Conditional Probability

Theory and Biological Applications

DRME

2025

Foundations of Probability

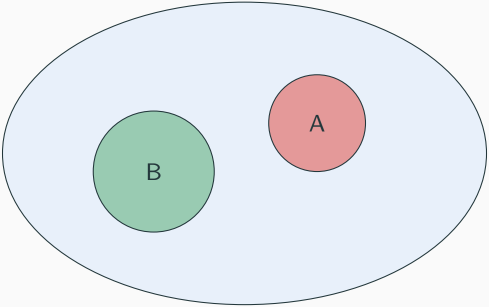
The Logic of Uncertainty

- **Mathematics** is the logic of certainty
- **Probability** is the logic of uncertainty
- Provides framework for quantifying doubt and updating beliefs

Wikipedia: Probability

Sample Space and Events

Ω: All possible outcomes



- Ω: Set of all possible outcomes
- Events: Subsets of Ω (e.g., "has disease allele")
- Pebble World: Each "pebble" = one possible outcome

Wikipedia: Sample Space

Basic Probability Rules

Naive Definition (Equal Likelihood)

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

Caution: Requires equally likely outcomes

Counting Rules

- Multiplication Rule: $n_1 \times n_2 \times \cdots \times n_r$
- Combinations: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Conditional Probability: Definition and Intuition

Why Conditional Probability?

Core Insight

All probabilities are conditional on available information

Biological Context: Genetic inheritance, disease risk, evolutionary relationships
Two Roles of Conditioning:

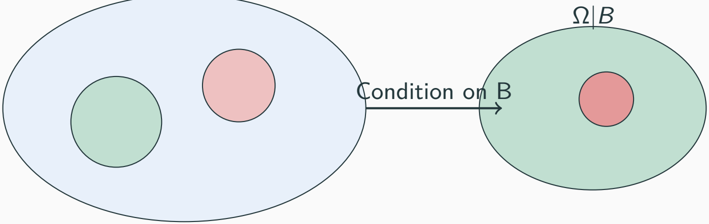
1. Updating beliefs with new evidence
2. Problem-solving strategy (break complex problems into simpler pieces)

Formal Definition of Conditional Probability

Definition (Conditional Probability)

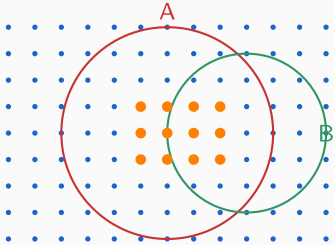
If $P(B) > 0$, the conditional probability of A given B is:

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



Wikipedia: Conditional Probability

Visualizing Conditional Probability: Pebble World



Three Step Process:

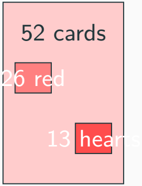
1. Start with all pebbles (mass $= 1$)
2. Discard pebbles not in B (mass now $= P(B)$)
3. Renormalize: divide all masses by $P(B)$ so total mass $= 1$

Example: Two Cards Problem - Setup

Problem: Draw two cards without replacement from a standard deck of 52 cards.

Events Defined

- A: First card is a heart ($\heartsuit$)
- B: Second card is red (either $\heartsuit$ or $\diamondsuit$)



Deck Composition

- 52 total cards
- 26 red cards (13 hearts + 13 diamonds)
- 13 hearts
- 13 diamonds

Question: Find $P(A|B)$ and $P(B|A)$

Step 1: Calculate $P(A \cap B)$

We need the probability that both events occur: first card is heart AND second card is red.

Method 1: Sequential Counting

- First card (heart): 13 choices out of 52
- Second card (red): Given first was heart, 25 red cards remain out of 51

Calculation

$$\begin{aligned} P(A \cap B) &= P(\text{1st heart}) \times P(\text{2nd red} \mid \text{1st heart}) \\ &= \frac{13}{52} \times \frac{25}{51} \\ &= \frac{1}{4} \times \frac{25}{51} \\ &= \frac{25}{204} \approx 0.1225 \end{aligned}$$

Note

After drawing a heart first, only 25 red cards remain (12 hearts + 13 diamonds) out of 51 total cards.

Step 2: Calculate $P(A)$ and $P(B)$

Probability of A

First card is a heart:

$$P(A) = \frac{13}{52} = \frac{1}{4} = 0.25$$

Probability of B

Second card is red:

$$P(B) = \frac{26}{52} = \frac{1}{2} = 0.5$$

Why $P(B) = 1/2$?

- By symmetry: no reason second card should favor any color
- More formally: $P(B) = P(\text{red on 2nd}) = \frac{26}{52} = \frac{1}{2}$

Step 3: Calculate $P(A|B)$

Conditional probability of first heart given second red:

Definition

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

Substitution

$$P(A|B) = \frac{\frac{25}{204}}{\frac{1}{2}}$$

Simplification

$$P(A|B) = \frac{25}{204} \times \frac{2}{1} = \frac{25}{102} \approx 0.2451$$

Interpretation

Given that the second card is red, there's about 24.5% chance the first was a heart.

Step 4: Calculate $P(B|A)$

Conditional probability of second red given first heart:

Definition

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

Substitution

$$P(B|A) = \frac{\frac{25}{204}}{\frac{1}{4}}$$

Simplification

$$P(B|A) = \frac{25}{204} \times \frac{4}{1} = \frac{25}{51} \approx 0.4902$$

Interpretation

Given that the first card is a heart, there's about 49% chance the second is red.

Alternative calculation: If first is heart, 25 red cards remain out of 51:

$$P(B|A) = \frac{25}{51} \approx 0.4902$$

Step 5: Compare and Interpret Results

Probability	Value	Interpretation
$P(A \cap B)$	$\frac{25}{204} \approx 0.1225$	Both events occur
$P(A\|B)$	$\frac{25}{102} \approx 0.2451$	First heart given second red
$P(B\|A)$	$\frac{25}{51} \approx 0.4902$	Second red given first heart
$P(A)$	0.25	Marginal probability of first heart
$P(B)$	0.5	Marginal probability of second red

Key Observations

1. **Asymmetry**: $P(A|B) \neq P(B|A)$ (0.2451   0.4902)
2. **Order matters**: Conditioning on different events gives different results
3. **Relationship**: $P(A|B) < P(B|A)$ because knowing second is red provides less information about first card than vice versa

Why $P(A|B) \neq P(B|A)$? Intuitive Explanation

Case 1: $P(B|A)$

- Know: First card is heart
- Remaining: 51 cards
- Red cards left: 25 (12 hearts + 13 diamonds)
- $P(B|A) = \frac{25}{51} \approx 0.49$

Case 2: $P(A|B)$

- Know: Second card is red
- First card: Unknown, could be any of 52
- Hearts among first cards: $13/52 = 0.25$
- But we must adjust for the fact that second is red
- $P(A|B) = 0.2451$

General Principle

$P(A|B)$ and $P(B|A)$ are fundamentally different questions:

- $P(B|A)$: Given specific info about first card, predict second
- $P(A|B)$: Given info about second card, infer about first

They are only equal if $P(A) = P(B)$, which is not true here (0.25 0.5).

Connection to Bayes' Rule

We can verify our calculation using Bayes' Rule:

Bayes' Rule Formula

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

Verification

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{\left(\frac{25}{51}\right) \times \left(\frac{1}{4}\right)}{\frac{1}{2}} \\ &= \frac{\frac{25}{204}}{\frac{1}{2}} = \frac{25}{102} \approx 0.2451 \end{aligned}$$

This matches our direct calculation!

Takeaway

The cards problem illustrates why we must be careful not to confuse $P(A|B)$

The Two Children Problem - Introduction

Classic Probability Puzzle

Two different families, two different questions about their children's genders.

Case 1: Mr. Jones

- Has two children
- **Specific information**: The **older** child is a girl
- **Question**: What is $P(\text{both girls})$?

Case 2: Mr. Smith

- Has two children
- **General information**: **At least one** is a boy
- **Question**: What is $P(\text{both boys})$?

Key Assumptions

- Binary gender (boy/girl)
- $P(\text{boy}) = P(\text{girl}) = 1/2$ for each child
- Children's genders are independent
- Order matters (older/younger distinction)

Step 1: Define the Sample Space

We need to list all possible outcomes for two children:

Complete Sample Space Ω
Four equally likely outcomes:

$$\Omega = \{BB, BG, GB, GG\}$$

Where:

- First letter = older child's gender
- Second letter = younger child's gender
- B = boy, G = girl

BB	BG	GB	GG
1/4	1/4	1/4	1/4

Each outcome has probability 1/4.

Step 2: Analyze Case 1 - Older Child is Girl

Mr. Jones: We know the **older** child is a girl.

Restricted Sample Space

From $\Omega = \{BB, BG, GB, GG\}$, keep only outcomes where older child is girl:

$$\Omega_{\text{Jones}} = \{GB, GG\}$$

BB	BG	GB	GG
1/4	1/4	1/4	1/4
		Kept	**Kept**

Note: We discard BB and BG because older child is not girl in those cases.

Step 3: Calculate Case 1 Probability

Conditional Probability Definition

$$P(\text{both girls} \mid \text{older is girl}) = \frac{P(\text{both girls AND older is girl})}{P(\text{older is girl})}$$

Step-by-step Calculation

1. $P(\text{both girls AND older is girl}) = P(GG) = \frac{1}{4}$
2. $P(\text{older is girl}) = P(GB \text{ or } GG) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
3. Apply formula:

$$P(\text{both girls} \mid \text{older is girl}) = \frac{1/4}{1/2} = \frac{1}{2}$$

Direct Counting Method

Within $\Omega_{\text{Jones}} = \{GB, GG\}$:

- 2 equally likely outcomes
- Only GG has both girls
- Probability $= \frac{1}{2}$

Answer for Mr. Jones: $\frac{1}{2}$ chance both children are girls.

Step 4: Analyze Case 2 - At Least One Boy

Mr. Smith: We know **at least one** child is a boy.

Restricted Sample Space

From $\Omega = \{BB, BG, GB, GG\}$, keep only outcomes with at least one boy:

$$\Omega_{\text{Smith}} = \{BB, BG, GB\}$$

BB	BG	GB	GG
1/4	1/4	1/4	1/4
Kept	**Kept**	**Kept**	

Note: We discard only GG because it has no boys.

Step 5: Calculate Case 2 Probability

Conditional Probability Definition

$$P(\text{both boys} \mid \text{at least one boy}) = \frac{P(\text{both boys AND at least one boy})}{P(\text{at least one boy})}$$

Step-by-step Calculation

1. $P(\text{both boys AND at least one boy}) = P(BB) = \frac{1}{4}$
2. $P(\text{at least one boy}) = P(BB \text{ or } BG \text{ or } GB) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$
3. Apply formula:

$$P(\text{both boys} \mid \text{at least one boy}) = \frac{1/4}{3/4} = \frac{1}{3}$$

Direct Counting Method

Within $\Omega_{\text{Smith}} = \{BB, BG, GB\}$:

- 3 equally likely outcomes
- Only BB has both boys
- Probability $= \frac{1}{3}$

Step 6: Compare and Contrast the Two Cases

	Case 1 (Mr. Jones)	Case 2 (Mr. Smith)
Information	Older child is girl	At least one is boy
Restricted Space	$\{GB, GG\}$	$\{BB, BG, GB\}$
Size of Space	2 outcomes	3 outcomes
Favorable Outcomes	$\{GG\}$	$\{BB\}$
Probability	$\frac{1}{2}$	$\frac{1}{3}$

Why Different Answers?

- **Specific vs. General Information**:
  - "Older is girl" pins down a specific child
  - "At least one boy" could refer to either child
- **Different Sample Space Sizes**:
  - Case 1 eliminates 2 of 4 outcomes
  - Case 2 eliminates only 1 of 4 outcomes

Common Misconceptions and Intuition

Incorrect Intuition

- "Both cases seem similar"
- "Shouldn't both answers be 1/2?"
- "Why does wording matter so much?"

Correct Intuition

- Information quality differs
- Specificity changes probabilities
- Wording reveals different evidence

Real-world Analogy

- **Case 1**: "My first child is a girl" $\rightarrow$ Tells you about a specific child
- **Case 2**: "I have at least one son" $\rightarrow$ Doesn't tell you which child

Mathematical Insight

Conditional probability depends on **how much** the conditioning event restricts the sample space:

- Strong restriction (eliminates many outcomes) $\rightarrow$ probability changes significantly
- Weak restriction (eliminates few outcomes) $\rightarrow$ probability changes less

Extensions and Related Problems

Variation 1: "The older child is a boy. What is P(both boys)?"

- Same logic as Case 1
- Answer: $1/2$

Variation 2: "At least one is a girl. What is P(both girls)?"

- Same logic as Case 2
- Answer: $1/3$

Variation 3: "I randomly select a child and it's a boy. What is P(both boys)?"

- Different from both cases!
- Answer: $1/2$
- Why? Because randomly selecting eliminates different outcomes

Key Takeaway

The exact wording of probability problems matters critically. Small changes in information can lead to different answers.

Wikipedia: Boy or Girl Paradox

Bayes' Rule and Law of Total Probability

The Multiplication Rule

Joint Probability Factorization

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Genetic Application: Compound Heterozygotes

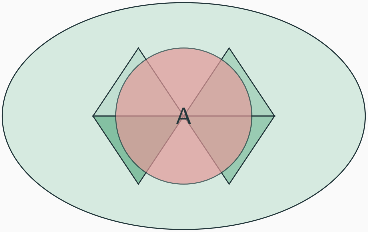
- Event A: Inherit mutant allele from mother
- Event B: Inherit mutant allele from father
- $P(\text{disease}) = P(A \cap B) = P(A|B)P(B)$
- For recessive disorders: $P(A|B) = P(A)$ if parents unrelated

Law of Total Probability (LOTP)

Theorem (Law of Total Probability)

If $B_1, B_2, \ldots, B_n$ partition Ω, then:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$



Wikipedia: Law of Total Probability

Bayes' Rule: Derivation and Statement

Theorem (Bayes' Rule)

For events A and B with $P(B) > 0$:

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

Derivation:

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \quad \text{(definition)} \\ &= \frac{P(B|A)P(A)}{P(B)} \quad \text{(multiplication rule)} \end{aligned}$$

Terminology:

- $P(A)$: Prior probability (before evidence)
- $P(A|B)$: Posterior probability (after evidence)
- $P(B|A)$: Likelihood of evidence if hypothesis true

Wikipedia: Bayes' Theorem

Rare Disease Testing - Problem Setup

Medical Diagnostic Testing Scenario

A new disease test is 99% accurate, but the disease is rare. What does a positive test result really mean?

Given Information

- **Disease prevalence**: $P(D) = 0.001$
- **Test sensitivity**: $P(+|D) = 0.99$
- **Test specificity**: $P(-|\neg D) = 0.99$
- **Question**: $P(D|+) = ?$

Terminology

- **Sensitivity**: True positive rate
- **Specificity**: True negative rate
- **False positive**: Healthy person tests positive
- **False negative**: Diseased person tests negative

The Puzzle

Test is 99% accurate, but disease affects only 0.1% of population. How reliable is a positive test?

Step 1: Define Events and Probabilities

Event Definitions

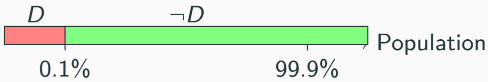
- D: Person has the disease
- $+$: Test result is positive
- $-$: Test result is negative

Given Probabilities

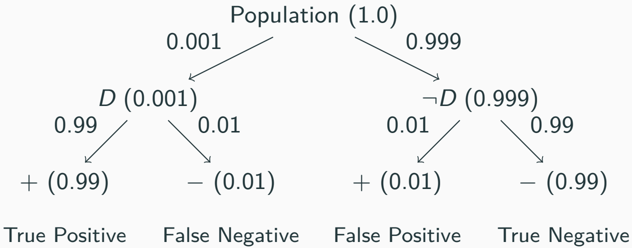
$$\begin{aligned} P(D) &= 0.001 \quad \text{(prevalence)} \\ P(+|D) &= 0.99 \quad \text{(sensitivity)} \\ P(-|\neg D) &= 0.99 \quad \text{(specificity)} \end{aligned}$$

Derived Probabilities

$$\begin{aligned} P(\neg D) &= 1 - P(D) = 0.999 \\ P(-|D) &= 1 - P(+|D) = 0.01 \\ P(+|\neg D) &= 1 - P(-|\neg D) = 0.01 \end{aligned}$$



Step 2: Visualize with Tree Diagram



Path Probabilities

- True positive path: $0.001 \times 0.99 = 0.00099$
- False positive path: $0.999 \times 0.01 = 0.00999$
- False negative path: $0.001 \times 0.01 = 0.00001$
- True negative path: $0.999 \times 0.99 = 0.98901$

Step 3: Calculate Probability of Positive Test

We need $P(+)$ to use in Bayes' Rule.

Law of Total Probability

$$P(+) = P(+|D)P(D) + P(+|\neg D)P(\neg D)$$

Substitution

$$P(+) = (0.99 \times 0.001) + (0.01 \times 0.999)$$

Calculation

$$\begin{aligned} P(+) &= 0.00099 + 0.00999 \\ &= 0.01098 \end{aligned}$$

Interpretation

- Only 1.098% of population tests positive
- But this includes both true positives and false positives
- Most positives come from healthy people being misdiagnosed

Step 4: Apply Bayes' Rule

We want $P(D|+)$: probability of having disease given positive test.

Bayes' Rule Formula

$$P(D|+) = \frac{P(+|D)P(D)}{P(+)}$$

Substitution

$$P(D|+) = \frac{0.99 \times 0.001}{0.01098}$$

Calculation

$$P(D|+) = \frac{0.00099}{0.01098} \approx 0.09016 \approx 9\%$$

Component	**Value**	**Meaning**
Numerator	0.00099	True positives
Denominator	0.01098	All positives
Result	0.09016	Only 9% of positives are true

Step 5: Alternative Calculation with Concrete Numbers

Consider Population of 1,000,000 People

- Diseased: $1,000,000 \times 0.001 = 1,000$ people
- Healthy: $1,000,000 \times 0.999 = 999,000$ people

Test Results Distribution

- **True positives**: $1,000 \times 0.99 = 990$ people
- **False negatives**: $1,000 \times 0.01 = 10$ people
- **False positives**: $999,000 \times 0.01 = 9,990$ people
- **True negatives**: $999,000 \times 0.99 = 989,010$ people

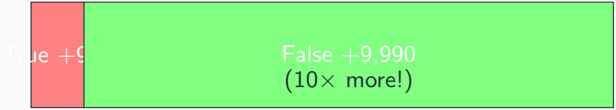
Step 6: Interpret with Concrete Numbers

Positive Test Results

- Total positives: $990 + 9,990 = 10,980$ people
- True positives: 990 people
- False positives: $9,990$ people

Probability Calculation

$$P(D|+) = \frac{\text{True positives}}{\text{All positives}} = \frac{990}{10,980} \approx 0.09016$$



Key Insight

Even with 99% accurate test:

- Only 9% of positive results are correct
- 91% of positive results are false alarms
- False positives outnumber true positives 10:1

Step 7: General Insights and Applications

Why This Happens

- **Base rate fallacy**: Ignoring disease prevalence (base rate)
- **Rarity amplifies errors**: When something is rare, even small error rates produce many false positives
- **Test quality matters less than prevalence**: For rare diseases, specificity is crucial

Mathematical Relationship

$$P(D|+) = \frac{\text{Sensitivity} \times \text{Prevalence}}{\text{Sensitivity} \times \text{Prevalence} + (1 - \text{Specificity}) \times (1 - \text{Prevalence})}$$

Real-world Applications

Medical screening (cancer, genetic disorders), Drug testing, Security screening (airport security, spam filters), Quality control in manufacturing

Takeaway

Always consider the base rate! A test's accuracy must be interpreted in context of how common the condition is.

Step 8: Sensitivity Analysis - What If...?

More Common Disease
$P(D) = 0.01$ (1%)

$$P(D|+) = \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.01 \times 0.99} = 0.5$$

50% chance with positive test

Better Specificity
$P(-|\neg D) = 0.999$ (99.9%)

$$P(D|+) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.001 \times 0.999} \approx 0.5$$

Nearly 50% chance

Perfect Test
$P(+|D) = 1$, $P(-|\neg D) = 1$

$$P(D|+) = \frac{1 \times 0.001}{1 \times 0.001 + 0 \times 0.99} = 1$$

100% chance

General Pattern

- For rare diseases, focus on improving specificity
- For common diseases, sensitivity matters more
- Perfect tests eliminate the base rate problem

This explains why screening tests for rare diseases often require confirmation with more specific tests.

Example: Random Coin Problem

Problem: You have a fair coin ($P(H) = 1/2$) and a biased coin ($P(H) = 3/4$). Pick one at random (50-50), flip it 3 times, get HHH. What is $P(\text{fair}|\text{HHH})$?
Solution:

$$\begin{aligned} P(F|HHH) &= \frac{P(HHH|F)P(F)}{P(HHH|F)P(F) + P(HHH|F^c)P(F^c)} \\ &= \frac{(1/8)(1/2)}{(1/8)(1/2) + (27/64)(1/2)} \\ &= \frac{8}{35} \approx 0.2286 \end{aligned}$$

Interpretation: After seeing HHH, probability it's the fair coin drops from 0.5 to about 0.23.

Extended Forms

Bayes' Rule with Extra Conditioning

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

LOTP with Extra Conditioning

$$P(B|E) = \sum_{i=1}^{n} P(B|A_i, E)P(A_i|E)$$

for partition $A_1, \ldots, A_n$.

Application: Sequential updating of beliefs

- Today's posterior becomes tomorrow's prior
- Coherency: Sequential updating gives same result as updating with all evidence at once

Conditional Probability as Probability

Conditional Probability Satisfies Probability Axioms

Theorem: For fixed E with $P(E) > 0$, $\tilde{P}(A) = P(A|E)$ satisfies:

1. **Non-negativity**: $\tilde{P}(A) \geq 0$ for all A
2. **Normalization**: $\tilde{P}(\Omega) = 1$
3. **Additivity**: For disjoint $A_1, A_2, \ldots$:

$$\tilde{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \tilde{P}(A_i)$$

Proof of additivity:

$$\begin{aligned}
\tilde{P}\left(\bigcup_i A_i\right) &= \frac{P\left((\bigcup_i A_i) \cap E\right)}{P(E)} \\
&= \frac{P\left(\bigcup_i (A_i \cap E)\right)}{P(E)} \\
&= \frac{\sum_i P(A_i \cap E)}{P(E)} = \sum_i \tilde{P}(A_i)
\end{aligned}$$

All Probabilities are Conditional

Fundamental Principle

There are no unconditional probabilities. All probabilities depend on background information K.

Example: $P(\text{Rain today})$ implicitly conditions on:

- Location, season, time of day
- Current weather conditions
- Meteorological models

Mathematical notation:

- $P(A)$ is shorthand for $P(A|K)$ where K is background knowledge
- The vertical bar is always there, even if not written

Biological Context:

- Mutation rates: Conditional on sequence context
- Disease risk: Conditional on genetic background, environment
- Evolutionary probabilities: Conditional on population size, selection

Independence

Definition of Independence

Definition (Independence)

Events A and B are **independent** if:

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

If $P(A) > 0$ and $P(B) > 0$, equivalent to:

$$P(A|B) = P(A) \quad \text{or} \quad P(B|A) = P(B)$$

Interpretation: A provides no information about B, and vice versa.

Genetic Independence

- **Unlinked loci**: Inheritance of alleles at different chromosomes
- **Hardy-Weinberg equilibrium**: Random mating assumption
- **Linkage equilibrium**: No association between alleles at different loci

Wikipedia: Independence

Properties of Independence

Theorem: If A and B are independent, then:

1. A and B^c are independent
2. A^c and B are independent
3. A^c and B^c are independent

Caution: Independence $\neq$ Disjointness!

- Disjoint: $P(A \cap B) = 0$
- Independent: $P(A \cap B) = P(A)P(B)$
- Can only be both if $P(A) = 0$ or $P(B) = 0$

Definition (Independence of Three Events)

Events A, B, C are **independent** if:

$$P(A \cap B) = P(A)P(B)$$
$$P(A \cap C) = P(A)P(C)$$
$$P(B \cap C) = P(B)P(C)$$
$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

Important: Need *all four* conditions!
Pairwise Independence $\neq$ Independence:

- Pairwise: first three conditions hold
- Independence: all four conditions hold

Example: Pairwise Independence $\neq$ Independence

Setup: Two fair, independent coin tosses

- A: First toss is H
- B: Second toss is H
- C: Both tosses have same result

Check pairwise independence:

$$P(A) = P(B) = P(C) = 1/2$$

$$P(A \cap B) = P(HH) = 1/4 = P(A)P(B)$$

$$P(A \cap C) = P(HH) = 1/4 = P(A)P(C)$$

$$P(B \cap C) = P(HH) = 1/4 = P(B)P(C)$$

So A, B, C are pairwise independent.
Check triple independence:

$$P(A \cap B \cap C) = P(HH) = 1/4 \neq P(A)P(B)P(C) = 1/8$$

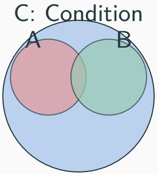
So NOT independent!

Conditional Independence

Definition (Conditional Independence)

Events A and B are **conditionally independent given** E if:

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



$A \perp\!\!\!\perp B \mid C$

Caution: Independence $\neq$ Conditional Independence!

- A, B can be independent but not conditionally independent given E
- A, B can be conditionally independent given E but not independent

Problem-Solving Strategies

Strategy: Condition on What You Wish You Knew

Approach:

1. Identify the unknown quantity
2. Determine what information would make the problem trivial
3. Condition on that information
4. Use Law of Total Probability to combine cases

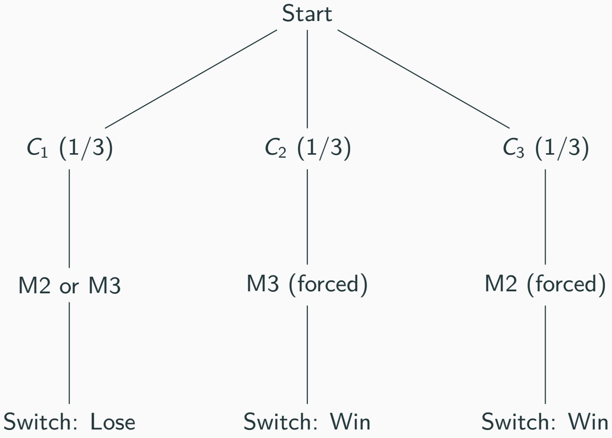
Monty Hall Problem

- 3 doors: 1 car, 2 goats
- You pick door 1
- Monty opens a door with a goat
- Should you switch?

Wish: I wish I knew where the car was!
Solution: $P(\text{win by switching}) = 2/3$

Wikipedia: Monty Hall Problem

Monty Hall Problem: Tree Diagram



Probability: $\frac{2}{3}$ chance of winning by switching

Strategy: First-Step Analysis

Method:

1. Consider the first event/step in a process
2. Condition on possible outcomes of first step
3. Set up recursive equations
4. Solve the system

Amoeba Problem

- Single amoeba
- After 1 minute: die (1/3), stay same (1/3), split into two (1/3)
- All future amoebas behave independently
- $p = P(\text{population eventually dies out})$

First-step analysis: $p = \frac{1}{3} + \frac{1}{3}p + \frac{1}{3}p^2$ **Solution**: $p = 1$ (certain extinction)

Gambler's Ruin: First-Step Analysis

Problem: Two gamblers A and B

- A starts with \$i, B with \$(N-i)
- Each bet: A wins with probability p, loses with probability $q = 1 - p$
- Game ends when someone has all \$N
- $p_i = P(\text{A wins}|\text{starts with } i)$

First-step analysis:

$$p_i = p \cdot p_{i+1} + q \cdot p_{i-1}$$

with boundary conditions $p_0 = 0$, $p_N = 1$.
Solution:

$$p_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)^N} & \text{if } p \neq 1/2 \\ \frac{i}{N} & \text{if } p = 1/2 \end{cases}$$

Genetics Applications

Example: Two Cards Problem (Genetics Version)

Genetic Counseling Scenario:

- Two alleles: N (normal), M (mutant)
- Parent genotype: NM (carrier)
- First child tested: Has mutation (M)
- Question: Probability second child has mutation?

Solution:

$$P(\text{2nd M}|\text{1st M}) = \frac{P(\text{both M})}{P(\text{1st M})} = \frac{(1/2)^2}{1/2} = \frac{1}{2}$$

First child

M	NM	$\boxed{\text{MM}}$
N	NN	NM

Key Insight

Knowing first child's status doesn't change probability for second child (Mendelian inheritance)

Strategy: Condition on What You Wish You Knew (Genetics)

Evolutionary Example: Coalescent Theory

- **Unknown**: Time to most recent common ancestor (TMRCA)
- **Wish I knew**: Number of generations since divergence
- **Condition on**: Population size, mutation rate
- **Result**: $E[T] = \sum_t P(T > t)$ conditioned on demographic history

Population Genetics Example

- Gene frequency in next generation
- Condition on current frequency
- Wright-Fisher model: $X_{t+1}|X_t \sim$ Binomial

Pitfalls and Paradoxes

The Prosecutor's Fallacy

Real Case: Sally Clark (1998) - two infant deaths

- Expert: $P(\text{SIDS}) = 1/8500$
- So $P(\text{two SIDS}) = (1/8500)^2 \approx 1/73$ million
- Conclusion: Probability of innocence is 1 in 73 million

Two Major Errors:

1. Assumed independence (SIDS might run in families)
2. Confused $P(\text{evidence}|\text{innocent})$ with $P(\text{innocent}|\text{evidence})$

Correct by Bayes':

$$P(\text{innocent}|\text{evidence}) = \frac{P(\text{evidence}|\text{innocent})P(\text{innocent})}{P(\text{evidence})}$$

$P(\text{innocent})$ is very high (double infanticide is also rare!)
Wikipedia: Prosecutor's Fallacy

The Defense Attorney's Fallacy

Scenario: Murder case, husband accused

- Evidence: Husband abused wife
- Defense: Only 1 in 10,000 abusive husbands murder their wives
- So $P(\text{guilty}|\text{abuse}) = 0.0001$, evidence irrelevant

Error: Fails to condition on all evidence!
Correct: Should compute $P(\text{guilty}|\text{abuse}, \text{wife murdered})$
Given:

$$\begin{aligned} P(\text{abuse}) &= 0.1 \\ P(\text{guilty}|\text{murder}) &= 0.2 \\ P(\text{abuse}|\text{guilty, murder}) &= 0.5 \\ P(\text{abuse}|\text{innocent, murder}) &= 0.1 \end{aligned}$$

Then:

$$P(\text{guilty}|\text{abuse, murder}) = \frac{0.5 \times 0.2}{0.5 \times 0.2 + 0.1 \times 0.8} = \frac{5}{9} \approx 0.56$$

Evidence is actually very relevant!

Simpson's Paradox

	Drug A		Drug B	
Subpopulation	Success	Total	Success	Total
Severe cases	30/100 (30%)	100	210/700 (30%)	700
Mild cases	90/100 (90%)	100	10/100 (10%)	100
Combined	**120/200 (60%)**	200	**220/800 (27.5%)**	800

Paradox:

- Drug A better for both severe and mild cases
- But Drug B has higher overall success rate

Why? Drug A used more on severe cases (harder to treat)
Wikipedia: Simpson's Paradox

Understanding Simpson's Paradox Mathematically

Let:

- A: Successful treatment
- B: Drug B is used
- C: Severe case

We have:

$$P(A|C,B) < P(A|C,B^c) \quad (30\% < 30\%)$$
$$P(A|C^c,B) < P(A|C^c,B^c) \quad (10\% < 90\%)$$

But:

$$P(A|B) > P(A|B^c) \quad (27.5\% > 60\%)$$

Reason: Different weights:

$$\begin{aligned} P(A|B) &= P(A|C,B)P(C|B) + P(A|C^c,B)P(C^c|B) \\ &= 0.3 \times 0.875 + 0.1 \times 0.125 = 0.275 \\ P(A|B^c) &= P(A|C,B^c)P(C|B^c) + P(A|C^c,B^c)P(C^c|B^c) \\ &= 0.3 \times 0.5 + 0.9 \times 0.5 = 0.6 \end{aligned}$$

Drug B's average weights easy cases more heavily!

Summary and Key Formulas

1. **Definition**: $P(A|B) = \frac{P(A \cap B)}{P(B)}$, $P(B) > 0$
2. **Bayes' Rule**: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
3. **LOTP**: $P(B) = \sum_i P(B|A_i)P(A_i)$ for partition $\{A_i\}$
4. **Extended forms**:
   - $P(A|B, E) = \frac{P(B|A, E)P(A|E)}{P(B|E)}$
   - $P(B|E) = \sum_i P(B|A_i, E)P(A_i|E)$
5. **Independence**: $P(A \cap B) = P(A)P(B)$
6. **Conditional Independence**: $P(A \cap B|E) = P(A|E)P(B|E)$
7. **Odds form of Bayes'**:

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

Problem-Solving Strategies

1. **Condition on what you wish you knew**
   - Monty Hall: Wish you knew where car is
   - Disease testing: Wish you knew true disease status
2. **First-step analysis**
   - Amoeba population: Condition on first minute
   - Gambler's ruin: Condition on first bet
3. **Use Bayes' rule when $P(B|A)$ easier than $P(A|B)$**
4. **Use LOTP to break complex $P(B)$ into simpler conditional pieces**

Common Pitfalls to Avoid:

1. Prosecutor's fallacy: Confusing $P(A|B)$ with $P(B|A)$
2. Defense attorney's fallacy: Not conditioning on all evidence
3. Assuming independence without justification
4. Misinterpreting Simpson's paradox

Sequential Updating of Beliefs

Prior | Posterior 1 | Posterior 2

$$P(A) \xrightarrow{B} P(A|B) \xrightarrow{C} P(A|B,C)$$

Step 1: Update with B $\Rightarrow$ **Step 2: Update with C**

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

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

Key Insight: Each posterior becomes the prior for the next update

Sequential Updating of Beliefs

Key Insight: Today's posterior becomes tomorrow's prior.
Coherency: Sequential updating gives same result as updating with all evidence at once:

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

where $P(A|B)$ is the "new prior" after observing B.

Conditioning is not just a technique

It's the fundamental way we incorporate information

Before Conditioning
Uncertain world

Conditioning Process
Zoom in on evidence

After Conditioning
Refined understanding

Further Reading

- **Textbooks**:
  - Blitzstein & Hwang, *Introduction to Probability*
- **Online Resources**:
  - Wikipedia: Probability
  - Wikipedia: Conditional Probability
  - Wikipedia: Bayes' Theorem
  - Wikipedia: Law of Total Probability
  - Wikipedia: Independence
  - Wikipedia: Monty Hall Problem
  - Wikipedia: Prosecutor's Fallacy
  - Wikipedia: Simpson's Paradox