sampling a Trump supporter from either Cambridge or Somerville.
  - We know the prop. of Trump supporters in each city from precinct data.
  - How to calculate the overall probability of B?
- A **partition** is a set of mutually disjoint events whose union is Ω.

Law of Total Probability



- The **law of total probability** (LTP) states if $A_1, \dots, A_k$ is a partition:

$$\mathbb{P}(B) = \sum_{j=1}^{k} \mathbb{P}(B \mid A_j)\mathbb{P}(A_j).$$

  - Overall probability = weighted sum of within-partition probabilities.
  - Weights are the probability of the particular partition.

A Mixture of Cities

- Randomly drawing voters from either Cambridge or Somerville:

A Mixture of Cities

- Randomly drawing voters from either Cambridge or Somerville:
  - Camb. had 50k voters and Somer. had around 40k, so: $\mathbb{P}(\text{Camb.}) = 0.56$ and $\mathbb{P}(\text{Somer.}) = 0.44$.

A Mixture of Cities

- Randomly drawing voters from either Cambridge or Somerville:
  - Camb. had 50k voters and Somer. had around 40k, so: $\mathbb{P}(\text{Camb.}) = 0.56$ and $\mathbb{P}(\text{Somer.}) = 0.44$.
- The state provides the following election results for each city:

A Mixture of Cities

- Randomly drawing voters from either Cambridge or Somerville:
  - Camb. had 50k voters and Somer. had around 40k, so: $\mathbb{P}(\text{Camb.}) = 0.56$ and $\mathbb{P}(\text{Somer.}) = 0.44$.
- The state provides the following election results for each city:
  - $\mathbb{P}(\text{Trump} \mid \text{Camb.}) = 0.066$

A Mixture of Cities

- Randomly drawing voters from either Cambridge or Somerville:
  - Camb. had 50k voters and Somer. had around 40k, so: $\mathbb{P}(\text{Camb.}) = 0.56$ and $\mathbb{P}(\text{Somer.}) = 0.44$.
- The state provides the following election results for each city:
  - $\mathbb{P}(\text{Trump} \mid \text{Camb.}) = 0.066$
  - $\mathbb{P}(\text{Trump} \mid \text{Somer.}) = 0.103$.

A Mixture of Cities

- Randomly drawing voters from either Cambridge or Somerville:
  - Camb. had 50k voters and Somer. had around 40k, so: $\mathbb{P}(\text{Camb.}) = 0.56$ and $\mathbb{P}(\text{Somer.}) = 0.44$.
- The state provides the following election results for each city:
  - $\mathbb{P}(\text{Trump} \mid \text{Camb.}) = 0.066$
  - $\mathbb{P}(\text{Trump} \mid \text{Somer.}) = 0.103$.
- To get the overall turnout rate, $\mathbb{P}(\text{Trump})$, we can apply the LTP:

$$\mathbb{P}(\text{Trump}) = \mathbb{P}(\text{Trump} \mid \text{Camb.})\mathbb{P}(\text{Camb.}) + \mathbb{P}(\text{Trump} \mid \text{Somer.})\mathbb{P}(\text{Somer.})$$

$$= 0.066 \times 0.56 + 0.103 \times 0.44 = 0.082.$$

QAnon

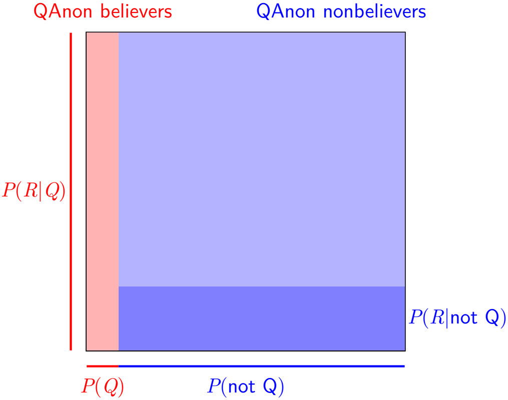
You meet a man named Steve and he tells you that he is a Republican. You have been interested in meeting someone who believes in the QAnon conspiracy theory. Given what you know about Steve, would you guess that he believes in QAnon or not?

QAnon

You meet a man named Steve and he tells you that he is a Republican. You have been interested in meeting someone who believes in the QAnon conspiracy theory. Given what you know about Steve, would you guess that he believes in QAnon or not?

- Common response: probably believes in QAnon since believers tend to be Republicans.
- **Base rate fallacy**: ignores how uncommon QAnon believers are!

Visualizing QAnon support



Chance a random Republican believes QAnon $= \frac{P(R|Q)P(Q)}{P(R|Q)P(Q)+P(R|\text{not Q})P(\text{not Q})}$

Bayes' rule

- Reverend Thomas Bayes (1701–61): English minister and statistician
- **Bayes' rule**: if $\mathbb{P}(B) > 0$, then:

Bayes' rule

- Reverend Thomas Bayes (1701–61): English minister and statistician
- **Bayes' rule**: if $\mathbb{P}(B) > 0$, then:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B \mid A)\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\mathbb{P}(A^c)}$$

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$
- Applying Bayes' rule is often called **updating the prior**.

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$
- Applying Bayes' rule is often called **updating the prior**.
  - $\mathbb{P}(\text{QAnon}) \to \mathbb{P}(\text{QAnon} \mid \text{Republican})$

Why is Bayes' Rule Useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$
- Applying Bayes' rule is often called **updating the prior**.
  - $\mathbb{P}(\text{QAnon}) \to \mathbb{P}(\text{QAnon} \mid \text{Republican})$
  - How does the evidence change the chance of the hypothesis being true?

Uses of Bayes' Rule

- **Medical testing**:

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- **Predicting traits from names**:

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- **Predicting traits from names**:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- **Predicting traits from names**:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- **Predicting traits from names**:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$
- **Spam filtering**:

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- **Predicting traits from names**:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$
- **Spam filtering**:
  - Want to know: $\mathbb{P}(\text{Spam} \mid \text{Email text})$

Uses of Bayes' Rule

- **Medical testing**:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- **Predicting traits from names**:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$
- **Spam filtering**:
  - Want to know: $\mathbb{P}(\text{Spam} \mid \text{Email text})$
  - Have: $\mathbb{P}(\text{Email text} \mid \text{Spam})$ and $\mathbb{P}(\text{Spam})$

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:
  - $\mathbb{P}(PT \mid C) = 0.8$: true positive rate

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:
  - $\mathbb{P}(PT \mid C) = 0.8$: true positive rate
  - $\mathbb{P}(PT \mid C^c) = 0.005$: false positive rate

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:
  - $\mathbb{P}(PT \mid C) = 0.8$: true positive rate
  - $\mathbb{P}(PT \mid C^c) = 0.005$: false positive rate
  - $\mathbb{P}(C) = 0.007$ rough prevalence of active COVID cases.

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT) = \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT \mid C^c)\mathbb{P}(C^c)$$

$$= (0.8 \times 0.007) + (0.005 \times 0.993)$$

$$= 0.011$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT) = \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT \mid C^c)\mathbb{P}(C^c)$$

$$= (0.8 \times 0.007) + (0.005 \times 0.993)$$

$$= 0.011$$

- Now plug in all values to Bayes' rule:

$$\mathbb{P}(C \mid PT) = \frac{\mathbb{P}(PT \mid C)\mathbb{P}(C)}{\mathbb{P}(PT)} = \frac{0.8 \times 0.007}{0.0106} \approx 0.53$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT) = \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT \mid C^c)\mathbb{P}(C^c)$$

$$= (0.8 \times 0.007) + (0.005 \times 0.993)$$

$$= 0.011$$

- Now plug in all values to Bayes' rule:

$$\mathbb{P}(C \mid PT) = \frac{\mathbb{P}(PT \mid C)\mathbb{P}(C)}{\mathbb{P}(PT)} = \frac{0.8 \times 0.007}{0.0106} \approx 0.53$$

- If false positive rate goes up to 1% $\rightsquigarrow \mathbb{P}(C \mid PT) \approx 0.36$

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
    - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

    - Sometimes written as $A \perp\!\!\!\perp B$

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

  - Sometimes written as $A \perp\!\!\!\perp B$
  - **Symmetric:** $A \perp\!\!\!\perp B$ equivalent to $B \perp\!\!\!\perp A$

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

  - Sometimes written as $A \perp\!\!\!\perp B$
  - **Symmetric:** $A \perp\!\!\!\perp B$ equivalent to $B \perp\!\!\!\perp A$
  - Events that are not independent are **dependent**.

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

  - Sometimes written as $A \perp\!\!\!\perp B$
  - **Symmetric:** $A \perp\!\!\!\perp B$ equivalent to $B \perp\!\!\!\perp A$
  - Events that are not independent are **dependent**.
- Important consequence: if $A \perp\!\!\!\perp B$ and $\mathbb{P}(B) > 0$, then:

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

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

  - Sometimes written as $A \perp\!\!\!\perp B$
  - **Symmetric:** $A \perp\!\!\!\perp B$ equivalent to $B \perp\!\!\!\perp A$
  - Events that are not independent are **dependent**.
- Important consequence: if $A \perp\!\!\!\perp B$ and $\mathbb{P}(B) > 0$, then:

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

  - Knowing B occurs has no impact on the probability of A.

Independence

- Heart of Bayes's rule: knowing B occurs often changes probability of A.
  - What if B provides no information? $\rightsquigarrow$ independence
- Two events A and B are **independent** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

  - Sometimes written as $A \perp\!\!\!\perp B$
  - **Symmetric:** $A \perp\!\!\!\perp B$ equivalent to $B \perp\!\!\!\perp A$
  - Events that are not independent are **dependent**.
- Important consequence: if $A \perp\!\!\!\perp B$ and $\mathbb{P}(B) > 0$, then:

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

  - Knowing B occurs has no impact on the probability of A.
  - Works other way too: if $\mathbb{P}(A) > 0$ and $A \perp\!\!\!\perp B$, then $\mathbb{P}(B \mid A) = \mathbb{P}(B)$.

Independence

- **Common misunderstanding: independent is different than disjoint!**
  - Mutually exclusive events **provide** information!

Independence example

- If we have a gathering of size n drawn randomly from the population of MA with a current COVID infection rate of 1.37%, what's the probability someone in attendance is infected?
- When seeing "prob. of at least one" $\rightsquigarrow$ work with complement:

$$\mathbb{P}(\text{At least one COVID case at gathering})$$

$$= 1 - \mathbb{P}(\text{No COVID cases at gathering})$$

Independence and random sampling

- How we draw the random sample matters:

Independence and random sampling

- How we draw the random sample matters:
  - Sample $n > 1$ with replacement $\rightsquigarrow$ independent events

Independence and random sampling

- How we draw the random sample matters:
  - Sample $n > 1$ with replacement $\rightsquigarrow$ independent events
  - Sample $n > 1$ without replacement $\rightsquigarrow$ dependent events

Independence and random sampling

- How we draw the random sample matters:
  - Sample $n > 1$ with replacement $\rightsquigarrow$ independent events
  - Sample $n > 1$ without replacement $\rightsquigarrow$ dependent events
- Sampling with replacement n for gathering:

$$
\begin{aligned}
&P(\text{No COVID cases at gathering}) \\
&\quad = P(\text{No COVID for Person } 1 \cap \cdots \cap \text{No COVID for Person } n) \\
&\quad = P(\text{No COVID for Person } 1) \cdots P(\text{No COVID for Person } n) \\
&\quad = (1 - 0.007)^n
\end{aligned}
$$

Independence and random sampling

- How we draw the random sample matters:
  - Sample $n > 1$ with replacement $\rightsquigarrow$ independent events
  - Sample $n > 1$ without replacement $\rightsquigarrow$ dependent events
- Sampling with replacement n for gathering:

$$\begin{aligned}
&P(\text{No COVID cases at gathering}) \\
&\quad = P(\text{No COVID for Person } 1 \cap \cdots \cap \text{No COVID for Person } n) \\
&\quad = P(\text{No COVID for Person } 1) \cdots P(\text{No COVID for Person } n) \\
&\quad = (1 - 0.007)^n
\end{aligned}$$

- Using the complement:

$$P(\text{At least one COVID case at gathering}) = 1 - (1 - 0.007)^n$$

Independence and random sampling

- How we draw the random sample matters:
  - Sample $n > 1$ with replacement $\rightsquigarrow$ independent events
  - Sample $n > 1$ without replacement $\rightsquigarrow$ dependent events
- Sampling with replacement n for gathering:

$$\begin{aligned}
&P(\text{No COVID cases at gathering}) \\
&\quad = P(\text{No COVID for Person 1} \cap \cdots \cap \text{No COVID for Person } n) \\
&\quad = P(\text{No COVID for Person 1}) \cdots P(\text{No COVID for Person } n) \\
&\quad = (1 - 0.007)^n
\end{aligned}$$

- Using the complement:

$$P(\text{At least one COVID case at gathering}) = 1 - (1 - 0.007)^n$$

  - $n = 5 \rightsquigarrow$ prob of 0.035

Independence and random sampling

- How we draw the random sample matters:
  - Sample $n > 1$ with replacement $\rightsquigarrow$ independent events
  - Sample $n > 1$ without replacement $\rightsquigarrow$ dependent events
- Sampling with replacement n for gathering:

$$\begin{aligned}
&P(\text{No COVID cases at gathering}) \\
&\quad = P(\text{No COVID for Person 1} \cap \cdots \cap \text{No COVID for Person } n) \\
&\quad = P(\text{No COVID for Person 1}) \cdots P(\text{No COVID for Person } n) \\
&\quad = (1 - 0.007)^n
\end{aligned}$$

- Using the complement:

$$P(\text{At least one COVID case at gathering}) = 1 - (1 - 0.007)^n$$

  - $n = 5 \rightsquigarrow$ prob of 0.035
  - $n = 100 \rightsquigarrow$ prob of 0.5

Conditional independence

- A and B are **conditionally independent** given E if

$$P(A \cap B \mid E) = P(A \mid E)P(B \mid E)$$

Conditional independence

- A and B are **conditionally independent** given E if

$$P(A \cap B \mid E) = P(A \mid E)P(B \mid E)$$

- Massively important in statistics and causal inference.

Conditional independence

- A and B are **conditionally independent** given E if

$$P(A \cap B \mid E) = P(A \mid E)P(B \mid E)$$

- Massively important in statistics and causal inference.
- **Warning:** independence $\neq$ conditional independence.

Conditional independence

- A and B are **conditionally independent** given E if

$$P(A \cap B \mid E) = P(A \mid E)P(B \mid E)$$

- Massively important in statistics and causal inference.
- **Warning:** independence $\neq$ conditional independence.
  - Cond. ind. $\not\Rightarrow$ ind.: flipping a coin with unknown bias.

Conditional independence

- A and B are **conditionally independent** given E if

$$P(A \cap B \mid E) = P(A \mid E)P(B \mid E)$$

- Massively important in statistics and causal inference.
- **Warning:** independence $\neq$ conditional independence.
  - Cond. ind. $\not\Rightarrow$ ind.: flipping a coin with unknown bias.
  - Ind. $\not\Rightarrow$ cond. ind.: test scores, athletics, and college admission.