# Chapter 2: Conditional Probability

Introduction to Probability, 2nd Edition
Blitzstein & Hwang

DRME

December 21, 2025

# Chapter 1: Foundations of Probability

# The Logic of Uncertainty

**Mathematics** is the logic of certainty; **Probability** is the logic of uncertainty.

- It provides a framework for quantifying doubt.
- It allows us to update beliefs as new data is acquired.

Wikipedia: Probability

# The Sample Space ($\Omega$)

The **Sample Space** is the set of all possible outcomes of an experiment.

- Outcomes must be distinct and mutually exclusive.
- An **Event** is a subset of the sample space.

Wikipedia: Sample Space

# The Naive Definition of Probability

If all outcomes in a finite sample space $S$ are **equally likely**:

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

- This relies on **symmetry** (e.g., fair dice).
- **Biohazard:** Do not assume equally likely outcomes without justification.

If an experiment has $r$ stages, with $n_1$ results for stage 1, $n_2$ for stage 2, etc.:

$$\text{Total Outcomes} = n_1 \times n_2 \times \cdots \times n_r$$

- This is the basis for most counting methods.

# Counting: Combinations

A **Combination** is a choice of $k$ elements from $n$ where **order does not matter**.

- Formula: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
- This is also called the **Binomial Coefficient**.

# Chapter 2: Conditional Probability

# Conditioning: The Soul of Statistics

**Key Idea**

All probabilities are conditional on background knowledge.
Conditional probability tells us how to update beliefs when we get new information.

**Example: Weather Forecast**

- Initial: $P(\text{Rain}) = 0.3$ (prior probability)
- See dark clouds: $P(\text{Rain}|\text{Dark Clouds}) = 0.8$ (posterior)
- Later: $P(\text{Rain}|\text{Dark Clouds, Barometer Falling}) = 0.95$

**Two Roles of Conditioning**:

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

Wikipedia: Conditional Probability

# Formal Definition of Conditional Probability

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

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

**Terminology**:

- $P(A)$: Prior probability (before evidence)
- $P(A|B)$: Posterior probability (after evidence)
- $B$: The evidence or conditioning event

**Example**: $P(A|A) = \frac{P(A \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$
Knowing $A$ occurred updates its probability to 1.

# Why Condition?

Conditional probability is the soul of statistics.

- It answers: How should we **update our beliefs** in light of evidence?
- All probabilities are technically conditional on background knowledge $K$.

## Example: Two Cards Problem

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

- $A$: First card is a heart
- $B$: Second card is red

Find $P(A|B)$ and $P(B|A)$.

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

$$P(A \cap B) = \frac{13 \times 25}{52 \times 51} = \frac{325}{2652} = \frac{25}{204} \approx 0.1225$$

(13 hearts for first card × 25 red cards remaining)

# Example: Two Cards Problem (Continued)

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

$$P(A) = \frac{13}{52} = \frac{1}{4} = 0.25$$
$$P(B) = \frac{1}{2} = 0.5 \quad \text{(by symmetry)}$$

Step 3: Apply the definition

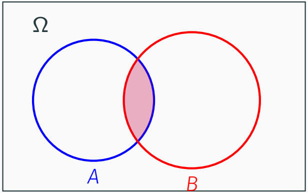
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{25/204}{1/2} = \frac{25}{102} \approx 0.2451$$

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

Key Observations:

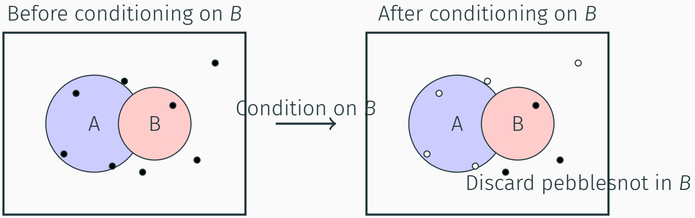
1. $P(A|B) \neq P(B|A)$ (don't confuse them!)
2. $P(B|A)$ can be found directly: if first is heart, 25 red remain out of 51

# Visualization: Pebble World




- Learning $B$ occurred **removes** all outcomes in $B^c$.
- We **renormalize** the mass of $B$ to equal 1.

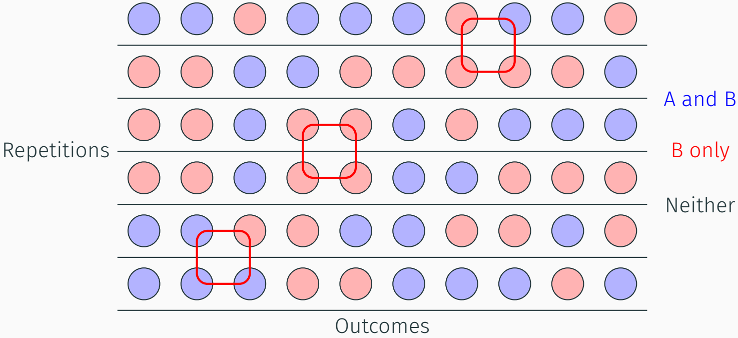
# Intuition 1: Pebble World Visualization




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

## Intuition 2: Frequentist Interpretation




**Interpretation**:

$$P(A|B) \approx \frac{\text{\# times both A and B occurred}}{\text{\# times B occurred}}$$

- Run experiment many times
- Circle repetitions where $B$ occurred
- Among circled, count how often $A$ also occurred

# The Two Children Problem

**Classic Puzzle** (Martin Gardner, 1950s):

1. Mr. Jones has two children. The **older** child is a girl. What is $P(\text{both girls})$?
2. Mr. Smith has two children. **At least one** is a boy. What is $P(\text{both boys})$?

**Assumptions**:

- Binary gender (boy/girl)
- $P(\text{boy}) = P(\text{girl}) = 1/2$ for each child
- Independent genders

**Sample space**: $\{GG, GB, BG, BB\}$, each with probability 1/4

# Solving the Two Children Problem

Case 1: Older child is a girl

$$P(\text{both girls}|\text{elder is girl}) = \frac{P(\text{both girls, elder is girl})}{P(\text{elder is girl})}$$

Step-by-step:

1. $P(\text{both girls, elder is girl}) = P(GG) = 1/4$
2. $P(\text{elder is girl}) = P(GG \text{ or } GB) = 1/4 + 1/4 = 1/2$
3. $P(\text{both girls}|\text{elder is girl}) = \frac{1/4}{1/2} = \frac{1}{2}$

Case 2: At least one is a boy

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

Step-by-step:

1. $P(\text{both boys, at least one boy}) = P(BB) = 1/4$
2. $P(\text{at least one boy}) = P(GB, BG, BB) = 3/4$
3. $P(\text{both boys}|\text{at least one boy}) = \frac{1/4}{3/4} = \frac{1}{3}$

# Why Are the Answers Different?

GG GB BG BB

GG GB BG BB

Case 1: Keep only GG and GB

GG GB BG BB

Case 2: Keep GB, BG, BB

Key Insight:

- ”Elder is girl” designates a specific child → eliminates BG and BB
- ”At least one boy” doesn’t specify which child → only eliminates GG
- The remaining sample spaces have different sizes and compositions

# The Multiplication Rule for Probabilities

For any events $A$ and $B$:

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

- This allows us to calculate joint probabilities using conditional ones.

# 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:

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

**Why useful?** Often $P(B|A)$ is easier to compute than $P(A|B)$.

Wikipedia: Bayes' Theorem

# The Law of Total Probability (LOTP)

If $A_1, \ldots, A_n$ partition the sample space:

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

- This is a **divide-and-conquer** strategy.

Wikipedia: Law of Total Probability

# Law of Total Probability: Statement and Proof

**Theorem (Law of Total Probability (LOTP))**

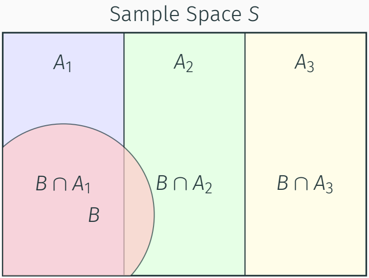
*Let $A_1, A_2, \ldots, A_n$ partition $S$ (disjoint, union = $S$, $P(A_i) > 0$). Then:*

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

Proof:

$$\begin{aligned}
P(B) &= P(B \cap S) \\
&= P\left(B \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) \\
&= P\left(\bigcup_{i=1}^{n} (B \cap A_i)\right) \\
&= \sum_{i=1}^{n} P(B \cap A_i) \quad \text{(disjointness)} \\
&= \sum_{i=1}^{n} P(B|A_i)P(A_i) \quad \text{(definition)}
\end{aligned}$$

# Visualizing the Law of Total Probability




**Interpretation**:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3)$$

$$= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$$

**Key Idea**: Chop $B$ into pieces using a partition, then add up the pieces.

# Example: Random Coin Problem

**Problem**: You have:

- Fair coin: $P(H) = 1/2$
- Biased coin: $P(H) = 3/4$

Pick one at random (50-50), flip it 3 times, get HHH. What is $P(\text{fair}|\text{HHH})$?

**Step 1: Define events**

- $F$: Coin is fair
- $F^c$: Coin is biased
- $A$: Get HHH (3 heads)

**Step 2: Known probabilities**

$$P(F) = 1/2, \quad P(F^c) = 1/2$$
$$P(A|F) = (1/2)^3 = 1/8$$
$$P(A|F^c) = (3/4)^3 = 27/64$$

# Example: Random Coin Problem (Continued)

Step 3: Apply Bayes' Rule with LOTP

We want $P(F|A)$:

$$\begin{aligned}
P(F|A) &= \frac{P(A|F)P(F)}{P(A)} \quad \text{(Bayes')} \\
&= \frac{P(A|F)P(F)}{P(A|F)P(F) + P(A|F^c)P(F^c)} \quad \text{(LOTP)} \\
&= \frac{(1/8)(1/2)}{(1/8)(1/2) + (27/64)(1/2)} \\
&= \frac{1/16}{1/16 + 27/128} \\
&= \frac{1/16}{1/16 + 27/128} = \frac{8}{8+27} = \frac{8}{35} \approx 0.2286
\end{aligned}$$

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

# Example: Rare Disease Testing

**Problem**: Disease affects 1% of population. Test is 95% accurate:

- Sensitivity: $P(\text{test}+|\text{disease}) = 0.95$
- Specificity: $P(\text{test}-|\text{no disease}) = 0.95$

If Fred tests positive, what is $P(\text{disease}|\text{test}+)$?

**Step 1: Define events**

- $D$: Has disease
- $T$: Tests positive

**Step 2: Known probabilities**

$$P(D) = 0.01, \quad P(D^c) = 0.99$$
$$P(T|D) = 0.95, \quad P(T|D^c) = 1 - 0.95 = 0.05$$

# Example: Rare Disease Testing (Continued)

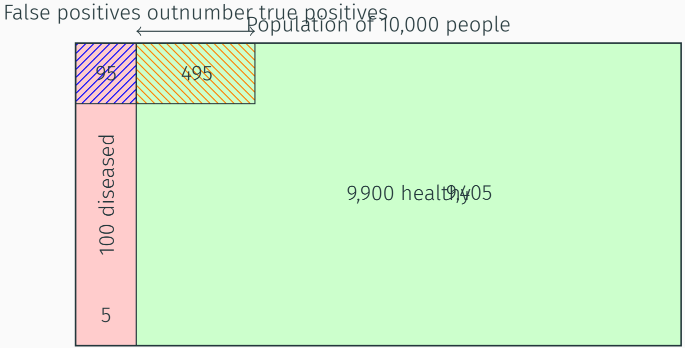
Step 3: Apply Bayes' Rule

$$
\begin{aligned}
P(D|T) &= \frac{P(T|D)P(D)}{P(T)} \\
&= \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} \\
&= \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.05 \times 0.99} \\
&= \frac{0.0095}{0.0095 + 0.0495} \\
&= \frac{0.0095}{0.0590} \approx 0.161
\end{aligned}
$$

Surprising Result: Only 16% chance of having disease despite positive test!

Why? Disease is rare, so false positives (5% of 99% healthy people) outnumber true positives.

# Visualizing the Rare Disease Problem




**Key Insight**: With rare diseases, test specificity (true negative rate) is crucial!

# Conditional Probability Satisfies Probability Axioms

**Key Insight**: For fixed $E$ with $P(E) > 0$, the function $\tilde{P}(A) = P(A|E)$ satisfies all probability axioms:

**Axiom 1**:

$$\tilde{P}(\emptyset) = P(\emptyset|E) = 0, \quad \tilde{P}(S) = P(S|E) = 1$$

**Axiom 2** (Countable 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**:

$$\begin{aligned}
\tilde{P}\left(\bigcup_{i=1}^{\infty} A_i\right) &= \frac{P\left((\bigcup_i A_i) \cap E\right)}{P(E)} \\
&= \frac{P\left(\bigcup_i (A_i \cap E)\right)}{P(E)} \\
&= \frac{\sum_i P(A_i \cap E)}{P(E)} = \sum_i \tilde{P}(A_i)
\end{aligned}$$

# All Probabilities are Conditional

**Philosophical Insight**

There is always background knowledge $K$ being conditioned on.

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

- Implicitly conditions on: location, season, time of day, etc.
- Actually means: $P(\text{Rain}|\text{today is Nov 1 in Boston, etc.})$

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

Implication: Any result about unconditional probability also holds for conditional probability (just add ”given $E$” everywhere).

# Extended Bayes' Rule and LOTP

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 $A_1, \ldots, A_n$.

Example: Random coin with extra evidence

- After seeing HHH, flip coin again
- $P(\text{4th H}|\text{first 3 H}) = P(\text{H}|F, \text{HHH})P(F|\text{HHH}) + P(\text{H}|F^c, \text{HHH})P(F^c|\text{HHH})$
- $\approx (1/2)(0.23) + (3/4)(0.77) \approx 0.69$

# 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 $P(B) > 0$, this is 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.

**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 $P(B) = 0$

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

**Proof for $A$ and $B^c$**:

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(B^c) \end{aligned}$$

**Intuition**: If $A$ gives no information about $B$, it also gives no information about $B^c$.

# Independence of Three Events

**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**:

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

but

$$P(A)P(B)P(C) = (1/2)^3 = 1/8$$

So NOT independent!

**Definition (Conditional Independence)**

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

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

**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
- $A, B$ can be conditionally independent given $E$ but not given $E^c$

**Example**: Two coin types, unknown which we have

- Conditional on coin type: tosses are independent
- Unconditionally: tosses are dependent (earlier tosses inform about coin type)

# Strategy: Condition on What You Wish You Knew

**Monty Hall Problem**:

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

**Wish**: I wish I knew where the car was!

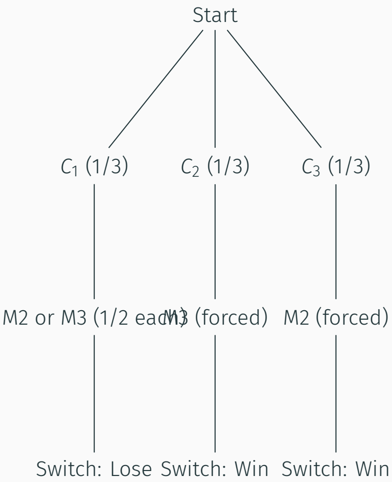
**Solution**: Condition on car location

$$
\begin{aligned}
P(\text{win by switching}) &= P(\text{win}|C_2)P(C_2) + P(\text{win}|C_3)P(C_3) \\
&= 1 \times \frac{1}{3} + 1 \times \frac{1}{3} = \frac{2}{3}
\end{aligned}
$$

(If car behind door 2 or 3, switching wins; if behind door 1, switching loses)

Wikipedia: Monty Hall Problem

# Monty Hall Problem: Tree Diagram




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

# Strategy: First-Step Analysis

Branching Process (Amoeba Problem):

- Single amoeba
- After 1 minute: die (prob 1/3), stay same (1/3), split into two (1/3)
- All future amoebas behave independently the same way
- What is $P$(population eventually dies out)?

Let $p = P(\text{extinction})$. Condition on first step:

$$\begin{aligned} p &= P(\text{extinction}|\text{die}) \times \frac{1}{3} \\ &\quad + P(\text{extinction}|\text{same}) \times \frac{1}{3} \\ &\quad + P(\text{extinction}|\text{split}) \times \frac{1}{3} \\ &= 1 \times \frac{1}{3} + p \times \frac{1}{3} + p^2 \times \frac{1}{3} \end{aligned}$$

(If splits, need both lineages to die out: probability $p \times p = p^2$)

# Solving the Amoeba Problem

From first-step analysis:

$$p = \frac{1}{3} + \frac{1}{3}p + \frac{1}{3}p^2$$

Multiply by 3:

$$3p = 1 + p + p^2$$

Rearrange:

$$p^2 - 2p + 1 = 0$$

Factor:

$$(p-1)^2 = 0$$

Solution:

$$p = 1$$

**Conclusion**: Population dies out with probability 1 (certain extinction).

**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
- What is $p_i = P(\text{A wins}|\text{starts with } i)$?

**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**:

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

# 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

| | Dr. Hibbert | | Dr. Nick | | |
|---|---|---|---|---|---|
| | Heart | Band-Aid | Heart | Band-Aid | Overall |
| Success | 70 | 10 | 2 | 81 | |
| Total | 90 | 10 | 10 | 90 | |
| Rate | 77.8% | 100% | 20% | 90% | |
| Overall | 80/100 = 80% | | 83/100 = 83% | | |

**Paradox**:

- Dr. Hibbert better at heart surgeries (78% vs 20%)
- Dr. Hibbert better at Band-Aid removal (100% vs 90%)
- But Dr. Nick has higher overall success rate (83% vs 80%)

**Why?** Dr. Hibbert does more risky heart surgeries (90%) vs Dr. Nick (10%).

Wikipedia: Simpson's Paradox

# Understanding Simpson's Paradox Mathematically

Let:

- $A$: Successful surgery
- $B$: Dr. Nick is surgeon
- $C$: Heart surgery (vs Band-Aid)

We have:

$$P(A|C,B) < P(A|C,B^c) \quad \text{(20\% < 78\%)}$$
$$P(A|C^c,B) < P(A|C^c,B^c) \quad \text{(90\% < 100\%)}$$

But:

$$P(A|B) > P(A|B^c) \quad \text{(83\% > 80\%)}$$

**Reason**: Different weights in weighted averages:

$$\begin{aligned} P(A|B) &= P(A|C,B)P(C|B) + P(A|C^c,B)P(C^c|B) \\ &= 0.2 \times 0.1 + 0.9 \times 0.9 = 0.83 \\ P(A|B^c) &= P(A|C,B^c)P(C|B^c) + P(A|C^c,B^c)P(C^c|B^c) \\ &= 0.778 \times 0.9 + 1.0 \times 0.1 = 0.80 \end{aligned}$$

Dr. Nick's average weights easy surgeries more heavily!

# Recap and Summary

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'**:

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

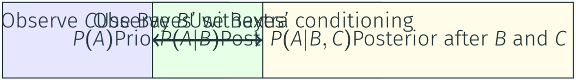
# Problem-Solving Strategies

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




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 $B$.

# Questions?

Complete R code for simulations available at
`http://stat110.net`