1:Basic Probability

Naijia Liu

Spring 2025

What is Probability?

- What is a reasonably safe gathering size in the age of COVID?

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:
  - Average incidence rate of COVID in MA: 1.37%.

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:
  - Average incidence rate of COVID in MA: 1.37%.
  - Gathering is a random sample from the population.

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:
  - Average incidence rate of COVID in MA: 1.37%.
  - Gathering is a random sample from the population.
  - Wrong assumptions $\Rightarrow$ wrong probabilities.

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:
  - Average incidence rate of COVID in MA: 1.37%.
  - Gathering is a random sample from the population.
  - Wrong assumptions $\Rightarrow$ wrong probabilities.
- Why care? Enables learning from data:

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:
  - Average incidence rate of COVID in MA: 1.37%.
  - Gathering is a random sample from the population.
  - Wrong assumptions $\Rightarrow$ wrong probabilities.
- Why care? Enables learning from data:
  - Are the data consistent with probabilities?

Probability and Assumptions

- Gathering size n: probability someone in attendance is infected?
- Assumptions:
  - Average incidence rate of COVID in MA: 1.37%.
  - Gathering is a random sample from the population.
  - Wrong assumptions $\Rightarrow$ wrong probabilities.
- Why care? Enables learning from data:
  - Are the data consistent with probabilities?
  - If not, assumptions are doubtful $\Rightarrow$ we've learned something.

Conjunction Fallacy

- Why mathematical probability? Our intuitions about probability are often terrible.

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Conjunction Fallacy

- Why mathematical probability? Our intuitions about probability are often terrible.

  > Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

- What is more probable?
  - Linda is a bank teller?
  - Linda is a bank teller and active in the feminist movement?

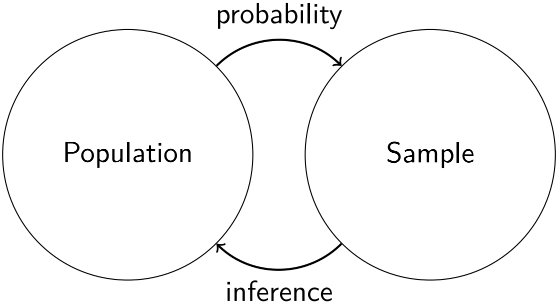
Conjunction Fallacy

- Why mathematical probability? Our intuitions about probability are often terrible.

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

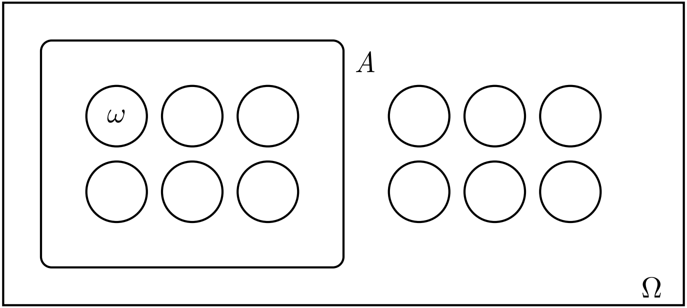
- What is more probable?
  - Linda is a bank teller?
  - Linda is a bank teller and active in the feminist movement?
- Majority chose the second, but this is impossible (conjunction fallacy).
- Learning mathematical probability avoids such mistakes.

Learning About Populations

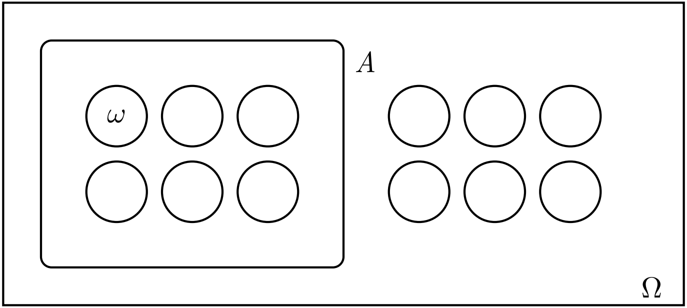


- **Probability:** formalize the uncertainty about how our data came to be.
- **Inference:** learning about the population from a set of data.

Sample Spaces and Events



Sample Spaces and Events



- Probability formalizes chance variation or uncertainty.
- Definitions:
  - Sample space Ω: Set of possible outcomes.
  - Event A: Subset of Ω.
- Example: Finite, countably infinite, or uncountably infinite sample spaces.

Naive Definition of Probability

- All outcomes equally likely:

$$P(A) = \frac{\text{Number of elements in } A}{\text{Number of elements in } \Omega}$$

- Justifications:
  - Symmetry: Fair coin, shuffled cards, etc.
  - Simple random samples. Ω
- Example: Probability of drawing a specific card from a deck.
  - Each card has equal probability:

$$P(4\clubsuit) = P(4\heartsuit) = \frac{1}{52}$$

Counting and Sampling

- **Multiplication rule:** if you have two sub-experiments, A with a possible outcomes and B with b possible outcomes, then in the combined experiment there are ab possible outcomes.

Counting and Sampling

- **Multiplication rule:** if you have two sub-experiments, A with a possible outcomes and B with b possible outcomes, then in the combined experiment there are ab possible outcomes.
- Example: what to watch where?

Counting and Sampling

- **Multiplication rule:** if you have two sub-experiments, A with a possible outcomes and B with b possible outcomes, then in the combined experiment there are ab possible outcomes.
- Example: what to watch where?
  - What to watch? Netflix or Hulu.

Counting and Sampling

- **Multiplication rule:** if you have two sub-experiments, A with a possible outcomes and B with b possible outcomes, then in the combined experiment there are ab possible outcomes.
- Example: what to watch where?
  - What to watch? Netflix or Hulu.
  - Where to watch? TV, tablet, or phone.

Counting and Sampling

- **Multiplication rule:** if you have two sub-experiments, A with a possible outcomes and B with b possible outcomes, then in the combined experiment there are ab possible outcomes.
- Example: what to watch where?
  - What to watch? Netflix or Hulu.
  - Where to watch? TV, tablet, or phone.
  - $2 \times 3 = 6$ possible outcomes (Netflix on TV, Hulu on phone, ...)

Counting and Sampling

- **Multiplication rule:** if you have two sub-experiments, A with a possible outcomes and B with b possible outcomes, then in the combined experiment there are ab possible outcomes.
- Example: what to watch where?
  - What to watch? Netflix or Hulu.
  - Where to watch? TV, tablet, or phone.
  - $2 \times 3 = 6$ possible outcomes (Netflix on TV, Hulu on phone, ...)
- Assumes number of outcomes in one experiment doesn't depend on the outcome of the other experiment.

Sampling objects

- **Sample with replacement:** Choose k objects from a set of n one at a time with replacement.
  - Any object may be selected multiple times.
  - There are n^k possible outcomes when order matters (multiplication rule).

Sampling objects

- **Sample with replacement:** Choose k objects from a set of n one at a time with replacement.
  - Any object may be selected multiple times.
  - There are n^k possible outcomes when order matters (multiplication rule).
- **Sampling without replacement:** Choose k objects from a set of n one at a time without replacement.
  - Chosen object can't be chosen again.

Sampling objects

- **Sample with replacement:** Choose k objects from a set of n one at a time with replacement.
  - Any object may be selected multiple times.
  - There are n^k possible outcomes when order matters (multiplication rule).
- **Sampling without replacement:** Choose k objects from a set of n one at a time without replacement.
  - Chosen object can't be chosen again.
  - Number of possibilities: $n(n-1)\cdots(n-k+1)$

Sampling objects

- **Sample with replacement:** Choose k objects from a set of n one at a time with replacement.
  - Any object may be selected multiple times.
  - There are n^k possible outcomes when order matters (multiplication rule).
- **Sampling without replacement:** Choose k objects from a set of n one at a time without replacement.
  - Chosen object can't be chosen again.
  - Number of possibilities: $n(n-1)\cdots(n-k+1)$

Sampling objects

- **Sample with replacement:** Choose k objects from a set of n one at a time with replacement.
  - Any object may be selected multiple times.
  - There are n^k possible outcomes when order matters (multiplication rule).
- **Sampling without replacement:** Choose k objects from a set of n one at a time without replacement.
  - Chosen object can't be chosen again.
  - Number of possibilities: $n(n-1)\cdots(n-k+1)$
- **Ex:** 2020 IA caucuses had 11 candidates. How many top-3 possibilities?
  - 11 first-place choices
  - 10 second-place choices among the remaining candidates
  - 9 third-place choices
  - Total: $11 \cdot 10 \cdot 9 = 990$ possibilities

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?
- If each possibility occurs c times, we can divide by c .

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?
- If each possibility occurs c times, we can divide by c .
- What if we wanted to book 3 IA caucus candidates for a debate?

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?
- If each possibility occurs c times, we can divide by c .
- What if we wanted to book 3 IA caucus candidates for a debate?
  - Multiplication rule would count both (Biden, Warren, Sanders) and (Sanders, Biden, Warren).

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?
- If each possibility occurs c times, we can divide by c .
- What if we wanted to book 3 IA caucus candidates for a debate?
    - Multiplication rule would count both (Biden, Warren, Sanders) and (Sanders, Biden, Warren).
    - But both of those make the same debate stage.

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?
- If each possibility occurs c times, we can divide by c .
- What if we wanted to book 3 IA caucus candidates for a debate?
  - Multiplication rule would count both (Biden, Warren, Sanders) and (Sanders, Biden, Warren).
  - But both of those make the same debate stage.
  - By the multiplication rule, there are

    $$3 \cdot 2 \cdot 1 = 6$$

    ways to arrange them.

Avoiding overcounting

- What if order doesn't matter? How many subsets of size k are there?
- If each possibility occurs c times, we can divide by c .
- What if we wanted to book 3 IA caucus candidates for a debate?
  - Multiplication rule would count both (Biden, Warren, Sanders) and (Sanders, Biden, Warren).
  - But both of those make the same debate stage.
  - By the multiplication rule, there are

    $$3 \cdot 2 \cdot 1 = 6$$

    ways to arrange them.
- **Binomial coefficient:** number of subsets of size k in a group of n objects:

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{(n-k)!k!}$$

The lady tasting tea

Lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

The lady tasting tea

Lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:
  - Prepare 8 cups of tea, 4 milk-first, 4 tea-first.

The lady tasting tea

Lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:
  - Prepare 8 cups of tea, 4 milk-first, 4 tea-first.
  - Present cups to friend in a **random** order.

The lady tasting tea

Lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:
  - Prepare 8 cups of tea, 4 milk-first, 4 tea-first.
  - Present cups to friend in a **random** order.
  - Ask friend to pick which 4 of the 8 were milk-first.

Assuming we know the truth

- The lady picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly?**
  - Only one way to choose all 4 correct cups.
- Chances of guessing all 4 correct is

$$\frac{1}{70} \approx 0.014 \text{ or } 1.4\%.$$

- $\rightsquigarrow$ the guessing hypothesis might be implausible.

Assuming we know the truth

- The lady picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly?**
  - Only one way to choose all 4 correct cups.
  - But $\binom{8}{4} = 70$ ways of choosing 4 cups among 8.

- Chances of guessing all 4 correct is

$$\frac{1}{70} \approx 0.014 \text{ or } 1.4\%.$$

- $\rightsquigarrow$ the guessing hypothesis might be implausible.

Assuming we know the truth

- The lady picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly?**
  - Only one way to choose all 4 correct cups.
  - But $\binom{8}{4} = 70$ ways of choosing 4 cups among 8.
  - Choosing at random $\approx$ picking each of these 70 with equal probability.
- Chances of guessing all 4 correct is

$$\frac{1}{70} \approx 0.014 \text{ or } 1.4\%.$$

- $\rightsquigarrow$ the guessing hypothesis might be implausible.

Birthday Problem

Birthday Problem

There are k people in a room. Assume each person's birthday is equally likely to be any of the 365 days of the year (no leap babies) and birthdays are independent. What is the probability that at least one pair of people have the same birthday?

Probability

- **A probability space** consists of:

Probability

- **A probability space** consists of:
  - Sample space Ω

Probability

- **A probability space** consists of:
  - Sample space Ω
  - Probability function $\mathbb{P}$ mapping events $A \subseteq \Omega$ onto the real line.

Probability

- **A probability space** consists of:
  - Sample space Ω
  - Probability function $\mathbb{P}$ mapping events $A \subseteq \Omega$ onto the real line.
- The function $\mathbb{P}$ must satisfy the following axioms:

Probability

- **A probability space** consists of:
  - Sample space Ω
  - Probability function $\mathbb{P}$ mapping events $A \subseteq \Omega$ onto the real line.
- The function $\mathbb{P}$ must satisfy the following axioms:
  1. (Non-negativity) $\mathbb{P}(A) \geq 0$ for every event A

Probability

- **A probability space** consists of:
  - Sample space Ω
  - Probability function $\mathbb{P}$ mapping events $A \subseteq \Omega$ onto the real line.
- The function $\mathbb{P}$ must satisfy the following axioms:
  1. (Non-negativity) $\mathbb{P}(A) \geq 0$ for every event A
  2. (Normalization) $\mathbb{P}(\Omega) = 1$

Probability

- **A probability space** consists of:
  - Sample space Ω
  - Probability function $\mathbb{P}$ mapping events $A \subseteq \Omega$ onto the real line.
- The function $\mathbb{P}$ must satisfy the following axioms:
  1. (Non-negativity) $\mathbb{P}(A) \geq 0$ for every event A
  2. (Normalization) $\mathbb{P}(\Omega) = 1$
  3. (Additivity) If a series of events, $A_1, A_2, \ldots$, are disjoint, then

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j).$$

Probability

- **A probability space** consists of:
  - Sample space Ω
  - Probability function $\mathbb{P}$ mapping events $A \subseteq \Omega$ onto the real line.
- The function $\mathbb{P}$ must satisfy the following axioms:
  1. (Non-negativity) $\mathbb{P}(A) \geq 0$ for every event A
  2. (Normalization) $\mathbb{P}(\Omega) = 1$
  3. (Additivity) If a series of events, $A_1, A_2, \ldots$, are disjoint, then

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j).$$

- Probability function assigns "mass" to regions of the sample space.

Interpretation of probabilities

- How do we interpret $\mathbb{P}(A)$? Huge debate about this in stats literature.

Interpretation of probabilities

- How do we interpret $\mathbb{P}(A)$? Huge debate about this in stats literature.
  1. **Frequentist:** $\mathbb{P}()$ reflects relative frequency in a large number of trials.

- Set debate aside: both viewpoints are helpful in different contexts.
  - Properties of probabilities exactly the same in either approach.
  - This class: focus on frequentist perspectives because it's pervasive.

Interpretation of probabilities

- How do we interpret $\mathbb{P}(A)$? Huge debate about this in stats literature.
  1. **Frequentist:** $\mathbb{P}()$ reflects relative frequency in a large number of trials.
     - ✓ Repeat a coin flip many times $\rightsquigarrow$ frequency of head ≈ 0.5.
- Set debate aside: both viewpoints are helpful in different contexts.
  - Properties of probabilities exactly the same in either approach.
  - This class: focus on frequentist perspectives because it's pervasive.

Interpretation of probabilities

- How do we interpret $\mathbb{P}(A)$? Huge debate about this in stats literature.
  1. **Frequentist:** $\mathbb{P}()$ reflects relative frequency in a large number of trials.
     - ✓ Repeat a coin flip many times $\rightsquigarrow$ frequency of head ≈ 0.5.
  2. **Bayesian:** $\mathbb{P}()$ are subjective beliefs about outcomes.
- Set debate aside: both viewpoints are helpful in different contexts.
  - Properties of probabilities exactly the same in either approach.
  - This class: focus on frequentist perspectives because it's pervasive.

Interpretation of probabilities

- How do we interpret $\mathbb{P}(A)$? Huge debate about this in stats literature.
  1. **Frequentist:** $\mathbb{P}()$ reflects relative frequency in a large number of trials.
     - ✓ Repeat a coin flip many times $\rightsquigarrow$ frequency of head ≈ 0.5.
  2. **Bayesian:** $\mathbb{P}()$ are subjective beliefs about outcomes.
     - ✓ How likely I think a particular event will be.
- Set debate aside: both viewpoints are helpful in different contexts.
  - ▶ Properties of probabilities exactly the same in either approach.
  - ▶ This class: focus on frequentist perspectives because it's pervasive.

Gambling 102

- What's the probability of selecting a 4 card from a well-shuffled deck?
  - "Well-shuffled" $\rightsquigarrow$ "randomly selected" $\rightsquigarrow$ all cards have prob. $1/52$.
- "4 card" event $= \{4\clubsuit \cup 4\spadesuit \cup 4\heartsuit \cup 4\diamondsuit\}$
- Union of mutually exclusive events $\rightsquigarrow$ use additivity:

$$\mathbb{P}(4 \text{ card}) = \mathbb{P}(4\clubsuit) + \mathbb{P}(4\spadesuit) + \mathbb{P}(4\heartsuit) + \mathbb{P}(4\diamondsuit) = \frac{4}{52}.$$

Some properties of probabilities

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
   - Probability of not A is 1 minus the probability of A.
   - Follows from $A \cup A^c = \Omega$ and $1 = \mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A^c)$.
2. If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$
   - Subsets of events have lower probability than the event.
   - Probability of 5 SCOTUS votes in the liberal direction is less than the probability of liberals winning the case ($A = \{5, 6, 7, 8, 9\}$).
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
   - Avoid "double-counting" the part where A and B overlap.
   - **Inclusion-exclusion**

Gambling

- A standard deck of playing cards has 52 cards:
  - 13 rank cards: (2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A)
  - In each of 4 suits: $(\clubsuit, \spadesuit, \heartsuit, \diamondsuit)$
- Hypothetical experiment: pick a card, any card.
- One possible outcome: picking a 4$\clubsuit$.
- Sample space:
  2$\clubsuit$ 3$\clubsuit$ 4$\clubsuit$ 5$\clubsuit$ . . . A$\clubsuit$
  2$\spadesuit$ 3$\spadesuit$ 4$\spadesuit$ 5$\spadesuit$ . . . A$\spadesuit$
  2$\heartsuit$ 3$\heartsuit$ 4$\heartsuit$ 5$\heartsuit$ . . . A$\heartsuit$
  2$\diamondsuit$ 3$\diamondsuit$ 4$\diamondsuit$ 5$\diamondsuit$ . . . A$\diamondsuit$
- An event: picking a Queen, $\{Q\clubsuit, Q\spadesuit, Q\heartsuit, Q\diamondsuit\}$

Social science examples

- Examples of substantively interesting **sample spaces**:
  - House elections: incumbents win or lose $\Omega = \{W, L\}$.
  - Supreme Court votes in a liberal direction: $\Omega = \{0, 1, \ldots, 9\}$.
  - Voter turnout: percent turnout in some county: $\Omega = [0, 100]$.
  - Duration of a war: any nonnegative number, $\Omega = [0, \infty)$.
- **Events** from these examples:
  - House election: incumbent wins, $A = W$.
  - Supreme Court: liberals win a Supreme Court case, $A = \{5, 6, 7, 8, 9\}$.
  - Voter turnout: over half the population votes, $A \in (50, 100]$.
  - Duration of war: war ends within five years, $A \in [0, 5]$.
- We want to know or model the probability of these events!

New events from old events

- Define events/sets $A, B \subseteq \Omega$.
- **Complement:** A^c ("not A")
  - All of the outcomes in Ω not in A.
  - Complement of picking a red card is picking a **black card**.
  - $\Omega^c = \emptyset$, where $\emptyset$ is the empty set—nothing happens.
- **Union** of two events $A \cup B$ ("A or B")
  - The event that A or B occurs.
  - Queen $\cup\clubsuit =$ any club card or any queen card.
- **Intersection** $A \cap B$ or just AB ("A and B")
  - The event that both A and B occur.
  - Queen $\cap\clubsuit = Q\clubsuit$ (queen of clubs).
- **De Morgan's laws:** $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$.
  - not $Q \cup \clubsuit$ must be not Q and not $\clubsuit$.
  - not $Q \cap \clubsuit = Q\clubsuit$ is either not a Q or not a $\clubsuit$.

Relationships between events

- A **implies** B when $A \subseteq B$.
  - $A = 4\clubsuit$ and $B = \clubsuit$ card.
- **Disjoint/mutually exclusive:** $A \cap B = \emptyset$.
  - $A = \clubsuit$ card, $B = \spadesuit$ card.
- **A partition** is a set of mutually disjoint events whose union is Ω.
  - $A_1 = \clubsuit, A_2 = \spadesuit, A_3 = \diamondsuit, A_4 = \heartsuit$
  - A_1, A_2, A_3, A_4 is a partition of a 52-card deck.